Dear Editor and Reviewer,

thank you for your comments. Your comments have revealed a number of shortcomings and come up with good ideas, which has led to an improvement in the quality of the article.

A linguistic correction of the whole article was performed by a professional company. We send a confirmation of the implementation of language corrections in the attachment.

Manuscript Number: FORECO-D-21-00760    
  
Veteran trees as a key factor for bird diversity in spruce-dominated production forests: implications for conservation management  
  
Dear Mr. Kebrle,  
  
Thank you for submitting your manuscript to Forest Ecology and Management. I have received one evaluation of your manuscript and am willing to proceed with a decision. I invite you to make major revisions and resubmit the paper by Jun 02, 2021. When revising your manuscript, please consider all issues mentioned in the reviewers' comments, and in the Reply to Reviewers note your changes and any disagreements with the reviewer.  
  
To submit your revised manuscript, please log in as an author at https://www.editorialmanager.com/foreco/, and navigate to the "Submissions Needing Revision" folder.    
  
Forest Ecology and Management values your contribution and I look forward to receiving your revised manuscript.  
  
Kind regards,      
  
T.S. Fredericksen    
  
Editor-in-Chief    
  
Forest Ecology and Management  
  
  
Editor and Reviewer comments:  
  
  
  
Editor: The paper needs a thorough review by a native English speaker or language editor. Just in the abstract, I have these suggested changes:  
  
Bird species richness increased with number of broadleaved veteran trees.  
Just one veteran tree/ha may have a positive effect on bird richness.  
Bird density in production forests with approximately 5 veteran trees/ha and reserves is similar.  
Five broadleaved veteran trees/ha can improve bird diversity in spruce production forests.  
  
  
Reviewer #1: Review of the paper paper #FORECO-D-21-760 submitted to Forest Ecology and Management.  
  
Dear authors,  
I have carefully read and commented the paper entitled "Veteran trees as a key factor for bird diversity in spruce-dominated production forests:implications for conservation management ». In this study, the authors use an interesting sampling design in spruce forests where they identified large (DBH>70cm) trees and test the effects of the number of trees on bird diversity indices. They worked both in production forests and (more marginally) in reserves. They found that the more the trees, the richest the community in general (total richness as well as generalist and specialist richness). Other variables also influenced the diversity of birds.  
The paper is interesting, its main asset being the nation-wide sampling design and the way it was controlled. I appreciated the way the authors managed to control for the number of large trees. However, the paper presents some inconsistencies that prevent it to be published in its present form. My major comments are:  
- Although not a native English speaker, I noticed several (but not all) language inconsistencies and at some point gave up to correct them. I recommend to make the paper reviewed by an English native;  
- The shortcut used throughout the manuscript between large trees / microhabitat bearing trees / veteran trees should be lowered a bit, since the authors only assessed large trees, but not the microhabitats, so not all large trees qualify as veteran trees in the end. Some points have been made here and there, but I have the feeling that the paper tend to oversell its results;  
- The analyses of count data are not adequate. Authors should use Poisson error distribution on scaled variables, and assess that the models are not overdispersed to comply with Poisson error assumptions. The scaled variables would allow to compared the magnitude of the results and asses if the large trees are really the dominant driver (in terms of magnitude). I am confident this will not change the results much;  
- They should also take into account the fact that plots are not of the same size between reserves and production forests, in addition to the systematic bias of tree composition in the reserves (broadleaf dominated), this may reinforce the discrepancy between the two modalities.  
  
Specific comments  
Highlights: « only one veteran tree… can have an effect on…"  
"Five veteran trees can improve…"

Thank you for your repair. We rewrote to according to your comments:

Bird species richness increased with number of broadleaved veteran trees.

Just one veteran tree/ha can have a positive effect on bird~~s~~ richness.

Bird density in production forests with approximately 5 veteran trees/ha and reserves is similar.

Five broadleaved veteran trees/ha can improve bird diversity in spruce production forest.

L24: the authors use plots either for sampling plots and for larger areas. 600ha is not really a plot, but rahter a site. Please clarify.

Thank you, we have changed the designation “Study plot” to “Study site” throughout the document.

L31-32: please rephrase "tree species diversity positively affected…"

L31-32: Furthermore, tree species diversity positively affected generalist species and negatively specialist species.

L34: I guess it is "≥" rather than "≤"?

We are sorry, this is a mistake. We fixed it on ≥ 5.

L40: what reference is "Forest"? If it is "FOREST EUROPE" (as it seems to be in the references), please cite the last publication of 2020. Spruce is not out of range all over Europe, please rephrase.

Thanks, we missed this error. It really is “FOREST EUROPE”. We traced the publication from 2020 and used in text.

L38-41: Many native forests in Europe have been transformed into even-aged production forests of such commercially attractive conifers as Norway spruce (*Picea abies*), which is in the high part of these places out of its original range (FOREST EUROPE, 2020; Klimo et al., 2000).

L45-48: Additionally, about three-quarters of forests in Europe are even-aged. These forests are beyond the regeneration phase and have not yet reached the mature phase. (FOREST EUROPE, 2020) These are usually forests between 20-80 years old (FOREST EUROPE, 2011).

L59: TreMs actually stands for "Tree related microhabitats". Please modify (at least once).

Thanks, we modified at L60.

L71: there are problems with the references, please check carefully

Thank you, corrected.

L76: affect

Corrected.

L82-86: the introduction of bioindicator and forest health is in my opinion not necessary here. Also, there are poor links between those sentences. Please reorganize and simplify.

L83-87: Given that, generalist species are more likely to be tolerant of environment conditions than specialist species. On the other hand, specialist species are more likely to be sensitive to extinction (Devictor et al., 2008; Richmond et al., 2005). Accordingly, the population trends of forest specialist are declining in Europe (Gregory et al., 2007).

L120: the introduction lacks clear hypotheses and the last paragraph sounds more like a conclusion. Please consider revision of this part.

Thank you. We revised this part of introduction and include the main hypotheses.

L143-145: unclear, what was limited in the 25ha? Please rephrase

Thank you for the notice. The text has been clarified.

L148-149: how do the authors assess the cover of spruce on a 100m-radius? Please detail

Thank you. Data from forest management plans were used to determine the presence of spruce and the suitability of the stand. In these plans, the species composition of the stand groups is given in percentages. However, a shortcoming of these data is that we know the proportion of the area occupied by the species, but we do not know the position of these trees in the stand group. Therefore, we used analysis of aerial photographs to obtain an accurate value of deciduous tree cover within a 100 m surrounding. We have specified in the text.

L154-172: how the authors have such a precise description of the 100m plots? Did they measured trees there? If so, how? Please detail

Thank you. Forest management plans were used to describe the stand. The data from forest management plans was used only for this description of forest composition and for localization of Study sites. Only trees over 70 cm DBH were measured. In the text we have added a data source.

L179: a verb is lacking here, please correct

Thank you. We've corrected it.

L196-197: why not using the length of edge per sampling plot, it would be more relevant than the mere distance to edge.

Thanks for the question. We choose this variable based on information, that the distance from forest edge can influence the bird richness (for example: Šálek et al. 2010; Hofmeister et al. 2017; Batáry et al. 2014). Furthermore, the use of edge lengths could be problematic for elongated shapes, where the length of the edge would be long, but due to the small area of ​​the clearing, the significance for birds itself is probably low. We should use the total area of clearing within sampling plot, however we found no effect on birds.

L210: "For analyse" is not correct, please modify. Specify "bird species"

Thank you, we've adjusted it.

L211: linear mixed models are not appropriate to analyse count data (species richness), for which the Gaussian distribution is not correct. The authors should use generalised linear models with Poisson error distribution, and verify the assumptions of the distribution (in particular, potential overdispersion). Please also note that in this case, there is no assumption on the normality of the results.

Thank you for your comment. We used generalized linear models with Poisson error distribution instead of the original linear mixed models.

L214: here again, "Study plot" is misleading, please modify to e.g. "site"

Thank you, we have changed the designation “Study plot” to “Study site” throughout the document

L216-219: R command makes no sense here (the authors did not even mention they are working in R yet). Either refer to a simplified formulation of the model (even discursive), or go to the mathematical formula.

Thank you for your comment. We changed the model description, excluding the explicit R formula and rather describing the model by words. Also, the R code of the analysis is accessible as an Appendix to the paper.

L231: same remark as above on the statistics  
The fact that the sampling points are not of the same surface area between production forests and reserves is never taken into account. The authors should consider the use of an offset to account for the differences in sampling effort.

Thank you for the comment. However, we believe this was a misunderstanding. Forest reserves did not enter the model with environmental variables (i.e., large trees, tree species diversity, etc.). In these models, we only focused on spruce-dominated production forests, and therefore, only plots located in the production forests were included in this model. All these plots had the same area. Then, we also wanted to compare the veteran tree densities and their impact on bird species richness in production forests with those in reserves. This was evaluated by other models, where categorized veteran tree density was the only predictor. Since we worked with tree densities instead of counts in these models, the differences of surface areas were effectively incorporated. It is also important to note that the plot area was only relevant for counts of trees, not for counts of bird species, as the species richness was always evaluated in 50 m radius around the sampling point, both in reserves and in production forests.

L272-292: it would be interesting to further comment the magnitude of the results (ie. How many species more in each category?).

Thank you for the comment. First, for models in production forests, we added to the main text of the Results values of species richness predicted by GLMMs for some representative values of the predictors, to see the magnitude of their effect. Also, we refer to the effect plots where the magnitude and nature of all predictor effects can be seen (Fig. 2). Second, for the models combining production forests and reserves, we added differences in predicted mean richness values for the veteran tree density categories from the value of the reference category (i.e. the lowest-density category). These values are now both in the main text and in the Fig 3.

L304: "a positive effect of the number of large living trees"

Thank you, corrected.

L321-322: this does not have much to do with the present study. I recommend to delete this marginal example.

Thank you, we deleted it.

L327: "injuries"

Thank you, corrected.

Section 4.1: I think the problem here is that the authors only designated large tree without accounting for TreMs on them, and the majority of the discussion is about TreMs. It is true that the larger the trees, the more the TreMs (e.g. Larrieu et al. 2018, Paillet et al. 2019 and many more), but there are also numerous large tree that do not bear TreMs. The discussion should be more focused on large trees per se, and less on TreMs that do not explain everything.

Thank you for your comment. We have edited this section and are discussing more about the large trees themselves.

The authors should also mention that broadleaves in what appears as an ocean of spruce represent a diversification of the niche that is exploited by numerous species (cf. the effects of dispersed aspens in conifer-dominated boreal forests, this is a bit what is said below but not that clearly).

Thank you. We have described this effect more clearly in the text.

L346: "chosen"

Thank you, corrected.

L348: see also Paillet et al (2019) for a larger scope.

Thank you, we have expanded the information from the recommended literature.

L377-378: this classification has not been mentioned before (in the M&Ms), please clarify. Also, D. leucotos has been spotted only once, so this does not make a clear proof that Reserves are more favourable to this species.

Thank you for letting us know. We have incorporated explanations of this species division directly into the text and added a source citation.

L392-395: I do not follow the reasoning, since the authors did not test explicitly for spatial aggregation of large trees in their dataset, but for the number (see the SLOSS debate and notably recent studies by L. Fahrig). So they could not claim that aggregated trees (in fact more numerous trees) are beneficial to the biodiversity of birds. This maybe true, but this is not what has been tested here.

Thank you for your warning and recommended literature. We have removed this part from the discussion.

Section 4.4: I would be curious about a comparison of the effects of the co-variables (e.g. scaled estimates), to see whether the effects of large trees is really more "important" than the other variables. This would replace estimate values in the Table 1.

Thank you for the comment. … Also, using scaled estimates is only valid for models with normally distributed errors. We believe that the magnitudes of individual predictor effects can be satisfactorily compared visually using effect plots.

L440-446: I also suspect an effect of sampling size here, but it is not very clear how this was assessed.

Thank you for the comment. We think (but are not sure) the concern here is about differing sampling plot sizes between production forests and reserves (and not about sample size). If so, we kindly emphasize that the bird survey was always done within 50 m radius around each sampling point, no matter if it was in a production forest or in a reserve.

L451-453: this also means that there are confounding effects between reserves and forest composition, so that both could not be disentangled.

*Conversely, due to a high proportion of broadleaved forests within 100 m of the survey midpoint in forest reserves (often 100 % of the area), the number of conifer forest specialists was lower than in production forests. In support of this, we found a negative effect of tree species diversity. This may reflect dominant conifer specialized bird species in assemblage.*

Thank you for the comment. We generally agree that it is hard to compare production forests and forest reserves especially because their differing species composition. Indeed, this was the very reason why in the main analysis, we only focused on production forests with comparable species compositions. We decided, however, to add an analysis combining these two forest types together, just to describe and quantify their difference in veteran tree density and its association with bird diversity. We are fully aware that it is problematic to derive any strong conclusion about drivers of bird diversity without incorporating species composition explicitly into the model. We believe, however, that we don’t do this, and that the different species composition of production forests and reserves is satisfactorily discussed in the Discussion. Our main conclusions are about production forests and we don’t aim at disentangling this complex interplay between forest species composition and the role of old trees in bird diversity.

Section 4.6: even if I do not doubt that most trees sampled could be considered "veteran", they are first of all "large" trees (as mentioned, 160 years old for a beech tree is not that old or veteran), so I think the shortcut is a bit too big here. These points have more or less been evoked in the introduction, and I think they could be summarized in a larger conclusion (the present conclusion is mainly a summary of the results of the study).

Thank you. You're probably right, using the term “veteran trees” is not the most suitable for our objects. We have replaced “veteran trees” with the term “large trees” throughout the article.

Figure 2: Please mention if co-variables were averaged for this representation. Also, the curves extrapolate beyond the maximum number of broadleaved trees. Please stop at the maximum unless the representation is not correct.

Thank you for the comment. Now, we explicitly state in the effect-plot figure captions that the co-variates were set to their average values. We also modified the plots so that now they don’t extrapolate outside the observed predictor’s ranges.

L780: arrows => whiskers.

Thank you, corrected.

Table 1: as mentioned, the estimates should be scaled to allow comparability and commented accordingly.

Thank you for the comment. As we explained previously, we didn’t scale the estimates as this practice is not appropriate for non-Gaussian models. The effects of individual predictors can be compared using effect plots.