Dear Editor

Please find enclosed the preprint entitled “**Combining statistical and mechanistic models to identify the drivers of mortality within a rear-edge beech population”, by** Cathleen Petit-Cailleux, Hendrik Davi, François Lefevre, Christophe Hurson, Joseph Garrigue, Jean-André Magdalou, Elodie Magnanou and Sylvie Oddou-Muratorio. This is the revision of the preprint previously assessed by Lucía DeSoto (as recommender) and three PCIEcology reviewers on the 13th of August.

We thank the recommender and the three reviewers for their detailed comments and suggestions on our manuscript. We provide here a thoroughly revised version of our manuscript taking into account most of these suggestions (changes are highlighted in blue in the main text), with a particular focus on the six main points raised by the reviewers and the recommender (in green below).

1. All reviewers raised that methods should be revised in order to better analyze data and fit the models. Also description of the variables used should be reviewed to make them well-defined. For instance, in Table 1 the time (and frequency) when the variables were recorded should be included.

First, methods were revised and/or completed. The statistical model for mortality at population level was thoroughly revised and simplified (see answer to point 2 below). Regarding the statistical model at individual level, we present in the new Appendix 4 an alternative approach (survival analysis) which account simultaneously for both levels of variability (individual and temporal) in our data set. Note that this approach gave the same results as the logistic regression model. Moreover, we found survival analysis less reliable than logistic regression due to departure from a major assumption. For these two reasons, we kept the logistic regression model in the main version of the manuscript (see detailed answer to reviewer 2).

Secondly, the variables used in the different methods were better described and defined. In particular, we added a new Figure 1 showing how the different variables were accounted for in the different approaches. We revised Table 1 as suggested.

2. One of the reviewers was concern about the population-level models and suggest the authors to focus on individual models, I agree with her and also think that the study lacks of replicates for population. Authors achieved results that may be only reliable for their specific population, and thus locally appealing.

In the revised version of the manuscript, we have tempered down the emphasis on the statistical model at the population level. We rather put forward two novelties of our study, which are (1) the demonstration of a significant relationship between the combination of response variables simulated by the PBM CASTANEA and the observed mortality rate at population-level and (2) the combination of statistical and PBM approaches at individual level which allow a better understanding of the factors modulating individual vulnerability to mortality.

To improve the population-level analyses, we removed the lengthy description of variables originally included in the population-level statistical model, and the automatic variable selection procedure; we rather selected a small number of variables (SPEI) based on ecological hypotheses, as suggested by reviewer 2 and 3.

3. Results are difficult to follow because of the large number of variables used. For instance, other reviewer made the helpful suggestion of using a table that should be not dismiss.

The methodological changes significantly reduced the number of variables used in the manuscript. In addition, as suggested by Reviewer 3, we have added a new table (included as the supplementary Table S2) to summarize the results. We hope that these changes make the "results" section easier to follow in the revised manuscript.

4. All reviewers agree that the introduction is fine but asked for more detailed discussion of the conceptual framework. Particularly, I found that introduction and discussion are vaguely interconnected.

In the revised manuscript, we better formulated the main two issues of the study at the end of the introduction, and structured the discussion section following these issues. We hope that the new Figure 1 also allows a better presentation of the conceptual framework of this study, and that the new Table S2 strengthened the connection between the introduction and discussion.

5. As authors aimed "to investigate patterns of mortality within a population located at the warm and dry ecological margin of European beech", I wondered about the expected results of the study and its explicit hypotheses.

Detailed hypotheses on the impact of individual variables on mortality are now presented in Figure 1 and at the end of the introduction.

6. Finally, the size of the numbers legends of the figures should be larger.

The size of the numbers in the figures have been enlarged in all the figures of the main document and in their legends, as well as in most supplementary online figures.

In addition, we join below a detailed answer to all reviewers' comments. We thank Lucía DeSoto, Lisa Hülsmann and the two anonymous reviewers for their comments on our manuscript, as well as for their numerous suggestions which helped us a lot during the revision process. The revised manuscript is ~ 7980 words long, excluding the 96 references, and it includes 4 figures and two tables, as well as supplementary materials and appendices. All authors have read and approved the material being submitted. This article is not being considered for publication elsewhere.

With kind regards,

Sylvie Oddou-Muratorio

Detailed answers to reviewers' comments

Reviewer 1 - anonymous

This manuscript uses statistical and process-based models to study mortality of Fagus sylvatica at individual and population scale. The main results are that drought and late frost events were the major environmental causes of mortality at the population scale, whereas loss of hydraulic conductance and carbon starvation were the physiological stresses connected to mortality. In individual level, high mortality was related to low growth rate, high crown defoliation, infection by a fungi, high competition, high loss of hydraulic conductance, higher frequency of late frosts, and early (statistical model) or late (process-based model) budburst.

The topic is important and the approach is very promising. However, it is also challenging to report in a clear and structured manner due to its complexity.

This still needs improvement before the manuscript is ready for publication. Also, there are several (small) technical mistakes and the language should be improved to be fluent.

Petit-Cailleux et al.: We have added to the main text a summary figure of the analyses carried out (Figure 1), as well as a summary table of the results that can be found in the appendix (Table S2).

We have also thoroughly revised the manuscript, and used the ReadyToPub service of to improve English langage (see the certificate enclosed).

I have one more general comment on the paper, and the detailed comments you find at the end of this letter. Statistical models were used to quantify the effects of climate, competition, tree size and health on mortality, and process-based models to find the physiological mechanisms among carbon reserves, hydraulic conductance and late frosts that explain these effects. I like the idea, but my largest confusion is that why did you not write open and discuss the processes that the PLC, carbon starvation and frost damages affect in the simulation model, and what is their role in the simulated mortality (not only measured mortality)? I think this needs to be clearly presented.

Petit-Cailleux et al.: We thank reviewer 1 for this remark, which show that some clarifications were needed regarding our approach.

We added on Figure 1 and in the Material and Methods section (L307-309) details on how climate (the input variable in CASTANEA) drive the evolution of the number of frost days (NLF) and frost damages, of photosynthesis and of xylem cavitation, which ultimately affect the evolution of the Biomass of Reserve (BoR) and PLC. Note that NLF, BoR, PLC are dynamic response variables simulated by CASTANEA, but that mortality itself is not simulated by CASTANEA in this study. Indeed, simulating mortality would require the knowledge of the thresholds in PLC, BoR and NLF triggering mortality. These thresholds could be inferred using CASTANEA and an inverse modeling approach, but it was out of the scope of this study (a sentence has been added in the discussion on this topic, L597-599).

Rather, in this study, we focused on the relationship between the simulated response variable, and the observed mortality rate (using simple correlations). This is also why we computed a composite vulnerability index integrating PLC, BoR and NLF: this index is not of direct use in the PBM, but it allows combining the different response to stress into a single predictor of mortality.

Also, I would be interested to see the number of late frost days also in the statistical model as it is mainly a climatic variable (although taking into account the timing of bud burst).

Petit-Cailleux et al : In the revised version of the manuscript, we clearly distinguished on the one hand the purely climatic variables (e.g., SPEI), which can directly be computed from climatic series, and on the other hand the tree response variable, which require accounting for individual tree physiology. Although the number of late frost days may looks like a climatic variable, it is indeed a response variable, as the sensitivity to frost in beech is maximal during burdburst (particularly at the stage where leaves spread out). This is why this variable is not included in the population-level statistical model (equation 8 in the revised manuscript).

Detailed comments of reviewer 1:

Comments from the reviewer (page refers to the first manuscript)			Reply (lines refers to the revised manuscript)	
Comment number	**Page**	**Comment**	**Reply**	**Lines**
Abstract and introduction				
36-37	2	"…mortality decreased with mean growth, and increased with crown defoliation…" Please write it open that mortality decreased with increasing mean growth and increased with increasing crown defoliation?	We rephrased as: "the individual probability of mortality decreased with increasing mean growth, and increased with increasing crown defoliation… "	42-43
40	2	Please explain why the frequency of late frost days is studied with the process model? Materials and MethodsMaterials and Methods In general, I think it is important to explain what are the main processes in the model affected by these variables (frost days, PLC and carbon starvation) – even if it could be found from earlier publications. This is essential to understand the studied mechanisms behind the mortality.	See the reply to the reviewer 1 's main comment	307-309 + Appendix 2
40	2	"earlier trees" does not mean anything, please write "individuals with earlier bud burst" or something similar.	replaced everywhere	
73	3	Remove the word "vessel"	we replaced "vessel" by xylem	77
128-132	5	I suggest not to use the term "iconic" here.	replaced by "important"	143
Materials and Methods				
151	5	You have listed "decline" as a monitored variable. But decline of what?	It is the decline of tree measured by the crown defoliation and the presence of fungi. It is now better stated	175
225	8	I think the index that was counted only for trees >3m away from the target tree is Competintra not Competintra+. Please check this.	We thank the reviewer to point out this mistake. We corrected it	Table 1
Table 1	8	In the table caption, you refer to tables 1A and 1B, but there is no subtables A	We added (a) and (b) to table 1	Table 1

		and B. So add A and B to the table or remove from the table caption.		
Table 1	8	How can Competintra minimum and maximum values be larger than those of Competintra+, although the latter one includes more trees (also trees closer than 3 m from the studied tree)? Please check this.	We thank the reviewer to point out this mistake. We corrected it	Table 1
240	8	You use altogether seven R packages (If I calculated correctly). This shows clearly the complexity of the paper.	We agree with reviewer 1 , although we think that complex analyses are required to tackle these questions	
254	9	Don't you have one to three variable per category (the equation below)? Please check this. Also all the equations should be numbered (eq. 1); please correct this throughput the whole manuscript.	We changed our method to evaluate the impact of climate on mortality rate. We numbered all the equations	345-361
259-263	9	Check the abbreviations of all variables in this paragraph, there are confusions. In the equation, you don't have MeanTDriestQ, but you have MeanTColdestQ. Also, you do not have SPEI_maxn in the equation at all. The sentence is also confusing, please rewrite.	We changed our method to evaluate the impact of climate on mortality rate.	345-361
278-279	10	You have competition index and DBH both in the same model. Competition index is derived from DBH, so they are highly correlated. This could have been avoided by using e.g. height as a variable describing tree size.	We agree with reviewer 1 that height might have been a better variable related to size to use in this model, but height was measured only in a subset of trees (~1200). We added in appendix 3.4 a model using height, all conclusions remain the same. Note also that the correlation between variables was accounted for by computing the variation inflation index, which showed that these correlations weakly affect our results (see Fig S5).	Appendix 3.4 and Fig S5
303	10	You need to write what is Oddsnormal in the equation; it is not given.	We added the formula for of Odds Normal.	390
308	11	Write "coarse roots"	Replaced	283
	11	I do not find information on what are the input parameters for the simulations? There are some climate variables, but which? Please write open clearly.	We added the climatic input variables in the climatic variables description section.	271-273
352	12	It is difficult for me to understand why you need such an index variable for a	See the reply to the reviewer 1 's main comment	

		process-based model, when you already include all the variables in the model that constitute the index. If there is a combination effect of these stresses, it should be taken into account in the model with interactions between processes.		
Results				
366	12	Please refer here to fig. 1. Based on the figure, the mortality rate is as high in year 2007 as it is in year 2010. Also in year 2004, the mortality is nearly as high.	We agree with reviewer 1, although the peak in 2007 may be a lagged effect of high mortality in year 2006. As for year 2004, we do not have mortality data on year 2003 which was one of the hottest and driest years recorded in France. We added in Appendix a table with the mortality rates values (table s1) .	table s1
370	12	Be careful with the variable abbreviations. Check from page 9 which abbreviations you used, and use them consistently (SPEI3*JJAn instead of SPEI*JJA; PDriestM instead of PdriestM).	All the abbreviations in the manuscript were checked.	
372	13	What does it mean in practice that the relative mortality rate was decreased by 1.15 and 3.03 times? The likelihood of mortality increased 1.15 and 3.03 times per how big change in SPEI_JJA and PDriestM?	We have changed this analysis and removed this part.	
397	13	The combined index did not correlate at all with observed mortality in year 2007. On that year, the mortality rate was high, but the index did not catch it. This should also be stated and discussed.	This is now explicitly stated on L 410 . We also added a comment L532-533 of the discussion	532 and 533
Fig. 1.	14	The mortality rate on the second y-axis is not percentages, but shares instead (multiply with 100 to get percentages). Also, use the same scale for the mortality in all sub-figures (now the second figure has different scale than the others). Also, I wonder why it is not shown how the different physiological stresses correlated with the simulated mortality (it only shown how they correlated with measured mortality)? I think it would be informative to show also this as it would show how severe is the effect of each physiological stress to the simulated mortality following the	As detailed above, we did not simulate mortality with CASTANEA. On the different panels of figure 2, the variations in observed mortality rates (in yellow) are the same. This is now stated in the legend. These mortality rates are now plotted as percentages (and not share).	Fig. 2

		processes and their connection built-in in the simulation model.		
410	15	You can remove "(i.e. dead individuals truly observed dead)" as this is obvious.	We removed it.	
412-413	15	See my comment related to page 13, line 372.	Done, we modified the sentence to make it more understandable: "Defoliation had the strongest linear effect on mortality: the relative probability of mortality increased by 1000 times for a one-unit increase in DEFw"	433-435
438-440	16	How about years 2009 and 2016 that were also good years?	Yes, they were good years too. We have made this clear. "The magnitude of the individual effects on tree vulnerability differed during a drought year (2006), a frost year (2010) and a good year (2008, 2014 or 2016)"	455-456
444	17	It is difficult to understand why large trees would have lower reserve of carbon than small trees. Can you explain this a bit more?	We added a sentence to explain this point	462-465
Discussion				
484	18	It does not make sense to give so many decimals in the mortality rates by Archambeau et al. (2019).	Corrected	493
507-508	18	Please check the end of this sentence; there is something wrong.	Rephrased:	510-512
522-528	19	Please rewrite this part. It is really difficult to follow.	Rephrased	538-542
535	19	I think plants often defoliate to prevent xylem embolism, not because of it (or probably both). Shedding the leaves helps to avoid cuticular transpiration.	We agree with reviewer 1 that this sentence was confusing and we removed it. The interactions between cavitation, defoliation and carbon starvation are discussed further in this paragraph.	549-552
589	21	I don't understand what is meant by saying that due to so many recurrent dry years the benefit of defoliation becomes negligible. Can you please write it open?	The sentence has been deleted and the discussion now focus on the difference in the defoliation in the observed data and in the simulations with CASTANEA	601-610

Reviewer 2 - Lisa Hülsmann

In this manuscript, Petit-Cailleux et al. explore and quantify drivers of tree mortality in *Fagus sylvatica* at the individual and population level using two complementary approaches: (1) fitting statistical models with inventory data from a rear-edge beech population, and (2) applying the mechanistic model CASTANEA to the same site.

As the authors nicely synthesize in their introduction, disentangling the multiple interacting factors that cause tree mortality and appropriately integrating them in dynamic vegetation models are key issues for forest ecology and climate change adaptation. This makes their study an interesting and timely contribution to the field of tree mortality research.

I have, however, methodological and conceptual concerns that restrict the validity of the current conclusions and would like to make several suggestions how the authors can increase the overall significance of their manuscript.

The following choices are, in my opinion, critical and can cause unreliable results: The authors (1) fit a beta-regression model for the population-level mortality rate that has only 14 observations but 28 predictors, while the data points are not even independent because of temporal autocorrelation,

(2) carry out strong model selection (for both empirical models) and then interpret p-values, which is statistically inadmissible,

and (3) unfortunately collapse their - in fact very valuable - dataset in two ways: on the one hand to population-level mortality rates and keep the temporal variability, and on the other hand to long term mortality probabilities but keep individual-level variability.

Regarding the conceptual framework, I would like to encourage the authors to discuss in more detail how they think mechanistic models should be employed to disentangle the drivers of tree mortality, taking into account that models can only reflect those processes that are implemented in the simulation framework and that the uncertainty for model parameters is typically very large.

In the following, I point out in more detail why I believe the analyses should be improved, what the authors could do instead and where more context and discussion is needed.

Petit-Cailleux et al.: (general response 1) Although we agree with Reviewer 2 on several points, we think that our approach was appropriate considering our data set. In the revised version of the manuscript, we attempted to better develop the hypotheses underlying our approach and highlight its limits, in regard with other possible approaches.

On the collapsing of the data set

First, we agree that collapsing a data set made of observations repeated across years at individual level would be a pity. However, in our data set, most variables at individual level (except for mortality and defoliation) were measured only once over the study period (see the revised Table 1) while climatic variables were only available at population level (as a single value for all the trees, every year). So aggregating the annual crown defoliation marks into the cumulated variable DEFw, and the annual observations of mortality into a probability of mortality over the period 2004-2016 was essentially a way to harmonize the temporal resolution of the variables available in this study. The main strength of this study is the large number of individuals measured, not the temporal survey of mortality and its drivers.

On other statistical approaches

Searching for other statistical approaches than regression models, we carefully considered survival analysis and Bayesian hierarchical models.

Survival analyses first appeared as an appropriate way to analyze the variation among individuals in temporal patterns of mortality. Note that, in our case, survival analysis typically consists in analyzing the relationship between the time that elapsed before mortality occurred on the one hand, and the measured covariates that may be associated with that quantity of time on the other hand. This is not exactly similar to measuring the effect of each covariate on the risk of mortality, as presented in the main part of the manuscript. We tried to use survival analyses to analyse our data set (see new online Appendix 4). However, we were confronted to a major departure from classical assumptions of survival analyses. Indeed, survival analysis following the classical cox model assume proportional hazards, ie the unique effect of a defoliation mark equal to 1 instead of 0 would for instance double the probability of mortality at every year t. However, the effect of defoliation on mortality is likely to vary among years depending on climate, and this cannot easily be taken into account in survival analyses.

Bayesian hierarchical models could also be an appropriate framework to model individual fixed effects on mortality related to the measured covariates, individual random effects on mortality due to unmeasured covariates (ie., genetic variability) and fixed population effects on mortality related to climate. However, we think that this framework ideally requires temporal series of both the response variable, mortality (which we have) but also of its drivers (which we have only for defoliation, since all the other driving factors were measured only once). We could not find an existing statistical package, software or routine that could be easily applied to our data set and account for these specificities. Of course, we could have contacted a biostatistician to develop our own Bayesian model, but we felt that this was out of the scope of our study,

These methodological issues are now discussed in the last part of the discussion.

On model selection and p-value

We think that there is a misunderstanding here. Model selection was not performed based on p-value, but on based on AIC in a Stepwise Algorithm. Then, once the model selected, effects were compared and tested based on the p-value, which is a very standard statistical approach.

Regarding the limitation of process-based model

We completely agree with L Hülsmann that process-based models can only reflect those processes that are implemented in the simulation framework. We revised the manuscript to better highlight which processes are accounted for and which are not (see new Figure 1).

Statistical mortality models

Problems of the beta-regression model of annual mortality rates for 14 years:

- The model can test for climate and competition effects only, while it is known that large trees may react differently to drought than small trees.
- The number of observations available for the population-level model is very small (n=14). The result is that the authors carry out a very strong model selection on 28 predictors. For predictive models, this may be a reasonable approach, but for inference this is a clear example of p-hacking. In this case, the discrepancy between available observations and tested predictors is so large that the resulting model is basically useless.
- The model doesn't account for temporal autocorrelation.

Petit-Cailleux et al.: (general response 2) We agree with these limitations (note that for the beta-regression model, only climate effects could be tested and not competition effects). Our initial objective was to test whether climatic predictors of mortality at population level were consistent with population-level predictions from the PBM CASTANEA. However, we agree that we face power limitations with the population-level statistical model. Hence, in the revised version of the manuscript, we removed the model selection.

Problems of the logistic regression for individual tree status at the end of the observation period:

- The authors don't make use of the 6 months resolution of the mortality surveys. This is unfortunate because, in my opinion, the high temporal resolution is the biggest strength of their dataset. This would also allow to disentangle the various drivers of mortality, which is actually what the authors promised in their introduction.

Petit-Cailleux et al.: See detailed response above. Note that the dataset is not a true 6-month monitoring of mortality. Even though mortality is surveyed twice every year (in spring and autumn), this corresponds to a single notation because mortality is "suspected" in autumn and "validated" in the early spring (by checking that suspected dead tree does not budburst).

- The model can test for the effect of tree characteristics and competition only, but not of climate.

Petit-Cailleux et al: We agree with L Hülsmann that climate effects cannot be included in the individual model, because we measured only the average climate at the study site, and assume that all individuals experience the same climate.

To improve the reliability and significance of their empirical models, I suggest that the authors (1) fit only one mortality model that keeps both levels of variability (individual and temporal) using half-annual mortality probabilities of individual trees as a response variable,

(2) jointly test for the effect of climate, competition and tree characteristics including also seasonal (winter/summer) effects and lag effects of drought and frost,

(3) adequately control for temporal autocorrelation, and

(4) choose a well-defined set of predictors based on ecological hypotheses so that model selection is unnecessary.

Petit-Cailleux et al .: (general response 3) Regarding points (1) and (2), as detailed in our general answer above, we do not know any available statistical approach which could easily account for both the annual effects of climate at population-level (based on a single measure for all individuals) and the mean effects of individual characteristics (measured as a cumulated/average value over years) on the risk of mortality at individual level.

(3) We are not sure to understand what temporal correlation relate to in our case. For instance, in the population-level model where we analysed the relationship between annual mortality rates and annual climate, the temporal autocorrelation of climate is accounted for implicitly (e.g. the temporal succession of two dry years). By contrast, lagged effects of climate are not considered (eg., effect of drought at year n-1 on mortality at year n), and would indeed be difficult to account for considering the low number of observations. Moreover, with a series with such a small number of points (14), it is difficult to detect a temporal autocorrelation. Finally all diagnostic statistic show a high validity and a good goodness-of-fit of the models.

We followed the suggestions of L Hülsmann for point (4), as detailed in the revised version of the manuscript L.343-362.

Mortality at the tree or the population level

The authors carry out both approaches at the tree and the population level. While I agree that population-level mortality rates are intuitive and helpful for reporting trends, I doubt that they are an ideal level for analyzing mortality drivers in inhomogeneous populations. Basically all mortality drivers may have different effects on small versus large trees. In addition, mortality rates can be derived from individual-level mortality probabilities. I suggest that the authors concentrate on individual mortality probabilities in both approaches.

The only case where I find the focus on mortality rates helpful is when comparing observed mortality rates against the proxies of physiological stress. I would see this as a rough test that the model can reflect those stresses that are relevant for observed mortality.

Petit-Cailleux et al . : We agree with L Hülsmann and re-structured the manuscript in this sense (see general answer to the editor).

Using mechanistic models to disentangle the drivers of mortality

The authors state that they aim to “decipher the respective roles of the drivers and mechanisms underlying tree mortality and understand their variability among individuals or years” (see discussion). To this end, they apply the process-based model CASTANEA and analyze three stress indicators: percentage of loss of conductance (PLC), biomass of reserve and number of late frost days. To test if these stress indicators are more or less realistic, they correlate them against the mortality rate. I think this should be okay, but the authors could do better in explaining their approach, the underlying assumptions and discussing its limitations.

Specifically, I am missing some critical notes on the limited number of mortality causes that are implemented in process-based models in the introduction and discussion. Using models to identify drivers of tree mortality can only reveal those processes that have been implemented in the models, while processes that are missing, e.g. competition and pathogens, cannot be reflected. The authors should also comment on how reliable the model parameterization is, i.e. how much they trust the simulation results.

I am not very familiar with the details of CASTANEA. From how the authors set up the analyses of the model simulations, I thought CASTANEA doesn’t explicitly simulate tree death, but physiological indicators only. In Line 562 however, the authors mention that CASTANEA cannot simulate “background mortality”. Maybe this refers to the inability to account for stresses due to resource competition? If CASTANEA also provides mortality probabilities, I think it would be great to make a similar individual-based mortality model for the model output, to compare how mortality is driven in the model compared to the observations.

Petit-Cailleux et al . : We thank L Hülsmann for these comments which helped us a lot revising the manuscript, and improve the presentation of CASTANEA model.

CASTANEA can simulate "background mortality" based on the self-thinning law and drought-related mortality from carbon starvation and hydraulic failure. Note that competition is modeled as the result of average stand density and not in a fully spatially explicit way. Note also that Online Appendix 2 present detailed result on the calibration and validation of CASTANEA.; actually, calibration was based on the comparison of simulated and observed ring width. Finally, in the simulations carried out here, we preferred to compare observed mortalities with a new index of composite vulnerability. In this natural reserve, the population is uneven, so the simulation of "background mortality" from the law of auto-clearing is not relevant. In addition, we do not know the NSC thresholds associated with mortality.

Detailed comments of reviewer 2:

Comments from the reviewer (page refers to the first manuscript)			Reply (lines refers to the revised manuscript)	
Comment number	Page	Comment	Reply	Lines
INTRODUCTION				
28	2	Consider rephrasing: "an exceptional annual monitoring of 4327 individual European beech trees". Is it annual or every 6 months?	See our answer above. We added details on the date of measurements (Table 1).	Table 1
29	2	Consider rephrasing: "We combined two types of approaches to analyze beech mortality: (1)…, and (2)…"	Replaced as suggested	35-37
36	2	"In the statistical models"	Replaced as suggested	40-41
38-39	2	Consider rephrasing: "The interaction between tree size and defoliation was significant, indicating a stronger increase in mortality probability due to defoliation for small than for tall trees."	Replaced as suggested	44-45
40	2	Unclear what "earlier" trees are.	We replaced "earlier tree" by "tree with early budburst" across the whole manuscript	
75-76	3	Consider rephrasing: "which can eventually become depleted, particularly…".	Replaced as suggested	81-82
97	4	Remove second time "reflush".	Removed	
108	4	Typically, empirical models are not for mortality rates (also not those cited here), but for mortality probabilities. Please correct.	Corrected	115,120
118	4	I don't really understand, why non-linearity is a problem, at least not as an inherent problem of empirical mortality models. Empirical datasets from inventories are typically quite large and allow testing rather complex functional forms including non-linear relationships. It is true that in some studies this possibility is not sufficiently explored, but in my opinion, this is not really a disadvantage of empirical mortality models. There are other more important limitations: low temporal resolution of data, reason of death is often not known (btw: is information on this recorded for the site used in this study?).	We agree with L.Hulsman, and tempered down this problem in the revised version of the manuscript. In this study, the mortality cause was recorded when obvious (lightning, fallen tree, etc) but not investigated.	125-127

128	5	The next two sentences sound contradictory. I get the point, but the phrasing can probably be improved.	Rephrased	135-139
132	5	To what kind of individual effects to you refer here? Be more precise.	Rephrased	140-141
139	5	Maybe it would be helpful for the non-European reader to specify where the rear-edges of beech are?	Done	L145 and L148
142	5	Consider rephrasing of the “have been associated in beech to decreased growth”.	Replaced	150-151
147	5	Remove “of”.	Done	
148	5	I really liked your introduction of Fagus sylvatica and why it is interesting to study it!	Thank you !	
151	5	You could mention the location of the site here already.	We added the location	175
Materials and Methods				
163-164	6	“South of beech distribution area” sounds a bit strange.	Rephrased: “ Located in the south of the beech range”	193
166	6	Are the temperature values absolute, daily or monthly minima/maxima?	They are daily minima and maxima reached. We have added some clarifications.	195
170	6	“removed”	Done	
173	6	At which spatial resolution were these additional variables recorded?	At population level	205-206
190	7	Consider renaming to “DBH2002”.	We renamed all the corresponding DBH.	
194	7	There may be several problems here. I think the equation is not correct. First, the radius should be squared. Second, relative basal area is typically a relative measure with respect to the basal area at t = 0: BA1-BA0 / BA0. The equation here is only annual basal area increment.	We thank L Hülsman for detecting this typo in the formula and name of the variable, it is now corrected	222
194 second		Finally, how do you deal with trees that died between 2002 and 2012? DBH measurements of dead trees are very unreliable, e.g. because the bark can fall off.	We cured the data set to remove records with $DBH_{2002} > DBH_{2012}$ possibly due to measurement error or decomposition. We trusted the large number of individual trees to compensate for this data set curation.	
205	7	I don’t understand the meaning of “ratio ordered” in this context.	Rephrased	244-245
219	8	Consider rephrasing “We computed this competition index in three ways:”.	Rephrased	257

220	8	Same “coppice” or same “species”? It may be helpful to introduce the idea and your definition of “same coppice” before using such definitions.	This is now defined at the beginning of the section.	209-210
226	8	Do you mean the “indices” or the “estimated slopes” reached a plateau?	The relationship between distance and competition indices becomes flat (ceiling) . We rephrased this sentence.	262-265
Table 1		Why are the values for COMPETintra+ smaller than those of COMPETintra?	We thank the reviewer to point out this mistake. We corrected it	Table 1
234-235	9	Consider rephrasing “… since 1976 and 1960 for temperature and precipitation/mean relative humidity, respectively”.	Rephrased as suggested	268
242		Check sentence structure.	Rephrased	Appendix 1
244-245	9	Does the number of days with negative temperature occurring after beech budburst account for the budburst date of individual trees?	The date of budburst of individuals is not precisely known. Note that we changed our model as a result of the evaluators' suggestions and we no longer take the variable into account.	
263	9	Remove second “is”.	Done	
265	9	I guess you mean predictors not factors.	Exactly, we rephrased.	370
275	10	What do you mean with qualitative explanatory variables? Categorical variables? I find this terminology a bit uncommon.	We changed qualitative variables by categorical variables.	372
294-296	10	True, but then you wouldn’t do model selection at all.	We used variable selection to remove the non significant ones.	
297	10	I think that odd ratios can only be compared if the continuous predictors in the model are scaled.	The glm function automatically scales the data before fitting the model. We stated it now explicitly	375
324	11	I am missing a sentence on the aim of using CASTANEA including which information is reported from the models and why. Or this should be outlined in “Simulation design”.	We added sentences in “Simulation design” to make explicit the aims of these simulations:	329-330, 341-342
330	11	Check the suitability of the verbs used. Estimate, calculate, compute and simulate don’t mean the same thing… I would also start with the equation of PLC and then explain its components. Starting with leaf water potential after the sentence on PLC and defoliation seems not intuitive.	The suitability was checked and we change the order of paragraphs as suggested.	282-328
349	12	What do you mean with “stresses without a priori”?	Rephrased	339-340

356-357 and 359	12	Consider rephrasing “randomly drawn from a normal distribution” and “randomly drawn from a uniform distribution”.	We rephrased and added details in the appendix 2	Appendix 2
Results				
367	13	Reference to Fig. 1 would be helpful.	Done (note that now, this is Fig.2)	396
371	13	It is Table 2 not 1.	Removed	Appendix 1
373	13	Why should the effects of these predictors only apply to some years?	Removed	
385		The results you present here are not only at the population level. You also report the range of individual PLC estimates. I think this is no problem, but it indicates that your differentiation between individual and population is not ideal.	We rephrased these results to highlight population-level response (because individual response is detailed in the last section of the results)	399-404
385		If you talk about an increase, you should mention from which level PLC has increased. The same for carbon reserves and late frost days in Line 388 and 391.	Rephrased	399-404
393-394		Consider rephrasing to “with 5 to 9 days and 5 to 7 days of late frost, respectively”.	Rephrased	399-404
396		Do you correct for temporal autocorrelation in the test?	See our general response 3 to L Hülsmann	
408		Significance cannot be interpreted when model selection was carried out beforehand (either automatic or manual).	See our general response 1 to L Hülsmann	
410	16	“dead individuals truly observed dead” sounds like the true positive rate. If this is the case, you should also report the true negative rate. If this is variance explained, your explanation in brackets is maybe not fully correct.	We added the true negative rate.	Appendix 3
412	16	Check again if continuous predictors were scaled.	They were scaled	
426		It is not the interaction that is true for low mean growth, but the difference in mortality probability between small and large trees was only evident for low mean growth.	Rephrased	447
437-441		It is unclear to which results you refer here. A reference to a figure or table is missing.	We added the reference to figure 3b.	446
Figure 3		The combination of color, symbol and filling is very confusing. Reduce degree of information or make more subpanels for key comparisons.	The figure was redrawn to avoid confusion.	Figure 4

		You may also replace years with “severe drought”, “no stress” and “late frost”.		
Discussion				
475-476		It is not surprising that a lot of variation is explained because the model assumptions are violated (temporal autocorrelation) and the number of observations is very small. In addition, the dataset is very homogenous compared to Greenwood et al.	Except for the leverage effect, all diagnostic statistics show a high validity and a high goodness-of-fit of the models	Online Appendix 1 and 3
480	18	Mortality rates always need to be reported in the context of the dbh threshold, because if more small trees are included the mortality probability will likely increase. This makes comparisons of rates with different dbh thresholds very difficult. You should account for this when comparing your mortality rates with previous publications.	This comment is now accounted for in the discussion	L500-502
497	19	Please specify what you mean with positive interaction.	Removed.	
506	19	Carbon starvation and hydraulic failure are typically debated in the context of drought. It is not clear how frost goes in here… You may reconsider the wording of “stresses vulnerability”.	We rephrased systematically to make explicit the stress we are talking about when mentioning “vulnerability”.	
512	19	Why couldn’t you account for the temporal dynamics across years?	See our general response 1 to L Hülsmann	
521	20	Be specific about which approach revealed a result. “Confirmed” is actually very strong. We know that empirical studies often don’t replicate.	Rephrased	539-540
519	20	I am missing some ecological explanations here.	We did not really understand this remark	
534	20	When talking about an effect, it is helpful to always mention on what the effect takes place (here mortality I guess).	Rephrased	549-550
545	20	Use “individuals with early budburst” instead of “early individuals” throughout the whole manuscript.	Done, throughout the whole manuscript.	
546	20	Vulnerable to what?	Rephrased into “prone to die”	559
575		You may exchange “interest” with “potential”.	We replaced as suggested.	589
581		To me, this complementarity is still vague. How do you suggest both approaches should be combined? Doing both and then compare results? Or something more integrative?	We add some details on the way to compare/combine the both approaches	596-601

591-592		The results of the population-level empirical model are not at all reliable. I am still wondering, how you can test the effect of early budburst in the population level model?	The population-level model has been revised. Moreover, the test of the impact of early budburst on mortality is performed only at the individual level.	226-229

Reviewer 3 - anonymous

Despite finding that this paper is a strong contribution, I have several comments which I think might help improve the quality of the manuscript.

1. I have had some difficulty understanding all results; some methods were not sufficiently developed in the Appendices in my opinion. For example, I could not find out why the Nstem competition index was chosen; I did not understand that from the data and explanations provided in Appendix 3.

2. For the choice of variables to use in the statistical models, two alternative approaches could have been used: (i) either rely on the PCA coordinates, to reduce the number of dimensions while not just choosing one variable correlated to each axis;

or (ii) make a model comparison, by using a full set of models (not only the stepAIC procedure, which I have sometimes found to identify really suboptimal models, as compared to testing all possible models); eg using R packages MuMIn or glmulti.

The added value would be that all possible models could be tested (or an intelligent subset of these models). Also, I would have found it useful to use climatic variables more related to those used in the process-based model (or at least, to present the results of the process-based model – e.g. the number of days with late frost.

3. Also, in my opinion, it is rather difficult to compare the results of both types of models. I think that a table looking like the one below would really help understand (well, I am not sure the data in the table is fully accurate, but I needed to make it to understand the results).

There are many abbreviations, probably a table mentioning all of them would be useful.

Petit-Cailleux et al . : We thank reviewer 3 for these useful suggestions. In the revised version of the manuscript:

1. we added in Online Appendix 3 a new table describing the selection procedure in order to allow comparison between models, and understand why some variables were not retained in the final model.
2. We added a new supplementary table summarizing the results as proposed by reviewer 3 (Table S2)
3. We completed Table 1 to give more information on the variables, and added the new figure 1 to better highlight the connections between them and their use in PBM and statistical models.

Additionally, in the course of the revision, we tried to use glmulti as suggested by reviewer 3. However, the model never converged, probably because of the combination of a large number of factors, and low number of observations.

Detailed comments of reviewer 3:

Comments from the reviewer ("Line" refers to the first manuscript)	Reply ("Line" refers to the revised manuscript)

Line	Comment	Reply	Line
	MATERIAL AND METHODS		
193-194	I did not understand the equation on lines 193-194 ; I think it should read MBAI MBAI i = π $DBH2012^2i - DBH^2i$ /(4N yearAlive, i).	We thank the reviewer for mentioning this error in the formula	222
Table 1	I did not understand why the max value for Compet Intra+ was inferior to the max value for Compet intra. Also the usit for MBAI should be cm2.year -1 (not cm -2 .year -1).	We thank the reviewer for mentioning these errors in the table.	Table 1
242-244	some more explanation on SPEI would be useful (do high values of SPEI indicate a wet year/period?)	We added some details in the revised manuscript"	275-280 Appendix 1
	At the very end of the methods section, an explanation is missing as to why these specific simulations were run.	Done.	329-330, 341-342
	Results		
373-377	needs more explanation. Indeed, you state that 2009, 2011-2016 experienced a winter drought; yet the driest quarter is also the warmest quarter in 2011, 2013, 2016. Maybe a graphic would help at that point?	We do not use this model anymore in the revised version of the manuscript	
391-394	does this apply to the late or early trees?	The ranges presented included trees with early and normal budburst. The simulation at population-level included 100 trees, add 10% of trees with early budburst (respectively 90% with late budburst)	Appendix 2
426 and Fig 2b	First word of the line should be "higher" not "lower". In the dataset, MBAI is always <2.8 cm2/yr (Table 1). Thus, how can this interaction be extrapolated to trees that grow 10 times faster than what is observed?	We thank reviewer 3 for detecting this error in Table 1; actually, MBAU varies between 0 and 95 cm^2/yr in the data set.	Table 1
443-444	I do not follow: On Fig 3 right panel, early trees seem to show higher reserve than normal trees? and the trees that do not defoliate also have higher reserves?	We changed the figures to make them easier to read. Individuals with early budburst have higher biomass of reserves (BoR) because they have a longer vegetative season. Trees able to defoliate have lower BoR because they photosynthesize less.	Fig. 4
	Discussion		
553	On the contrary, on Fig S6 it seems that early trees have higher carbon stocks than normal	We usually prefer not to apply statistical test on simulations resulting from	

	trees. A statistical test would be useful at that point.	mechanistic models. (As stochasticity is not taken into account in CASTANEA, increasing the number of simulation would easily make significant any test)	
Supplementary Information			
Table A1.1	Are values for AnP shown on a monthly basis? Does NLF indicate the number of frost days suffered by the early or by the late leaves?	We thank the reviewer for mentioning these errors in the table. Finally, we did not keep this variables. NLF is the number of frost days suffered by all trees. We calculated it from the mean budburst date of the population	
Section 1	Needs more explanations	We changed the method employed	Appendix 1
Section 3	Needs more explanations	We added some precisions about the model selection.	Appendix 3

Additional requirements of the managing board:

As indicated in the 'how does it work?' section and in the code of conduct, please make sure that:
-data are available to readers, either in the text or through an open data repository such as zenodo (free), dryad (to pay) or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.
Petit-Cailleux et al.: We have added a section "Data availability" L631-632

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.

Petit-Cailleux et al.: We used classical statistical methods, and described their use in the material and methods section. All the R packages used are mentioned in the text.
We have added a section "Supplementary material" mentioning that the process-based model CASTANEA is an open-source software available on capsis website: http://capsis.cirad.fr/

-Details on experimental procedures are available to readers in the text or as appendices.

Petit-Cailleux et al.: We provided all the details concerning the measurements protocols

-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."

Petit-Cailleux et al.: we have added the section "Conflict of interest disclosure"

References:

Dobbertin, Matthias and Peter Brang. 2001. "Crown Defoliation Improves Tree Mortality Models." *Forest Ecology and Management* 141(3):271–84.

Vidal, Jean-Philippe, Eric Martin, Laurent Franchistéguy, Martine Baillon, and Jean-Michel Soubeyroux. 2010. "A 50-Year High-Resolution Atmospheric Reanalysis over France with the Safran System." *International Journal of Climatology* 30(11):1627–44.

AUTHOR
SERVICES PROVIDER

Editing Certificate

This is to certify that the manuscript detailed below was Translated and/or Edited by a Scientific Editor from ReadyToPub and that it was edited for proper English language by an English native speaker or by a professional in English Linguistics.

Manuscript title

Combining statistical and mechanistic models to identify the drivers of mortality within a rear-edge beech population

Authors list

Cathleen Petit-Cailleux, Hendrik Davi, François Lefèvre, Joseph Garrigue, Jean-André Magdalou, Christophe Hurson, Elodie Magnanou, and Sylvie Oddou-Muratorio.

Date issued

2019-11-03

Verification code

PNLPX-0P976-6K633-M02KG-Q7ISX

Please authenticate the service provided at www.readytopub.com/certificate by using the verification code given above.
ReadyToPub does not have any interference with the scientific content presented in the manuscript. Our Scientific Editors advise and recommend major improvements to be included in the manuscript, however, it is of the sole responsibility of the Authors to accept to include those improvements in the final version of the manuscript.
For more information about the service provided please contact our Editorial Team at info@readytopub.com.