Point-by-point response to referee comments (4)

We would like to thank the reviewer and editor for their continued help in improving this manuscript.

Our responses to individual comments are highlighted in blue, with any proposed changes highlighted in red. Note that red page numbers refer to the location in the updated marked-up manuscript.

Editor comments

E1. Line 55 - consider adding "such as basin size and type"

We have amended the text as follows:

L53. …. and site characteristics that affect the pollen source area, such as basin size and type (Bradshaw and Webb, 1985; Prentice and Webb, 1986; Prentice, 1988; Sugita, 1993).

E2. Table 1 - in my view would better fit under Chapter 2, where a subchapter could summarize these methodological baselines.

We have moved this into the Methods section. We have created a new sub-section, 2.4 Comparison of Reconstructions. We have moved the text describing the methods used to extract information from the Serge et al (2023) and Zanon et al. (2018) reconstructions into this new section. The text of the new section is as follows:

L295. Our new approach shares some features with the methods used in previous reconstructions (Table 1), but is less data-demanding than the REVEALS-based technique and does not require a priori decisions about the selection of analogues. We compare our final predictions to both modern and fossil reconstructions of tree cover by Serge et al. (2023) and Zanon et al. (2018). Modern is defined as the interval 100 cal. BP and the present day in Serge et al. (2023) and between 125 cal. BP and present for Zanon et al. (2018). We make comparisons to the Serge et al. (2023) reconstructions based on the 31 taxa originally used by Githumbi et al. (2022) since Serge et al. (2023) show that this produces better results than using the expanded data set of 46 taxa.

Table 1: Key elements of the reconstruction technique from this study, the REVEALS approach (i.e. Serge et al., 2023) and MAT (i.e. Zanon et al., 2018)

	This paper	**REVEALS (Serge)**	**MAT (Zanon)**
Training data	Modern pollen data Modern tree cover	NA	Modern pollen data Modern tree cover
Training model	Regression based model linking modern pollen to modern tree cover	NA, although defined relationship between RPP and FS per taxa from modern data underpins technique	Calibration of modern pollen assemblages to tree cover

Downcore data requirements	Pollen data; Site characteristics	Pollen data; Site characteristics; RPP and FS per taxa	Pollen data
Main Assumptions and Challenges	Regression model applicability and included variables	Accuracy of RPP and FS values; Limited set of taxa	Number of analogues used (commonly 3-5); Threshold of similarity; Non-analogues
Scale	Site-based	Typically 1° where sites are located	Site-based; (Zanon: spatio-temporal interpolation)

For each of the 1° grid cells in Serge et al. (2023), tree cover was calculated from the sum of the appropriate vegetation types. Time series of the change in median tree cover were constructed using median tree cover corresponding to the pollen source area of each of our individual modern reconstructions. As the Serge et al. (2023) and Zanon et al. (2018) data is available in gridded format, comparison with our site-based predictions is not straightforward. Where the site location source areas straddled multiple grid cells, a median was calculated, weighted by the proportion of grid cell coverage using R package exactextractr (function: exact_extract) (Baston, 2023; version 0.10.0). The tree cover time series for the Zanon et al. (2018) and Serge et al. (2023) data were initially constructed using all of the extracted tree cover values for each of our model training site locations. However, since there can be multiple sites within some of these grid cells, we tested whether this affected the comparisons by taking an average of extracted tree cover values for locations sharing the same grid cell values from Zanon et al. (2018) or from Serge et al. (2023), and using this to create new time series for these two reconstructions.

E3. Can you also mention versions of R and packages used?
We have included R and package versions in the text and references.

E4. I think that the mismatch between the timing of tree cover increases in your data (Figs 7, 8, 9) and those of Serge (or Zanon) is sometimes 2000 years. You did not discuss sufficiently what was wrong with your or their approach.
Differences between the reconstruction methodologies, and the lack of primary (non-interpolated) data for previous reconstructions, make strict comparisons between the three approaches difficult or to diagnose what is wrong with the previous reconstructions. We have pointed out the differences, including for example the double peak in tree cover in the Zanon et al reconstruction, in the Results section (Lines 525 to 564). We have expanded the text in the Discussion to make it clear what the potential sources of these differences are, as follows:

L576. There are some differences between our reconstructions and those from other studies. Firstly, the maximum tree cover from our reconstructions is around 5–10% less than the maximum calculated from the other reconstructions. This could reflect the conservative nature of our modern-day tree cover model, which underestimates tree cover at the high end despite the application of quantitative mapping adjustment to model predictions. However, Zanon et al. (2018) also underestimate tree cover at high levels of tree cover. Alternatively, the difference may reflect the exclusion of higher elevation records from fossil dataset in order to minimise the impact of upslope pollen transport, which was not done in the other two reconstructions and would tend to reduce overall median tree cover. Secondly, the timings of peak tree cover vary between the reconstructions, with the MAT-based estimate peaking 2000 years earlier and the REVEALS estimate ca. 500 years later than shown in our reconstruction. The small difference between the peak timing shown by REVEALS and our reconstructions likely reflect differences in coverage through time and differences in the binning procedure. This may also partially explain the difference with the MAT-based reconstructions. However, this reconstruction also shows a marked decline between ca. 8000 and 7000 yr BP and a second peak in tree cover slightly after the peak shown by our reconstructions. It is difficult to assess the robustness of the reconstructed decline, which is strongly affected by a single point, but this would affect the overall shape of the curve and hence the timing of peak cover. Thirdly, the increase in tree cover in the early Holocene is less rapid in our reconstruction than the other reconstructions. This could reflect an underestimation of tree cover because of the lower tree SI values in our reconstructions, but the rapid increase in tree cover in the MAT- and REVEALS-based reconstructions is more likely to reflect an overestimation of tree cover because of long-distance pollen transport into the more open landscape characteristic of the early Holocene. The major difference at the pan-European scale is the reduction in tree cover from ca. 2000 cal. BP to present, which is less marked in our reconstructions and more consistent with observed tree cover. The observed tree cover values used in the model construction exclude areas dominated by land-cover types such as built areas or areas dominated by crops. We account for this in defining modern source areas in our model. Not accounting for changes in these other land-cover types, which through anthropogenic land use have increased substantially over the past 1000 years (Klein Goldewijk et al., 2017) would result in a steeper decline in tree cover, as seen in the other two reconstructions.

Referee comments

R1. I appreciate the effort of testing the influence of the pollination syndrome and the influence of Pinus instead of % Needleleaf pollen. I am surprised by the negative result, which leaves me puzzled as to why % Needleleaf pollen is the best predictor. Here and also for the Shannon index I would still find it interesting to carefully discuss the results and reflect on how the modern situation may differ from the past for which this regression is applied.

We included a brief consideration of our assumption of stationarity between tree cover and the explanatory variables in the Discussion section, but we have now expanded this and commented on the importance of %needleleaf and species diversity as predictors, as follows:

L687. Our simple modelling approach yields a reasonably robust picture of changes in tree cover through the Holocene, largely consistent with known changes in climate. We have shown that both %needleleaf and the SI are important predictors of tree cover. These measures are, to some extent, surrogates for pollen productivity and factors affecting pollen transport distance as explicitly addressed in the REVEALS-based reconstructions. Thus, their importance in the model is not surprising and we have demonstrated that the final model is able to capture modern day patterns. However, our approach is predicated on the assumption of stationarity between tree cover and the explanatory variables through time, as indeed are the other statistical reconstruction techniques considered here. This may be problematic for variables such as elevation, where changes in elevational lapse rates (Mountain Research Initiative EDW Working Group, 2015) or atmospheric circulation patterns (Bartlein et al., 2017) could affect the relationship, but is less likely to be an issue for explanatory variables that reflect biophysical controls on pollen transport and deposition such as basin type or proportion of needleleaf trees. Tree diversity is somewhat more problematic, since the influence of long-distance transport into the more open landscapes likely to have been characteristic of the colder, drier climate at the start of the Holocene may not be adequately captured in the modern training data. The relatively slow rate of the initial increase in tree cover may be a reflection of this. This could be explored using macrofossil of sedimentary DNA data, to discriminate between local diversity and potential long-distance contamination of the SI index. Nevertheless, the overall pattern of changes in tree cover during the Holocene appear to robust and explicable, supporting the idea that modern relationships between tree cover and the explanatory variables provide a reasonable basis for reconstructions.

R2. Now there is a lot of information in the supplements which could be better integrated into the main manuscript.

We agree that there is a lot of information in the Supplements, but this documents the multiple additional tests that we have made in the light of the reviewers' comments. We have outlined these results in the appropriate place in the main text, but we think that it is appropriate to keep the methodological testing in the supplement, since these results are largely negative and this supports the key methodological decisions that we have made in making the reconstructions. This allows us to be more confident about our results and conclusions, but we think including this material in the main text would obscure the main message.

R3. L. 52: Gil-Romera et al. – year missing

Thank you for noticing this omission, this has been amended.

R4. L. 105 ff: The table could be improved. Specifics: NA for training model MAT, RPP and FS per taxa is more similar to Training model then to downcore as these are assessed by modern situations.

We have revised Table 1 and moved this table into the Methods section (see response E2)

R5. L. 174: This is a posterior interpretation that SI may be sensitive to long distance cover no prior reasoning.

We agree and have modified this statement, since we have text in the Results section that discusses the impact of the Shannon Index on the model and how this could be interpreted in terms of either local species diversity or long-distance transport, as follows:

L175. We also calculated the Shannon Index (SI) of tree species diversity from the pollen data.

We have also modified the following sentence:

L251. We also included tree SI, as a way to evaluate potential impacts of very localised tree cover or long-distance transport from a single taxa or few taxon influencing the recorded AP%.

R6. L. 437: SI changes during the early Holocene due to the new establishment of trees generating a more divers forest with little change in the importance of long distance transported pollen.

The influence of arboreal SI in our model is based on the relationships between tree species diversity and tree cover in the modern day. In more open areas, longer distance pollen may upwardly bias tree cover values within the record. During the earlier Holocene, both the reconstructions by Serge and Zanon, as well as our reconstructions, suggest lower general tree cover. Increases in tree SI through time may therefore reflect both (1) a general increase in tree species diversity, and (2) a more open environment generally, which implies long distance transport may have a bigger impact on inflating tree cover values. Within the manuscript we cautiously suggest that tree cover in general may be slightly underestimated (i.e. suggesting that (1) may be more important), but in some areas the issue of long-distance transport may be more important, suggesting a slight overestimation of tree cover.

Additional changes

Please rename the supplementary figures according to the example: Supplementary Figure 1 -> Figure S1, etc.

We did not refer to the Supplementary Figures and Tables in the main text, but rather to the section dealing with each issue. We have re-named the sections in the Supplementary, and now refer to these in the main text as Supplementary Information, Section S1 etc. The figures and tables in the Supplementary Information are all labelled sequentially and have been renamed (Table S1, Figure S1 etc). We will list the figures and tables in each section at the beginning of the Supplementary Information file. We will also add references to these supplementary figures and tables in the appropriate place in the main text (e.g. Supplementary Information, Section S1, Table S1).