Peer Community In
Ecology

Protected Areas effects on biodiversity: a test using bird data that hopefully will give ideas for much more studies to come

Paul Caplat based on peer reviews by *Willson Gaul* and 1 anonymous reviewer

In the face of worldwide declines in biodiversity, evaluating the effectiveness of conservation practices is an absolute necessity. Protected Areas (PA) are a key tool for conservation, and the question "Are PA effective" has been on many a research agenda, as the introduction to this preprint will no doubt convince you. A challenge we face is that, until now, few studies have been explicitly designed to evaluate PA, and despite the rise of meta-analyses on the topic, our capacity to quantify their effect on biodiversity remains limited. This study by Cazalis et al. [1] uses the rich dataset of the North-American Breeding Bird Survey and a sound paired design to investigate how PA change bird assemblages. The methodological care brought to the study in itself is worth the read, and the results are insightful. I will not spoil too much by revealing here that things are "complicated", and that effects – or lack thereof – depend on the type of ecosystem, and the type of species considered. If you are interested in conservation, bird communities, species life-history, or like beautiful plots: go and read it.

References:

[1] Cazalis, V., Belghali, S., & Rodrigues, A. S. (2019). Using a large-scale biodiversity monitoring dataset to test the effectiveness of protected areas at conserving North-American breeding birds. bioRxiv, 433037, ver. 4 peer-reviewed and recommended by PCI Ecology. doi: [10.1101/433037](https://dx.doi.org/10.1101/433037)

Reviews

Evaluation round #2

Authors' reply, 24 March 2019

Thank you for reviewing a second time our manuscript. We are happy to see that you have appreciated the modifications we applied and that you now only have minor comments. We have applied every text modification suggested. Here are our responses to the three other comments:

"L347: I got a little confused with your results notation ("c=", and I apologize if I missed this at first place... I would truly suggest to use the conventional notation for statistical results (have a look at `http://abacus.bates.edu/~`ganderso/biology/resources/writing/HTWstats.html), which is to report Estimate ± S.E. and p-value at least (e.g. estimate ± SE = -0.046 ± 0.007, P = 0.143). Standard Errors are missing in both your Results section and the Tables (in the core ms. and the SI). Please make sure to add them." As suggested, we have added to each result (in the text and in tables, both in the main text and SI) values of standard error using the format estimate ± SE, P-value.

"L379: following what I say above, confidence intervals have to be presented in as results in the Tables, as we can't only rely on what you tell us as being "large confidence intervals" (although we get a sense of that in Fig.2) " I agree it was missing in the last version. However, as SE values are mentioned now, I think that adding CI values would make a huge table with redundant information. I realise that SE and CI are not completely redundant in the case of phylogenetic model because of the bootstrapping but they are strongly correlated. Then, unless you think it is critical, I would prefer to keep tables without CI, considering that SE and the plot give enough information to the reader.

"L381: please provide coefficient estimates here, and Table or Fig. you want the reader to refer to. This has to be a general comment for the Results (and more specifically for the species-level analyses): make sure to provide coefficient estimates to support your description and/or refer to Table or Figure appropriately. " We have added estimates, P-values and SE to these results and clarified the reference to appendix.

Decision by **Paul Caplat**, posted 19 March 2019

Encouraged revision of the preprint "Using a large-scale biodiversity monitoring dataset to test the effectiveness of protected areas in conserving North-American breeding birds"

Dear authors, we have completed a new round of reviews, and I am pleased to say I am ready to recommend your manuscript. Before I do so, could you try to apply some of the minor edits recommended by the reviewer?

Reviewed by anonymous reviewer 1, 11 March 2019

I read this second version of the manuscript from Cazalis et al., in which they investigate the effectiveness of protected areas at conserving N-A breeding bird assemblages, still with a great interest.
I acknowledge the thorough work done by the authors, to reply and account for the comments and suggestions I and the other reviewer made in the first round of reviews. I think the authors greatly improve their manuscript, which seems now suitable for publication.
I only have minor comments for the authors.
Minor comments:
L191: you may want to add the version of the dataset you used (e.g. "version 2017.0").
L220: I suggest "the 'Weather' file" and you add a reference link to this and/or the reference for the Citation

Pardieck et al. 2017.

L232-237: maybe "should reduce this bias" or "is expected to reduce".

L295: I suggest "modelled as a function of a one-way interaction between the proportion of [...] and the type of vegetation structure".

L345: I suggest you add the reference Ives & Garland 2010 (as "(following Ives and Garland, 2010)") for the bootstrap procedure.

L347: I got a little confused with your results notation ("c=", and I apologize if I missed this at first place... I would truly suggest to use the conventional notation for statistical results (have a look at `http://abacus.bat es.edu/~`ganderso/biology/resources/writing/HTWstats.html), which is to report Estimate ± S.E. and p-value at least (e.g. estimate ± SE = -0.046 ± 0.007, P = 0.143). Standard Errors are missing in both your Results section and the Tables (in the core ms. and the SI). Please make sure to add them.

L379: following what I say above, confidence intervals have to be presented in as results in the Tables, as we can't only rely on what you tell us as being "large confidence intervals" (although we get a sense of that in Fig.2).

L381: please provide coefficient estimates here, and Table or Fig. you want the reader to refer to.

This has to be a general comment for the Results (and more specifically for the species-level analyses): make sure to provide coefficient estimates to support your description and/or refer to Table or Figure appropriately.

L393: replace "aberrant results" by "outlier results".

L394: Suggest rephrasing " Including non-native species in the analyses had little effect on the results at the assembly level".

L397: "did not change the results".

L420: Please consider rephrasing the caption of your Figure. I suggest "Effect of human-affinity on species' responses to protected areas within forest routes".

L491-495: I suggest splitting this long sentence in two pieces. L493: start a new sentence with something like "Instead, it only measures the effects PAs can have..."

L521: "expected from"

Evaluation round #1

Authors' reply, 06 February 2019

Download author's reply
Download tracked changes file

Decision by **Paul Caplat**, posted 06 December 2018

Encouraged revision of the preprint "Using a large-scale biodiversity monitoring dataset to test the effectiveness of protected areas in conserving North-American breeding birds"

The preprint "Using a large-scale biodiversity monitoring dataset to test the effectiveness of protected areas in conserving North-American breeding birds" has been reviewed by two experts in the field. They agreed on the value of the research presented, and the soundness of the overall approach: the manuscript addresses an important question of conservation ecology "Are protected areas effective?", and brings new insights into how to answer that question, specifically by accounting for environmental variation across sites to allow better quantification of effects. Both reviewers outlined a number of concerns, mostly around adequate presentation of the research, and some methodological aspects, and I would encourage to take them into consideration

to improve the manuscript. Particularly, I agree with the reviewers that some terms could be made more precise (e.g. "protection effectivness" being different from a "good designation", and the choice of considering only native species made more explicit in the terminology). The relevance of species richness as a metric for change is worth discussing, as is detectability (though I agree with the reviewer that a simple mention might be enough). Overall more information on the methods (sample size, raw data distribution and motivation for winsorization, motivation for having two phylogenetic models) need to be included. Detailed suggestions can be found in the review, about a number of aspects including writing style and references, which I won't comment on here but seem like good options to improve the manuscript.

Reviewed by **Willson Gaul**, 01 November 2018

Cazalis et al. assess the effectiveness of protected areas for preserving birds, using both assemblage-level and individual species-level analysis. The main contribution of this paper seems to me to be the explicit consideration of land cover type when comparing protected and unprotected areas, such that protected areas are compared to unprotected areas of the same land cover type. At the individual species level, the study evaluates the different effects of protected areas on species with different habitat requirements and human tolerance, highlighting which species are likely to be most benefited by protected areas. Overall I find the study interesting and useful, especially because it evaluates whether protected areas are more successful for some habitat types and some types of species – this seems like it could be of direct interest to applied managers. I also find this study interesting in the context of other recent overviews of biodiversity change that, like this study, have found little change in assemblage-level metrics but greater change in species-level metrics [1][2][3]. I think this paper would benefit from more explicit discussion of other studies that have found little assemblage-level change, and what implications that has for using assemblage-level metrics (particularly species richness) to measure biodiversity change. I have one concern (discussed below) about the authors' decision to exclude non-native species from analysis, and I have a concern about "winsorizing" the data in the statistical methods (also discussed below). Finally, I have listed some minor comments about word choice and sentence structure; these minor sentence-level issues are a slight barrier to smooth reading, but generally did not interfere with my ability to understand the article.

The introduction and/or discussion could benefit from addressing other studies that have found small changes or no changes in species richness over time and in response to disturbance [1][2][3]. This study compares species richness of protected and unprotected sites, with the assumption that species richness would decline more in unprotected than in protected sites because of local extinctions and population declines in non-protected areas. This would lead to an observed pattern of lower species richness in unprotected areas. However, Supp & Ernest [3] reviewed published studies and found that species richness did not respond strongly to disturbance. They suggested that "community-level measures are poor indicators of change..." [3]. Supp & Ernest [3] did find stronger species-level responses to disturbance, which is in line with the results of this study. I think it would be useful to put this study and its results into the context of broader biodiversity trends, and especially to mention other studies that, like this study, have found that differences in individual species metrics (e.g. individual species abundance changes, species turnover) are larger than differences in species richness [1][2][3].

If I understand correctly, this study excluded non-native species from analyses (line 210). It may be reasonable to limit the study to native species if the goal is to evaluate the effectiveness of PAs in preventing local or regional extinctions and population declines. However, excluding non-native species does not make sense if the objective is to study the effectiveness of PAs in conserving ecological communities, because communities are changed by the addition of non-native species. The abstract of this study states that the study investigated "protected areas effectiveness in conserving bird assemblages" (line 22) and this is also mentioned in the discussion (line 418). In fact, I think this study investigates protected areas' effectiveness in conserving *native bird species*, even if the assemblages in which those native species live have changed due to the addition of non-native species. Excluding non-native species could be masking important changes in the ecological

community.

I think it would be helpful for the authors to clarify whether this study is aiming to study native species conservation or entire community assemblages. If the aim is to study community assemblages (e.g. line 22), then I think all species detected should be included in the assemblage analyses. If non-native species are excluded from analyses, then I would find it clearer if the terms "species richness" and "abundance" were changed changed to "native species richness" and "abundance of native species" throughout the text. This will clarify exactly what aspects of the community are being studied. In particular, claims like that made in line 488 that "...we measured PAs effectiveness as the difference in abundance or richness between protected and unprotected sites" are not quite correct if non-native species have been excluded. The actual difference in species richness or abundance between sites is unknown in this study, and may differ from the difference between native species richness or native species abundance.

I am concerned about the winsorizing of abundance values for species (line 217). If I understand the methods correctly, individual species abundance is analyzed using GAMs with a Poisson error distribution. Did the authors check whether the data fit distributional assumptions before or after winsorizing, e.g. by looking at QQ plots or histograms of residuals [4]? In general, it is preferable to model the original data values using an appropriate error distribution rather than transforming or changing extreme values. If the authors still decide to winsorize the data, then they should justify this by reporting whether the winsorized data fit a Poisson error distribution better than the un-transformed data did. I do not know how much the choice of whether to winsorize the data affects the results, but in general I think it would be preferable to model these data using their true values and an appropriate error distribution, rather than changing those values to the value of the chosen quantiles.

References:

[1] Dornelas, M. et al. (2014) Assemblage time series reveal biodiversity change but not systematic loss. Science 344, 296-299.

[2] McGill, B. J. et al. (2015) Fifteen forms of biodiversity trend in the Anthropocene. Trends in Ecology & Evolution 30, 104-113.

[3] Supp, S. R. & Ernest, S. M. (2014) Species-level and community-level responses to disturbance: a cross-community analysis. Ecology 95, 1717-1723.

[4] Wood, Simon N. 2017. Generalized Additive Models: An introduction with R (2nd edition). Chapman and Hall/CRC.

Minor comments For spelling or word change suggestions, I have put the proposed new word in **bold**.

Line 17-19: "which whereby..." is odd wording. Remove the word "which"?

Line 27-28: Wrong sentence structure. Perhaps delete "are the one avoiding human activities" so that the sentence reads "At the species level, we found that species that avoid human activities tend to be favoured by protected areas."

84-89: I think the comparison to population trend-based methods can be shortened. A criticism of trend-based methods is not a major development of this study. The example (starting at "Hence..." and ending at "...20 times more important" can be removed.

96-137: These 2 paragraphs are the most informative and important part of the introduction. The choice of counter-factuals and how that influences the interpretation of measures of PA effectiveness seems to be a major theme of this study. These 2 paragraphs are a well-written, concise overview of the issues.

132: Re-cite the meta-analyses here? They were originally cited in line 98, but I ended up scrolling up through the text looking for those citations.

184: "effectiveness **in** selecting the most interesting sites..."

185: "effectiveness in creat**ing** ..."

202: "presence **of** PAs ..."

204: "starting stop of each route" is confusing. Maybe better to say "**first stop** of each route".

210: I find the phrase "non-native" species more familiar than "non-indigenous species" but this is probably

just a matter of convention. The authors could change it or leave it as it currently is.

216: Consider at least mentioning detectability and the difference between abundance and detections. Unless they have already corrected for detection, what the authors did here was sum detections, not abundances. Using this metric for analysis assumes that detectability was constant across routes and protected and unprotected areas. This assumption should be mentioned.

220: Include which software was used (e.g. QGIS, ArcGIS, R).

228: Include a reference for the quoted IUCN definition of a protected area.

244: Regarding "effectiveness can vary with protection level", has the introduction already cited studies showing this? Perhaps provide a reference here.

250: "(**May** to June)"

262: "... not analysed because **they were** too scarce."

Fig 3: The negative relationship between human-affinity and PAEfor is hard to see on the scatter plot. Drawing the regression line from the LM model reported in Table 1 might help make the relationship more visible.

390-505: Overall, the Discussion is interesting and well written.

422: The claim that forest PAs have "more forest-typical bird assemblages" would be stronger if all species (including non-native species) were included in analysis.

452: "changed" should be "**changes**"

507-508: This sentence seems weaker than the rest of the Discussion. It does not add anything new.

Reviewed by anonymous reviewer 1, 04 December 2018

Download the review