

We thank the reviewer for their feedback on this article. Our responses to individual comments are in blue, with any proposed changes to the text in red.

General comment

Estimates of forest cover during the Holocene inferred from pollen data may be useful for e.g. climate modelling (e.g., REVEALS-based forest cover was used in regional climate modelling for Europe in Strandberg et al. 2014 in CP, 2022 in QSR, 2023 in CP). The REVEALS-based estimates of plant cover have the advantage (above MAT and pseudobiomization) to provide cover for each of the plant taxa used in the reconstruction, which allows to then calculate the cover of various groups of taxa that may be very useful in e.g. climate modelling. In e.g., Strandberg et al. studies, evergreen trees were separated from summer-green trees. Nonetheless, it is most useful to produce as many reconstructions as possible using different methods to test a) the effect of the method itself on the result, if results differ between methods, and b) the effect on e.g., climate of the differences between the various forest/tree cover products. Therefore, producing pollen-based estimates of forest/tree cover using different methods is of value even if the comparison can be done only for forest cover versus open-land cover and cannot be done for more detailed land-cover units. The study by Sweeney et al. is thus welcome.

I have read the review of Thomas Giesecke and agree with his concerns and comments. Therefore I won't repeat the questions and comments provided by Thomas but wish to add some additional points, questions, and information/references for consideration by the authors. One of my major concerns relates to the discussion on the comparison between the reconstructions from this study's model and the REVEALS model. The authors should acknowledge that their model reconstructs FOREST cover based on a specific definition of forest, i.e. A definition among OTHERS, while the REVEALS model estimates TREE cover. In Serge et al. REVEALS reconstructions TREES include the taxa *Buxus*, *Ericaceae*, *Juniperus*, *Phillyrea* and *Pistacia* that may not belong to the definition of FOREST you are using in your model. This should be considered in the discussion, most importantly for the Mediterranean region. I also comment on this issue in the comments below.

Our major goal in this paper was to develop a site-based model of tree cover using readily available data, and to test this in Europe and compare it with other available reconstructions as a measure of robustness. It was not our intention to provide a detailed explanation or critique of the other methods. We agree, however, that there should be more care taken to acknowledge the difficulty in comparing the datasets, given the different scopes, methods and data presentation for each. We have addressed this issue in our responses to the detailed comments below.

Detailed comments – a mix of major and minor ones

R1: 61-68: The REVEALS reconstructions by Trondman et al. (2015) and Githumbi et al. (2022) were produced for the study of land use as a climate forcing (biogeophysical forcing) in Europe during the Holocene using climate models (papers by Strandberg et al.). The scale of the reconstructions (1 degree) and protocol used was motivated primarily by the primary aim of these reconstructions. These authors produced estimates of plant cover for 25 respectively 31 taxa. These data can be freely accessed in PANGAEA (references to be found in Githumbi et al.). As said above, Strandberg et al. used the total cover of three groups of taxa, open land taxa, evergreen tree taxa and summer-green tree taxa. This is only one possible use of these datasets. Similarly, the REVEALS reconstruction by Serge et al. (2023) was produced for specific use in the European Terranova project (terranova-itn.eu). The focus of these reconstructions was NOT to just reconstruct open land versus forest cover. **The above should be clarified here.**

We agree that the aims of different studies may be different and that this may affect methodological choices, such as spatial scale or what is being reconstructed (e.g. tree cover, PFTs, individual taxon abundances). However, our purpose here is NOT to discuss the purpose of each study or indeed how this affected the methodology. Rather, we are trying to explain what pan-European datasets are available that can be compared to our reconstructions. We have modified this text, in response to comments by Reviewer 1 and Reviewer 2, as follows:

<L56>. Several different techniques that have been applied to reconstruct regional and sub-regional vegetation in Europe using pollen such as biomization/pseudobiomization (e.g. Fyfe et al., 2015; Binney et al., 2017) or the application of MAT using plant functional types (e.g. Davis et al., 2014). Other studies have made reconstructions combining different approaches (e.g. Roberts et al., 2018) and by combining pollen-based reconstructions with simulated potential vegetation (Pirzamanbein et al., 2014). However, the most recent quantitative pan-European pollen-based reconstructions of Holocene vegetation changes have been made using the Landscape Reconstruction Algorithm (LRA) REVEALS approach (Sugita, 2007b, a) or the Modern Analogue Technique (MAT) (Overpeck et al., 1985; Guiot, 1990; Jackson and Williams, 2004; Zanon et al., 2018). The REVEALS method calculates regional vegetation cover based on modelled relationships between pollen abundance, estimated differences in species level pollen productivity and pollen transport, and differences in site characteristics. Initially used at individual sites or small regions (e.g. Gaillard et al., 2010; Nielsen et al., 2012; Marquer et al., 2014), REVEALS was first applied at a pan-European scale by Trondman et al. (2015) and later extended with additional sites, taxa and an improved temporal resolution by Githumbi et al. (2022). The most recent analysis by Serge et al. (2023), is based on 1607 records for 500-year intervals before 700 cal. BP and for the subsequent intervals of 700-350 cal. BP, 350-100 cal. BP and 100 cal. BP- present. They tested the impact of including 13 additional taxa (total n=46) on the vegetation reconstructions, producing maps of landcover and species abundance at record-

containing 1° grid cells. In contrast, the MAT approach reconstructs past vegetation based on identifying modern analogues of fossil pollen assemblages, on the assumption that samples found in the fossil record that share a similar composition to those found in present-day pollen assemblages will have similar vegetation. Zanon et al. (2018) applied MAT to 2,526 individual fossil pollen samples from Europe to generate interpolated maps at 250-year intervals at 5 arc-minute resolution through the Holocene.

R2: 70-87: data demanding versus less data demanding methods: I would not “classify” methods as such. They are ALL data demanding, if RPP is not necessary something else is needed, in your case a good remote-sensed data on tree cover! Then all methods have their assumptions (very thoroughly stated in the case of REVEALS, Sugita 2007a) and difficult decisions to take in terms of data handling and interpretation. The only difficult issue with REVEALS is the need of RPPs. Obtaining RPPs is indeed time consuming although it can be done relatively fast if some money for field work and “man/woman power” is available. This was possible in China for which there is to date many RPP values for most of the major taxa; these were produced within the last 6-7 years. Moreover, this work could be realized in much less time with the technology at hand today, i.e. drones for plant surveys and automatic pollen counting. MAT, PB and your new method use other ways to account for the inter-taxonomical differences in RPP. All methods assume that RPP were stable through time. All methods are challenging, and they all have their pros and cons. None of them can be judged as more or less robust. We should rather use them as “ensembles”, as is done with climate model simulations using several different models. Four reconstructions using four different methods/models (Europe) can be considered as an “ensemble”. More reconstructions would of course be better; I am sure new models will constantly be created and more parameter data will be produced (e.g. RPP). My advice is therefore to avoid evaluation of the methods “against” each other but rather do a synthesis evaluation considering/acknowledging all pros and cons of each method. Here (lines 76-80), do not speak of data demanding method (REVEALS) versus other methods. Similarly, in abstract, discussion, and conclusions, do not grade methods. I do not think there is ONE best method/model.

We are not trying to classify the methods and we agree that all the methods require data. However, there is a difference between methods for which all the necessary data is available (i.e. regional modern pollen training data sets, good remote-sensing products) and methods such as REVEALS which require and are sensitive to regional data on RPPs. As Marie-Jose rightly points out these RPP data could of course be produced, given sufficient people-power and funding, but there are many regions of the world for which they currently not available and are unlikely to be available in the near future. Since our focus here is on the potential issues involved in application of these techniques, we will rewrite this paragraph as follows:

<Line 77>. The REVEALS approach requires, and is sensitive to, estimates of relative pollen productivity (RPP) and pollen fall speeds (FS) for individual species (Bunting and Farrell, 2022;

Githumbi et al., 2022; Serge et al., 2023). Landscape-level reconstructions are problematic if RPP and FS information are not available for relatively common taxa (Harrison et al., 2020). RPP values have been estimated for common taxa in Europe and China, and there are a limited number of studies from North America (see e.g. Wiczorek and Herzschuh, 2020) but are not readily available for other regions of the world. The MAT technique requires a large modern pollen data set for training purposes, but such data sets are now available for all regions of the world. However, the application of MAT involves a number of arbitrary decisions including the choice of analogue threshold (i.e. how similar modern and fossil assemblages must be to be considered analogous), and the number of analogues used (Jackson and Williams, 2004). Techniques designed to minimise the number of samples for which no analogues are found, such as grouping species into plant functional types (PFTs), introduce further uncertainties since the allocation of pollen taxa to PFTs is often ambiguous (Zanon et al., 2018).

Additional references:

Harrison, S. P., Gaillard, M.-J., Stocker, B. D., Vander Linden, M., Klein Goldewijk, K., Boles, O., Braconnot, P., Dawson, A., Fluet-Chouinard, E., Kaplan, J. O., Kastner, T., Pausata, F. S. R., Robinson, E., Whitehouse, N. J., Madella, M., and Morrison, K. D.: Development and testing scenarios for implementing land use and land cover changes during the Holocene in Earth system model experiments, *Geosci Model Dev*, 13, 805–824, <https://doi.org/10.5194/gmd-13-805-2020>, 2020.

Wiczorek, M. and Herzschuh, U.: Compilation of relative pollen productivity (RPP) estimates and taxonomically harmonised RPP datasets for single continents and Northern Hemisphere extratropics, *Earth Syst Sci Data*, 12, 3515–3528, <https://doi.org/10.5194/essd-12-3515-2020>, 2020.

R3: 106-108: there are more vegetation classes that could be considered as “non-natural” such as planted forest, cultivated trees, even grazed land if strongly fertilized isn’t “natural”, as well as ley. You perhaps rather mean “non-pollen producing land” in the case of crops like most cereals (except rye) that do not release pollen as long as they are in the field, i.e. before collected and treated to get the grains. But there are other crops. Do the remote-sensed data provide details on the crops. Do all the crop areas you deleted correspond to areas that do not produce much pollen?

In our response to a comment from Reviewer 1, we amended the text to exclude the “non-natural” terminology, given its potential ambiguity as follows:

<L.106>. A composite map of modern tree cover for the region 12°W to 45°E and 34-73°N was generated by averaging annual percentage tree cover data from Copernicus annual land cover maps from 2015 to 2019 (Buchhorn et al., 2020a, e, d, c, b), after removing cells dominated (> 50%) by

other land-cover classes, including bare ground, built up areas, moss or lichen, permanent water, snow, and crops (Fig. 2A).

R4: 126: Githumbi et al. did not exclude large bogs, which would certainly have been better. But the decision came from Trondman et al. 2015 in which it was decided not to exclude sites based on this criterion; instead, cells including large bogs were emphasized as providing less reliable results. The correct reference for large bogs being not recommended for REVEALS reconstructions, even when multiple sites are used for one reconstruction, is Trondman et al., 2016 in VHA.

We did not mean to imply that Githumbi et al. (2022) excluded bogs, but rather that they urged caution in using these records, specifically "...REVEALS estimates of plant cover using pollen assemblages from large bogs should only be interpreted with great caution ...". Whether they should or should not be used for REVEALS-based reconstructions is not really the issue. The key point here is that we include taxa that grow on bogs and we felt that this could bias our tree cover

reconstructions. In response to a comment from Reviewer 1, we have amended this text as follows:

<Line 126>. However, bog records with a radius $\geq 400\text{m}$ were excluded from the analysis. Githumbi et al (2022) indicated that caution was necessary in interpreting REVEALS vegetation reconstruction estimates based on large bogs and, given that we included taxa that grow on bog surfaces in our analysis (see below), we excluded large bogs to reduce the potential for these to bias the regional vegetation reconstructions.

However, given that we do not wish to comment on the applicability for REVEALS-based reconstructions but are simply justifying the exclusion of large bogs for our reconstructions we will further modify this to:

<Line 126>. However, bog records with a radius $\geq 400\text{m}$ were excluded from the analysis because we included taxa that grow on bog surfaces in our analysis (see below), and the exclusion of large bogs reduces the potential for these taxa to bias the regional vegetation reconstructions.

R5: 130: I would also refer here to Marquer et al. 2020, QSR 228-106049.

We will include this reference, as follows:

<Line 128>. Finally, since upslope pollen transport is known to increase the proportion of non-local pollen at high-elevation sites (Fall, 1992; Ortu et al., 2008, 2010), and the complex topography of mountainous areas also impacts pollen transport (Markgraf, 1980; Bunting et al., 2008; Marquer et al., 2020; Wörl et al., 2022), we excluded 236 site records above 1000m.

Additional reference

Marquer, L., Mazier, F., Sugita, S., Galop, D., Houet, T., Faure, E., Gaillard, M.-J., Haunold, S., de Munnik, N., Simonneau, A., De Vleeschouwer, F., and Le Roux, G.: Pollen-based

reconstruction of Holocene land-cover in mountain regions: Evaluation of the Landscape Reconstruction Algorithm in the Vicdessos valley, northern Pyrenees, France, *Quat. Sci. Rev.*, 228, 106049, <https://doi.org/10.1016/j.quascirev.2019.106049>, 2020.

144-154:

- **R6:** It is true that the area represented by a REVEALS reconstruction of plant cover using pollen data from large lakes might correspond to Z_{max} as defined in Sugita 2007a “distance within which most pollen comes from”, i.e. the maximum distance from a large lake from which pollen is coming and deposited in that lake (my words). However, this might not be true for small sites; in fact, we do not know as this has never been tested because single small sites are not appropriate for REVEALS applications. Small sites can only be used for REVEALS reconstructions if they are used together, i.e. within a grid cell, and a mean REVEALS estimate is calculated. That’s fine you use Z_{max} for all your sites, but **you should acknowledge that you ASSUME that Z_{max} is also the area represented by a REVEALS reconstruction using pollen from a small lake, but this has never been tested.** A longer discussion on the “spatial scale” of a REVEALS reconstruction can be found in Li et al. 2020 in ESR, page 5, upper left column: *“In theory and practice, however, the strict definition of the pollen source area is difficult for REVEALS application. Sugita (2007a) defined it as the area within which most of the pollen comes from. Simulations and previous empirical studies (e.g. Sugita, 2007a, b; Hellman et al., 2008a; Sugita et al., 2010; Mazier et al., 2012) have indicated that, when the radius of the source area defined varies from 50 km to 400 km, the REVEALS results of regional vegetation reconstruction do not change significantly. The basin size is potentially important for REVEALS-based estimate of regional vegetation because differences in basin size among sites can lead to a significant site-to-site variation in the pollen assemblages. However, as long as the multiple study sites are located within a region that satisfies the first assumption as described above (no gradients in the overall vegetation composition), the averaged REVEALS estimates effectively represent the regional vegetation composition as demonstrated in Hellman et al., 2008a. The accuracy of the reconstructed vegetation against the observed vegetation composition was assessed for areas of 50 km × 50 km and 100 km × 100 km around each site in two regions of southern Sweden. The pollen records used are from 5 large lakes in each region, thus 10 lakes in total, that vary in size between 76 ha and 1965 ha. The results support the main conclusions and implications for the REVEALS application based on the theory and the simulations described in Sugita (2007a). Such evaluation is an essential step for credible application of the REVEALS model.”*

We are at a loss to see how these comments refer to the cited paragraph, which describes the use of the source area formula developed by Prentice (1985). The choice of the 70% threshold was also taken from that paper. We used median fall speeds derived from Githumbi et al (2022) and Serge et al (2023) because the Prentice (1985) formulation used species specific values which were not

available for all taxa included in our analysis. We will make this point clearer by modifying the text as follows:

<Line 144>. The source area for each record, and hence the appropriate area for the calculation of mean tree cover, was calculated using Prentice's (1985) source area formula for 70% of pollen, and lake or bog area from the SMPDS. The original source area formula used species-specific FS values but, since these were not available for all the taxa used in our analysis, here we used the median FS (0.03) from Githumbi et al. (2022) and Serge et al. (2023) since the tree cover map represents the broad species community around each record location.

- **R7:** Why use 70% of pollen rather than 80% or 90%?

Please see response to R6. Note that we tested the 70% criterion and found that it gave plausible results.

- **R8:** Why use the median FS 0.03 and not simply estimate the distance for the lightest pollen type you use in your reconstruction, which would be Z_{max} ? If you use 0.03 you get a Z_{max} for a vegetation composed of taxa such as *Pinus*, *Ulmus*, *Buxus*. But your total vegetation is composed of taxa with pollen grains that come from much longer distances (most herbs but also several common tree taxa with lower FS values). Why do you assume/think that this isn't important to define the area around each of your sites, i.e. the area that is represented by the pollen assemblages in those sites?

Our approach is based on matching the site source area to the underlying tree cover maps. The choice of a median FS is a pragmatic attempt to reflect the average vegetation composition surrounding the site area. We take your point that by using this median we may not be fully capturing the source of the lighter pollen, especially in more open environments. But taking a source area that is too large would also be problematic, as heavier pollen at the fringes of this area would not reach the site or be reflected in the pollen assemblage.

R9: 200-203: I do not fully understand how you handled the Serge et al. REVEALS estimates of tree cover for comparison with your reconstruction. There is doubtless a problem in comparing single site reconstructions within the site's "70% Z_{max} " with REVEALS reconstructions representing at minimum the area of the 1-degree grid cell including the pollen sites used in the reconstruction. Wouldn't it be fairer (for each method!) to calculate the median tree cover from your single site reconstructions covering +/- one or several 1-degree grid cells of the REVEALS reconstruction before comparing the results? The REVEALS estimates are mostly based on several sites in each grid cell. If you compare your results for each site with the median forest cover from several grid cells in the REVEALS reconstruction, you'll compare the vegetation cover between two areas of different sizes, smaller size for your reconstruction than for the REVEALS reconstruction (size of several 1-degree

grid cells). Please, acknowledge this issue. – Now I see in Supplement S6 that you seem to have done what I am suggesting above as fairer, although I am not sure.

Comparing the different reconstruction values is a challenge, given the different methods and scope of the reconstructions. The ideal comparison would be to use site-based estimates for each since this allows for potential variability between nearby sites that can be masked when using median values for a 1° grid cell, as in the REVEALS estimates. Supplement S6 investigates the implications of this to some extent, by averaging tree cover values for records that occur in the same grid cells. But this still leaves the problem of differences between reconstructions potentially being the result of having different numbers of sites in each grid cell. We will try to clarify this, both in this paragraph (L200-207), and in new text in the discussion, as follows:

<Line 200>. 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). 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 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.

We will modify the text in the discussion as follows:

<Line 363>. Our reconstructions show that tree cover peaked in the mid-Holocene period, with median tree cover ca. 40% greater than at the beginning of the Holocene. This general pattern is shown by the REVEALS and MAT reconstructions, and is also visible in plant functional type (Davis et al., 2015) and pseudo-biomization reconstructions of vegetation cover (Fyfe et al., 2015). Despite the similarities in median values between our reconstructions and those calculated from the REVEALS and MAT reconstructions, our site-based estimates are not fully comparable with the gridded estimates provided by the REVEALS reconstruction and the gridded and interpolated values provided by the MAT reconstructions. Nevertheless, the similarities give some support to the overall robustness of our reconstructions.

Figures 3 and 4:

- **R10:** I am not astonished that your reconstruction performs better than other reconstructions given that you use the same source of forest data to establish your model and test it – Note that I understand you haven't used the same sites to create the model and test the model, of course!!! – Your model is entirely dependent on the forest data you have used, and there must be good chances that your predictions will be relatively good when compared with the same data source of forest cover.

We compared the reconstructions to the Copernicus data set because this is considered to be the best available. There is no guarantee that the final model would reproduce the training data, as you point out. We used leave-one-out cross validation to test the model performance (Methods section 2.2).

- **R11:** Further, REVEALS (Fig. 4C) is closer to your reconstruction (Fig. 3B) than Zanon (Fig. A4) in terms of the spread of the points between low and high predicted % cover. Here, as said above, I am not sure to understand how you made the comparison between your results and Serge et al. The points in Figures 3 and 4 represent your sites. **I do not see how the REVEALS values can be constrained to represent the plant cover for your “pollen source area” (i.e. Z max for each of your site) as these REVEALS estimates are valid for at least the 1-degree grid cell including the sites used for the reconstruction, i.e. a much larger area; it is NOT valid for a much smaller area.** You write on line 269-271 “*The correlation between etc.... is only 0.5. This is partly caused etc..... but even when taking this into account etc.... were still lower (0.59)*”. **From Supplement S6 I understand you took the scale of reconstructions into account, but I am not sure how.** In any case it increases the correlation, which is good news! **Could you please clarify.**

For each site location from our reconstruction, we extracted the median tree cover value from each time window for the area around each site. If a site source area is fully contained within an individual grid-cell, then the value for that cell is returned and subsequently compared to the observational value. If the source area straddles several grid cells, the value returned is the median, calculated to take into account the proportion of each grid cell value covered by the site-specific source area. Supplement 6 shows the implications of restricting the sites to one per grid-cell, to potentially limit some of the double counting of values from the REVEALS (and MAT) reconstructions. Here, where sites share the same grid-cell, the average of the median value for each site within a grid-cell is calculated. Please see our response to R9 for changes to the method text to clarify this.

- **R12:** Does your model produce error estimates on the predicted forest cover? If so, why are they not provided in the graphs; similarly, why not use the error estimates provided with the REVEALS-based forest cover?

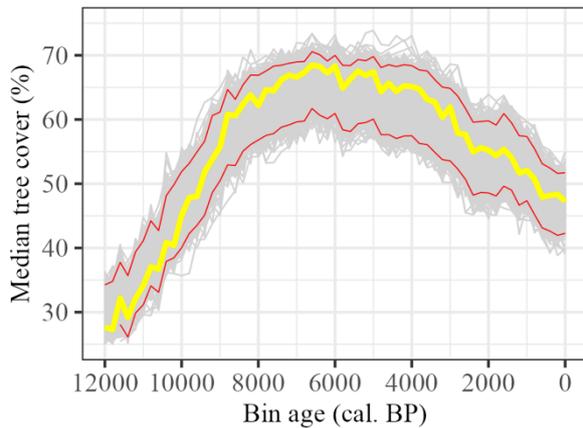
We didn't include error estimates for the site predictions. However, we agree that this would be a useful inclusion, both for the site estimates and as another way to reflect uncertainty in the median reconstructed tree cover. We have included the following additional text and supplementary figure:

<Line 214> ...binned in 200-year bins. Standard error estimates for site predictions were calculated through the application of a bootstrapping approach, with 1000 resamples of the model training data used to generate models, and equivalent quantile mapping adjustment, which were then applied to the fossil pollen data.

<Line 288> This same general pattern is shown when considering changes in mean tree cover (Supplementary Information: S7), different LOESS smoothing (R package *locfit* Loader, 2020) of the median tree cover value (Fig. 5B), and based on median tree covers reconstructed using the model bootstraps used to generate reconstruction standard errors (Supplementary Information: S8).

S8: Median reconstructed tree cover, with bootstrapped models

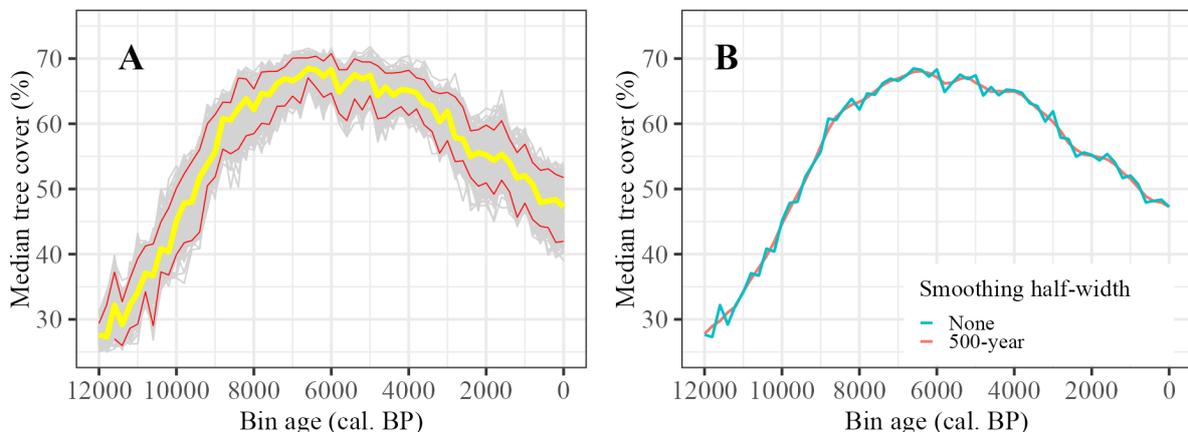
In order to generate reconstruction standard errors, the predictive model that linked observed tree cover to modern pollen data was generated 1000 times by bootstrapping the modern pollen data. These models were also used to generate the equivalent number of quantile mapping adjustments, by relating model predictions using the full dataset to observations. Together these elements were used to produce 1000 different reconstructions of tree cover for each fossil data sample, with prediction standard error calculated by sample and averaged by 200-year bin. As well as using the bootstrapped reconstructions to generate the standard error, we can use the median of these bootstraps as another way of assessing the confidence in the median reconstruction. Supplementary Figure 6 shows the median tree cover estimate, together with the bootstrapped medians of tree cover based on the different models and quantile mapping adjustment generated. Although the bootstrapped medians follow the same general pattern as the reconstructed median, maximum tree cover values for the reconstruction are generally on the higher side compared with the bootstrapped medians, implying that some training samples may have a larger influence on the generated model than others.



Supplementary Figure 6: Median reconstructed tree cover for Europe from 12,000 to 0 cal. BP, with 95% confidence intervals for models generated through 1000 bootstrap resamples of model training data

R13: Figure 5B: Can't see the different colors for the smoothing half-width in the graph, probably because the curves are very similar? Find another way to illustrate this, or have only -none and 500-year, and comment in the text for the other two?

The curves are very similar and we agree it is redundant to have the 1000 and 2000-year half-widths within in the figure. We will adjust the figure and amend the text as follows:



<Line 287>. This same pattern is shown when considering changes in mean tree cover (Supplementary Information: S7) and LOESS smoothing (R package *locfit* Loader, 2020) of the median tree cover value (Fig. 5B)

R14: Figure 7: Could the high values of forest cover in Zanon 10-8 k be due to over-representation of *Pinus* and *Betula*? In your reconstruction there is the same tendency.

Since *Pinus* and *Betula* are included in all three reconstructions, it is difficult to see why their over-representation between 10-8 ka in the Zanon et al. reconstruction would be due to this. However, diagnosing the cause of this peak is beyond the scope of this paper. We do not see the double peak shown by Zanon et al. in our reconstructions, and our reconstructions in general show lower values of tree cover than Zanon et al, despite including *Pinus* and *Betula*.

R15: 370-376: You write: “*This could reflect the conservative nature of our moder-day tree cover model*”. YES, I think this might well be the major reason to this difference. Can’t this also explain the same “phenomenon” in the MAT reconstruction (Zanon)?

Yes, as we mention in the subsequent sentence, Zanon et al. (2018) also indicate that they likely underestimate tree cover at higher levels of tree cover.

379-385:

- **R16:** I do not understand this reasoning. First, you did not account for all “anthropogenic land use”, you only excluded non-pollen producing areas, including crops (see also my comment above on methods and in relation to crops). You write “We account for this (land use/” non-natural vegetation”) in defining modern source areas in our model, since the pollen only provides evidence of the natural vegetation.” This sounds very odd. Pollen provides evidence of natural vegetation AND human-influenced vegetation, i.e., not only crop cultivation but also grazing, managed woodland, planted forests etc.... It is the purpose of the REVEALS model and MAT and pseudo-biomization (PB). The purpose is NOT to reconstruct natural forest cover, but the “actual” forest cover. **Please clarify, this section is very confusing. What you did was to exclude non-pollen producing areas, which is good (I only have concerns about crops, see comments above). You did NOT exclude anthropogenic land use.**

We agree that this text could be clarified. We have amended the text as follows:

<Line 379>. 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 the maps of 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.

- **R17:** Note that in Roberts et al 2018 (Scientific Reports, Figure 2), the REVEALS and PB (closed sum) reconstructions agree with each other and agree with the Corine remote-sensed forest cover. Roberts et al. (2018) write “*A second means of testing the different forest reconstructions is to examine how well they match modern forest cover for the same grid cells. While remotely sensed estimates of forest might be expected to offer the clearest results and the most rigorous test, in fact the Corine and Forest Map 2006 data have strongly different outcomes, i.e. 45% and 29% modern forest cover, respectively. This inconsistency partly reflects the ontological question of “what is a forest?” Corine uses distinct land-cover classes, and land classified as forest may*

include some open areas as the minimum required crown cover for a forested class is only 30%. The Forest Map 2006 is based on a minimum 50% tree crown cover with 5m used as a minimum height of trees. It also highlights the epistemological problem that differences in spatial resolution of measurement can fundamentally alter results³⁰, in this case between 25m and 100m measured spatial resolution. An alternative data source for modern forest cover derives from surface pollen samples. We have transformed the surface pollen data set for Europe²⁷ using both variants of the PBM, which leads to modern forest cover estimates of 49% (PBMsc) and 54% (PBMlcc). Overall, most estimates of modern forest cover for the grid cells used by Trondman et al. 10 are between 45% and 49%; that is, close to that reconstructed for the 100 to –65 BP REVEALS time window.”

This is the kind of issues you should acknowledge in this discussion. Please see my major comment under “General comment” above, i.e. REVEALS reconstructs tree cover, while your model reconstructs FOREST cover based on a specific definition of FOREST.

In fact, our reconstruction is of tree cover not forest cover. In response to a comment by reviewer 2, we have amended the text throughout to make this clear and have added an explanation of what we mean by tree cover in the Introduction.

419-424:

- **R18:** In this discussion, you do not attempt to explain why REVEALS estimates > 65% tree cover around 6-5.5 k BP while both your model and Zanon’s MAT predict forest cover < 50% (40-45%) 8.5-5 ka (your model) respectively 9.5- 4 ka BP (Zanon). I would emphasize here the major difference between MAT, your model and REVEALS in terms of WHAT is reconstructed. REVEALS estimates TREE cover and NOT FOREST cover. The authors using REVEALS may define taxa as trees in various ways. You need to consider what taxa are defined as trees in Serge et al. and think about whether these taxa may belong to land-cover types your remote-sensed data define as non-forest vegetation.

Please see response to comment above.

- **R19:** You write (412-424) “.....*the more rapid decline in tree cover during the last millenium shown , and shown more dramatically in the Zanon And Serge reconstructions, is more difficult to explain - Human influence on the landscape MAY help explain etc.....*”. I do not understand why you are so careful/doubtful on whether human impact may explain tree/forest cover decline from 6k, 5.5 k, 4k (depending on the region) and more so from 2k BP. This is documented and has been tested in a large number of publications by palaeoecologists, archaeologists, historians, etc... over past decades; I can’t see what is problematic or controversial with this. In Strandberg et al. 2023 (CP), Figure 1 (based on Githumbi et al. REVEALS reconstruction, see figure copied below) shows clearly the increase in mean and median tree cover in three major biomes of Europe (note that we did not separate the Atlantic

region, which would indeed have been interesting) from mid Holocene, accelerating around 2 k BP, and more so around 1 k BP. Strandberg et al. (2023) write: “*The recent pollen-based reconstruction of land cover in Europe (spatial resolution of 1°; Githumbi et al., 2022) suggests that the earliest of the two major deforestation episodes before the start of the Modern period (1500 CE (0.45 ka) – present) took place between ca. 4 and 2.5 ka, i.e. the period during which the Bronze Age culture expanded from southeastern (Turkey, Greece) to central and western Europe (Mediterranean area included) and northern Europe (Champion et al., 1994; Coles and Harding, 1979). The second deforestation episode (before the Modern time deforestation) occurred ca. 0.9–0.5 ka, during the Middle Ages (ca. 500 (1.45 ka)–1500 CE in most of Europe, started 1050 CE (0.9 ka) in northern Europe) (Fig. 1a). The difference in open land cover between 4 and 2.5 ka of ca. 10 % (in either mean or median cover; Fig. 1a) is assumed to represent deforestation of Europe by Bronze Age cultures. This change in the land cover of Europe was also explained by deforestation for agriculture in the study of Marquer et al. (2017). If we consider ...etc., the Bronze Age deforestation corresponds to an increase in open land cover by 200 % since 4 ka. etc. ... (Githumbi et al., 2022).*” And “*The time around 3 ka (the Bronze Age) was also pinpointed as the time when “the planet [was] largely transformed by hunter-gatherers, farmers, and pastoralists”, as suggested by an archaeological global assessment of land use from 10 ka to 1850 CE (ArchaeoGLOBE Project, 2019).*” There is no doubt that deforestation caused by land use did strongly influence forest/tree cover in Europe over the last ca. 3000 years.

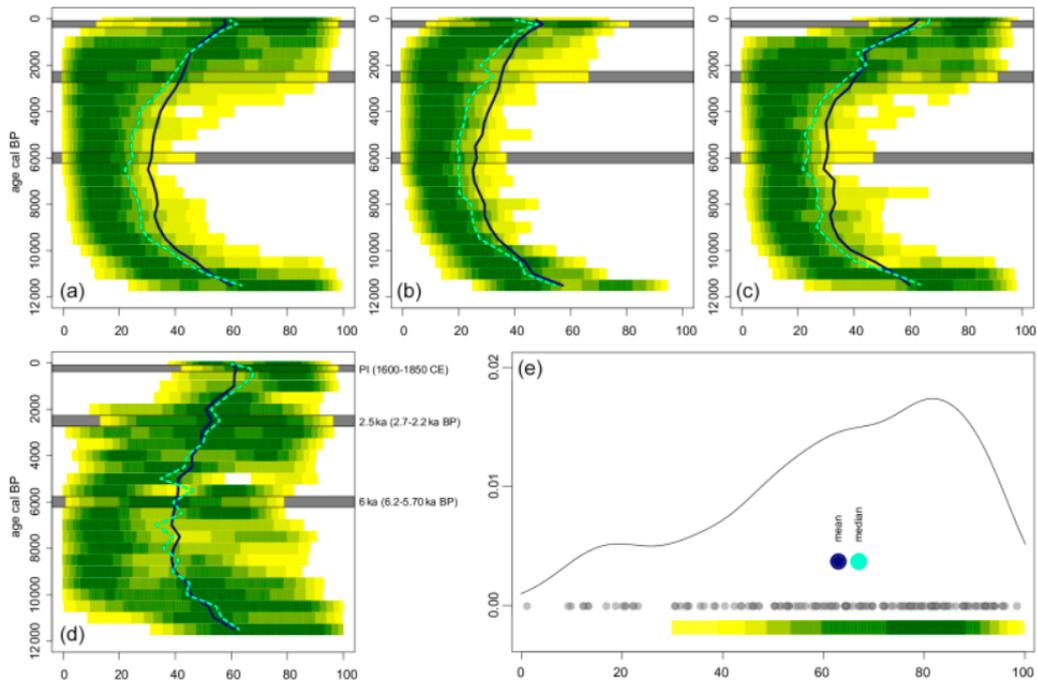


Figure 1. Density plots of REVEALS-based open land (% cover) for (a) all European grid cells, (b) the boreal zone, (c) temperate Europe, and (d) Mediterranean Europe. Full black lines represent the mean value of open land cover, and dashed green lines represent the median value. Panel (e) is illustrative of how the density plots are derived for one time window, showing data values for all grid cells (opaque grey circles), the derived density function as a curve and expressed on the colour gradient, and the mean and median values. Key time intervals (6, 2.5 ka, and the PI period) are highlighted on panels (a-d) by the grey bars.

There have indeed been many studies that have attributed observed changes to human activities, but these are largely based on correlations between two variables, e.g. vegetation changes and archaeological records. Correlation is not causation. It is true that the Strandberg et al. paper cited above compares REVEALS-based reconstructions with potential vegetation reconstructions based on using a model (LPJ-GUESS), but these are (a) only as good as the vegetation model itself, (b) conditioned by the climate (and CO₂) inputs used to drive the model, and (c) do not account for climate-induced changes in vegetation disturbance such as fire. Note that over the past 12,000 years changes in CO₂ of themselves would have had a non-negligible impact on the vegetation through changing water-use efficiency. The only way to establish rigorously the causes of reconstructed changes in vegetation during the Holocene would be to model them quantitatively using robust reconstructions of all the variables that could impact the vegetation, i.e. climate changes, CO₂ changes, changes in fire regimes and changes in human activity. This is a goal that partially underpins our efforts here to reconstruct tree cover. We are cautious about attributing any of the reconstructed changes to human activities until this quantitative assessment is made, although we acknowledge (given the seemingly robust evidence of human population changes over the last millennium available from HYDE) that at least on that time frame human activities look to be a plausible explanation.

Abstract and conclusions:

R20: Please revise the abstract and conclusions following the revisions you might consider making in response to the comments above. Among other, the statements “*our approach is more robust and less data-demanding than previously applied methods*” (abstract) and “*Our simple approach produces ... etc.... using more complex methods, and thus provide a less data-demanding approach... etc... of the world*” (conclusions) should be revised. There are no less/least data-demanding method/model and/or best (most robust) method/model; there are several possible methods/models that all are data-demanding and have their pros and cons.

Please see our response to R2 above and the proposed modification to the text in the Introduction.

We will modify the last sentence of the abstract as follows:

The reconstructed patterns of change in tree cover are similar to those shown by previous reconstructions, but our approach is relatively simple, only requires readily available data and could therefore be applied to reconstruct tree cover globally.

We will modify the last sentence of the conclusions as follows:

Our simple approach produces similar reconstructions of the trends in tree cover during the Holocene reconstructed using more complex methods, and since it only requires readily available data could be used to reconstruct tree cover in other regions of the world.