

## **Reviewer background/context**

I am an English-As-Second-Language researcher, with background in conservation genomics. My expertise is limited regarding this kind of analysis; however, I am very interested in this field since I also work with conservation. Although I cannot offer much in-depth comments on the specifics of the methods, I am hoping to add the perspective of a broader audience, especially non-technical stakeholders, that are tangentially interested in this topic and will be able to learn the key takeaways of this work.

## **Summary**

Ferraille et al. (bioRxiv doi.org/10.1101/2023.05.09.539999, submitted to PCI Ecology) conducted a study evaluating the use of different databases for Systematic Conservation Planning (SCP) based on Species Distribution Modeling (SDM). Three databases were chosen for this analysis, each comprising different scales and following distinct protocols for data collection and validation: French Natural and Landscape Information System (SINP); Global Biodiversity Information Facility (GBIF), and French biodiversity monitoring scheme (Vigie Nature). These were assessed using three different territories in France that vary in ecosystems, bioclimates and anthropogenic pressures; and focused on taxonomic groups commonly used in impact studies. The database evaluation followed three levels:

1. Database:
  - a. Number of observations for SDM modeling
  - b. Database scale
  - c. Methods for generating pseudo-absence
2. SDM with Aves/Pappilionidae:
  - a. Calibration
  - b. Performance
  - c. Overlap
3. Species community:
  - a. Species representation before/after implementation of workflow

The authors found that databases differ greatly in their number of observations for different taxonomic groups, which will have significant impacts for downstream analysis. For the calibration of models, larger, national scale databases performed better and complemented local databases, while the latter worked better for evaluating performance. Also, although there was high overlap between areas for conservation priorities, the ranking of habitat suitability was highly variable, indicative of sampling biases.

Overall, I thought this manuscript to be an important consideration of methods applicable to anyone interested in conservation planning. Below, I offer my considerations that I hope will be helpful to authors.

## **Major comments:**

My main difficulty with the text was understanding the different classification levels used for the databases. This made reading and interpreting the results challenging, and it was hard for me to connect the objectives with the workflow. I would strongly recommend standardizing terminology throughout the text and potentially creating a table or glossary with the terminologies, their definition and what they refer to specifically. For example, I am still unsure of what “local”, “regional” and “national” mean in this context. Are those terminology used to describe the databases and their level of coverage, or is it referring to how the dataset was processed and analyzed?

For example, you first introduce SINP as a database structured at “regional scale” (line 133), but then in the workflow diagram it is highlighted as “local” (and there is no “regional” scale present also). However, in table 2, I understood these terms to be used as parameters for subsetting the data. And while there is a definition for “local” and “France”, there is none for “regional” (and this is not introduced in the text either—at least I could not find it). Another example is in the use of the term “national database”, which in line 151-152 is introduced as being both GBIF and Vigie Nature (which I assumed is related to their level of coverage); then in line 176 is introduced as something else “(i.e. GBIF, Vigie Nature and National databases)”, which seems to be referring to a third category (!?). It looks like there is some ambiguity with the terms that makes it hard to know what exactly is being tested and analyzed.

Besides the table/glossary of terms, another suggestion I think would be helpful would be to align the objectives of the study with the workflow diagram. So, for example, point out on the Analysis section: “Step 1.2 – Objective 1”. This structure could also be replicated in the text when describing the methods and results.

Because I had some difficulty to understand clearly the objectives and methods of this study, I do not want to devote too much time with minor comments at this point. Once these major issues are clarified, I think I will be in a better position to provide more feedback. However, I present here some specific notes I think are worthwhile addressing at this stage.

#### **Minor comments**

Line 104-106: This phrase is repeated.

Line 158: Did you mean “spatial” or “sampling”? In your appendix it looks like you are referring to sampling bias, and I think it merits a replacement of word in the main text. It’s more specific and just for this one word is not worth it to make your reader consult the appendix.

Line 168-172: These terminologies need to be clarified. Are they referring to the spatial extent of the observations (e.g. the territories versus the country)? Or are they related to the coverage of the different databases for the same territories (shown in Fig. 1)? I think the second is what makes the most sense in this comparison, am I right?

Line 179: Replace “seems to be sufficient” with a more specific term. Phrased like this sounds like you were going on a hunch, but I imagine that it was more methodical than that and they are following the thresholds mentioned above, right?

Line 190-193: Again, the terminology needs some clarification here and throughout the document to reduce ambiguity.

Line 204: Replace “biogeographic” with “geographic”. I do not think this is the proper term here, since biogeography relates to the interplay between biology and geography, and you are listing environmental variables.