Sylvie Oddou-Muratorio
URFM INRAE
Avenue de l'aérodrome
84000 AVIGNON, France
sylvie.muratorio@inrae.fr

Lucia de Soto

Recommender for PCIEcology

Avignon, the 18th of June 2020

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 Petit-Cailleux et al. This is the second revision of the preprint previously assessed by Lucía DeSoto (as recommender) and three PCIEcology reviewers on the 13th of August, and then on the 7th of February 2020.

We thank the editor and the three reviewers for their comments and overall positive feedback on our revised manuscript. We agree with the three main points summarised by the editor (in green below), and we further revised our manuscript to account for these points (changes are highlighted in blue in the main text, downloaded in bioRxiv as Petit-Cailleux_etal_PCIEcology_main_version3_markedup.pdf).

1) To make the study more readable, the introduction should be shortened and focus on the objective of the manuscript.

In the revised version, we shortened the introduction by removing the long paragraph on the mechanisms of mortality in response to drought.

2) The population-level approximation is not convincing, because only one population, and 13 years, were available (consequently 13 data). Furthermore, data are auto-correlated between years. I recommend the authors to remove this model (methods and results) from de main text. It can be kept in the supplementary results if it is needed for the CVI discussion.

We answer below in detail on the issue of autocorrelation. However, we agree with the editor and the reviewer that the beta-regression statistical model at population scale can be removed from the main article. In the revised version of the manuscript, it is included in the new Appendix 2, and used for the CVI discussion.

3) I also agree that inferred models must keep all the explained variables following the rationale of the study. Please, change it accordingly or justify it.

As recommended, we kept all explained variables in the logistic model selected in the revised version. We detailed in the material and method the new model.

I received the reviews of the revision of your manuscript entitled "Combining statistical and mechanistic models to unravel the drivers of mortality within a rear-edge beech populations."

The three reviewers suggests some minor changes that will help to increase the quality of the manuscript. They also highlight some minor issues that should be corrected or discussed.

We enclosed below our detailed answers to reviewers'comments.

We hope that this revised version will be suitable for recommendation by PCIEcology. The present manuscript is not in consideration in any other journals of media.

In behalf of all the authors

Sylvie Oddou-Muratorio

Detailed response to reviewers' comments.

Reviewed by anonymous reviewer, 2019-12-09 13:51

In this article, the authors investigate the causes of mortality of European beech, using a large dataset of yearly measurements of more than 4000 trees in an unmanaged forest located at the rear edge of the distribution range of the species.

They use two types of models

- regression models to assess the relative contributions of drought, competition, growth, size, phenology, defoliation events, pathogenic fungi to individual-level probability of death.

- a process-based model to investigate : i) at the population level, how modelled, climate-driven carbon reserves, loss of conductance, late frosts contribute are correlated to observed mortality (at the population level). They find that none of the three variables is, but that their combination is well correlated to observed yearly mortality. ii) at the individual (tree) level, how differences in size, phenology, ability to defoliate would affect these response traits (and mortality) depending upon climate

This is a revised version of an earlier manuscript. I find it much clearer than previously and I think that the referees' comments were adequately addressed. I have 1-2 minor points but am otherwise really happy with this manuscript (which has really improved upon revision):

- I am not quite sure that equation6 is homogeneous.

PC et al: Thank you to have noticed this mistake in the equation, we have corrected it in the revised version of the manuscript.

- One perspective to include late frosts as a climatic variable, would be to use the variable NLF (as computed from CASTANEA) as a predictive variable in the statistical model.

PC et al: We also thought about this possibility, but we decided not to mix approaches in order to compare the results of different methods. Anyway, this statistical model is not used anymore I the main version of the revised manuscript.

- for the compound vulnerability index, it might be useful to design a multiplicative index (something like PLC * [NLF/max NLF allowing survival] / [BoR / max BoR ensuring survival])

PC et al: Indeed a multiplicative index would allow to take into account interaction effect between the three components of the CVI. However, such a multiplicative CVI would not account for a threshold effect of one component that would lead to mortality. For example with a PLC-value equal to 1 (leading to mortality) and a NLF-value equal to 0 (no vulnerability to frost), the multiplicative CVI value would also be null; this would mean that the cumulated vulnerability is negligible, whereas risk of mortality by hydraulic failure would be quite high.

We explored another way to improve the cumulated vulnerability index, expressing it as the sum of the risk of mortality due to each of the three processes:

$$CVImulti_n = \frac{NLF_n}{max(NLF)} + \left(1 - \frac{NLF_n}{max(NLF)}\right) \times \left(\frac{PLC_n}{max(PLC)}\right) + \left(1 - \frac{NLF_n}{max(NLF)}\right) \times \left(1 - \frac{PLC_n}{max(PLC)}\right) \times \left(\frac{BoR_n}{max(BoR)}\right)$$

with rBoR'=1-rBoR

Testing the correlation between this probabilistic cumulated vulnerability index and the observed rate of mortality for the 13 years, we found a positive correlation of 0.48 that was marginally significant (p-value of 0.09). This suggests that our results are robust to the computation of the coupound vulnerability index.

Reviewer 2 - *Lisa Hülsmann*

Reviewed by Lisa Hülsmann, 2020-01-08 15:38

I would like to thank the authors for their extensive modifications of the manuscript and the additional clarifications that make the findings more accessible. The authors now point out the novelty of their study more clearly and have better explained how the two model types can be combined. The new overview figure and results table (why in appendix? We add it in the main text) and modified introduction lay out the approach and findings more comprehensibly. Finally, the authors have addressed some, but not all, statistical and conceptual issues (see further comments below).

I am still not fully convinced that the approach presented is the best possible: As suggested during the first review round, the authors have fitted a survival model using individual-level mortality data (see Appendix 4). Unfortunately, they use the same set of predictors but include no climatic variability into the model, although jointly testing for individual tree characteristics and climate would be - in my opinion the - biggest advantage of such an approach. I cannot fully understand why the authors conclude that climate cannot be incorporated. Doing so will probably also solve the issue of non-proportional hazards – e.g. when using interactions with climate. Note that I am aware that climate is equal for all trees and that not all individual-level variables are available for every year. A solution could be to simply keep these variables constant or to interpolate between years where applicable. I agree that this is not an ideal solution but still think this is not a reason for collapsing the response data a priori (see previous review round). I think an individual-based model with climatic effects (also in interaction) could make the whole study more meaningful, but I leave it to the authors if such a model should be part of this manuscript.

PC et al: We have tested survival models methods for data violating the hazard ratio assumption and incorporating both climate and year into a single statistical model as suggested by LH. To that aim, we followed the methods used in In and Lee (2018, 2019) and Zhang et al. (2018). However, with frequentist methods climate effect cannot be disentangled from the year effect. We detail below the different models we tested and their conclusion. More details on survival analysis can also be found in Appendix 4.

(1) Simple survival model (as in Online Appendix 4) but adding climate

time_model <- coxph(Surv(Start, Stop, DeadEvent) ~

FungiEvent+ DEF_cumulated + poly(DBH$_{2002}$,2)+ Budburst + MBAI + Nstem + SPEI12_fix + SPEI_dryVg1mt + cluster(Ind), data= longi_table)

With *SPEI12_fix* the long term drought index and *SPEI_dryVg1mt* the short term drought-index.
This model cannot be fitted, because it is impossible to dissociate the year effect (period of time in the surv fonction) from a climate effect. (Of course, we tried also with one climate variable in the model and it gives the same result).

(2) Stratified model:

time_stepmodel <- coxph(Surv(Start, Stop, DeadEvent) ~ Fungi_Event + Budburst + MBAI + Nstem + DEFw+ strata(class_DBH$_{2002}$) + strata(Climate) + cluster(Ind), data= longi_table)

The diagnosis of the model shows again a deviation from the Hazard ratio assumption as in Appendix 3.

(3) It is not possible to fit a glm on this kind of longitudinal data (mortality), due censored data on the right.

In the current state of our knowledge, it is impossible for us to take climate and individual characteristics over time into account in the same frequentist model. Moreover, since the simpler logistic regression model allows us to answer our main question which is "What are the individual characteristics influencing the probability of mortality?, we decided to keep the individual model as it is now and we added to the discussion the perspective and advantages of developping a more spatio-temporal integrated model.

Further, I see two issues that could be improved in the new version:

(1) The focus and the research questions of the study could be more precise. This applies in particular to the abstract but also to the end of the introduction. It may also be helpful to slightly streamline the introduction to better match the research aims (e.g. I don't think that the carbon starvation versus hydraulic failure discussion is that important for the findings). In general, the introduction could be more concise, while the end of the discussion may profit from a few concluding sentences including the main findings and their relevance.

PC et al: As recommended, we reformulated more precisely the research question in the abstract. We also shortened the introduction by removing the section on the causes of mortality in response to drought. We now hope to be more concise. We have also added sentences summarizing the main findings and their relevance at the beginning of the discussion.

(2) The focus on population- versus individual-level processes is my opinion not an ideal way of structuring the findings. Mortality rates are only the result of individual level processes and even if one must collapse the data to a lower resolution (here mortality rates of a population) to identify mortality drivers, this is not an advantage but a compromise. Therefore, I would find it more compelling to structure the findings around the different types of mortality drivers itself rather than the level at which they become effective.

PC et al: We agree with LH and revised the manuscript accordingly. We now structure our finding considering one the one hand models investigating inter-annual variation in mortality rate (i.e., only the PBM CASTAEA in the main text) and on the other hand, models investigating inter-individual variation in the probability of mortality (the logistic regression model and CASTANEA)

Statistical feedback:

I appreciate the new population-level model that tests for predefined influences and thus suffers less from data dredging. However, I would like to mention that every model selection (both manual or automatic) is in general not a good idea if the aim is inference and not prediction, even if more data is available. I know that this is a common approach in ecology, but the conclusions from reduced models are wrong. This is particularly true for p-values which become smaller and more significant in model selection. In case of the individual-level model, p-values are very small so even in the full model the effects are likely significant. Another problem of model selection is that effect sizes and directions may be wrong, if important confounders are removed from the model during the selection process. In conclusion, model selection should be avoided in case of inference and is recommended for prediction only. For more details see e.g. Georg Heinze, Christine Wallisch et al., Variable selection – A review and recommendations for the practicing statistician. Biometrical Journal 60 (2018), S. 431-449 and D. J. Lederer, S. C. Bell et al., Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Ann Am Thorac Soc 16 (2019), S. 22-28.

PC et al: We followed these recommendations for the statistical model presented in the main part of the manuscript (i.e., the logistic model). We made only one exception for the four competitions indexes, which were strongly correlated one with each other (because are built in similar way); hence, for the competition variable only, we used model selection to select the best model with the competition index giving the lowest R^2. We left all the variables and interactions we wanted to test, this did not change our general conclusions. However, we got a minor issue of collinearity between DBH and other variable in interaction with DBH (see appendix Table A3.2).

Time series are typically characterized by autocorrelation, e.g. the value of this year depends on the one of last year. One assumption of calculating correlations and regression models is that the observations are independent, which is not true for time series. Ideally, this should be accounted for when correlating mortality rates and PBM outputs and fitting the population-level mortality model. Note that this is not done with using climate as a predictor. Nevertheless, I agree that fitting autocorrelation is probably difficult considering the rather short lengths of the time series.

PC et al: We agree with LH that time-series are usually characterized by temporal autocorrelation. This why, in the diagnostic of the population model, we paid a particular attention to the residuals patterns, which turned out to be random. Taking into account the warranted doubt of the randomness of the residuals, we tested their temporal autocorrelation with the Durbin-Watson test using the dwtest function from “lmtest” package (Zeileis and Hothorn, 2002). We found no significant temporal autocorrelation in the population model (beta-regression with DW = 1.9759 with a p-value = 0.5035). However, as recommended, we removed the population model from the main article.

Moreover, we also think that it is not always desirable to correct for any possible temporal autocorrelation when studying the correlation between PBM outputs and observed mortality rates. Indeed, in the PBM, what happens in year n-1 has an impact on year n and this is what allows us to understand the impact of stress combinations on the vulnerability of trees.

Comments from the reviewer (page refers to the first manuscript)		**Reply (lines refers to the revised manuscript)**	
Line	**Comment**	**Reply by PC et al.**	**Line**
35	I would argue that some of this is already known but their remains high uncertainty. So maybe relax this statement a bit.	Rephrased: “still need to be better assessed”	36
36	After the intro and research gap sentence, I typically expect a sentence that defines the research aim of the study.	We introduced a new sentence: “This study investigates mortality in a rear-edge population of European beech (*Fagus sylvatica* L.) using a combination of statistical and process-based modelling approaches”	37
38	Which decline is meant here?	We replaced “decline” by “defoliation” (we meant canopy decline)	40
42-43	To me it is not clear why it is important to emphasize that this is a feature of the population level. The processes are expected to occur on the individual level, and I don’t think it is helpful to emphasize the difference here as it may make the reader to assume that the processes are not important at the individual level. I think the model type is more important than the level of observation.	This is because our previous version of the manuscript focussed on population- versus individual-level processes. As this is no longer the case, we rephrased in this sense to remove these details.	
74-85	This part may be shortened.	We removed the unnecessarily long explanations on the physiological mechanisms leading to mortality under drought.	
110	stand size?	We removed stand	
111	“level” is missing after “individual-“.	corrected	102
127	I think the decrease was not dramatical. “can decrease” is maybe more correct.	corrected	117
174	Do you mean patterns or drivers?	Drivers, we corrected in this sense	120
180	“When and how” sounds very vague to me.	Rephrased : “How do climatic factors and physiological processes drive temporal variation in the mortality rate ?”	154
180,182,185	I think “questions” is better than “issues”.	Replaced	153
436	I think this is also relative mortality probability!	Thank you , we added “relative“	405
501	ideally, one should report mortality rates together with the size threshold.	We agree with this point of view. However, adding these figures would make the text much longer. This is why we simply chose to add the last sentence of the paragraph “we cannot rule out that these different mortality estimates …)”	493

512	The conclusions about hydraulic failure and carbon starvation are not based on empirical evidence but rather reflect the believes about these processes and how they were implemented in CASTANEA. I don't think that only because the composite index is slightly correlated with mortality rates this is enough to say that both processes indeed cause mortality.	We're not sure we fully understand your point. Indeed, we do not conclude that our results prove that both processes cause mortality in our case, but they just support other studies which advocate that a combination of hydraulic failure and carbon starvation are more likely to cause mortality than just a single of these two processes.	504
538	From what do you know conclude that the trees die from drought and frost? How about competition?	In this sentence, we did not conclude anything, but simply introduced the idea that this study focused primarily on drought and frost (since the PBM does not account properly for competition). Also, note that in our statistical model at the individual level, competition was the factor with the least influence on the probability of mortality. In the revised version of the manuscript, to account for the fact that competition still plays a role, we modified this sentence as follows "to climatic hazards (drought and late frost) and to two biotic pressures (the presence of a fungus and competition)."	518
590-591	What is meant with "high precision local mortality predictions"?	We forgot the "in", which gives "high precision in local mortality predictions" in the corrected sentence	573
591-592	"A weaker ability to generalize" is also mentioned as a disadvantage of PBMs in the intro. This is a bit confusing…	PBMs have the usual weakness of not taking into account the variability of individual responses and biological interactions, but the strength of predicting the average population response in new environments. By contrast, while statistical models can take into account the inter-individual variability of responses and intrinsic biological interactions, they do not allow for the generalization of proximal causes of these responses in other environments than those studied. We rephrased the sentence of the introduction "which makes it difficult to have a precise projection in other populations where the model has not been precisely calibrated and validated." as : "Most often, calibration is made using the average parameter value known at species level, and therefore does not account for possible inter-individual variability of ecophysiological processes, and for its effect on response trajectory (Berzaghi et al. 2019)."	127
601	You may add a reference to Cailleret, M., Bircher, N., Hartig, F., Hülsmann, L., and Bugmann, H.. 2020. …	The reference was added	
613	If "disaggregate" means comparing the contribution of different factors, I would argue that this is not possible with the results of this	We agree with LH that we could not fully achieve this goal here, and rephrased the sentence as "should ultimately allow". Moreover, we rephrased with "disentangle" instead of "disaggregate.	595

	study. To me this seems possible only if the lower level (individual level) model includes also climate variables.		
633	I think the discussion should end with a few conclusions.	We added a summary of our results at the very beginning of the discussion, and preferred ending it with perspective for future studies.	476
TableS2	This could also be an element of the main article. Be more precise about the effects of drought on the mortality rates. What does “when and how” means here? NLF, PLC, BoR and CVI and their correlation with mortality rates could also be a third question, e.g. “Which indicators of tree decline that can be derived from PBM are linked to tree mortality rates?”	As suggested by LH, we now included this Table in the main article. We rephrased the titles of the table, addressing LH suggestions	New Table 3

Reviewer 3 - *anonymous*

Reviewed by anonymous reviewer, 2019-12-03 13:57

This is a revised manuscript. The authors have improved the manuscript based on the review comments and answered all our comments. Figure 1 and Table 1 are important additions to the manuscript to clarify the measured variables and used approaches. Also simplifying the statistical analysis at population scale made the manuscript more clear. However, I think the manuscript still needs some rather minor improvements before it is ready for publication.

Comments from the reviewer (page refers to the first manuscript)		**Reply (lines refers to the revised manuscript)**	
Comment number	**Comment**	**Reply by PC et al.**	**Lines**
38	I am not sure what "decline on mortality" means? Maybe just "mortality".	We changed decline by "defoliation and fungi presence."	40
42-43	I think it is important to express it clearly in the abstract that mortality rate at population level was associated with all processes simulated with the PBM: conductance loss, carbon reserve depletion and occurrence of late frosts. In the current form, the reader might get a wrong impression that these tree processes were identified among many others to be the most important ones.	Rephrased: "The combination of all these simulated processes were found associated with the temporal variations in the population mortality rate. The individual probability of mortality decreased with increasing mean growth, and increased with increasing crown defoliation, earliness of budburst, fungi presence and increasing competition, in the statistical model."	44-47
160	Rephrase "the right text in square boxes"	Rephrased : "The square boxes indicate the measured factors and response variables considered in statistical models. Boxes with rounded corners indicate stress-related output variables"	159
164-166	You make two simulations with CASTANEA: one at population level using a simulation of a stand with 100 trees (and compare the results with measured mortality), and another at individual level investigating the differences in physiological responses between individuals with different characteristics. I think this is not clear from Figure 1 and its figure caption. Also, I would consider including Table S2 into the manuscript. This would clarify things even further.	We completed the legend of Figure 1 to make this point more clear. Also, we included the former Table S2 in the manuscript as the new table 3.	Fig 1 & Table 3
171	Percentage loss of conductance (delete the extra "of" throughout the manuscript).	Done (five found)	
215-216	The division of variables into endogenous and exogenous is not clear to me. For example, stem diameter is affected by both endogeneous (i.e. age) and exogeneous (i.e. competition) factors. How do you define this division, and do you really need such a division?	We agree with reviewer 3 that we do not really need this distinction between endogeneous and heterogeneous factors here. We have rephrased the sentences and removed it from table 2.	102
Table 1	In the mean, min and max, you use varying amount of decimals. Please be consistent.	Corrected in table 1	Table 1
244	What does an "ordered variable" mean here?	This was a bad formulation this variable is simply quantitative	228
264	"all indices reached a ceiling after this distance value" is odd language. Please rephrase.	Rephrased : "plateaued"	248
334-335	It is mentioned here that CASTANEA was validated based on ring width patterns and reference to Appendix 2 is given. I think the results of the validation should be summarized here. It should be said that although growth simulated with CASTANEA correlated with observed growth, CASTANEA overestimated growth strongly, and simulated a decreasing trend in growth over time that was not visible in observations.	We summarized the main results of the validation and explanation: "We found a significant positive correlation between ring width observed and simulated (p-value << 0.01). Although CASTANEA tended to overestimate growth at the beginning of the simulated period, and simulated a decreasing trend in growth over time that was not visible in the observed data. This is likely to be due to a bad estimation of population density prior to the monitoring period (see details in Appendix 1). "	366-370
Figure 2	Use the same axis range for the mortality rate in all subfigures (it now varies from -1 to 3.5, to 0	We modified figure 2 so that the y-axis for mortality in sub-figures a,b,c, varies from 0 to 3.5. However,	Figure 2

	to 4.5 depending on the figure).	in order to compare the joint variations in the mortality rate and in the CVI, we found it better to keep the zeros of the two axes at the same level. This is why the mortality rate axis varies from -1 to 3.5 on sub-figure d.	
418-425	This section is about statistical analysis on population scale, and has been greatly condensed since the first version. This decision is good. However, in case any statistical analysis on population scale is maintained, some results should be shown in the manuscript. Now there is nothing (only in the appendix).	In the revised version of the manuscript, we removed also the description off the statistical analysis at population scale from the main text.	
439,448	subscript Nstem	Done	410
486	Either “shed new light” or “gave new insight”	“S” remove	476
502	remove “automatically”	Done : “(including smaller trees increases the mortality rate)”	495
508	remove two extra commas	Done: “of drought and late-frost stresses”	501
517	replace “or” with “and”	Done	Appendix 2
533	high mortality in 2007	Done	513
569	What do you mean by mechanical increase of water need?	Rephrased: “which also increases their water needs due to the increase of transpiration”	551
597-599	If you compare measured mortality with variables simulated with CASTANEA, how do you there finely integrate the statistical and simulation approach? Please rephrase.	Rephrased by removing the sentence “can be further integrated”	580

References

Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Berzaghi F, Wright IJ, Kramer K, et al (2019) Towards a new generation of trait-flexible vegetation models. Trends Ecol Evol 1–15. https://doi.org/10.1016/j.tr ee.2019.11.006

In J, Lee DK. Survival analysis: Part I - analysis of time-to-event. *Korean J Anesthesiol*. 2018;71(3):182-191. doi:10.4097/kja.d.18.00067

In J, Lee DK. Survival analysis: part II - applied clinical data analysis. *Korean J Anesthesiol*. 2019;72(5):441-457. doi:10.4097/kja.19183

Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CGM. Time-varying covariates and coefficients in Cox regression models. *Ann Transl Med*. 2018;6(7):121. doi:10.21037/atm.2018.02.12