

Dear reviewers and recommender,

Thank you for having carefully read my preprint and for the numerous insightful comments you made. I have considered and answered all your suggestions and corrections to make the paper as clear, transparent and accurate as possible.

All minor corrections were taken into account and are highlighted in the comparison document entitled “article_PDG-Arena_20240209_v_20240620”. Here after, I copied, pasted (in italic blue) and sometimes regrouped the major criticisms of the review. I described (in black font) how I modified the article to integrate each of them. In the final section, I describe additional modifications to the preprint to improve the manuscript quality. For each modification, a sample of the comparison document is given to clearly show the changes involved.

I hope you will now find this revision acceptable for publication in PCI Forest & Wood Sciences.

Best regards,

Camille Rouet, on behalf of all co-authors

Answer to the Review by Erwin Dreyer (recommender)

- Description of the repository dataset and inclusion of other dataset (LAI, soil texture...)

I had also a look at the data presented in the Zenodo repository. Although they are easily accessed, I had the feeling that they would require a careful description of the different data sets (I was unable to find it) in order to facilitate potential reuse.

Concerning the Github/Zenodo repository, we firstly wrote an exhaustive description of the files in the root readme.md file. We also added a readme file accompanying the Rdata file to describe the simulation dataset, which include the simulations results as well as the inventory data. Finally, we added a table that includes the plots LAI that were used during simulations (Sec. 2.3) and a table that describes the soil texture of the plots. The original dataset that made it possible to compute the crown radii based on individual diameters per species were provided by Hendrik Davi. Unfortunately we can't provide it yet.

The repository can be found at <https://doi.org/10.5281/zenodo.12191049>

- Title

The title has been slightly modified:

PDG-Arena: An ecophysiological model for characterizing tree-tree interactions in heterogeneous [and mixed](#) stands

- Abstract

(abstract) please define 'better performance'

(abstract) OK, but are such models the only approach to the question is predicting/simulating the growth and development of mixed stands?

The abstract was improved in order to describe how the model performance was evaluated and now includes numerical results (R² and net mixing effect):

Abstract

1 In the context of [the](#) ongoing climate and biodiversity crises, mixed forest
2 stands are increasingly considered as a sustainable management alternative to
3 monocultures. We developed a new individual-based and process-based for-
4 est growth model, PDG-Arena, to simulate mixed forest functioning and test
5 ecophysiological interactions among trees in mixed stands. The model builds
6 upon ~~of a the~~ validated ecophysiological stand-scale model [CASTANEA](#) and in-
7 tegrates tree competition for light and water. We evaluated the ~~simulation-per-~~
8 formance of PDG-Arena ~~using annual by comparing the simulated growth with~~
9 [annual dendrochronological](#) growth data from ~~39-37~~ common beech and silver
10 fir monospecific and mixed plots in the French Alps. PDG-Arena showed ~~similar~~
11 ~~performance as the validated stand-scale model a slightly better performance~~
12 [than CASTANEA](#) when simulating even-age and monospecific forests ~~—and~~
13 ~~significantly better performance when—~~ (r^2 of ~~32.1 versus 29.5%~~). ~~When~~ using
14 structure-diverse and species-diverse inventories ~~—it—~~, PDG-Arena performed
15 [better than CASTANEA in pure beech \(38.3 versus 22.9%\) and mixed stands](#)
16 [\(40.5 versus 36.3%\), but not in pure fir stands \(39.8 versus 42.0%\)](#). The new
17 [model](#) also showed a significant positive effect of species mixing on gross pri-
18 mary production [\(+5.5%\)](#), canopy absorbance ~~and transpiration~~ [\(+11.1%\) and](#)
19 [transpiration \(+15.8%\)](#). Our results thus show that tree-level process-based
20 models such as PDG-Arena, formally simulating interspecific interactions, ~~are~~
21 ~~needed to better can serve as a valuable tool to~~ understand and simulate the
22 ~~functioning carbon, light and water dynamics~~ of mixed stands.

Keywords: ecophysiology, process-based modeling, mixed forest, competition, ~~biodiversity diversity~~, overyielding, drought, ray-tracing, French Alps

- Maximum water shortage

(p. 39 fig B.7) how was this defined? Why not use the more traditional REW (relative extractable water) used in the water balance model? or any other unit, but please justify it.

A justification of the choice for this variable was added:

412 ration rate and ~~water shortage level~~ [maximum water shortage](#) (defined as the
413 maximum difference reached during simulation between the current and full use-
414 ful reserve, in mm). ~~The NBE~~ [We chose the maximum water shortage because,](#)
415 [in comparison to the relative extractable water \(REW\), it is expressed in absolute](#)
416 [and is therefore independent of the site depth.](#) NME was tested against the null
417 hypothesis using two-sided Wilcoxon signed rank test.

- Appendix “Computing Leaf Mass per Area”

I found this paragraph to be rather confuse. Please clarify and be careful with the scales: LMA is a leaf trait (mean values may be computed for sections in the canopy). LAI is a canopy related trait.

LMA is a property of individual leaves; how is it then integrated to crown level? is it used to compared species, or the degree of dominance of a given species?

Thank you for your comments. We have completely restated the Appendix B.1 in order to be more accurate and avoid any possible confusion:

674 Appendix B. Supplementary description of PDG-Arena

675 Appendix B.1. Computing ~~of~~ Leaf Mass per Area

676 The Leaf Mass per Area (LMA) is a ~~key physiological parameter defining~~
677 ~~leaf-level trait defined as~~ the mass per unit area of leaves (g/m^2). LMA varies
678 both in time during leaf growth and in space: leaf mass gain is indeed favored by
679 ~~the light level~~ local irradiance, resulting in an exponentially decreasing distribution
680 of LMA across the canopy from top to bottom. In the CASTANEA model, which
681 assumes that the stand is homogeneous and monospecific, the LMA ~~decay follows~~
682 ~~an exponential distribution according to an attenuation coefficient $kLMA$ for~~
683 ~~each species:~~ follows an exponentially decreasing function (Davi et al., 2008a):

$$LMA(LAI_{above}) = LMA_0 \times e^{\underline{kLMA \times LAI_{above}} - \underline{kLMA \times LAI_{above}}} \quad (\text{B.1})$$

684 LAI_{above} is ~~given by the position of the considered layer within the canopy~~
685 ~~The average LMA within a layer is then obtained by integrating $LMA(LAI_{above})$~~
686 ~~within the layer vertical boundaries~~ the Leaf Area Index that accounts only for the
687 leaves in the canopy above the considered leaf. LMA_0 and $kLMA$ depend on
688 the species and describe the decrease in LMA within the canopy, which itself
689 depends on the decrease in light intensity within the canopy. Then, the average
690 LMA within a layer is obtained by integrating $LMA(LAI_{above})$ within the layer
691 vertical boundaries.

692 In the case of ~~the PDG-Arena model~~, the canopy is more structurally complex
693 than in CASTANEA and can include several species ~~with different $kLMA$~~ . ~~Then,~~
694 ~~the LMA of each crown is defined according to its position within the global~~
695 ~~canopy.~~ The LMA at a given position of a tree is defined taking all trees into

696 account and using the same ~~equation as B.1. Here, formula as in~~ Equation B.1.

697 LAI_{above} is computed ~~as the sum of the LAI from the different crowns by counting~~
698 ~~only the leaves of the canopy~~ that are located above the considered ~~layer of~~
699 ~~leaves~~ leaf. It should be noted that the model is not completely accurate given
700 that the parameter $kLMA$ ~~is species-dependent~~ and LMA_0 are those of the
701 species of the considered leaf, although the leaves taken into account in LAI_{above}
702 potentially come from another species. However, this method does represent the
703 phenomenon of light attenuation which is specific to each individual.

Answer to the review by Harald Bugmann (reviewer)

- Name of the model

(note that it is not explained anywhere why this is "Arena").

Thank you very much, we had not realized this omission. We added an explanation for the name of the model :

100 Here we present ~~PDG-Arena~~ a new individual-based ~~and~~ process-based ~~and~~
101 forest growth model, PDG-Arena (the arena represents the stand, a place where
102 trees compete and more generally interact). Our model was developed to observe

- Discussion about the simplifying assumptions in the model

The development of PDG-Arena is described coherently and in a transparent manner. Yet, there is some discrepancy between the plea of the authors that mechanistic, ecophysiological-based formulations need to be sought, and the fact that multiple simplifying assumptions are made that are countering this endeavor somewhat; for example that trunks are ignored for the calculation of radiation although it is known that they can absorb a considerable fraction of the sunlight, and are also important for backscattering; or the assumption that there is no differentiation between trees (in terms of sizes or species) regarding rooting behavior, which likely has strong implications for water competition in real stands. Undoubtedly any model needs to be based on simplifying assumptions. Some of them are addressed in the Discussion section (e.g. the rooting problem), but they are not put in context with the results of the study, unfortunately. Others are not addressed at all.

The models are made up of a number of simplifying hypotheses, and it is always difficult to choose, for the benefit of readability and comprehension, those that can be presented in the manuscript and those that refer to previous publications. In line with your expectations, we have therefore chosen to go further in their presentation. To answer these discrepancies, we thus developed further the paragraph discussing the pros and cons of process-based models:

71 spiration, phenology, etc.). For some authors (Fontes et al., 2010; Cuddington
72 et al., 2013; Korzukhin et al., 1996), process-based models ~~because of their~~
73 ~~supposed greater versatility,~~ seem more relevant for simulating ecosystem func-
74 tioning undergoing climate change because they can theoretically be applied to

75 a larger range of environmental conditions than empirical ones. As a result, they
 76 now play an important role in research into the ~~functioning and predicting of~~
 77 ~~forest ecosystem dynamics (Gonçalves et al., 2021)~~ ecophysiological functioning
 78 and prediction of forest dynamics (Gonçalves et al., 2021; Barbosa et al., 2023)
 79 . However, process-based models are more difficult to parameterize and rely on
 80 more assumptions about the ecological functioning of forests (e.g., the hypothesis
 81 that growth is primarily driven by photosynthetic activity, Fatichi et al., 2014).
 82 When it comes to ~~simulate~~ ~~simulating~~ mixed stands, models that simulate ele-
 83 mentary processes ~~theoretically have a better ability~~ ~~are expected~~ to reproduce
 84 the mechanisms that lead to interspecific interactions, bringing us closer to un-
 85 derstanding them (Forrester and Bauhus, 2016).

Concerning the radiation interception by trunks, we forgot to mention it in the original preprint. Even though the interception by trunks is ignored during the SamsaraLight process, it is simulated later in PDG-Arena by retrieving a fraction of the absorbed radiation by the soil:

274 After interception by a crown, the ray continues its course until it reaches
 275 either a new crown or a ground cell to which the remaining energy of the
 276 ray is transmitted. A proportion of absorbed radiation ϵ is uniformly removed
 277 from soil cells to represent the light extinction from trunks, assuming a random
 278 arrangement of trees:

$$\epsilon = 1 - \exp\left(-\frac{\sum_i TS_i}{S}\right) \quad (6)$$

279 where S is the stand area and $\sum_i TS_i$ is the sum of the trunk shade surface of
 280 individual trees. TS_i depends on the DBH and height of each tree i (supposing
 281 a cylindrical shape of the trunk), as well as on the hourly sun angle $\beta(h)$:

$$TS_i = DBH_i \times \frac{height_i}{\tan(\beta(h))} \quad (7)$$

Furthermore, in the new version we now discuss the implications concerning the absence of phenotypic plasticity of the crowns shape and size, as well as the absence of competition for nutrients:

551 ~~no crown plasticity could emerge in the simulation~~ ~~Jucker et al. (2015)~~ Canopy
 552 packing is commonly decomposed into two mechanism: the phenotypic plasticity
 553 of the shape and size of crowns and the vertical stratification (i.e. the occupation
 554 by crowns of different vertical strata). Although it is likely to play a role
 555 in the functioning of mixed stands (Pretzsch, 2019; Dieler and Pretzsch, 2013),
 556 phenotypic plasticity was not implemented yet in PDG-Arena. Thus, our model
 557 can only simulate the vertical stratification of crowns, but not their morphological
 558 adaptation to their local competitor (see, for example, Jonard et al., 2020 and

559 [Morin et al., 2021](#)), potentially leading to an underestimation of overyielding.
560 The observed overyielding in the French National Forest Inventory for beech-fir
561 mixtures (20%, [Toïgo et al., 2015](#)) is greater than the one we simulated. In
562 addition to canopy packing, the real-life overyielding in mixed stands can also
563 be explained by reduced competition for nutrients. Indeed, nutrient content in
564 above-ground biomass and the nitrogen concentration of leaves are likely to be
565 increased by species mixing ([Richards et al., 2010](#)). However, competition for
566 nutrients was not integrated in PDG-Arena since its main objective was to build
567 an individual-based model upon the physiological processes that already exist in
568 [CASTANEA](#).

We also discuss the consequence of the absence of interspecific differentiation concerning the zone of water uptake:

569 In addition, species mixing increased the yearly water shortage ~~due to in-~~
570 ~~creased transpiration (Figure C.9) Figure C.9) at equivalent LAI.~~ This confirms
571 the idea that the nature of the diversity-functioning relationship in forests strongly
572 depends on the limiting resources ([Forrester, 2014](#)). According to our simula-
573 tions, promoting diverse stands could maximize light interception ~~[Jucker et al. \(2015\)](#)~~
574 ~~and growth~~ but would also increase transpiration, which would be detrimental
575 in ~~water-stressed sites. The use of an individual-based and sites with limited~~
576 ~~water reserves. In reality, an increase in water use in mixed stands could be~~
577 ~~counter-balanced by a reduced competition for water between trees of different~~
578 ~~species (Schume et al., 2004). Although an interspecific differentiation between~~
579 ~~the water uptake depth has been observed for some species (Schwendenmann et al., 2015)~~
580 ~~, our model can not simulate this mechanism yet. A comprehensive knowledge of~~
581 ~~each species water uptake depth is still in construction but could be integrated in~~
582 ~~process-based model such as PDG-Arena, in combination with the measurements~~
583 ~~of physiological traits in mixed stands could help better understand the relationship~~

584 ~~between tree diversity, stand productivity and resistance to water stress. models~~
585 ~~in the near future (Bachofen et al., 2024). Concerning the horizontal distance of~~
586 ~~tree water uptake, few data exist at the moment. The assumption of horizontally~~
587 ~~homogeneous water uptake in our model is justified by the small surface area of~~
588 ~~the simulated plot.~~

- Date of the inventory and coring of trees

There are larger worries regarding the data used for benchmarking the simulations. The description of the data is not very clear. I presume that an initial inventory of the 39 plots was made in 1996 (this is why the simulations are starting in 1996), but no further inventories were executed (this is not clarified anywhere). It then appears that the growth of the trees was analyzed using dendrochronology (l. 213 [273?]). It would be important to know whether all 1177 stems (l. 277) were

cored indeed (this would be a huge amount of work!), or whether this was done only for a subset of the trees; but this is not stated anywhere.

It is true that our description of the inventory and data collection was not sufficiently exhaustive. The preexisting data set contains only one inventory per plot and was performed from 2014 to 2016. During the same period, part of the trees was cored to infer the growth of the 1996-2013 period. We've rewritten this section to explain this and how we extrapolated the growth to non-cored tree:

309 2.2. Data set

310 ~~The simulations were evaluated at plot scale using dendrochronological data~~
311 ~~obtained on beech, fir and beech-fir stands from the French pre-Alps. To evaluate~~
312 ~~the simulations, we used an existing data set~~ (GMAP forest plot design, Jourdan
313 et al., 2019, 2020) ~~. The data set includes composed of 39 plots of 10-m~~
314 ~~radius beech, fir and beech-fir plots sampled between 2014 and 2016. Plots are~~
315 distributed on three sites ~~from the French pre-Alps~~ (Bauges, Ventoux, Vercors) ~~as~~
316 ~~described in Table 2, and represents the annual growth dynamics, which are~~
317 ~~described in~~ Table 2. ~~They consist in a 10 m radius area in which the position,~~
318 ~~height, crown base height, age, diameter and species of each tree with a DBH~~
319 ~~greater than 7.5 cm were collected once.~~

320 ~~Out of 1177 stems, 731 were cored to assess the growth dynamics over the~~
321 ~~18-year period 1996-2013. Wood volume increments are obtained by multiplying~~
322 ~~the individual~~ (Jourdan et al., 2019). ~~Growth of non-cored stems was inferred on~~

323 ~~the assumption that basal area increment over basal area was constant for a given~~
324 ~~species and site. To be comparable with the model output, basal area increments~~

- Bias in the computation of the wood volume increment

More importantly, tree height was assumed to be constant (l. 208), i.e. no height increment was simulated over 17 years (1996-2013), whereas growth was assumed to be BAI times tree height, which induces a strong bias in the simulation results as well as in the data (assuming that woody increment was calculated from the data using the same assumption - this is not stated anywhere, though). But if height was constant indeed, why not compare BAI (obs) against BAI (sim)? This would be a much more straightforward comparison (let's not even think about issues such as form factors etc.).

In this article, we chose to infer the wood volume increments from the basal area increments in order to get closer to the model output. As the model is carbon-based, its output can be converted into volume gain without needing strong assumption. However, there is indeed a gap between volumes and basal areas. As pointed out by the reviewer, using fixed height to compute the tree volume increment during an 18

year period is insufficient. Thus, we have therefore modified the past tree heights, which are now calculated by using a height-DBH relationship based on the data set (on the year of inventory):

324 species and site. To be comparable with the model output, basal area increments
325 were converted in wood volume increments. To do that, we inferred past tree
326 heights by using values of past DBH and the relationship between measured
327 height and DBH. Past DBH were reconstructed using basal area increments ~~by~~
328 ~~each tree height. Finally~~ and measured DBH. Then, a model was fitted on trees of
329 the same species and site to evaluate the relationship between measured height
330 and DBH (see Appendix A). This model was used to compute past height based
331 on reconstructed past DBH.

The height-DBH relationship is specified in Appendix A:

669 **Appendix A. Height-diameter relationship**

670 For each group of trees of the same species and site, a linear model (Equation A.1)
671 was fitted on the logarithms of their measured height (in m) and DBH (in cm)
672 as shown in Figure A.5. The slope and intercept parameter a and b as well as
673 the coefficients of determination r^2 are shown in Table A.5 for each group.

$$\log(\text{height}) = a \times \log_{10}(\text{DBH}) + b \quad (\text{A.1})$$

Table A.5: [Parameters of the height-DBH model described in Equation A.1.](#)

| Site | Species | a | b | r² |
|-------------------------|-------------------------|----------------------|-----------------------|-------------------------------|
| Bauges | Beech | 0.69 | 0.33 | 0.78 |
| Bauges | Fir | 0.81 | 0.065 | 0.86 |
| Ventoux | Beech | 0.62 | 0.31 | 0.62 |
| Ventoux | Fir | 0.72 | 0.097 | 0.81 |
| Vercors | Beech | 0.78 | 0.13 | 0.87 |
| Vercors | Fir | 0.83 | 0.033 | 0.90 |

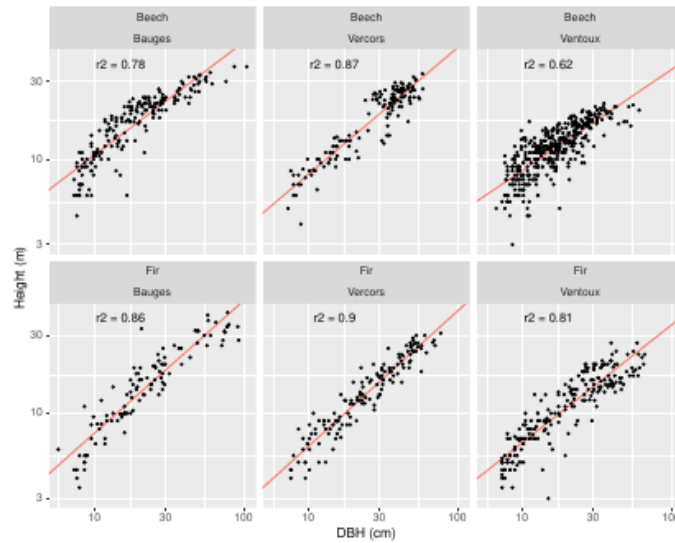


Figure A.5: [Relationship between measured height and DBH. The red line indicates the model fitted on logarithmic values.](#)

Moreover, we integrated species-specific form factor coefficients to take into account the non-cylindrical shape of the trees:

332 [Wood volume increments were computed by multiplying each tree basal area](#)
 333 [increment with its inferred past height and \$\Phi\$, a form factor coefficients which](#)
 334 [takes into account the non-cylindrical shape of the trunks \(Deleuze et al., 2014\)](#)
 335 [. On the one hand, PDG-Arena was evaluated using wood volume increments at](#)
 336 [individual scale. On the other hand, we used the wood volume increments per](#)
 337 ~~stand to evaluate the simulations~~ [at stand scale to evaluate both PDG-Arena and](#)
 338 [CASTANEA.](#)

- Performance of the models

The simulation results nicely show the patterns hypothesized by the authors. Yet, they left me concerned. On the one hand, the effects induced by moving from CASTANEA to the three simulation studies with PDG-Arena (using three

assumptions re. stand structure) are really very minor (Tab. 3 of the manuscript), with the r^2 always being >0.965 . Hence we are talking about minute effects here.

Indeed, these results are very close. So as not to be ambiguous, we now highlight in more details the similarity of the two models in the discussion:

509 We showed that PDG-Arena was able to reproduce the behavior of CAS-
510 TANEA when simulating regularized inventories with no species interactions.
511 Thus, the increase in complexity of PDG-Arena, ~~made necessary~~ required in order
512 to simulate the functioning and interactions of distinct trees, was not at the cost
513 of decreased performance at stand scale. Even when using original inventories
514 (i.e. integrating the diversity in structure and species), the stand scale ~~results~~
515 of PDG-Arena were very correlated to those of CASTANEA. This is explained
516 by the fact that both models are based on LAI, which remains identical for each
517 stand between simulations. Still, PDG-Arena, in comparison to CASTANEA, ~~is~~

On the other hand, the performance of any of the simulations appears poor to very poor (Tab. 4 of the manuscript) compared to measured (?; cf. above - what was actually measured, and how?) wood volume increment. Clearly PDG-Arena has the best performance, but this does not exceed an r^2 of 0.5 and does not go below an error of 34%, which is clearly not satisfactory. Thus, the question arises whether these results can actually be trusted. It appears that both models would need to undergo additional development to be able to represent stand growth over a period of 17 years accurately, particularly since the initial situation was taken exactly from the measurements (inventory).

It depends on what is meant by “being trusted”. We agree that the overall performances are not sufficient for, for example, instruct a local forest management plan. We think this is a limit of most process-based forest models developed to operate over wide areas. Surely the results of such a model is not to be taken at face value, but only as a tool to help explain observed patterns or to anticipate changes, based on the hypotheses contained in the model.

Furthermore, the results are strongly dependent on the input data as well as on the reference data. As described in the following section, we performed new simulations using LAI from a remote sensing source. The resulting shift in the model performance highlights the dependence of the simulation to the input data, in particular the LAI, which is hard to obtain (a fortiori retrospectively). Also, the simulations would probably have been more reliable if they were tested against carbon flux measurements. Unfortunately, flux tower installations are rare on mixed stands.

Therefore, the objective of this work is not to have the best growth prediction in absolute but to evaluate the relative effect of switching from the stand scale to the individual scale in terms of performance and functioning. This is why the focus of the study is mostly on the comparison between the CASTANEA and PDG-Arena models and between the types of inventory used with PDG-Arena. That said, we acknowledge this limit of our results. Thus, we have chosen to remove the word “accurately” from the conclusion, which was not consistent with what the results showed:

613 **5. Conclusion**

614 ~~The~~ new individual-based model PDG-Arena we developed ~~can accurately is~~
615 ~~able to~~ simulate the interactions between trees in monospecific and mixed stands
616 and predict their productivity ~~based on an explicit tree inventory~~. Compared

In order to highlight the ability of PDG-Arena to simulate the growth of individual trees, we added an evaluation of PDG-Arena at the tree level. Here are the concerned sections:

New sentence in the “model evaluation” of the Results section:

398 variability. ~~Additionally, PDG-Arena with O inventories was evaluated at the~~
399 ~~individual scale, by computing the r^2 and MAPE of the simulated versus measured~~
400 ~~wood volume increment per tree for each group of the same site, type of stand~~
401 ~~(beech, fir or mixed) and species.~~

New paragraph in the Results section:

471 Figure C.8 ~~show the simulated versus measured wood volume increment at~~
472 ~~the tree scale using PDG-Arena and original inventories (O). The r^2 ranged from~~
473 ~~20% to 64% depending on the set of trees, with a mean at 47%. The MAPE~~
474 ~~ranged from 50% to 146%, with a mean of 82%. (Table C.7).~~

New paragraph in the Discussion:

526 ~~The performance of both CASTANEA and~~ ~~At the individual scale,~~ PDG-
527 Arena ~~at predicting the variability of fir stands productivity remained poor ($r^2 <$~~
535 ~~data necessarily did not capture~~ ~~explained half of the variability of tree growth,~~
536 ~~showing that it can capture the competitive status of each tree based on their~~
537 ~~leaf surface, height and position. However, the mean absolute error was often~~
538 ~~large and systematic, indicating that the model lacks calibration for each site.~~

- Use of LAI values from remote sensing

In that same context, the authors mention that LAI data are from the recent drought period (l. 216 and 418) whereas the inventory data are from nearly 30 yrs ago. This may explain the divergence between LAI values of ca. 3 (single-sided, I presume) and the very high BA values (ca. 50-60 m²/ha), which are hard to reconcile and may lead to strong inconsistencies in a model that is driven by radiation, and thus LAI.

Thank you for that pertinent comment. We have taken it into account: considering the discrepancy between the growth period (1996-2013) and the LAI measurement

period (2022-2023), we chose to use values of LAI based on remote sensing data for a period that fit with the measured growth. The description in the manuscript has been modified in accordance:

246 The ~~attenuation coefficient k depends on species, radiation domain, type of~~
 247 ~~radiation (direct, diffuse) and beam height angle. Its value is determined using~~
 248 ~~reverse engineering of SAIL, the radiation sub-model in CASTANEA, as described~~
 249 ~~in Appendix B.2. stand LAI was retrieved using each plot coordinates and the~~
 250 ~~1 km resolution SPOT/PROBA-V remote sensing data set (Baret et al., 2013)~~
 251 ~~. We computed the average value of the yearly maximum LAI observed over the~~
 252 ~~1999-2013 period. During the radiation balance computation, LAI are at their~~
 253 ~~maximum. However, a fraction of the absorbed radiations per tree is removed~~
 254 ~~daily depending on their current phenological state (see Appendix B.4).~~

This modification in our protocol substantially changed the simulation results, but the pattern is conserved, with PDG-Arena being overall better than CASTANEA (except for fir stands):

Table 4: Evaluation of the performances of PDG-Arena and CASTANEA on the 37 stands. Coefficient of determination (r^2 , in %) and Mean Absolute Percent Error (MAPE, in %) were computed for the simulated versus measured yearly wood volume increment per stand over the period 1996-2013. Inventories are characterized as: 'RNRM' (regularized ~~with no and monospecific, i.e. without~~ species interactions), 'RSR' (regularized, but with species interactions) and 'O' (original inventories).

| Set | Model | Inventories | r^2 | MAPE |
|------------|-----------|-------------|----------------------|----------------------|
| All stands | CASTANEA | RNRM | 17.6-29.5 | 41.0-40.6 |
| | PDG-Arena | RNRM | 18.4-32.1 | 43.0-40.5 |
| | PDG-Arena | RSR | 19.0-32.5 | 43.2-41.8 |
| | PDG-Arena | O | 20.9-34.2 | 40.5-40.4 |
| Mixed | CASTANEA | RNRM | 40.2-36.3 | 36.4-30.1 |
| | PDG-Arena | RNRM | 40.3-37.6 | 37.8-30.7 |
| | PDG-Arena | RSR | 43.1-36.3 | 38.9-33.1 |
| | PDG-Arena | O | 50.1-40.5 | 34.1-31.5 |
| Beech pure | CASTANEA | RNRM | 22.0-22.9 | 53.1-55.3 |
| | PDG-Arena | RNRM | 21.6-25.0 | 51.6-57.4 |
| | PDG-Arena | RSR | 21.6-24.7 | 51.9-57.9 |
| | PDG-Arena | O | 36.2-38.3 | 47.0-53.9 |
| Fir pure | CASTANEA | RNRM | 7.8-42.0 | 41.5-33.7 |
| | PDG-Arena | RNRM | 12.5-51.9 | 38.5-29.6 |
| | PDG-Arena | RSR | 11.5-50.1 | 37.8-30.4 |
| | PDG-Arena | O | 12.9-39.8 | 40.0-33.0 |

Also, the use of higher LAI leads to more differences between the PDG-Arena and CASTANEA simulations using RM inventories, which is now explained as follows:

419 3.1. Comparison of ~~the simulation modalities~~ PDG-Arena and CASTANEA
420 Using ~~regularized inventories with no species interactions (RN)~~ regular and
421 ~~monospecific inventories (RM)~~, CASTANEA and PDG-Arena showed similar pre-
422 dictions for the stand-level GPP, ~~as represented in Figure 3. The~~ with a coefficient
423 of correlation ~~between the two models was estimated at 99.6%. Moreover, as~~
424 ~~shown in Table 3~~ at 99.8%. However, the GPP simulated by PDG-Arena was
425 ~~in average 4.2% greater than that of CASTANEA (Figure 3). As shown in Ta-~~
426 ~~ble 3, which compares the 4 modeling situations based on the coefficient of~~
427 ~~determination, correlation, simulations from~~ PDG-Arena was closer to ~~those of~~
428 CASTANEA when using regularized ~~stands and when species interactions were~~
429 ~~disabled~~ inventories (R) on the one hand and when using regularized monospecific
430 ~~inventories (RM) on the other hand.~~

It is not clear whether LAI was set to a constant across the simulation time, using the 2022/2023 values. This would be a highly debatable assumption.

It is true that the yearly maximum LAI was fixed during the simulation (now from SPOT/PROBA-V 1999-2013 period). However, the LAI does have an intra-annual variability inside the model simulation. The variation of LAI firstly affects the physiological processes but it also affects the absorbed radiations, as described in Appendix B.4.

- Alternative measurements

Importantly, one aspect that would be key for simulations across multiple years with a dynamic model is not addressed in the results shown in this study at all: it appears that all that is tested is diameter increment (and indirectly, volume increment using a static tree height), but not the development of any other tree characteristics such as height, the height of the crown base, leaf mass or leaf area, let alone allocation to belowground compartments (for which admittedly there would most likely not be any data). It would have been highly insightful to compare the simulated stands at the end of the inventory period (2013 - was there really no inventory carried out at that point in time?) against measurements. Focusing the comparison on simulated GPP among models and woody increment between models and a tree ring reconstruction leaves many questions unanswered.

It's a pertinent remark and we agree that evaluating the model using measurements related to the other tree compartments would have been highly valuable. However, the data set on which the study is based does not contain repeated measurements about tree height or leaf area. Therefore, we only evaluated the growth using yearly tree ring measurements and the retrospectively estimated tree heights. We are sorry it is not possible to add more data for the evaluation.

- Description of the virtual inventory types

l. 289: Although I am intrigued by the approach taken here to generate synthetic stands that systematically miss certain aspects (compared to reality) such as species or species interactions, I found the description very difficult to follow and am not really sure I have understood. The setup O is straightforward, and in the setup RS tree positions were "regularized" (not sure this is a proper English term, but I understand). But what is the essence of the difference between RS and RN? This remains elusive to me, even after repeated reading of l. 293-297. Pls re-consider.

To be clearer, we've tried to simplify things, especially the nomenclature. We renamed the RS, RS and O inventory sets as RM, R and O types. RM for regularized monospecific inventories, R for regularized inventories (that can be, or not, multispecific), and O for original inventories. Also, we have rewritten the description of inventory types as follows:

353 2.3. Simulation plan

354 Using field inventories, we generated three sets of virtual inventories for
 355 PDG-Arena, following three levels of abstraction, denoted ~~RN, RS~~ RM, R and
 356 O. The first set represents regularized ~~inventories with no species interactions~~
 357 ~~(RN monospecific inventories (RM))~~: for each species of each stand, we generated
 358 a new inventory with equally spaced trees of the same species, age, diameter and
 359 height. ~~The For mixed stands, the~~ simulation results using ~~regular monospecific~~
 360 ~~inventories generated from the same stand were then~~ RM inventories were assem-
 361 bled relatively to the proportion of each species basal area. ~~RN~~ RM inventories
 362 can then be used to simulate the growth of multispecific stands ~~while~~ ignor-
 363 ing species interactions. The second set represents regularized inventories ~~with~~
 364 ~~species interactions (RS)~~: (R), in which trees of different species can coexist but
 365 trees of the same species share the same age, diameter and height. ~~Plus, trees~~
 366 Trees in R inventories are regularly spaced in a random order, independently of
 367 the species. Lastly, original inventories (O) include the information of the real life
 368 ~~dataset~~ data set, that is: species, position, diameter and height of every individual
 369 trees. For each type of inventories representing the same stand (regularized or
 370 not, with or without species interactions), the mean quadratic diameter, volume
 371 per tree and tree age per species and the basal area were conserved.

Answer to the Review by the anonymous reviewer

- Competition for nutrients

1) Competition for water and light is implemented, but there is no competition for nutrients, even if the authors themselves mention it in the introduction as being important (line 45). Shortly mentioning why this has not been done (and maybe possible consequences?) would be helpful. Maybe either in relation to mentioning ‚limiting resources‘ (line 436), or when discussing possible improvements of the model (lines 451-465)?

Indeed, this is a point that would have been interesting to explore. However, it's a complex task that wasn't part of the objective of this study and would probably merit a dedicated article. A section of the discussion is now dedicated to the absence of competition for nutrients and its consequences (see the section of this document “Discussion about the simplifying assumptions in the model”).

- Global change

2) What is meant by global change? (e.g. line 49)

We refer to the rapid changes in the planetary system associated with the Anthropocene, including climate change, nitrogen pollution, land use and biodiversity loss. We added a reference to the article of González de Andrés (2019):

1. Introduction

1 Understanding how forest ecosystems function is a crucial step for develop-
2 ing forest management strategies adapted to the challenges of ~~global change,~~
3 ~~particularly~~ climate change (Bonan, 2008; Lindner et al., 2010; Trumbore et al.,
4 2015) and more generally global change (González de Andrés, 2019). In this

- Modelling of the competition for water

3) If there is limited water available, is all water divided equally among the trees? Or is water uptake connected to transpiration? In other words, how does competition for water take place? (section 2.1.2)

That is a good question. Transpiration is computed based on the absorbed radiation and the stomatal conductance. When water becomes limiting, stomatal conductance is negatively affected through a reduction factor, which will stop the transpiration. PDG-Arena follows the same process, but since the individual transpirations are simulated before the stand soil water balance, it can happen that trees transpire more than what the soil can provide. In this case, the reduction factor on transpiration will take effect during the following water balance, which is the next day. This part should

be reconsidered to prevent trees from transpiring if water is not sufficient. Yet, since the error is caught the next day, this should not significantly affect the results.

4) There is no vertical or horizontal differentiation of soil water availability (lines 160-164). For the lack of horizontal differentiation an explanation is added, but not for the lack of vertical differentiation (although it is discussed in the discussion). I suggest either discussing/mentioning both here (lines 163-164), or neither (and discuss both in the discussion).

The discussion has been enriched with a paragraph about the consequence of this simplification. The justifications for the lack of horizontal and vertical differentiation are now in the same paragraph (see the section of this document “Discussion about the simplifying assumptions in the model”).

- Net Biodiversity Effect

5) Calling mixing two different tree species „biodiversity“ is a bit of a stretch. After all, one can think of many more factors being of importance for biodiversity that all contribute to „maintaining key ecosystem services“ (line 5-6). I’m fully aware that it is difficult to quantify biodiversity or to somehow distill it out of model results and I agree that using the number of different tree species is a good first step/approximation, but I would be careful with the terminology. I propose to change the ,net biodiversity effect (NBE)‘ to net mixing effect, or something along those lines, or otherwise clearly state that number of tree species is used as a proxy for biodiversity.

Originally, the term “Net Biodiversity Effect” was used in biodiversity-functioning experiments that took place in grasslands. Later on, it was also adopted in several diversity-functioning study in forests (for example, <https://onlinelibrary.wiley.com/doi/full/10.1111/ele.12357> and <https://www.nature.com/articles/s41559-020-01329-4> and <https://academic.oup.com/jpe/article/10/1/158/2966831>). However, we agree with the criticism and renamed it Net Mixing Effect. Additionally, we changed in the manuscript every mention of “biodiversity” to “diversity” or “tree species richness”.

6) Since the paper introduces a new model, I think it would be nice to shortly discuss the limits/opportunities of the model. For what kind of setups can it be used? What are the size limitations of the plots? Can it be used for more than two different tree species? Can trees completely cover each other (undergrowth)? Etcetera

We agree with this suggestion and added a description of the model entry in the conclusion:

613 5. Conclusion

614 The new individual-based model PDG-Arena we developed ~~can accurately is~~
615 able to simulate the interactions between trees in monospecific and mixed stands
616 and predict their productivity based on an explicit tree inventory. Compared
617 to CASTANEA, PDG-Arena showed improved predictive capability for beech and
618 mixed beech-fir forests. The model can simulate the growth of small-sized stands
619 (less than 1 ha), of regular or irregular structure, and composed of trees of similar
620 or different species (given that the species are parametrized in CASTANEA). As
621 PDG-Arena simulates the competition for water and light between trees with no
622 preconceived ideas about the direction of interspecific interaction (from compe-
623 tition to complementarity), it can be used to test specific hypotheses about
624 mixed forests and better understand the diversity-functioning relationship in
625 forests under contrasted scenarios. For example, ~~one could explore the following~~
626 ~~outstanding~~ the model could be used to explore following open questions, keep-
627 ing in mind that the answers are largely ~~dependent on the species identities~~
628 ~~(Ratcliffe et al., 2015) and on each resource scarcity in a given environment~~
629 ~~(Forrester et al., 2017a)~~ species-specific and environment-dependent (Ratcliffe et al.
630 : is overyielding more likely to occur in less productive sites ? (Toigo et al., 2015)
631 (Toigo et al., 2015)? Can overyielding increase water stress in mixed stands ?
632 ~~(Forrester et al., 2016)~~ (Forrester et al., 2016)? Are mixed stands more resilient
633 to drought (Grossiord, 2018)? Lastly, being ~~made-built~~
634 demo-genetics model, PDG-Arena is suitable to evaluate the evolutionary dynam-
635 ics of functional traits of a population under various biotic (stand composition,
636 density and structure) and abiotic (soil, climate) constraints, as intraspecific di-
637 versity is a major adaptive force in natural tree populations (Lefèvre et al., 2014;
638 Oddou-Muratorio et al., 2020).

Additional modifications

- Management of outliers

Having performed new simulations with an alternative source for the value of LAI, we have excluded the hypothesis that poor LAI measurements could explain the discrepancy between simulated and measured values for silver fir plots in the Bauges. Therefore, we decided to exclude 2 outlier plots from the performance analysis:

431 3.2. ~~Modeling-Model~~ performance

432 The simulated versus measured stand wood volume increment for the 39
433 stands are reported in Figure C.6 for the CASTANEA model using RM inventories
434 and in Figure C.7 for the PDG-Arena model using O inventories. Two fir stands

435 from the Bauges site, denoted haut_sp_2 and bas_sp_4, stand out from
436 the point cloud, with measured growths of 1995 and 1562 cm³/m², while the
437 simulated growth did not exceed 973 m³/m² for CASTANEA and PDG-Arena.
438 Simulations using values of LAI measured in 2022 using Terrestrial Laser Scanning
439 (unpublished data from one of the author, C. Rouet) were done and showed the
440 same discrepancy with growth measurements for these two stands. As their
441 inclusion in the analysis affects the overall results, these stands were discarded
442 from the following analysis (see Table C.6 for the performance analysis that
443 includes all stands).

For the sake of transparency, the performance results with the entire plot set is given in appendix of the document:

Table C.6: Evaluation of the performances of PDG-Arena and CASTANEA without discarding outliers. Coefficient of determination (r^2 , in %) and Mean Absolute Percent Error (MAPE, in %) were computed for the simulated versus measured yearly wood volume increment per stand over the period 1996-2013. Inventories are characterized as: 'RM' (regularized and monospecific, i.e. without species interactions), 'R' (regularized, but with species interactions) and 'O' (original inventories).

| Set | Model | Inventories | r^2 | MAPE |
|------------|-----------|-------------|-------|------|
| All stands | CASTANEA | RM | 25.7 | 41.7 |
| | PDG-Arena | RM | 26.5 | 41.6 |
| | PDG-Arena | R | 26.4 | 42.8 |
| | PDG-Arena | O | 24.0 | 41.7 |
| Mixed | CASTANEA | RM | 36.3 | 30.1 |
| | PDG-Arena | RM | 37.6 | 30.7 |
| | PDG-Arena | R | 36.3 | 33.1 |
| | PDG-Arena | O | 40.5 | 31.5 |
| Beech pure | CASTANEA | RM | 22.9 | 55.3 |
| | PDG-Arena | RM | 25.0 | 57.4 |
| | PDG-Arena | R | 24.7 | 57.9 |
| | PDG-Arena | O | 38.3 | 53.9 |
| Fir pure | CASTANEA | RM | 18.0 | 38.4 |
| | PDG-Arena | RM | 24.8 | 34.9 |
| | PDG-Arena | R | 23.7 | 35.6 |
| | PDG-Arena | O | 19.1 | 38.6 |