Dear reviewer 2,

We very much appreciate your review of our manuscript, which identifies similar themes to reviewer 1. We are very happy to address each of your comments below, which have helped improve the manuscript.

Kindest regards.

Referee 2:

General comments:

This study addresses the challenges of uncertainty in palaeoecological data derived from core-type samples. To quantify information loss due to environmental uncertainties and observer errors in palaeoecological research, the researchers employed a virtual ecological approach. They generated synthetic, error-free core-type samples of pseudoproxies and systematically introduced uncertainties related to core mixing, sub-sampling, and proxy quantification. The Fisher information and Principal curves statistical methods were used to assess the effect of different uncertainties imposed to the error-free samples. Their findings showed that increasing sub-sampling intervals had the most significant impact on both statistical methods. Moreover, the combination of increased sub-sampling intervals and decreased proxy counts per sub-sample showed the strongest influence on the analyses. Notably, principal curves were less affected by uncertainties related to proxy counting and sub-sampling compared to Fisher Information, which is more sensitive to short-term variability and driver interactions.

This paper presents a novel and interesting approach to deal with environmental, and observer introduced uncertainty in palaeoecology and its objectives are within the scope of CP. The objectives of the study are clearly outlined, the results and their interpretation support the conclusions. However, the methods section is very complicated and needs to be re-structured to aid readers who are not familiar with virtual ecology (like myself) to understand it a bit better. The way the methods are written, making the reproduction of the study rather difficult. On the other hand, the results and discussion section flow a lot better, and they are clearly written. I would recommend major revisions on the presentation of the methodology before publishing to CP.

Specific comments:

1. I agree with the previous comment that the methods' section needs some more clarification. At the moment it is very difficult to follow the workflow of the study. There is a lot of terminology involved, and the authors should try and explain it in a more simple way. It would be very useful to have a methods' figure that explains the whole processes and maybe another diagram explaining how the pseudoproxies are formed. Currently, there is no direct linkage between each method step, which makes it very challenging for the reader to follow the paper. For example, I

started understanding the methodology only when I proceed to the results section and I had to go back to the methods to re-read it.

The main concern from both reviewers was clarity in the methods section. We have substantially edited the methods section with additional figures and text to improve clarity and flow of the paper. The steps of the VE approach are now more clearly linked with the methods section. And the description of the modelling process is more explicitly linked with the PSM conceptual framework. While unavoidably technical in nature, we believe that the understandability and accessibility of paper is much improved. We thank the reviewers for their suggestions here. Two additional figures (now Figures 2 and 3) have been included to illustrate the methods, and a conceptual diagram of the proxy system model process (now Figure 1).

2. A little more clarification regarding the replicated models is needed. For example, why is there a specific number of replicas (31)? Is it a decision the authors took or is there another explanation? Why is there a 5000 time-steps? These might be parameters that are well known to readers who are familiar with virtual ecology studies but are not clear to all readers.

Lines 194-196 have been included to clarify the choice of time-span and replicates.

3. Both virtual sub-sampling and virtual proxy counts are referred as “observation models”. Could there be a sort of discrimination on those two uncertainty drivers?

We have kept sub-sampling and proxy counting under the broad category of observation models, but have made clearer the distinction between the degradation model and the two observation models (and the two observation models themselves). Specifically, in what is now Figure 1 and also throughout the text in section 2.2.

4. It would be helpful to know what kind of simulated pseudoproxies were used in the study. As explained in the discussion (443-457) some proxies are more sensitive to environmental sensors, so knowing what kind of pseudoproxies were used in the simulating models will make them more transparent and interpretive.

Lines 68-70: have been included at the end of the introduction to clarify the types of proxy we focus on.

5. The limitations of the study are mentioned for the first time on the conclusion. A brief discussion of those limitations should be given in the discussion section along with a few examples where this approach may not be suitable. Virtual ecology as a tool to quantify uncertainty on palaeoecology is a novel approach and therefore authors should explain in more detail the benefits as well as the limitations of their method.

We have reformatted the discussion and conclusion so that limitations are included, and expanded on, in the discussion with some examples. We now more explicitly discuss the relevance to empirical studies, and the limitations.

Technical corrections:

- Line 237: “2.4. Principal curves” should be: “2.3.2. Principal curves”

- Line 254: “2.4.1. Feature analysis” should be: “2.4. Feature analysis”
- In Figure 3A and Figure 3B the heatmaps should have more space horizontally.

All technical corrections have been addressed. Space between plots in Figure 3 has been maximised by moving axis labels and sub-plots as much as possible while maintaining sub-plot size.

Citation: https://doi.org/10.5194/egusphere-2024-3845-RC2