UAB

Universitat Autònoma
de Barcelona

Dipòsit digital
de documents
de la UAB

This is the **accepted version** of the article:

This version is avaible at https://ddd.uab.cat/record/218310

A new approach for optimal discretization of plant functional types in a

process-based ecosystem model with forest management: a case study of

temperate conifers.

Marc Peaucelle[1], Valentin Bellassen[2], Philippe Ciais[1], Josep Peñuelas[3,4], Nicolas Viovy[1]

[1] Laboratoire des Sciences du Climat et de l'Environnement, CEA CNRS UVSQ, Gif-sur-Yvette, France

[2]Centre d'Economie et de Sociologie appliquées à l'Agriculture et aux Espaces Ruraux, INRA, Dijon, France

[3]CREAF, Cerdanyola del Vallès, Barcelona 08193, Catalonia, Spain

[4]CSIC, Global Ecology Unit CREAF -CSIC-UAB, Cerdanyola del Vallès, Barcelona 11 08193, Catalonia, Spain

Marc Peaucelle: marc.peaucelle@lsce.ipsl.fr (corresponding author)

Valentin Bellassen: valentin.bellassen@dijon.inra.fr

Philippe Ciais: philippe.ciais@lsce.ipsl.fr

JosepPeñuelas: josep.penuelas@uab.cat

Nicolas Viovy: nicolas.viovy@lsce.ipsl.fr

Running title: Defining the optimal number of PFTs using a trait-based approach.

Key words: PFT classification, conifers, forest management, functional traits, hierarchical classification.

Word Counts*: Abstract – 289, Main body – 5263, References – 1631, Biosketch – 57, Figures – 368, Tables – 90.

References: 42

Figures: 5

Tables: 3

30 **ABSTRACT**

31 **Aim**

32 Dynamic global vegetation models (DGVM) use a discretization of forest vegetation based on

33 plant functional types (PFT). The physiological and ecological parameters used to model a

34 given PFT are usually fixed, being defined from point-based observations while model

35 applications are often grid-based. This rigid approach is causing spatial biases in the results of

36 DGVM-simulated productivity and biomass-related variables. We aim to overcome this

37 limitation with a new approach that uses a hierarchical classification of forest PFT parameters

38 from traits retrieved from the literature and from the TRY global database of plant traits. This

39 approach is applied to temperate conifers in the DGVM ORCHIDEE-FM, which has previously

40 been shown to produce systematic biases in the simulation of biomass and biomass increments.

41 **Location**

42 Temperate coniferous forests; France

43 **Time period**

44 2005-2012

45 **Major taxa studied**

46 Gymnosperms

47 **Methods**

48 The five major coniferous species in France, *Abies alba*, *Picea abies*, *Pinus pinaster*, *Pinus*

49 *sylvestris* and *Pseudotsuga menziesii*, were grouped objectively into PFTs within the

50 ORCHIDEE-FM DGVM using a hierarchical classification based on 12 key attributes related

51 to photosynthesis, phenology and allometric relationships.

Results

We show that the single PFT covering all temperate coniferous forest that was used by default in ORCHIDEE-FM could be replaced by two representative sub-categories defined by grouping species-level data, without necessarily having to adopt a set of parameters for each species. The definition of new temperate conifer PFTs with this approach allows us to reduce the spatial heterogeneity by 40% in average in model-measurement misfit for stand volume, growth and stand density at the regional scale.

Main conclusions

The proposed approach to improve the representation of plant functional types in DGVMs, while keeping the number of different PFTs manageable, is promising for application to regions where a single PFT can correspond to a number of different species.

INTRODUCTION

Dynamic global vegetation models (DGVMs) simulate the biogeochemical and hydrological processes involved in ecosystem functioning. They model the controls on the fluxes of carbon, energy and water from local up to regional and global scale. DGVMs also simulate the response of processes and state variables to environmental drivers. Most of these models use a discretized representation of vegetation using Plant Functional Types (PFT), with generally a dozen different PFTs representing the world's vegetation. By definition, each PFT represents a "mean" type of vegetation defined by a set of traits (translated to model parameters). While the use of a limited number of PFTs is practical to for grid-based simulations at large scale, this rather rigid representation of the vegetation does not allow a robust analysis of regional gradients in productivity and biomass, and leads to over-simplify processes controlling vegetation dynamics. It is therefore a source of uncertainty in model predictions (Alton, 2011).

The research question of how to deal with the uncertainties resulting from the discrete and rigid representation of PFTs in DGVMs has been addressed by increasing the number of PFTs, or by introducing a continuous spatial variation of vegetation parameters, e.g. using plant trait datasets (Kleidon & Mooney, 2000; Reu *et al.*, 2011; Van Bodegom *et al.*, 2012; Wullschleger *et al.*, 2014; Verheijen *et al.*, 2015). A continuous parameterization of parameters allows spatial variability of parameters within a given PFT (Van Bodegom *et al.,* 2012; Pavlick *et al.,* 2012; Verheijen *et al.,* 2013). However, this method needs a large amount of data for model calibration. Although plant trait databases are building up an increasingly large bank of relevant data, the quality and spatial representativeness of these data might not yet be sufficient to derive robust relationships at the global scale, and even less so at regional scale, such as in our case for temperate coniferous forest in Europe.

In this context, increasing the number of PFTs seems a reasonable compromise. Naudt *et al.* (2015) defined one PFT by dominant European tree species. This strategy is relevant for forests

4

88 in Europe where only a few dominant tree species prevail, but it is impractical in higher

89 biodiversity systems such as grasslands or tropical forests, in which intermediate levels of PFT

90 representation are needed.

91 Past studies mostly focused on biogeographical or climatic envelopes to define new PFTs for

92 DGVMs (Nemani & Running, 1996; Bonan *et al.,* 2002; Harrison *et al.,* 2010) and often sought

93 to find an optimal and/or minimal subset of traits to give the best classification (Pillar, 1999;

94 Pausas & Lavorel, 2003; Pillar & Sosinski, 2003; Bernhardt-Römermann *et al.,* 2008).

95 Boulangeat *et al.* (2012) already argued for the necessity to improve PFTs in DGVMs by

96 regrouping representative species with a classification based on key biological characteristics.

97 Here, we propose an approach to refine the PFT representation in typical DGVMs through the

98 case study of improving the representation of the spatial variability of the parameters of the

99 PFT "temperate needleleaf forest" (TNF) across Europe, using the ORCHIDEE-FM DGVM.

100 The ORCHIDEE-FM DGVM ("ORganising Carbon and Hydrology In Dynamic Ecosystems –

101 Forest Management"; Krinner *et al.,* 2005; Bellassen *et al.,* 2010) simulates the growth, and the

102 competition between individuals leading to density-driven mortality within a forest stand, and

103 optionally an idealized management practice of plantation forests, assumed to be even-aged.

104 The results from this DGVM was evaluated for temperate broadleaf and needleleaf forest in

105 Europe against, among other data sources, biometric data obtained from numerous plots from

106 national forest inventories by Bellassen *et al.,* (2011). Although simulations of biomass and

107 growth were satisfactory for broadleaf forests at the regional scale, the model was not able to

108 reproduce regional patterns for temperate needleleaf forests. Bellassen *et al.* (2011) found

109 regional biases in conifer forests standing volume and growth as a function of forest age, which

110 could not be corrected by simply optimizing parameters, unlike for broadleaf forests

111 (Supplementary Fig.S1). They concluded that the main reason for this model systematic error

112 was the use of a unique conifer PFT associated with a single set of parameters. This suggests

113   that the model suffers more weakness due to its too coarse discretization of vegetation, than

114   from its internal structure and parameter values.

115   Therefore, in this study, we seek to establish if TNF species can be objectively sub-divided into

116   sub-groups with coherent set of parameters for inclusion in the ORCHIDEE FM DGVM. We

117   search for coherent parameters related to eco-physiological attributes, namely photosynthesis,

118   phenology, allometry and differences in management practice.

119   **METHODS**

120   We proposed a new approach for improving the representation of PFTs in ORCHIDEE-FM

121   based on literature review and field data analysis. This method combines three steps:

122   **1)** We first identified 12 key *sensitive* eco-physiological parameters, i.e., parameters that

123   have important influence on the model output. These parameters can be related to

124   photosynthesis, phenology, allometry and management parameters for TNF forests. The

125   definition of *sensitive* is based on previous work that optimized parameters against net

126   ecosystem exchange (NEE) CO_2 and evaporation flux measurements at TNF sites with

127   eddy-covariance technique.

128   **2)** Next, we retrieved a range of values for each sensitive parameter obtained in step-1,

129   based on a literature review and the TRY global trait database (Kattge *et al.,* 2011). This

130   search was restricted to the five most abundant TNF species present in France (Vallet

131   *et al.,* 2006): Silver fir (*Abies alba),* Norway spruce (*Picea abies*), Maritime pine (*Pinus*

132   *pinaster*), Scots pine (*Pinus sylvestris*) and Douglas fir (*Pseudotsuga menziesii*).

133   **3)** Based on the results of 1) and 2), we showed that TNF species can be objectively

134   regrouped using a hierarchical classification into two homogeneous clusters that have

135   parameter similarities. Implemented in the ORCHIDEE model, we showed that the new

136   PFTs improve the spatial representation of conifer stands. The consistency between

137   PFTs based on parameter values from the literature and PFTs used by ORCHIDEE was

138   assessed by comparing the parameter values of each cluster defined with independent

139   photosynthesis and phenological data from the literature to the values obtained from the

140   calibration of ORCHIDEE using eddy covariance (EC) evaporation and NEE

141   measurements (Santaren *et al.,* 2007; Kuppel *et al.,* 2014).

142 **Description of the ORCHIDEE-FM model**

143 ORCHIDEE is a process-based vegetation model that can be run point-based or grid-based

144 (Krinner *et al.,* 2005). In the standard version of ORCHIDEE (v.1.9.6) an average forest is

145 simulated as a mature stand in near equilibrium between growth and mortality. For given

146 vegetation, soil type and climatic condition, ORCHIDEE simulates the carbon, water and

147 energy budgets on a half-hourly time step (Fig. 1).

148 Bellassen *et al.* (2010) added a forest management module (FM; Fig. 1) to ORCHIDEE

149 (v.1.9.6) to simulate self-thinning and management. The FM module replaces an "average tree"

150 representation with a statistically distributed population of individual trees and includes rules

151 for both natural mortality and human interventions. This forest management version of the

152 model is called ORCHIDEE-FM (v.1.9.6-rev.30). It simulates carbon stocks and detailed stand

153 structure of forests of varying ages (Bellassen *et al.,* 2011) which can be directly compared with

154 forest inventory data or ecological site biometric observations. The FM equations are fully

155 described by Bellassen *et al.* (2010).

156

157 **Selection of *sensitive* eco-physiological parameters**

158 The starting point for improving the description of TNF forests is the determination of the

159 variability of so called sensitive or influential eco-physiological parameters among the TNF

160 species. Among all the parameters of ORCHIDEE-FM, we defined as influential those whose

161 variations have a significant effect on the modelled stand biomass. The selection of influential

162 parameters was carried out by comparing information from previous sensitivity tests made by

163 Bellassen *et al.* (2010) to the results of an optimization of ORCHIDEE eco-physiological

164 parameters made by Kuppel (2012) using eddy-covariance measurements. Both methods allow

165 the identification of sensitive parameters, given different "target" state variables. Bellassen *et*

166 *al.* (2010) identified "slow" parameters that control stand density and biomass dynamics on

167 time scales of decades (hereafter B-parameters), while Kuppel (2012) identified mainly "fast"

168 parameters that control physiological processes controlling fluxes (hereafter P-parameters), on

169 time scales of days to seasons. Both fast and slow parameters interact with each other in

170 determining ecosystem carbon state variables in the model equations (Fig. 1). This interaction

171 is non-symmetrical, however. Slow parameters controlling long-term forest growth and carbon

172 stocks have a small impact on hourly to seasonal fluxes, whereas fast parameters do influence

173 decadal net primary productivity (NPP) and thus a change from their default value is expected

174 to modify the decadal stand-scale growth, competition and mortality, and thus the long-term

175 changes of biomass.

176 The method used to select each parameter is fully described in Appendix S2 (see Table 1 for

177 parameters names). The results are summarized here. Three B-parameters were selected as

178 being influential: allometric coefficients that govern the tree biomass-to-circumference ratios

179 (b_{bc}), self-thinning parameters ($ln(\alpha_{st})$ and β_{st}) and the threshold of stand density at which forest

180 management thinning is set to occur (rdi_target) for managed forests. Six P-parameters were

181 found influential: the maximum leaf-level limited carboxylation rate normalized at 25°C

182 (V_{cmax_opt}), the optimal temperature for photosynthesis (T_{opt}), the slope of the Ball-Berry linear

183 relationship between stomatal conductance (g_{slope}) and net assimilation rate, the specific leaf

184 area (SLA) and the critical leaf age for leaf senescence ($L_{age,crit}$).

185 Processes in which the different parameters identified in this study are involved are illustrated

186 in Fig. 1 and listed in Table 1. The default values of these parameters for the TNF PFT in

187 ORCHIDEE are as defined by Krinner *et al.* (2005) and by Bellassen *et al.* (2010) and listed in

188 Table 2. Optimized values obtained from inversion of the fast parameters against EC

189 measurements are from Santaren *et al.* (2007) and Kuppel *et al.* (2014).

Literature review of key parameters

191 We gathered species-specific data for the values of each sensitive parameter listed in Table 1.

192 A search of the literature (up to 2012) was conducted for the predominant TNF species found

193 in France. In addition of this literature search, a number of values were obtained through the

194 TRY database (Kattge *et al.,* 2011), especially for *SLA*. The number of values found for each

195 parameter is indicated in Fig. 2 and Table 2. The literature was first searched for each species

196 and each parameter; the search was then restricted to temperate regions (Köppen classification;

197 Peel *et al.*, 2007). TRY values that were also in the literature were eliminated to avoid

198 duplication; and when multiple measurements were available at a given site, the median value

199 of the reported values was used. Care was taken to ensure the physiological meaning of the data

200 found in the TRY database corresponded to the model parameter definitions. The use of mean

201 values reduced the size of the dataset and prevented any bias from over-sampling a few better-

202 measured sites.

203 For all parameters identified as influential, we tried to minimize the uncertainty due to different

204 measurement methods by filtering or correcting data whenever possible. Values were not

205 corrected according to environmental variability, nor species provenance or age. All references

206 were searched for each parameter described in the following sections and a detailed description

207 of the methodology used to homogenize values for each parameter can be found in Appendices

208 B and C in the Supporting Information.

209 Even though it is a *sensitive* parameter according to Kuppel *et al.* (2012), the number of data

210 reviewed for g_{slope} was insufficient to perform a separate analysis per species within the TNF

211 PFT (Fig. 2). We thus decided to extract another parameter from the literature, the maximal

212 stomatal conductance (g_{smax} – even though it is not directly used in ORCHIDEE). We defined

213 g_{smax} as the maximum conductance measured at one site at any measurement period. The model

214 is sensitive to the value of g_{slope}, but because we have no evidence that g_{smax} is also a key

215 parameter, the following analyses were performed either including or excluding g_{smax}.

216

Clustering of species based on parameters values

218 For all parameters, except *rdi_target*, T_{opt} (see Table 1 for definitions) and g_{slope}, because of

219 insufficient data, the inter-species and intra-species parameter variabilities were calculated

220 using Kruskal–Wallis one-way analysis of variance and the Mahalanobis distance (distance

221 between species taking into account the parameter variance; Mahalanobis, 1936):

$$D(x) = \sqrt{(x - \mu)^T \sum^{-1} (x - \mu)} \tag{1}$$

222 where x is a vector of data (here different parameters) with p columns, μ the mean of the

223 distribution of length p, $\sum$ the covariance matrix ($p*p$) of the distribution. These methods have

224 the advantage of providing a statistical test of similarity. Their drawback is the impossibility of

225 obtaining a unique set of species clusters based on the variability of all the parameters.

226 As all parameters are quantitative, we used a clustering algorithm to objectively define

227 homogeneous groups of species based on Ward's method (1963). In this method, all parameter

228 median values are centred and scaled to calculate dissimilarity (distance, eq.2).

$$D(C1, C2) = \sqrt{\frac{n1n2}{n1 + n2}} d(\mu1, \mu2) \qquad (2)$$

229 Where $D(C_1, C_2)$ is the distance between clusters C_1 and C_2, n_1 and n_2 the number of data, and

230 $d(\mu_1, \mu_2)$ the distance between their respective centroids.

231 Data treatment was entirely performed with R software (R Team 2011).

ORCHIDEE-FM Simulations

233 First, we analysed previous regional simulations from Bellassen *et al.* (2011) in the light of the

234 new PFT classification by comparing model/measurement errors from Bellassen *et al.,* (2011)

235 to the distribution of the newly defined PFTs.

236 In addition, four sets of simulations were performed to assess the gain (or loss) of information

237 obtained by different sub-divisions of TNF species, each based on coherent parameter values.

238 The first set (REF) uses the default parameter values and processes of ORCHIDEE-FM as

239 described by Bellassen *et al.* (2010), that is parameters are the same for all TNF forests. The

240 second configuration (S1) use one PFT to represent TNF forest but this PFT is parameterized

241 according to results of the literature review. The third set (S2) was performed using the new

242 PFTs corresponding to species groups obtained from the hierarchical clustering of parameters.

243 The fourth simulation (S3) was performed with the parameters of each TNF species (5 in total),

244 each species being modeled as a separate PFT. All the simulations are evaluated against the

245 French NFI observations interpolated at a 0.25° spatial resolution. We simulated 50 years old

246 stands. To minimize bias between simulations and observations, we averaged the simulation

247 outputs over a 10 year period (between years 45 and 55) and only trees with a diameter higher

248 than 7.5 cm were used.

249 The same protocol and pedo-climatic inputs as in Bellassen *et al.* (2011) were used for all the

250 simulations. No historical change in species distribution or forest area was taken into account

251   in the simulations. The performance of each simulation was assessed against observations by

252   using the Spearman's rank correlation coefficient for three different variables: the average stand

253   aboveground volume, volume increment and stand density.

254

RESULTS

Synthesis of parameter values

257   We found differences in the quantity of data collected for the different parameters. For instance,

258   336 values were collected for *SLA* while only one value per species could be obtained for T_{opt}

259   and *rdi_target* (Table 3). There is also a range in the number of data available per species. The

260   values for *Pinus sylvestris* and *Picea abies* are more numerous than for *Abies alba*, *Pinus*

261   *pinaster* and *Pseudotsuga menziesii* in particular for V_{cmax}, *SLA* and g_{smax}. The mean coefficient

262   of variation (CV) for parameters V_{cmax}, *SLA and* g_{smax} are relatively high (0.35, 0.43 and 0.71,

263   respectively). This spread of parameter distributions could be the result of environmental

264   conditions, bias in measurement method or representation.

Differences between species

266   Species was found to be a statistically significant explanatory variable for differences in P-

267   parameters. According to the Kruskal-Wallis one-way analysis of variance, the values of

268   V_{cmax25}, J_{max}/V_{cmax}, $L_{age,crit}$ and *SLA* were found to be significantly different between species (p-

269   value<0.05). Variance analysis shows a significant inter-species variability for most P-

270   parameters, whereas B-parameter differences are statistically not significant between species.

271   For parameters V_{cmax}, J_{max}/V_{cmax}, *SLA and Longevity*, the Mahalanobis distance between species

272   is found to be higher than for g_{smax} (Fig.S2 in Appendix S1). However for B-parameters, the

273   distance between species is relatively lower than for P-parameters. All *Pinus* species fall into

12

274 the same group for P-parameters, which supports grouping these species into one group for
275 modelling carbon stocks and fluxes.

276 **Differences in parameters between literature values and eddy covariance optimized**
277 **values**

278 We observe discrepancies between literature values and values optimized from eddy covariance
279 data for several parameters. V_{cmax25} values observed in the literature for *Pinus pinaster* and *Pinus*
280 *sylvestris* (90.31 and 106.61µmol m^{-2} s^{-1}) are higher than optimized values obtained by Kuppel
281 *et al.* (2014) from 20 conifer sites (viz., 41.7 and 31.9 µmol m^{-2} s^{-1}; Table 3). By contrast, for
282 non-pine species, V_{cmax25} estimated from literature data (at leaf scale) are found to be within the
283 range of the results from Kuppel *et al.* (2014). For T_{opt}, only values obtained for *Pinus sylvestris*
284 are close to the values optimized by Kuppel *et al.* (2014), but the median value considering all
285 TNF species is twofold higher in the literature data than in Kuppel *et al.* (2014) (32.3°C
286 compared to 17.5°C). *SLA* literature values are half those of Kuppel *et al.* (2014) (10 mm² mg^{-1}
287 ^{1}), excepted for *Pseudotsuga menziesii* for which we found an *SLA* of 7.08 mm² mg^{-1}. For
288 $L_{age,crit}$, only the literature value for *Pinus sylvestris* (34.0 months) is close to Kuppel *et al.*
289 (2014) (35.6 months). Too few data were available to compare g_{slope} for all species, however
290 the single literature value reviewed for *Pinus pinaster* (10.05) is in the range of the inversion
291 results of Kuppel *et al.* (2014) (8.8) and Santaren *et al.* (2007) (8.1). The values of g_{slope} obtained
292 from the literature for *Picea abies* (5.94) and *Pinus sylvestris* (5.59) are smaller than values
293 found by the Bayesian calibration of Kuppel *et al.* (2014). Results from the Santaren *et al.*
294 (2007) and Kuppel *et al.* (2014) optimizations are summarized in Table 3.

Differences between literature and values derived from forest inventory data for

B-parameters

In addition to the literature search, B-parameters linked to biomass dynamics (b_{bc}, $ln(\alpha_{st})$ and β_{st}) were re-determined by fitting to the French national forest inventory (NFI) data (IFN, 2012; refer to Appendix S2 for a complete description). Fitted values of b_{bc} are from this study, while values of $ln(\alpha_{st})$ and β_{st} are from Charru *et al.* (2012) (also obtained from French NFI). For these three B-parameters, most of the values fitted using the French NFI data were found to be out of the range of the literature values (blue dots in Fig. 2). Values for b_{bc} and β_{st} are larger for *Pinus* compared to literature values, but the opposite is found for $ln(\alpha_{st})$. For *Picea abies*, values of b_{bc} and β_{st} are smaller than literature values and larger for $ln(\alpha_{st})$. Exceptions are *Pseudotsuga menziesii* for which $ln(\alpha_{st})$, $\beta_{st.}$ and b_{bc} are in the same range between the two approaches, and *Abies alba* for parameters $ln(\alpha_{st})$ and $\beta_{st.}$

Species clustering

The first clustering of species, based on all parameters, is called global classification (Fig. 3a). It results in a first group composed of *Abies alba, Picea abies* and *Pseudotsuga menziesii* and a second group with *Pinus sylvestris* and *Pinus pinaster*. Two potential biases were tested for: 1) parameter redundancy – the inclusion of different parameters related to the same eco-physiological process may give too much weight to this process in the classification – and 2) the type of parameter used for the classification, because P- and B-parameters are related to different processes and may have uneven weights in the classification. To test for these possible biases, three other classifications were performed. A classification based on P-parameters only was found to be equivalent to the global classification, yet with a reduced distance between groups and a longer distance between *Pseudotsuga menziesii* and *Picea abies*, *Abies alba* (Fig. 3b). Another classification based on B-parameters resulted in grouping of *Pseudotsuga*

319 *menziesii* and *Picea abies* on one side*, Pinus sylvestris, Abies alba* and *Pinus pinaster* on the

320 other side (Fig. 3c).

321 To decrease redundancy, a second classification was based on all parameters except $ln(\alpha_{st})$, T_{optj}

322 and J_{max}/V_{max} to minimize redundancy of these parameters with β_{st}, T_{optv} and V_{cmax25} respectively

323 (Figs 3d and 3e). The two main clusters of this second classification are found to be the same

324 as in the (first) global classification, and appear even more clearly since the distances between

325 species within each cluster are relatively smaller than in the global classification. If we separate

326 again between the reduced parameter set for the P-and B-parameters (hereafter rP and rB), the

327 classification result using rB-parameters only is equivalent to the one using rP-parameters only.

328 From this we conclude that two new PFTs describing temperate conifers can be defined, namely

329 PFT1 with *Abies alba*/*Picea abies*/*Pseudotsuga menziesii* and PFT2 with *Pinus pinaster*/*Pinus*

330 *sylvestris*.

331 **Model evaluation**

332 We reprocessed the results of Bellassen *et al.* (2011) with the new classification. The spatial

333 distribution of the new PFTs within the TNF group is illustrated in the Supplementary

334 Information (Fig. S3). It allows us to explain qualitatively the spatial distribution of differences

335 between observed and simulated volumes and volume increments (Fig. 4). From the

336 observations available in the French NFI dataset we separated pixels where only one PFT was

337 present from pixels where the two PFTs coexist. Pixels which show an overestimation of the

338 volume increment correspond to a predominance of PFT2 (Fig. 4a); and the opposite, pixels

339 showing an underestimation of the volume increment correspond to a predominance of PFT1.

340 This distinction is also visible on pixels where the two PFTs are present (Fig. 4b).

341 Figure 5 illustrates the spatial correlation between observed and simulated volume, volume

342 increment and stand density with 1, 2 or 5 PFTs. Considering TNF as a single PFT with

343   parameters values retrieved from the literature (S1) degrades the spatial representation of

344   volume and volume increment compared to the reference simulation (REF). Increasing the

345   number of PFTs improves the spatial correlation (r) between observed and simulated total

346   aboveground volume (r_{REF}=0.27, r_{S2}=0.73), volume increment (r_{REF}=0.44, r_{S2}=0.56) and

347   average stand density (r_{REF}=-0.16, r_{S2}=0.61). However, further increasing the number of PFTs

348   up to one PFT per species in S3 does not improve the results compared to S2, excepted for the

349   volume increment (r_{S2}=0.56, r_{S3}=0.7).

DISCUSSION

350

The species clustering method
351

352   The main result of this study is that the temperate coniferous species planted in French forests

353   can be grouped into a few distinct PFTs from our clustering analysis method based on *sensitive*

354   parameter similarity (Fig.3). The use of a reduced set of influential parameters, designed to

355   prevent redundancy of information across parameters, leads to a similar grouping of species

356   when considering either P- or B-parameters, or both, to group species. This suggests that the

357   parameter classification method proposed here can determine groups of species. The significant

358   inter-species variability observed for most P-parameters, but not for B-parameters could be

359   partly explained by the fact that some B-parameters related to forestry (allometric parameters

360   or *rdi*) are determined from managed stand data, where management strategy can mask species

361   effects. By contrast, species effects are present in P-parameters.

362   Moreover, all eco-physiological parameters are driven by environment and evolutionary

363   selection. Inter-specific variability should be statistically distinguishable from environmentally

364   induced variability. For example here, too few data are available to screen environmental effects

365   or to perform robust statistical tests on all parameters. Thus co-linearity between environment

366   and trait values can potentially bias the classification. However, most PFTs in DGVMs are

367 currently parameterized without screening out possible environmental effects, and their

368 parameterization is usually based on far fewer references than the 156 data used here (Zaehle

369 *et al.,* 2005; Alton, 2011; Wullschleger *et al.,* 2014). The method used in this study is thus

370 relevant for use with DGVMs other than ORCHIDEE-FM.

371 Past studies sought to define PFTs by finding an optimal set of traits (Pillar & Sosinski, 2003;

372 Bernhardt-Römermann *et al.,* 2008; Boulangeat *et al.,* 2012). In contrast, in this study we started

373 from the modelling viewpoint and sought to define PFTs by selecting key functional traits

374 related to parameter values. We managed to define two different PFTs from TNF: one group is

375 formed by pines: *Pinus pinaster* and *Pinus sylvestris*, and the other by spruces and firs: *Abies*

376 *alba*, *Picea abies* and *Pseudotsuga menziesii*. We can see that the classification results in the

377 separation of species with *pinus* trees clustered together, arguing for a genus-based

378 classification. Liston *et al.* (1999) already demonstrated clear phylogenetic differences between

379 pines and spruces. However, all eco-physiological parameters are driven by environment and

380 strong selections have been operating for a long time on these species. Thus, the grouping of

381 pines in a same cluster may be the result of parallel acclimatization of species sharing the same

382 bio-climatic conditions.

383 This proposed re-classification of the TNF of France into two PFTs seems to explain the spatial

384 distribution of simulated misfits of the volume and volume increment in Bellassen *et al.* (2011)

385 (Fig.4, Sup.Fig.3), even within pixels where the two PFTs are present (Fig. 4b). Moreover, the

386 use of two PFTs improves the spatial correlation between observed and simulated stand

387 variables by more than 40 %, while the use of one PFT by species does not bring more precision.

388 This result tends to confirm our method and that the differences between simulated and

389 observed stands are mostly attributable to the PFT representation rather than a potential bias

390 linked to local soil-climate conditions. Past studies highlighted the discrepancies between the

391 leaf and the canopy scale, but also the structural uncertainties of current DGVMs, which

392 therefore cannot be parameterized with leaf-scale observations (Bonan *et al.,* 2012). Difficulties

393 in parameterizing one unique PFT can also arise from the heterogeneity of species within this

394 PFT.

Inconsistencies between parameter values from eddy-covariance and literature data

396 The inconsistency between the literature review and eddy-covariance (EC) based optimization

397 results from Kuppel *et al.* (2014) implies that the Bayesian optimization based on variables like

398 fluxes (which depend on many parameters) is inadequate because it does not correctly represent

399 physiological processes. The optimization method may benefit from the growing availability of

400 trait observations in the TRY database. However, as highlighted by Bonan *et al.* (2012), it also

401 suggests that some fundamental processes in ORCHIDEE-FM are lacking or poorly

402 represented, thus leading to compensations in the parameterization.

403 We have no information about the correlation of optimized parameter errors related to

404 temperate needleleaf forest in Kuppel *et al.* (2014). However, Kuppel (2012) showed strong

405 posterior error correlations between the carboxylation rate, the needle lifespan and the optimal

406 temperature for example when considering all multi-site optimizations and all PFTs (see

407 Fig.S13 of Kuppel (2012) for the correlation matrix of parameters errors). These correlations

408 occur because of the inability (or equifinality, i.e. different combinations of parameter values

409 can lead to the same result) of EC data to discriminate the value of each single parameter. Traits

410 optimized in Kuppel *et al.* (2014) can be biased due to equifinality or a poor representation of

411 some processes, thus leading to an over-parameterization when constrained by EC data.

412 In the other hand, parameters optimized by Kuppel *et al.* (2014) represent traits integrated at

413 the canopy level, while measurements gathered in this study were performed on individual

414 leaves. Therefore, the comparison between optimized and measured traits should be made in

415 the light of the different biases specific to each method and scale.

416   These differences between leaf-scale measurements of parameters from literature data and the

417   values optimized using ecosystem-level observations of fluxes of NEE, H, and E can be

418   attributed to the lack of physiological consideration in the mathematical procedure used to

419   optimize parameters from ecosystem-scale flux data. Despite the constraint imposed by setting

420   a realistic prior error value for each parameter in the inversion (based on expert judgment), the

421   minimization of a Bayesian cost function produces a posterior value based on purely

422   mathematical criteria. Such a value may work well locally, but might be disconnected from eco-

423   physiological reality (Pelikan 2005; Ward *et al.,* 2010). This discrepancy can arise from the

424   erroneous representation of eco-physiological processes in ORCHIDEE, but also

425   compensations or co-variations between parameters dealing with same eco-physiological

426   processes that cannot be separated by NEE measurements alone. Another point revealed by this

427   study is the potential bias introduced by giving equal weight to all the geographic sites used for

428   optimization. Most of the sites (20 in total) optimized in Kuppel (2012) were *Picea abies* and

429   *Pinus sylvestris* whereas the only site optimized by Santaren *et al.* (2007) was *Pinus pinaster*.

430   In the end, parameter values of optimizations were allocated to the same PFT, ignoring the

431   different species. However we show here differences between species for key parameters inside

432   the same plant functional type. The problem highlighted here is an over-representation of one

433   species in the optimization. Unless this over-representation reflects the real species distribution

434   on the ground, it will bias the final value retained for the global PFT. Especially in the Northern

435   Hemisphere, where we expect an increase in conifer plantations to meet the objectives of

436   climate change mitigation, the non-representation of the different capacity of species will lead

437   to the wrong estimation of carbon stocks and fluxes (Vallet *et al.,* 2009; Matthews *et al.,* 2014).

438   Our results thus highlight the importance of confronting the results of "brute force" parameter

439   calibration methods using mathematics and EC data with a more refined approach based on

440   eco-physiological knowledge before using them for model improvement.

441 The use of a literature review as performed in this study can provide better optimizations by

442 applying appropriated ranges of variability for each parameter. One of the limitations of a

443 literature search however, is the lack of data harmonization related to differences in

444 environmental conditions and methodologies. We tried to correct for some methodological

445 effects but too few data were available to allow a correction based on biotic and abiotic

446 environmental as well as management factors, which leads to large spread in parameter

447 estimates (Fig. 2). Results obtained for each parameter and the associated uncertainties are

448 debated in detail in the extended discussion (Appendix S4).

449

450 **CONCLUSION& PERSPECTIVES**

451 In this study we proposed a new method combining literature review and hierarchical

452 classification to constrain sensitive parameters in a DGVM model. The model used was

453 ORCHIDEE FM and our results are model-specific, but the approach is applicable to any other

454 DGVM with a discrete representation of PFTs. We proposed to group species into coherent

455 groups based on parameter similarity for refining the number of PFTs. This method was applied

456 to temperate conifers in France for 5 dominant species, and resulted in the successful derivation

457 of two PFTs with coherent parameters. The parsimonious choice of two new PFTs to describe

458 TNF improved the spatial representation of conifers stands carbon dynamics by more than 40%

459 compared to the reference simulation where all TNF have the same parameter values. However

460 using a more refined representation with PFT per species did not improve the simulations, and

461 did not degrade them either. Our approach of coherent parameter classification is thus a good

462 compromise between the too rigid simple PFT representation and a more complicated fully

463 trait-based continuous parameterization of vegetation parameters in DGVMs. Thanks to the

464 increasing availability of trait observation in global databases, our method could be applied to

465 refine other PFTs, for instance the "temperate broadleaf forests" PFT which currently groups

466 together in a model like ORCHIDEE species with different characteristics such as fast growing

467 pioneer species (*Populus sp.*) and slow growing dominant species like oak (*Quercus sp.*)

468 In addition, we showed clear discrepancies between observed parameter values and values

469 optimized against eddy covariance flux measurement. This result highlights the importance of

470 confronting the results of "brute force" parameter calibration methods with a more refined eco-

471 physiological knowledge. We argue that this warrants further research for a better

472 understanding of the definition of parameters, their characterization and the adequacy of point-

473 based field observations to parameterize grid-based models.

474

483

484 Data Accessibility

485 All data and sources used in this study are listed in Appendix S3.

486

487

Biosketch

M. Peaucelle did his PhD at the Climate and Environment Science Laboratory (LSCE, France). His research focuses on improving the representation of temperate forest in the ORCHIDEE DGVM. In particular, he is interested in improving the representation of plant functional types by a more continuous approach exploring different ways to include the plasticity of plant functional traits.

Supporting information

494

495 Appendix S1: Supplementary tables and figures (list below)

510 **References**

511 Alton, P.B. (2011) How useful are plant functional types in global simulations of the carbon,

512 water, and energy cycles? *Journal of Geophysical Research: Biogeosciences*, 116.

513 Bellassen, V., Le Maire, G., Dhôte, J.F., Ciais, P. & Viovy, N. (2010) Modelling forest

514 management within a global vegetation model—Part 1: Model structure and general behaviour.

515 *Ecological Modelling*, 221, 2458–2474.

516 Bellassen, V., le Maire, G., Guin, O., Dhôte, J.F., Ciais, P. & Viovy, N. (2011) Modelling forest

517 management within a global vegetation model—Part 2: Model validation from a tree to a

518 continental scale. *Ecological Modelling*, 222, 57–75.

519 Bernhardt-Römermann, M., Römermann, C., Nuske, R., Parth, A., Klotz, S., Schmidt, W. &

520 Stadler, J. (2008) On the identification of the most suitable traits for plant functional trait

521 analyses. *Oikos*, 117, 1533–1541.

522 Van Bodegom, P.M., Douma, J.C., Witte, J.P.M., Ordoñez, J.C., Bartholomeus, R.P. & Aerts,

523 R. (2012) Going beyond limitations of plant functional types when predicting global

524 ecosystem–atmosphere fluxes: exploring the merits of traits-based approaches. *Global Ecology*

525 *and Biogeography*, 21, 625–636.

526 Bonan, G.B., Levis, S., Kergoat, L. & Oleson, K.W. (2002) Landscapes as patches of plant

527 functional types: An integrating concept for climate and ecosystem models. *Global*

528 *Biogeochemical Cycles*, 16, 5–23.

529 Bonan, G.B., Oleson, K.W., Fisher, R.A., Lasslop, G. & Reichstein, M. (2012) Reconciling leaf

530 physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the

531 Community Land Model version 4. *Journal of Geophysical Research: Biogeosciences*, 117.

532 Boulangeat, I., Philippe, P., Abdulhak, S., Douzet, R., Garraud, L., Lavergne, S., Lavorel, S.,

533 Van Es, J., Vittoz, P. & Thuiller, W. (2012) Improving plant functional groups for dynamic

534 models of biodiversity: at the crossroads between functional and community ecology. *Global*

535 *Change Biology*, 18, 3464–3475.

536 Charru, M., Seynave, I., Morneau, F., Rivoire, M. & Bontemps, J.-D. (2012) Significant

537 differences and curvilinearity in the self-thinning relationships of 11 temperate tree species

538 assessed from forest inventory data. *Annals of Forest Science*, 69, 195–205.

539 Harrison, S.P., Prentice, I.C., Barboni, D., Kohfeld, K.E., Ni, J. & Sutra, J.-P. (2010)

540 Ecophysiological and bioclimatic foundations for a global plant functional classification.

541 *Journal of Vegetation Science*, 21, 300–317.

542 IFN. 2012. "Raw Inventory Data." http://inventaire-forestier.ign.fr/.

543 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha,

544 S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W.,

545 Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. & Joseph, D. (1996)

546 The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological*

547 *Society*, **77**, 437–471.

548 Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby,

549 M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P., Van BODEGOM,

550 P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., Anand, M., Atkin,

551 O., Bahn, M., Baker, T.R., Baldocchi, D., Bekker, R., Blanco, C.C., Blonder, B., Bond, W.J.,

552 Bradstock, R., Bunker, D.E., Casanoves, F., Cavender-Bares, J., Chambers, J.Q., Chapin Iii,

553 F.S., Chave, J., Coomes, D., Cornwell, W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka,

554 W., Elser, J., Esser, G., Estiarte, M., Fagan, W.F., Fang, J., Fernández-Méndez, F., Fidelis, A.,

555 Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V.,

556 Green, W.A., Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S.,

557 Joly, C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M.H.,

558 Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F., Lenz,

559 T., Lewis, S.L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M.D., Manning, P., Massad,

560 T., Medlyn, B.E., Messier, J., Moles, A.T., Müller, S.C., Nadrowski, K., Naeem, S., Niinemets,

561 Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V.G., Onoda, Y., Ordoñez, J.,

562 Overbeck, G., Ozinga, W.A., Patiño, S., Paula, S., Pausas, J.G., Peñuelas, J., Phillips, O.L.,

563 Pillar, V., Poorter, H., Poorter, L., Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch,

564 S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A.,

565 Sosinski, E., Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram,

566 M., Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E. & Wirth,

567 C. (2011) TRY – a global database of plant traits. *Global Change Biology*, 17, 2905–2935.

568 Kleidon, A. & Mooney, H.A. (2000) A global distribution of biodiversity inferred from climatic

569 constraints: results from a process-based modelling study. *Global Change Biology*, 6, 507–523.

570 Krinner, G., Viovy, N., Noblet-Ducoudré, N. de, Ogée, J., Polcher, J., Friedlingstein, P., Ciais,

571 P., Sitch, S. & Prentice, I.C. (2005) A dynamic global vegetation model for studies of the

572 coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19.

573 Kuppel, S. (2012) Assimilation de mesures de flux turbulents d'eau et de carbone dans un

574 modèle de la biosphère continentale, Versailles-St Quentin en Yvelines.

575 Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F. & Richardson, A.D. (2012)

576 Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*,

577 9 (10), 3757–3776.

578 Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L. & Cescatti, A.

579 (2014) Model–data fusion across ecosystems: from multi-site optimizations to global

580 simulations. *Geoscientific Model Development*, 7 (6), 2581–2597.

581 Liston, A., Robinson, W.A., Piñero, D. & Alvarez-Buylla, E.R. (1999) Phylogenetics of *Pinus*

582 (*Pinaceae*) based on nuclear ribosomal DNA internal transcribed spacer region sequences.

583 *Molecular Phylogenetics and Evolution*, 11, 95–109.

584 Mahalanobis, P. (1936) On the generalised distance in statistics. *Proceedings National Institute*

585 *of Science*, 2, 49–55.

586 Matthews, R., Sokka, L., Soimakallio, S., Mortimer, N., Rix, J., Schelhaas, M.-J., Jenkins, T.,

587 Hogan, G., Mackie, E., Morris, A. & Randle, T. (2014) Review of literature on biogenic carbon

588 and life cycle assessment of forest bioenergy. Final Task 1 report, The Research Agency of the

589 Forestry Commission, Farnham, UK.

590 Naudts, K., Ryder, J., McGrath, M.J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray,

591 G., Bönisch, G. & Campioli, M. (2015) A vertically discretised canopy description for

592 ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes.

593 *Geoscientific Model Development*, 8, 2035–2065.

594 Nemani, R. & Running, S.W. (1996) Implementation of a hierarchical global vegetation

595 classification in ecosystem function models. *Journal of Vegetation Science*, 7, 337–346.

596 Pausas, J.G. & Lavorel, S. (2003) A hierarchical deductive approach for functional types in

597 disturbed ecosystems. *Journal of Vegetation Science*, 14, 409–416.

598 Pavlick, R., Drewry, D., Bohn, K., Reu, B. & Kleidon, A. (2012) The Jena Diversity-Dynamic

599 Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial

600 biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences*, 10,

601 4137–4177.

602 Peel, M.C., Finlayson, B.L. & McMahon, T.A. (2007) Updated world map of the Köppen-

603 Geiger climate classification. *Hydrology and Earth System Sciences*, 11 (5), 1633–1644.

604   Pelikan, M. (2005) Hierarchical Bayesian Optimization Algorithm. *Studies in Fuzziness and*

605   *Soft Computing*, 31–48.

606   Pillar, V.D. (1999) On the identification of optimal plant functional types. *Journal of Vegetation*

607   *Science*, 10, 631–640.

608   Pillar, V.D. & Sosinski, E.E. (2003) An improved method for searching plant functional types

609   by numerical analysis. *Journal of Vegetation Science*, 14, 323–332.

610   Reich, P.B., Rich, R.L., Lu, X., Wang, Y.-P. & Oleksyn, J. (2014) Biogeographic variation in

611   evergreen conifer needle longevity and impacts on boreal forest carbon cycle projections.

612   *Proceedings of the National Academy of Sciences of the United States of America*, 111, 13703–

613   13708.

614   Reu, B., Zaehle, S., Proulx, R., Bohn, K., Kleidon, A., Pavlick, R. & Schmidtlein, S. (2011)

615   The role of plant functional trade-offs for biodiversity changes and biome shifts under scenarios

616   of global climatic change. *Biogeosciences*, 8, 1255–1266.

617   Santaren, D., Peylin, P., Viovy, N. & Ciais, P. (2007) Optimizing a process-based ecosystem

618   model with eddy-covariance flux measurements: A pine forest in southern France. *Global*

619   *Biogeochemical Cycles*, 21.

620   Team, R.D.C. (2011) R: A Language and Environment for Statistical Computing, Vienna,

621   Austria.

622   Vallet, P., Dhôte, J.-F., Moguédec, G.L., Ravart, M. & Pignard, G. (2006) Development of total

623   aboveground volume equations for seven important forest tree species in France. *Forest*

624   *Ecology and Management*, 229, 98–110.

625   Vallet, P., Meredieu, C., Seynave, I., Bélouard, T. & Dhôte, J.F. (2009) Species substitution for

626   carbon storage: Sessile oak versus Corsican pine in France as a case study. Forest Ecology and

627   Management, 257, 1314–1323.

628   Verheijen, L., Brovkin, V., Aerts, R., Bönisch, G., Cornelissen, J., Kattge, J., Reich, P., Wright,

629   I. & Bodegom, P. van (2013) Impacts of trait variation through observed trait-climate

630   relationships on performance of a representative Earth System model: a conceptual analysis.

631   *Biogeosciences*, 10 (8), 5497-5515.

632   Verheijen, L.M., Aerts, R., Brovkin, V., Cavender-Bares, J., Cornelissen, J.H.C., Kattge, J. &

633   van Bodegom, P.M. (2015) Inclusion of ecologically based trait variation in plant functional

634   types reduces the projected land carbon sink in an earth system model. *Global Change Biology*,

635   21, 3074–3086.

636   Vetter, M., Churkina, G., Jung, M., Reichstein, M., Zaehle, S., Bondeau, A., Chen, Y., Ciais,

637   P., Feser, F., Freibauer, A., Geyer, R., Jones, C., Papale, D., Tenhunen, J., Tomelleri, E.,

638   Trusilova, K., Viovy, N. & Heimann, M. (2008) Analyzing the causes and spatial pattern of the

639   European 2003 carbon flux anomaly using seven models. *Biogeosciences*, **5**, 561–583.

640   Ward, J.H. (1963) Hierarchical Grouping to Optimize an Objective Function. *Journal of the*

641   *American Statistical Association*, 58, 236–244.

642   Ward, B.A., Friedrichs, M.A.M., Anderson, T.R. & Oschlies, A. (2010) Parameter optimization

643   techniques and the problem of under determination in marine biogeochemical models. *Journal*

644   *of Marine Systems*, 81, 34–43.

645   Wullschleger, S.D., Epstein, H.E., Box, E.O., Euskirchen, E.S., Goswami, S., Iversen, C.M.,

646   Kattge, J., Norby, R.J., van Bodegom, P.M. & Xu, X. (2014) Plant functional types in Earth

647   system models: past experiences and future directions for application of dynamic vegetation

648   models in high-latitude ecosystems. *Annals of Botany*, 114, 1–16.

649   Zaehle, S., Sitch, S., Smith, B. & Hatterman, F. (2005) Effects of parameter uncertainties on

650   the modeling of terrestrial biosphere dynamics. *Global Biogeochemical Cycles*, 19.

Tables

Parameter	Description	Equation used in ORCHIDEE-FM	Processes involved
V_{cmax_opt}	Optimal maximum rate of carboxylation (μmol m^{-2}s^{-1})	$$Vc = \frac{Vcmax_{opt} \times Ci}{Ci + Kc\left(1 - \frac{Oi}{Ko}\right)}$$ Vc= RUBISCO activity (μmolCO2 m^{-2}s^{-1}) Oi= intracellular oxygen concentration Kc and Ko=Michaelis-Menten constants of RUBISCO activity	Impacts the gross primary production and substantially the stomatal conductance
T_{opt}	Optimal temperature of photosynthesis (°C)	$$Vc = Vcmax_{opt} \times \frac{(Tair - Tmax)(Tair - Tmin)}{(Tair - Tmax)(Tair - Tmin) - (Tair - Topt)^2}$$ Tair= Air temperature (°C) Tmin, Tmax= minimal and maximal temperature of photosynthesis (°C)	Impacts the temperature dependency of the photosynthesis, and daily and seasonal variations of the GPP. It is the same equation used for V_{jmax} (maximal rate of the RuBP regeneration)
g_{slope}	Slope of stomatal conductance in the Ball-Berry model (unitless)	$$gs = gmin + gslope \times A \times \frac{hr}{Ca}$$ gs= stomatal conductance (mol m^{-2} s^{-1}) A = net assimilation of the photosynthesis (mol m^{-2} s^{-1}) hr = relative humidity (unitless) Ca = boundary layer CO$_2$ concentration (μmol mol^{-1}) g_{min} = empirical coefficients (mol m^{-2} s^{-1})	Parameter which modifies the slope of the linear relationship between the stomatal conductance and the assimilation. It substantially impacts the estimation of the NEE and E.
SLA	Specific length area (mm² mgC^{-1})	$$LAI = SLA \times Bm_{leaves}$$ LAI= Leaf area index (m² leaves m^{-2}ground) Bm_leaves= leaves biomass (mgC)	Parameter which determines the Leaf Area Index and impacts the GPP.
$L_{age,crit}$	Critical leaf age for leaf senescence (month)	$$Vc = Vcmax_{opt} \times e_{rel}\left(\frac{af}{Leaf_{agecrit}}\right)$$ af= leaf age Leaf_agecrit= mean leaf lifespan e$_{rel}$=relative efficiency	Parameter which determines the mean leaf lifespan. It impacts the photosynthetic activity and the senescence of leaves.

Vj_{max}/V_{cmax}	Maximum flux of electron / maximum rate of carboxylation ratio (unitless)	$$Vjmax = 2 \times Vcmax$$ Vjmax = maximal rate of the RuBP regeneration (μmol m^{-2}s^{-1})	The parameter impacts the estimation of the GPP and substantially impacts the stomatal conductance.
a_{bc}	Coefficient of biomass-circumference allometry	$$M = a_{bc} \times D^{b_{bc}}$$ M = total aboveground tree dry biomass (kg) D = diameter at breast height (cm)	Determines the circumference of trees and impacts thinning events and the exportation of biomass
b_{bc}	Coefficient of biomass-circumference allometry		
α_{st}	Coefficient of self thinning equation	$$dens_{max} = \frac{\alpha_{st}}{Dg^{\beta st}}$$ $dens_{max}$ = stand maximum density in ind ha^{-1} (individuals per hectare) Dg = quadratic mean diameter (meters).	Determines the maximal density of trees in the stand. Impacts thinning events and consequently the biomass distribution among trees
β_{st}	Coefficient of self thinning equation		
rdi_target	Targeted value or relative density index (unitless)	$$thinning\,if\,rdi = rdi_{target} \pm \delta rdi$$ rdi = relative density index δrdi= tolerable variation of rdi around rdi_target	Determines the moment when the thinning occurs. Impacts the biomass and exportation

652

653   Table 1: List of parameters, equations and processes involved in ORCHIDEE-FM and selected for this study.

654

655

656

657

|	Number of data from literature / TRY					
Parameter	*Abies alba*	*Picea abies*	*Pinus pinaster*	*Pinus sylvestris*	*Pseudotsuga menziesii*	Total
V_{cmax25}	13 / 0	30 / 3	13 / 0	15 / 2	20 / 0	93
T_{opt}	2 / 0	2 / 0	2 / 0	2 / 0	2 / 0	10
g_{slope}	0 / 0	7 / 6	2 / 0	8 / 0	0 / 0	23
g_{smax}	10 / 0	20 / 1	21 / 0	18 / 1	15 / 0	86
SLA	3 / 23	13 / 10	20 / 3	24 / 224	10 / 6	336
$L_{age,crit}$	2 / 2	7 / 2	7 / 0	17 / 4	4 / 2	47
J_{max}/V_{cmax}	19 / 0	25 / 2	7 / 0	11 / 1	5 / 0	70
b_{bc}	3 / 0	39 / 0	3 / 0	22 / 0	12 / 0	79
α_{st}	2 / 0	6 / 0	3 / 0	8 / 0	4 / 0	23
Rdi_target	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	5

658

659 Table 2: Number of observations retrieved for each parameter from the literature and the TRY database. Refers to Table 1 for the description of
660 each parameter.

661

662

Parameter	Description	Current value	Santaren et al., 2007	Kuppel et al., 2014	Median value from literature								Unit
		TNF	TNF	TNF	A.alb	P.abi	P.men	P.syl	P.pin	PFT1	PFT2	TNF	
V_{cmax25}	Optimal maximum rate of carboxylation	35	41.7	31.94	46	46.77	38.35	106.61	90.3	45.5	100.2	61	µmol m^{-2} s^{-1}
T_{opt}	Optimal temperature of photosynthesis	25	-	17.49	36.81	31.58	29.94	27.56	38.84	31.78	33.2	32.34	Vmax °C
		25	-	17.49	33.20	21.76	27.3	19.89	36.87	27.42	28.38	27.8	Jmax °C
g_{slope}	Slope of stomatal conductance in the Ball-Berry model	9	8.1	8.841	-	5.94	-	5.59	10.05	-	-	-	unitless
gs_{max}	Maximal stomatal conductance	-	-	-	3.41	3.88	4.12	3.9	3	3.89	3.78	3.89	mm s^{-1}
SLA	Specific length area	4.63	-	10	5.51	5.4	7.08	4.73	3.1	5.6	4.73	4.85	mm^2 mg^{-1}
$L_{age,crit}$	Critical leaf age for leaf senescence	29.5	-	35.6	98.7	102.33	65.4	34	51	78	42.5	51	Months
J_{max}/V_{cmax}	Maximum flux of electron / maximum rate of carboxylation ratio	2	-	-	2.2	2.2	3.01	2.1	2.5	2.2	2.1	2.2	unitless
a_{bc}	Coefficient of biomass-circumference allometry	7.03b$_{bc}$$^{-4.76}$	-	-	-	-	-	-	-	-	-	-	kgDM
b_{bc}	Coefficient of biomass-circumference allometry	2.295	-	-	2.52	2.36	2.4035	2.2355	2.13	2.37	2.234	2.355	ln(kgDM) ln(m)$^{-1}$
α_{st}	Coefficient of self thinning equation	12,2	-	-	12.548	12.7925	12.3165	12.48097	12.544	12.4995	12.544	12.5	ln(ind ha^{-1})
β_{st}	Coefficient of self thinning equation	1.6	-	-	1.7505	1.664	1.6255	1.7915	1.815	1.664	1.815	1.7295	ln(ind ha^{-1}) ln(m)$^{-1}$
Rdi_target	Targeted value or relative density index	0.75	-	-	0.59	0.54	0.63	0.44	0.59	0.56	0.64	0.56	unitless

663 Table 3: Default (ORCHIDEE current value), optimized (Santaren *et al.*, 2007; Kuppel *et al.*, 2014) and observed (from the literature and the
664 TRY database) values for each parameter used in this study. TNF correspond to the current Temperate Needleleaf evergreen Forest plant
665 functional type. (-): no data available.

666 **Figures**

667
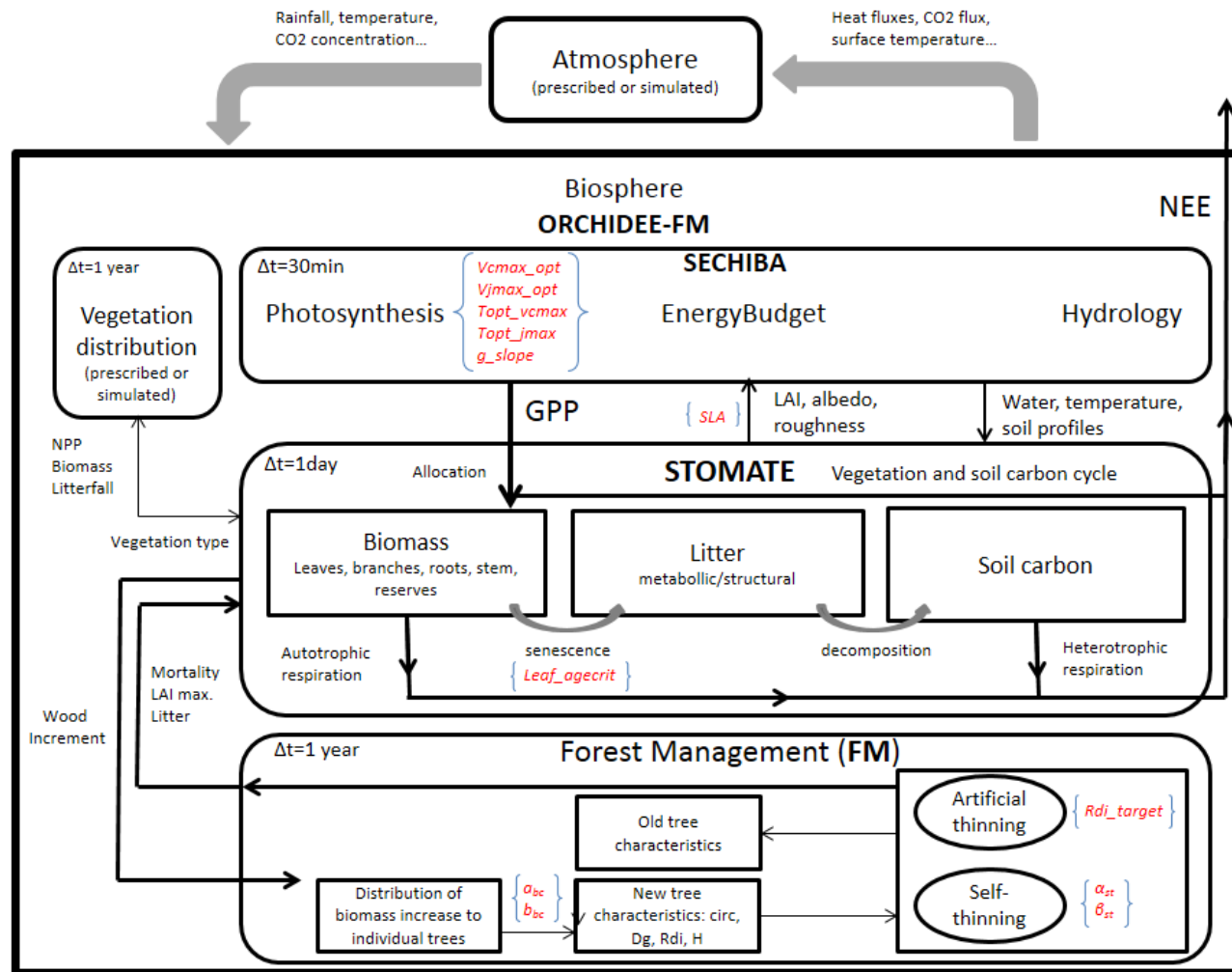
Fig.1: Representation of the ORCHIDEE-FM model and related processes. Round corner boxes correspond to ORCHIDEE modules (bold). The main simulated processes and the corresponding time step (Δt) are listed for each module. In the STOMATE module are listed the different carbon compartments, while the FM module illustrates the management processes simulated by ORCHIDEE. Arrows represent shared variables, processes and the different links between modules. Parameters used in this study are listed in parenthesis (Table 1).

672
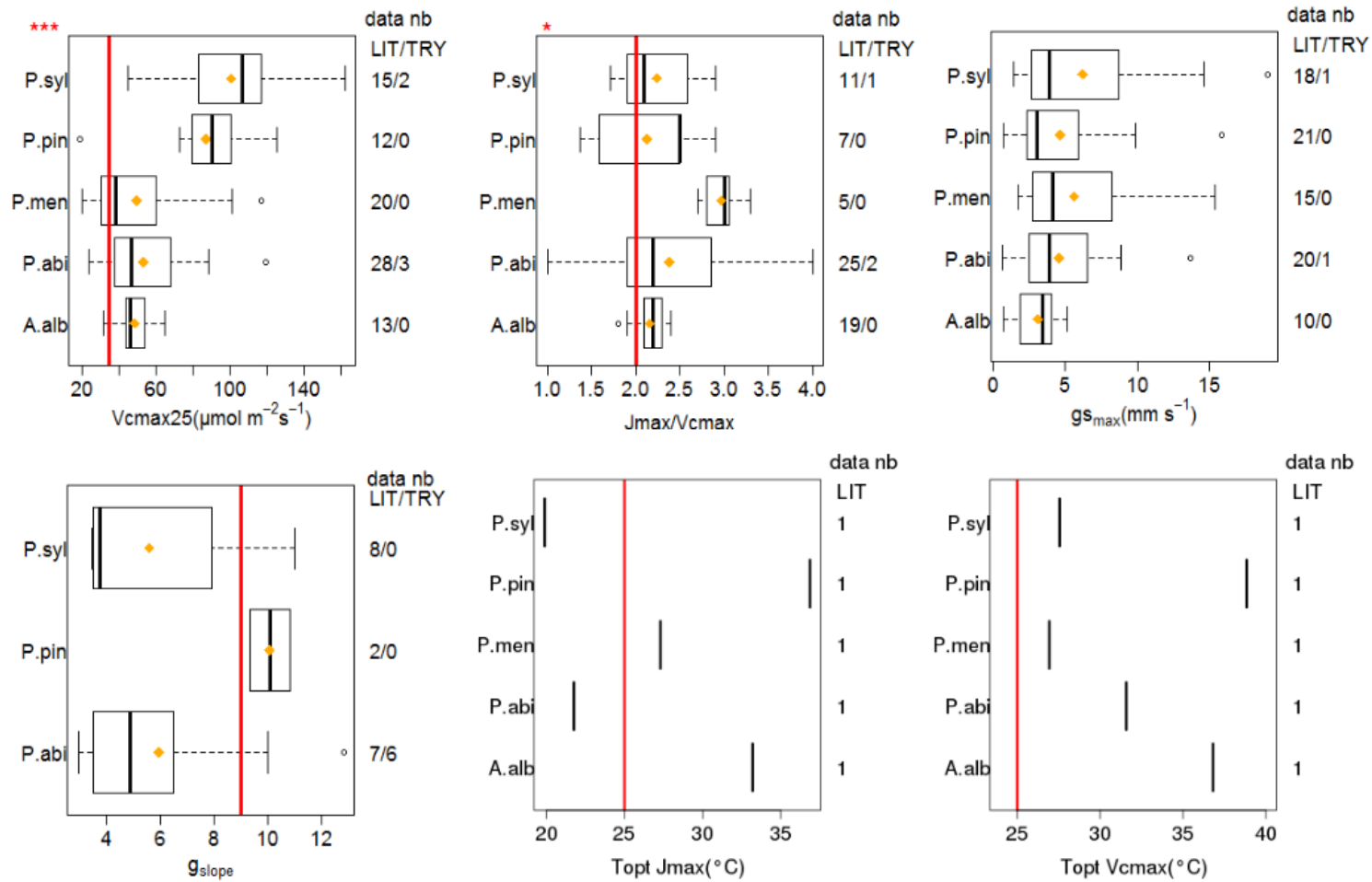
Fig.2:Boxplot representation of parameters for each species (median, first and third quartile, first and last deciles, black open circles represent values outside 1.5 times the interquartile range above the upper quartile and below the lower quartile). Refers to Table 1 for the description of each parameter. Orange diamond represent mean values; blue dots represent values fitted on French national forest inventories data. Number of data available for each species is specified on the right of each boxplot. LIT/TRY correspond to the number of data from literature and from the

677 TRY database. The red line represents the current ORCHIDEE value of the TNF PFT. Parameters with significant Kruskal-Wallis test : '*' p-
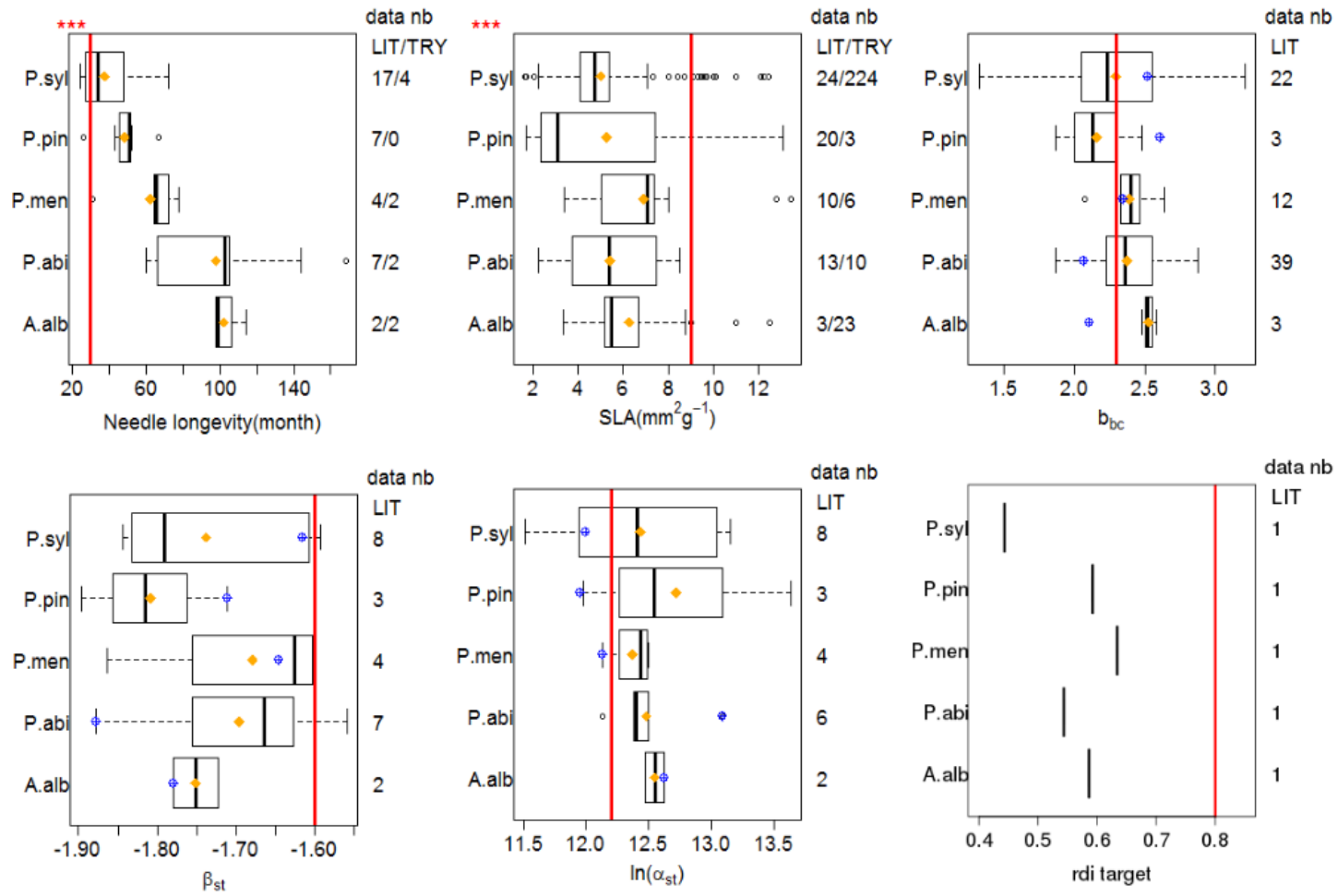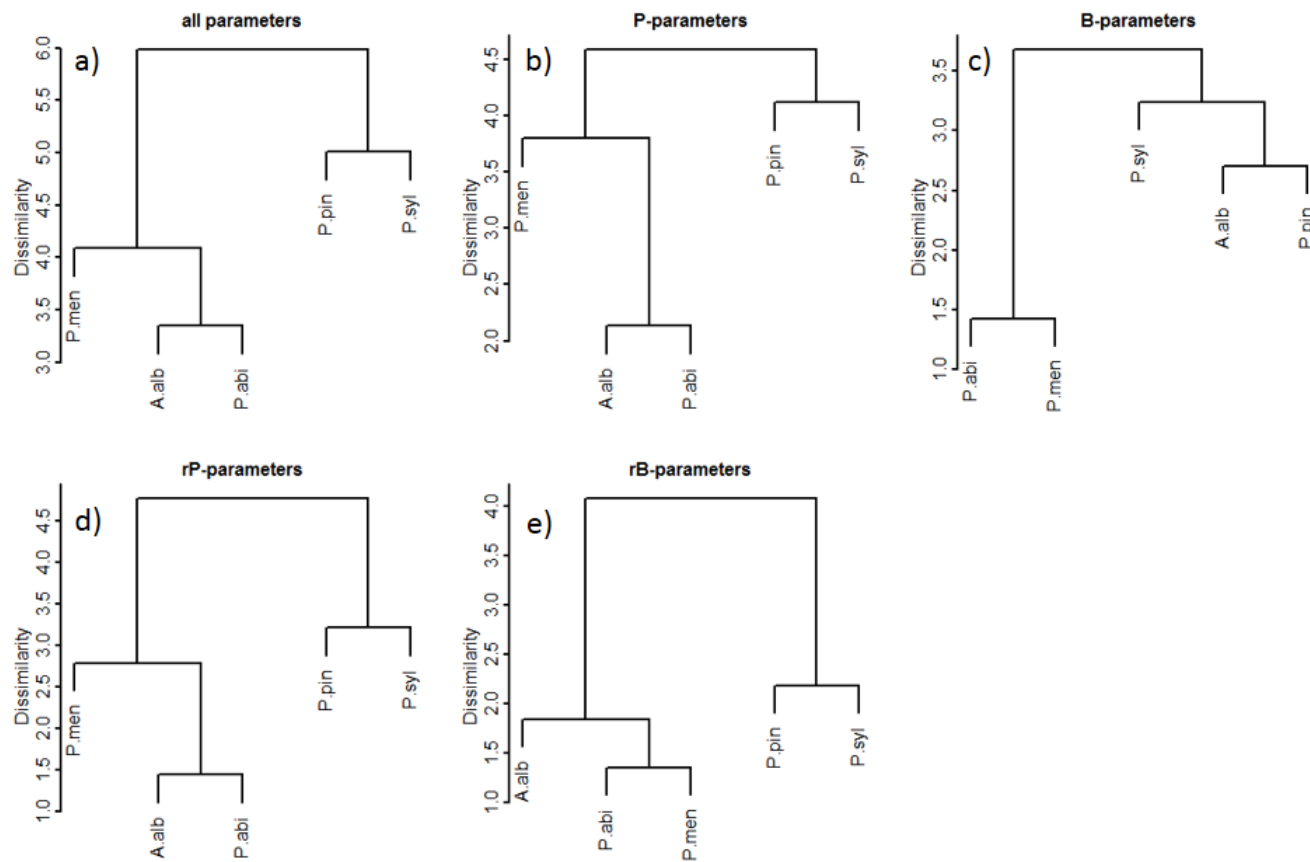678 value < 0.05; '***' p-value < 0.01

679

680

Fig.3: Clusters of species (*Abies alba*/ *Picea abies*/ *Pseudotsuga mensiezii*/ *Pinus sylvestris*/ *Pinus pinaster*) obtained with different parameter selections. a) with all parameters; b) with photosynthesis-parameters (P); c) with biomass and management-parameters (B); d) with a reduced set of P-parameters (rP) and e) with a reduced set of B-parameters (rB) and f) with significantly different parameters between species.
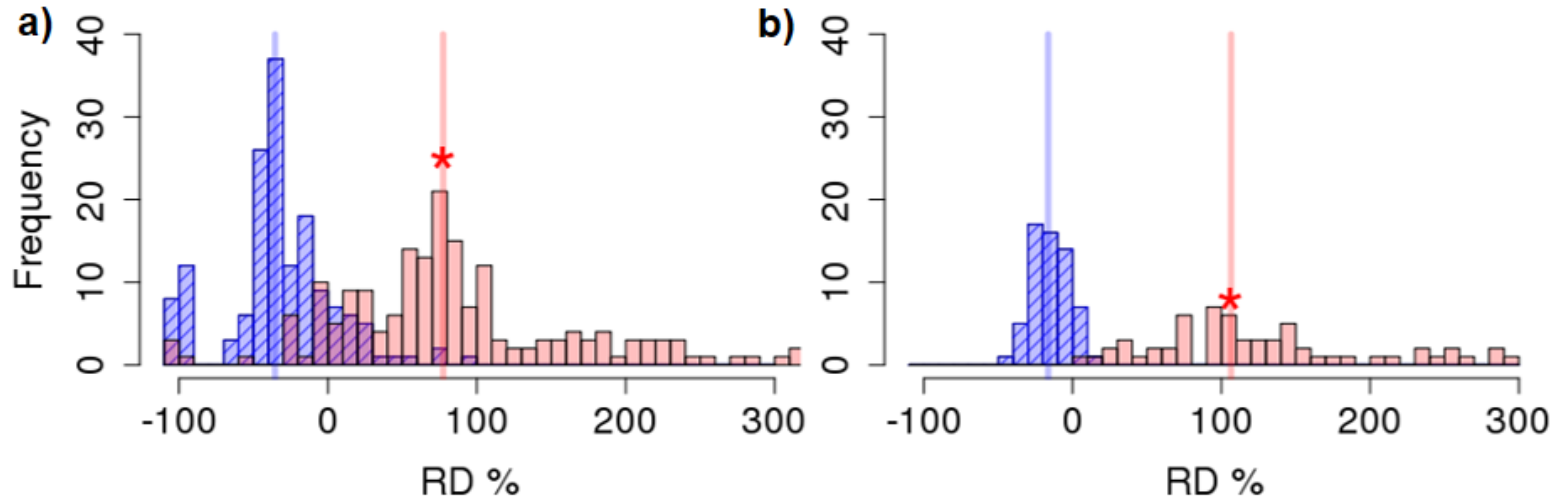
687

688   Fig.4: Distribution of the mean relative difference ("RD %") between observed and simulated volume increment in Bellassen *et al.,* (2011)

689   according to the new PFT classification for a) pixels where only one PFT is present and b) where both PFTs are present on the same pixel. Blue

690   striped frequencies represent PFT1 (*Abies alba/ Picea abies/ Pseudotsuga mensiezii*), red frequencies PFT2 (*Pinus sylvestris/ Pinus pinaster*).

691   Vertical lines represent the median relative difference for each PFT. The asterisk indicates a significant difference between the two means at

692   *P<0.05*

693

694
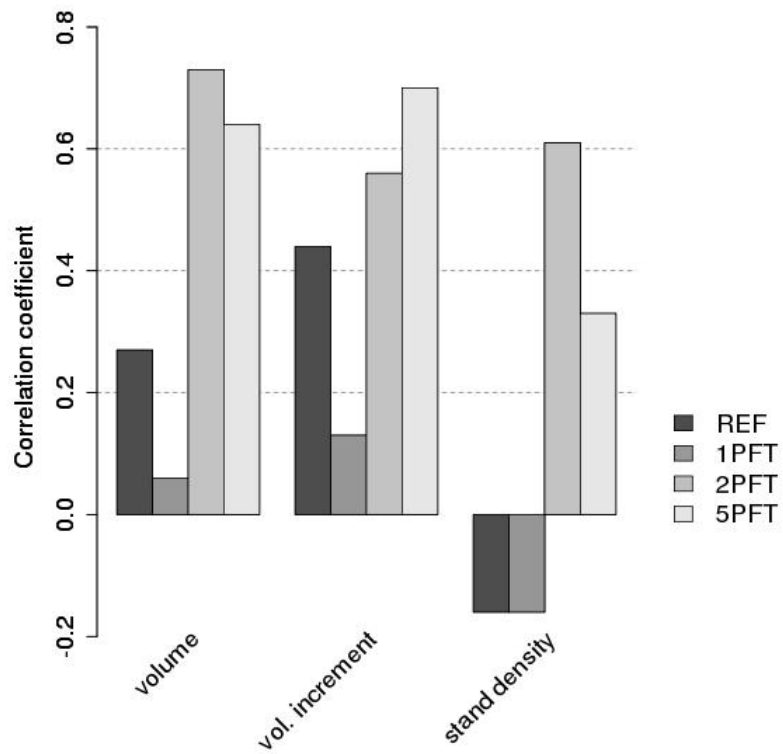
Fig.5: Spearman's correlation coefficient between observed and simulated volume, volume increment and stand density for simulations: REF (standard), S1 (1 PFT), S2 (2 PFTs) and S3 (5 PFTs). Results are for 50y old stands over France