Letter to the editor

Dear Kirsten Elgers,

We thank you for the opportunity to revise our manuscript once more. While all three reviewers had interesting suggestions, none criticized how we applied the REVEALS model to reconstruct vegetation from the input pollen data and we were able to accommodate a majority of recommendations.
Following suggestions from all three reviewers, we added a comparison of our reconstructed tree cover with tree cover reconstructed by Serge et al. (2023) and found that our tree version has a significantly lower reconstruction error than the previous reconstruction. We also restructured and expanded our discussion to explore methodological shortcomings and differences to previous reconstructions, reasons for continued tree cover overestimation, and data set application. As suggested by both Thomas Giesecke and the anonymous reviewer, we decided to move the calculation of the 80% pollen source area to the supplementary materials (S5). An additional supplement (S6) also compares different rasterization methods and shows marginal differences..
We also corrected an error in the validation and reconstruction figures, where a 2x4° rasterization was mistakenly used instead of the intended 2x2° format. Importantly, this adjustment does not impact dataset validity. We believe that these and other minor changes improved the manuscript. We answer Reviewer comments in detail in the individual responses.
The most recent version of the dataset can be found on Zenodo (https://doi.org/10.5281/zenodo.13902921) as well as code for data set production (https://doi.org/10.5281/zenodo.10191859) and dynamic rasterization (https://doi.org/10.5281/zenodo.13902976). We will update the dataset currently deposited on PANGAEA once reviewers agree to the current version (https://doi.pangaea.de/10.1594/PANGAEA.961588). It is for this reason that we currently link to Zenodo in our manuscript instead of PANGAEA, but we will change this before publishing.

Best regards
Laura Schild and Ulrike Herzschuh

Reply to Anonymous reviewer

General Reply

Dear Reviewer,

Thank you very much for your valuable comments and suggestions. We are pleased to inform you that we have addressed all of your concerns and have implemented the majority of your suggestions.

To address your main concerns:

- **14 ka cutoff**: We now provide a clear justification for this cut-off, explaining its relevance to the climatic stability starting from this period.
- **80% PSA calculations**: As per your suggestion, we have moved these calculations to the supplementary materials.
- **Comparison with Serge et al.**: We have expanded the manuscript to include a direct comparison of our dataset with the previously published European REVEALS reconstruction by Serge et al.

Additionally, we have improved the structure of the discussion and expanded upon the validity of the dataset, ensuring a clearer presentation of our findings. We also corrected an error in the validation and reconstruction figures, where a 2x4° rasterization was mistakenly used instead of the intended 2x2° format. Importantly, this adjustment does not change the main messages of the paper. Once again, thank you for your thoughtful input. The dataset (https://doi.org/10.5281/zenodo.13902921), code (https://doi.org/10.5281/zenodo.10191859), and rasterization script (https://doi.org/10.5281/zenodo.13902976) can be found on Zenodo. We are confident that your suggestions have enhanced the quality and clarity of the manuscript and thank you for your review.

Best regards,
Laura Schild and Ulrike Herzschuh

In-depth replies

Major Comments

1. No justification is given for the 14 ka cutoff in the time series; please elaborate. Is this a result of historic climate values, the length of the sedimentary records used or both? 14 ka is ending up in the deglacial period, which is a bit of a push considering the abrupt climatic change at the onset of the Holocene. The prior LegacyPollen dataset paper (Herzschuh et al., 2022) states this as such, calling it a deglacial period/transition. Thus, at 14 ka, pollen productivity may not be

accounted for with Holocene interglacial PPEs. The authors should consider stopping at the Holocene boundary.

The climatic conditions during the past 14 ka BP were relatively stable which is why we consider it as a cutoff for the REVEALS reconstruction (Shakun et al., 2012). Additionally, the main peak of vegetation turnover had already happened during the late glacial, so that we can assume that forested landscapes, similar to today's landscapes, had already been established (Li et al., 2024; Mottl et al., 2021). We have added an explanation in the manuscript.

Li, C., Dallmeyer, A., Ni, J., Chevalier, M., Willeit, M., Andreev, A. A., Cao, X., Schild, L., Heim, B., and Herzschuh, U.: Global biome changes over the last 21,000 years inferred from model-data comparisons, EGUsphere, 1–26, https://doi.org/10.5194/egusphere-2024-1862, 2024.

Mottl, O., Flantua, S. G. A., Bhatta, K. P., Felde, V. A., Giesecke, T., Goring, S., Grimm, E. C., Haberle, S., Hooghiemstra, H., Ivory, S., Kuneš, P., Wolters, S., Seddon, A. W. R., and Williams, J. W.: Global acceleration in rates of vegetation change over the past 18,000 years, Science, 372, 860–864, https://doi.org/10.1126/science.abg1685, 2021.

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, Nature, 484, 49–54, https://doi.org/10.1038/nature10915, 2012.

taxonomic harmonization for consistency of records. Reconstruction chronologies may, therefore, differ slightly from previous reconstructions due to this revised age modeling. Spatial data coverage of records in the reconstruction is dense in Europe (~~1275~~ 1287 records) and North America (~~1016 records~~1040) and sparsest in Asia (~~441~~446) (see Fig. 1). The records' sample density decreases with age (see Fig. 2). Only samples dated to 14 ka BP or younger were used to ensure that the climatic conditions of recorded vegetation were similar to the modern climate.

2. Source Area: The selection of an 80% source area is justified by the following: "The primary objective of this calculation is to provide a clear understanding of the scale of the source area for users unfamiliar with pollen data. It highlights the regional nature of lacustrine pollen data and demonstrates the influence of lake size on this source area", which is a valid comment. However, no justification is given for using 80% pollen source areas; why not 70%, 90% or 95% etc? Is the 80% source area just a reflection of the value set for comparing the GPM to the LSM in Theuerkauf et al. (2016) based on Prentice and Webbs' (1986) somewhat arbitrary statement that "…significant amounts of pollen can be derived from far beyond the source area for, say, 80% of the pollen grains" found at a site.
   Also, it would be worthwhile to mention that when using the GPM source, areas have been found to vary greatly between taxa based on their individual fall speeds (Theuerkauf et al., 2016).
   The upper end of the 80% source area presented here is 762 km, which seems rather large.
   If the purpose of the 80% source area is to show that pollen source area increases with basin radius (Fig 5), I think the words from this section would be better spent elsewhere. The inclusion of an 80% source area here is confusing, and the manuscript may improve with this information being supplementary information instead.
   More importantly, no explanation of the dispersal model used to calculate the pollen source area is given, assumed to be the Gaussian plume with unstable conditions? The choice of dispersal model would substantially affect the results of this 'source area' calculation, hence this should be noted in the discussion. As paper is currently lacking in novelty, a good improvement would be to test dispersal models (e.g. changing settings in the GPM, but also trying the Lagrangian Stochastic model) in the validation for these regions. These is somewhat a low-hanging fruit, as the R code for REVEALSinR allows to change dispersal model, but the PPEs would need to be recalculated with the same settings to match the various attempts.
   - Following your suggestion we decided to move the calculation and description of 80% pollen sources areas to the supplementary materials (S5, please find it with the new

manuscript upload). We still want to touch upon your points raised above here.
The threshold of 80% is indeed an arbitrary choice. However, we find this reasonable for this illustrative measure. We account for the fact that different area have different source areas by choosing the area where the median of pollen input for all taxa is 80%. This means that lakes with the same areas but different compositions could have different 80% source areas. However, we found basin area to have the bigger impact.
The dispersal model chosen plays an important role in the application of REVEALS and therefore the calculation of 80% pollen source area. This is why we state it in Table 2. Finally, we do not believe that the testing of different dispersal models and modes is within the scope of this data description paper, which ESSD states should “[...] describe original research data, databases, or combined datasets derived from them”. [..] Although examples of data outcomes may prove necessary to demonstrate data quality, extensive interpretations of data – i.e. detailed analysis as an author might report in a research article – remain outside the scope of this data journal.” (https://www.earth-system-science-data.net/about/manuscript_types.html).

3. No comparison to prior European scale reconstructions is presented, e.g. (Serge et al., 2023), which would be very helpful to position where the methodology differs/falls in comparison to prior reconstructions.
    - We added a comparison with Serge et al. to our manuscript. Differences mainly exist in the choice of RPP values, arboreal taxa, and the aggregation procedure. Especially the latter is now described in detail in our methods section.
      We compare our reconstructed tree cover to Serge et al.’s and find that our tree cover tends to be lower. A comparison of validations shows a significant lower MAE with our dataset compared to Serge et al.’s dataset. Please see the relevant new manuscript sections pasted below.

Methodological rasterization differences:

This method of temporal and spatial averaging differs from several previous REVEALS applications. Pollen counts are often summed in temporal bins prior to running REVEALS to increase pollen counts and reduce uncertainty (Trondman et al., 2015; Githumbi et . However, temporally averaging after the REVEALs application, as implemented by us, increases the flexibility of the dataset with the trade-off of potentially increased uncertainty. Rasterization has previously been perfomred by using a weighted average taking into account the basin size of the original record (Trondman et al., 2015; Githumbi et al., 2022; Serge et al., 2023). However, the most recent REVEALS-based North American vegetation reconstruction uses the same arithmetic mean as described above (Dawson et al., 2024b). When comparing our method of temporal and spatial aggregation to that used by previous European reconstructions (e.g. Serge et al., 2023), we also found no significant differences in the validation of reconstructed tree cover (see S6).

Comparison with Serge et al. (methods):

~~Using REVEALS , radii of~~ Additionally, we compare our REVEALS reconstruction to the most recently published REVEALS reconstruction in Europe by Serge et al. (2023, version: RPPs.st1). We average our reconstruction in the same grid and temporal bins as used by Serge et al. to compare the reconstructed tree cover between both reconstructions. To get the total tree cover, we sum evergreen and summergreen tree cover values in Serge et al.'s dataset, while excluding broadleaved summergreen temperate warm shrubs (BSTWS) and broadleaved evergreen xeric shrubs (BEXS). We validate the previous reconstruction and our reconstruction in the most recent time slice available in Serge et al.'s reconstruction (-65 to 100 BP, https://doi.org/10.48579/PRO/J5(with the remote sensing forest cover and compare validations. Unfortunately, direct validation could only be performed with the most recent time slice available online, rather than the historical time slice used in the validation by Serge et al., which limits the ability to reproduce their validation results exactly. We do not apply any openness correction here as we do not have comparable 80~~% pollen source areas were calculated for large lakes(see Fig. ??) . The radii indicate in which area 80%~~

~~of the deposited pollen originated from (see Section 2.2.2) and yield an understanding of which area the pollen record is representative of, which is especially useful when individual time series from large lakes are being used for analyses. The 80~~% pollen source areas ~~are roughly a function of basin size (see Fig.??)and range between 155 km and 762 km. The median 80% pollen source radius is 225 km including all large lakes~~ available for the records used in Serge et al. (2023). The reconstruction by Serge et al. differs in the temporal as well as spatial aggregation routine, as described above. Definition of arboreal taxa varies, a different RPP-value set was used, and the amount of total records included is higher than in our reconstruction (Serge et al.: 1607, LegacyVegetation: 1287).

Comparison with Serge et al. (results):

The comparison between our reconstruction and tree cover reconstructed in Serge et al. (2023) shows that LegacyVegetation (this publication) tends to have a lower tree cover independent of sample age. Serge et al. tend to overestimate forest cover even more than LegacyVegetation which leads to a much lower mean absolute error in LegacyVegetation compared to Serge et al. (Fig. 10). The MAE for LegacyVegetation is slightly higher than presented in Fig. 7 due to the difference in spatial resolution and the lack of openness correction.

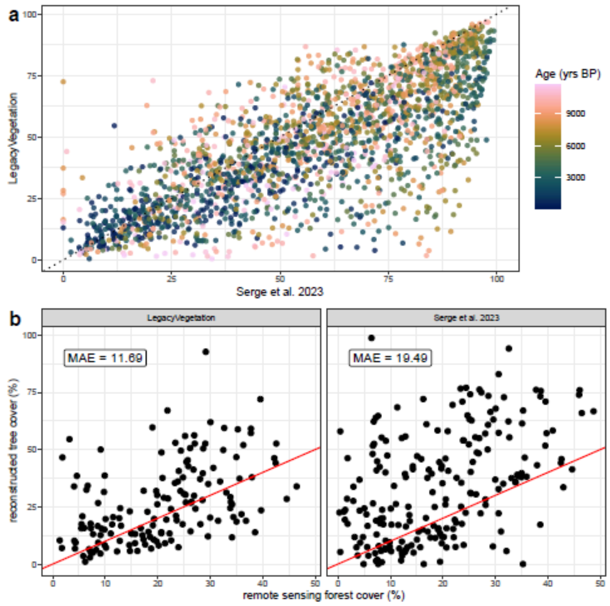


Figure 10. (a) Comparison between LegacyVegetation (this publication) and the tree cover from Serge et al. (2023) and (b) validations with modern, remote-sensing forest cover for both data sets.

4. Section 3.2 is slightly arbitrary/worded strangely; trends in pollen and REVEALS are broadly the same, and the composition of the vegetation is what changes with PPEs adjusting the pollen values. Cyperaceae is a strange choice to include for the bogs and swamps as it would be locally deposited around the coring location. Whilst the Peatland setting in REVEALSinR would assume that the Cyperaceae originates from outside of the immediate area surrounding the core, many of such pollen grains would be local.
    - The section is supposed to be rather illustrative, showing the reader the compositional nature of the dataset produced and what differences between the original pollen record and the REVEALS estimate could look like. We illustrate continental mean compositions and highlight only the 8 most common taxa. So while peatlands are included, they only contribute to this mean together with small and large lakes. We highlight that the reconstruction is difficult to validate on a taxon-level and that we believe it is associated with higher uncertainty in the discussion. We also edited this Section 3.2 slightly to highlight its illustrative character (see tracked changes below).

~~well as Poaceae and Cyperaceae with increased coverages.~~ REVEALS was used to reconstruct quantitative vegetation cover. Here we ~~compared~~ illustrate a comparison between these reconstructed compositions to the original pollen composition. Differences in composition between ~~Pollen~~ pollen data and REVEALS are apparent for all continents of the Northern Hemisphere. Some clear examples include: increases of Cyperaceae in all continents, decreases of ~~Betula~~ *Betula* in Europe, decreases of ~~Pinus~~ *Pinus* in all continents, and increases of ~~Acer~~ *Acer* in North America with the application of REVEALS and its intended correction of ~~taxon-sepcific~~ taxon-specific biases (see Fig. 4).

5. Figure 11 reveals that outside of Europe, continental vegetation reconstructions are not yet robust, likely due to the density of sedimentary records in North America and Asia or the implementation of continental averaged PPEs or another factor not explored here but nonetheless are better than unmodelled pollen reconstructions.

- This is correct. We highlight and discuss these limitations for North America and Asia in our edited discussion. We point out a lack of RPP studies in North America and hypothesize a higher regional variability of RPP values in Asia and North America than in Europe as potential reasons for higher reconstruction errors. Please see the tracked changes below.

~~Reconstruction results are also~~ However, continental differences are evident in the quality of tree cover reconstruction, with Europe showing a significantly larger reduction in errors compared to other regions. North America and Asia exhibit larger reconstruction errors in the REVEALS estimates, though these are still lower than those derived from tree pollen percentages. Notably, regions such as the Great Lakes, the Labrador Peninsula, and Pacific Northwest display particularly high errors in tree cover reconstruction. Asia, characterized by sparser coverage, presents fewer large errors. This highlights the need for improved vegetation reconstruction especially in North America and Asia. The reason for this reduced performance could lie in a lack of RPP studies, especially in North America, or in a significantly higher regional variability of RPP values compared to Europe. While differences in validation outcomes across varying spatial resolutions are marginal (see S4), some variability is observed when different grids are employed, highlighting spatial heterogeneity in reconstruction success. Despite these caveats, overall trends appear consistent, with acceptable correlation coefficients, though absolute values in certain regions remain challenging to interpret with confidence as tree cover continues to be overestimated in all continents.

6. The discussion section would benefit from sub-sectioning, perhaps between (1) more methodological issues/limitations arising from the selected method and potential solutions or a forward outlook on how a northern hemisphere may become as robust as those seen prior for just continental scale European reconstructions. (2) more generalised and outlooking insights, e.g., those in relation to the validation and training of climate vegetation models, as mentioned early on in the manuscript (Abstract and Introduction). Section 2 may require more nuance and restraint when related back to some of the clearer limitations of the presented method, neatly summarised by the reconstruction error being much greater than Europe for North America and Asia in Figure 11.
   - We edited the discussion to follow a clearer structure and include a more extensive discussion of methodological limitations and data set validity.
   - Please see the new manuscript version for the edited discussion.

7. Moreover, there is a lack of discussion surrounding the role of human agency on forest cover through the Holocene, which is likely one of its biggest drivers.
   - Our aim is not really to discuss the reasons for forest cover trends, but rather the differences between Pollen and REVEALS estimates and the remote sensing forest cover used in our validation. In this context, we now highlight the anthropogenic impact on modern forest cover leading to potential mismatches between “modern” pollen-based tree cover (younger than 100 BP) and modern remote-sensing forest cover (2000 to 2015). Please see the tracked changes below for this edit.

is evident in the validations. Furthermore, pollen-based estimates are derived from records that span a much longer timescale than the modern forest cover data available, even though modern timeslices are used for validation. Increased anthropogenic impact could exacerbate discrepancies between pollen-based and ~~other plant inventories or compositional remote sensing products could offer better validation.~~ remote-sensing estimates. This could contribute to the overestimation of forest cover, which persists in all continents. Additionally, these modern and arguably unnatural vegetation conditions may not correspond to past vegetation and may therefore have reduced significance for the reconstruction of past, natural landscapes.

8. Case sensitivity in the axes and legends of figures should be consistent. Italics of species-genus names need to be implemented throughout.

- We checked and implemented this.

Other Comments by Line:

1. L6: Why is the last 14 ka selected as the cutoff, 11.7 ka or 11 ka requiring less justification as clearer alternatives?
    - We have added an explanation for this cutoff in the manuscript (see tracked changes below).

taxonomic harmonization for consistency of records. Reconstruction chronologies may, therefore, differ slightly from previous reconstructions due to this revised age modeling. Spatial data coverage of records in the reconstruction is dense in Europe (~~1275~~ 1287 records) and North America (~~1016 records~~1040) and sparsest in Asia (~~441~~446) (see Fig. 1). The records' sample density decreases with age (see Fig. 2). Only samples dated to 14 ka BP or younger were used to ensure that the climatic conditions of recorded vegetation were similar to the modern climate.

2. L8-9: does not read correctly, should be, e.g. "The pollen source area where 80% of the pollen originated within was calculated for large lakes (>50ha)".
    - Removed from main manuscript and added to supplementary materials (S5).
3. L40-45: called the Fagerlind Effect.
    - Added.

in the pollen record, while those with low pollen productivity and less effective dispersal are underrepresented. These factors, together with the compositional nature of pollen data, result in a non-linear relationship between pollen and vegetation~~(Prentice and Webb III, 1986)~~, titled the Fagerlind effect (Prentice and Webb III, 1986; Fagerlind, 1952). Approaches such

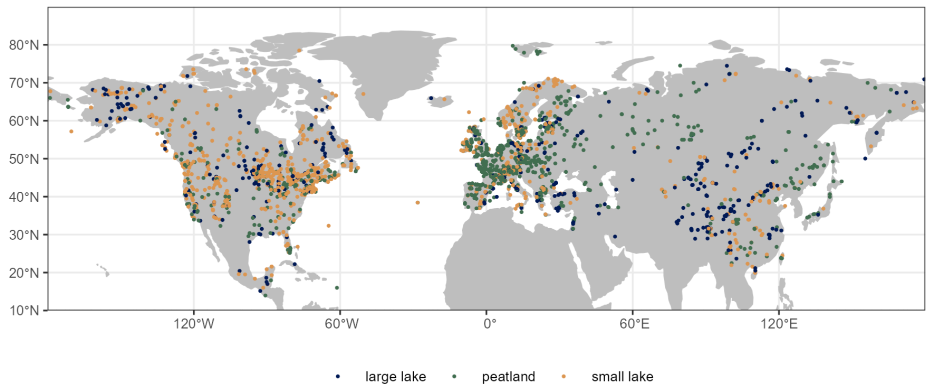
4. L44: The R-value and ERV models are constituents of the LRA. The REVEALS model is not merely a "refinement" per se of the ERV. The ERV is a key part of the REVEALS model incorporated in the PPE calculation stage of the LRA.
    - We reworded this section. Please see tracked changes below.

tion~~(Prentice and Webb III, 1986)~~, titled the Fagerlind effect (Prentice and Webb III, 1986; Fagerlind, 1952). Approaches such as the R-value model (Davis, 1963; Webb et al., 1981) and the extended R-value model (Parsons and Prentice, 1981) were created to address this issue and were ~~refined with~~ later included into Sugita's (2007) model for "Regional Estimates of Vegetation Abundance from Large Sites" (REVEALS). By accounting for taxon-specific RPP and fall speed values, as well as basin-

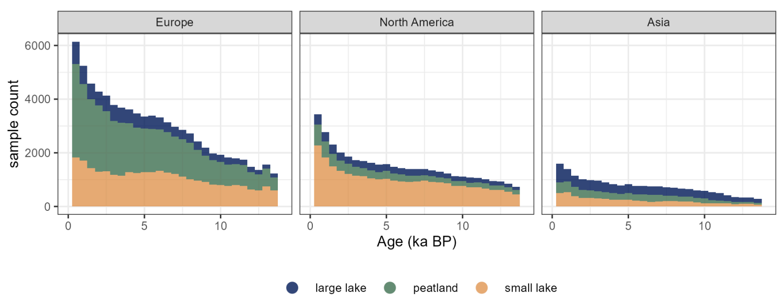
5. L53: space needed after Githumbi et al. (2022)
    - Implemented.
6. L75: why are Australasian and Latin American pollen databases included here? The reconstruction is for the Northern Hemisphere. See the end of the next comment in relation to this.
    - These were related to the LegacyPollen dataset and were now removed to avoid confusion. (screenshot of tracked changes)

The pollen data synthesis LegacyPollen2.0 (Li et al., 2024b) includes 3680 temporally resolved records (time-series) distributed globally. Data were collected from individual publications and the Neotoma Paleoecology Database which includes data from the European Pollen Database, ~~the QUAVIDA data base for Australasia, the Latin American Pollen Database, the African Pollen Database~~ and the North American Pollen database ~~(Flantua et al., 2015; Fyfe et al., 2009b; Giesecke et al., 2014; Lézine et al., 2021; Rowe~~ (Fyfe et al., 2009b; Giesecke et al., 2014; Whitmore et al., 2005; Williams et al., 2018). An overview of Neotoma records included in LegacyPollen 2.0 and this reconstruction can be found in S1.

7. L79: why are some records “marine in origin”? Is this a typo? If it relates to the LegacyPollen 2.0 dataset creation, this should be covered in the LegacyPollen2.0 publication, not here it causes confusion.
   - These were related to the LegacyPollen dataset and were now removed to avoid confusion.
8. Fig1: consider separating this figure into four map panels, three regions and one hemispheric map as shown currently or reduce the point size slightly.
   - We reduced the point size.



9. Fig2: consider making the bars 3 stacked colours between the three categories of sites, large lakes, peatland and small lakes, or adding these as another facet element to the figures.
   - Implemented.



10. L85: REVEALS does not need to be defined again.
    - Removed definition.

The REVEALS model ~~("Regional Estimates of Vegetation Abundance from Large Sites")~~ estimates quantitative vegetation coverage from pollen assemblages using site and taxon-specific parameters (Sugita, 2007). Based on wind speed and taxon-

11. L100: what version of the DISQOVER package is being used also, is REVEALSinR a function rather than a package itself?
    - We changed the package name and added the version used.

We use REVEALSinR from the DISQOVER package in R to implement REVEALS (Theuerkauf et al., 2016, Version 0.9.13, https://github . It mainly differs from the original program by Sugita (2007) in the process of error calculation. REVEALSinR includes repeated model runs with random error added to RPP values and pollen counts (see Table 2 for the number of variations). The resulting distribution of REVEALS results allows for an estimation of the standard deviation of vegetation cover per taxon. The program by Sugita (2007), however, derives error estimates with a hybrid method from a variance-covariance matrix of PPE and Monte Carlo simulations. For further details on the REVEALS model see the original publication Sugita (2007) ~~or~~

12. L104-130/Tab2: how many model runs (n) occurred for each site? It is not mentioned but is a key parameter of the REVEALSinR function defaulting to 1000; see the example given in the function information in the most recent DISQOVER package release version available here: https://github.com/MartinTheuerkauf/disqover.

```
REVEALSinR(
pollen,
params,
tBasin,
dBasin,
dwm = "lsm unstable",
n = 1000,
regionCutoff = 1e+05,
ppefun = rnorm_reveals,
pollenfun = rmultinom_reveals,
writeresults = FALSE,
verbose = TRUE
n        number of model runs per time slice, by default 1000
```

- All static parameters used in the REVEALSinR function are listed in Table 2 (see below). The amount of pollen count and RPP variations (n) is set to 2000 in our REVEALS runs.

Table 2. Static model parameters and model settings for REVEALS runs using REVEALSinR (Theuerkauf et al., 2016).

Parameter	Values and settings used in model run
atmospheric model	unstable atmosphere
dispersal model	gaussian plume
wind speed	$3m \times s^{-1}$
maximum extent of regional vegetation (Z_{max})	1000 km
number of RPP and pollen count variations (n)	2000
peatland basin area (for missing sizes)	31.41 ha
lake basin area (for missing sizes)	49 ha
function to randomize pollen counts	rmultinom_reveals

13. L125: To avoid confusion with the usage of (n), as mentioned previously, which is the number of model runs in the REVEALSinR implementation, consider the notation for the mixing value to something else.
    - We changed the notation to m.

standard value which can be found in Table 2 together with several constant parameters set in REVEALSinR. Lastly, we also reduced computational effort in REVEALSinR by implementing a maximum number of steps in the lake model used to model mixing in the basin. The number of steps was set to 500 unless m falls below that maximum value for $m = basin\,radius/10$ for basins with a radius of at least 1000 m and $m = basin\,radius/2$ for basins with a radius smaller than 1000 m.

14. L160: why was 80% chosen to see the major comment.
    - Please see our reply to your major comment on this.
15. Sec3.1: 762km is rather large as the 80% source area, especially if Zmax is set to 1000km (Tab 2 region cutoff value). Similarly, the lower 155km is still quite large, even for the GPM (Theuerkauf et al., 2016), which generally has larger pollen source areas for taxa.

  - We only show the 80% pollen source areas for large lakes here. As they depend both on Zmax (rather large in our case) and the dispersal model this can be the case. Additionally, the pollen composition influences the 80% source area, by increasing it when far-dispersing taxa are dominant or decreasing it when short-dispersing taxa make up a majority of the sample.
  - Following your suggestion above, we have decided to move the calculation and results of the 80% pollen source area to the supplementary materials (S5).

16. L168-9: Betula, Pinus, Acer. Italicisation and case sensitivity need checking throughout the manuscript, including figures and captions.
    - Implemented.
17. Fig 7: the key agebin (0 ka – modern) would benefit from being presented as point data rather than a time series. Four bar charts with the three series: pollen vs REVEALS vs remote sensing, might be a good idea. The time series is interesting but is not the impactful point.
    - We chose a time series here to illustrate continental and Northern Hemispheric trend, while adding the modern remote sensing forest cover to already indicate, the potential error in this reconstruction. As we go into a lot more detail with comparing modern pollen-based, REVEALS estimated and remote sensing tree and forest cover, we believe that this Figure does not have to focus on the modern age bin.
18. L184-185: implies that REVEALS is not working well in North America and Asia for continental averages. The difference should be larger and closer to the remote sensing for the modern modelled pollen values.
    - Yes, our REVEALS application is indeed not able to reduce the error more in Asia and North America, even though the error is reduced compared to tree pollen percentages. We chose to highlight and discuss this more in our revised manuscript (please see tracked changes below).

~~Reconstruction results are also~~ However, continental differences are evident in the quality of tree cover reconstruction, with Europe showing a significantly larger reduction in errors compared to other regions. North America and Asia exhibit larger reconstruction errors in the REVEALS estimates, though these are still lower than those derived from tree pollen percentages. Notably, regions such as the Great Lakes, the Labrador Peninsula, and the Pacific Northwest display particularly high errors in tree cover reconstruction. Asia, characterized by sparser coverage, presents fewer large errors increasing the overall continental reconstruction error. This highlights the need for improved vegetation reconstruction, especially in North America and Asia. The reason for this reduced performance could lie in a lack of RPP studies, especially in North America, or in a significantly higher regional variability of RPP values compared to Europe. While differences in validation outcomes across varying spatial resolutions are marginal (see S4), some variability is observed when different grids are employed, highlighting spatial heterogeneity in reconstruction success. Despite these caveats, overall trends in tree cover appear consistent, with acceptable correlation coefficients, though absolute values in certain regions remain challenging to interpret with confidence as tree cover continues to be overestimated in all continents.

19. Sec3.4: while MAE is useful, I would consider using additional measures like dissimilarity to help identify/quantify the difference between REVEALS vs remote sensing, pollen vs remote sensing, and split into the three continents (Jackson and Williams, 2004; Overpeck et al., 1985). It would be possible to estimate critical values.
    Jackson, S.T., Williams, J.W., 2004. MODERN ANALOGS IN QUATERNARY PALEOECOLOGY: Here Today, Gone Yesterday, Gone Tomorrow? Annu. Rev. Earth Planet. Sci. 32, 495–537. https://doi.org/10.1146/annurev.earth.32.101802.120435

Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative Interpretation of Fossil Pollen Spectra: Dissimilarity Coefficients and the Method of Modern Analogs. Quat. res. 23, 87–108. https://doi.org/10.1016/0033-5894(85)90074-2

- While dissimilarity measures are useful to compare multivariate data, they are not applicable here as we are comparing only one variable (reconstructed tree cover, i.e. a univariate comparison). Multivariate comparisons would be great to validate REVEALS applications further. However large-scale standardized data is missing to do so. We also highlight the need for such data for better validation in the discussion (see manuscript excerpt below). In addition to the MAE, we also show the distribution of error in the boxplot in Figure 8 to give more information surrounding the error distribution.

Another challenge lies in validating the compositional reconstruction results. It remains uncertain whether RPP values have remained stable over time, and historical compositional data are not only scarce but also likely too recent to test this assumption effectively (Baker et al., 2016). Validating modern compositional reconstructions on large spatial scales is therefore difficult. ~~Moreover, there is uncertainty surrounding the success of the compositional reconstructions.~~ As global compositional vegetation data ~~is~~ are not readily available, ~~using remote sensing forest cover poses~~ remote sensing of tree cover serves as the best option for validation. ~~Even with an accurate forest cover reconstruction, uncertainties persist~~ But even with accurate tree cover reconstructions, uncertainties remain regarding the abundance of individual taxa due to the aggregated nature of the ~~forest~~ tree cover measure. To address this issue, global syntheses of ~~forest and other~~ tree and plant inventories or compositional remote sensing products could ~~offer better validation. Another challenge lies in validating the results with past vegetation data. It is uncertain whether RPP values have remained stable over time, and historical compositional data are not only scarce but likely too recent to test this assumption (Baker et al., 2016). Vegetational compositions~~ provide more robust validation. Additionally, vegetational compositions derived from sedimentary ancient DNA ~~could provide a solution. Local aDNA~~ (sedaDNA) offer a promising avenue for comparing past vegetation data. Local quantitative sedaDNA vegetation signals could be averaged across multiple records ~~within a pollen source area to generate a comparable reconstructed vegetation composition using a different proxy and to compare~~ to compare with pollen-based results ~~(Niemeyer et al., 2017)~~(Niemeyer et al., 2017; Capo et al., 2021).

20. L202: Europe shows the best results, but trees are still being overrepresented here in comparison to the remote sensed cover; what is driving the overrepresentation of trees here and beyond? This is not unpacked in great enough detail in the discussion.
    - We expanded on this in the Discussion. (Please see tracked changes below.)

~~However, the reconstructions are associated with some of the limitations of sedimentary pollen data . This includes age uncertainty, temporal mixing, and irregular spatial and temporal resolution of records. Age uncertainty is already treated as best as possible through consistent age modeling of the pollen dataset (Li et al., 2022a, 2021). Nevertheless, in general, replicating sediment and peat cores could provide more accurate estimates~~ Validating pollen-based tree cover estimates with remote sensing-derived forest cover also presents a challenge. One key issue is the inherent errors associated with remote sensing forest cover data. While validation using other sensors is possible, only a limited subset of the available data is cross-validated with Lidar data, which itself is characterized by limited spatial coverage (Sexton et al., 2013). A critical limitation of surface reflectance methods, as used in the Landsat-based forest cover, is their reliance on a 2D perspective, primarily capturing the forest canopy. This means that the understory is often not detected, resulting in an incomplete representation of the forest structure. In contrast, pollen-based estimates provide a more comprehensive, stratified view of the vegetation, as they

18

---

incorporate all contributing taxa, not just the tree canopy. Despite this broader scope, pollen data and REVEALS estimates tend to emphasize trees more than other vegetation types consistently as is evident in the validations. Furthermore, pollen-based estimates are derived from records that span a much longer timescale than the modern forest cover data available, even though modern timeslices are used for validation. Increased anthropogenic impact could exacerbate discrepancies between pollen-based and remote-sensing estimates. This could contribute to the overestimation of forest cover, which persists in all continents. Additionally, these modern and arguably unnatural vegetation conditions may not correspond to past vegetation and may therefore have reduced significance for the reconstruction of past, natural landscapes.

21. Fig9: REVEALS reconstruction validations for Europe can be compared with the prior validations of (Serge et al., 2023).
    - We added this comparison in our manuscript. Please see relevant parts of the Methods, Results and Discussion sections in our reply to your general comment above.
22. Fig11: Strongly suggests that robust continental reconstruction in North America and Asia is not possible yet, even when validated just on the arboreal layer.
    - As stated in our reply to you “other comment” No. 18, we expand on this in the Discussion.
23. Sec4: see the main comments section.
    - We edited the Discussion to follow a clearer structure as suggested. Please see the tracked changes document.
24. L232: a comparison of the results of Serge et al. (2023) is warranted. Should the forest cover for Europe be better if not the same/similar?
    - We added this comparison as stated in our reply to your “other comment” no. 21. We find that our reconstruction is closer to remote sensing forest cover than the previous Serge et al. reconstruction and state some possible reasons. See more details in our reply to your general comment no 3.
25. L255-260: DNA comments do not make sense in terms of vegetation reconstructions. A reconstruction implies quantity. SedDNA/eDNA data today remains point data, which provides no information on the quantity of vegetation around a site, merely presence and absence.
    - Quantitative estimates of past vegetation from sedimentary ancient DNA are possible with meta barcoding and have become a more common method to reconstruct past

environments. We added an additional reference here (https://www.mdpi.com/2571-550X/4/1/6). (see tracked changes)

could provide more robust validation. Additionally, vegetational compositions derived from sedimentary ancient DNA ~~could provide a solution. Local~~ (sedaDNA) offer a promising avenue for comparing past vegetation data. Local quantitative aDNA vegetation signals could be averaged across multiple records ~~within a pollen source area to generate a comparable reconstructed vegetation composition using a different proxy and to compare~~ to compare with pollen-based results ~~(Niemeyer et al., 2017)~~ (Niemeyer et al., 2017; Capo et al., 2021).

26. L269-271: These sentences about the reliability of data are more methodological considerations that belong in the methods or a separate discussion unpacking the limitations of the presented method.
    - As stated in our reply to you major comment and “other comment” no. 23. We edited the Discussion to follow a clearer structure.

Reply to Giesecke

General reply

Dear Thomas Giesecke,

Thank you very much for your valuable comments and suggestions. We believe we were able to implement the majority of your recommendations.

- We validate our reconstructed modern vegetation using observed modern forest cover. Validation is a standard practice applied in many datasets, including those published in *Earth System Science Data* (ESSD).
- Regarding your question on why REVEALS still overestimates forest cover, we have expanded our discussion to highlight the differences between remote sensing-based forest cover and pollen-based forest cover. We also address potential deficiencies with RPP values in this context. We now compare our results directly with those of Serge et al., focusing on validation results, where we observe a significantly smaller error in our reconstruction.
- As suggested, we have moved the 80% source area calculation to the supplementary materials.
- Furthermore, we restructured the discussion to expand on dataset usage and challenges.
- Regarding the current dataset version, we believe that there may have been some confusion regarding the repository. We made the new version of the dataset available on Zenodo and will update the final version on PANGAEA, due to the longer processing times at PANGAEA. The dataset deposited on Zenodo does not include reconstructions from marine records or samples before 14 ka BP.

The dataset (https://doi.org/10.5281/zenodo.13902921), code (https://doi.org/10.5281/zenodo.10191859), and rasterization script (https://doi.org/10.5281/zenodo.13902976) can be found on Zenodo. Once again, thank you for your insightful feedback, which has contributed to the overall improvement of the manuscript. We are confident that the changes strengthened the clarity and depth of our work.

Best regards,
Laura Schild and Ulrike Herzschuh

In-depth replies

General comments

Reading the revised manuscript Schild et al. I am glad to see that all of the controversial analyses were dropped. However, the full initial data is still available online. The remaining manuscript has a strong focus on testing the REVEALS estimates of forest cover against remote sensing data. Like before, I feel this is a research question that should be individually addressed in a research and not in this data paper, unless the aim of the comparison is the justification of the chosen RPPE. If this topic is kept as part of this publication it needs more attention: the authors need to describe how they made sure that the samples used in the comparison represent the last 100 years; clearly state what was compared; differences of using averages of small lakes or a large lake should be evaluated; and the spatial differences nicely displayed in Fig 11 should be discussed. Fig 11 also indicates that there are grids with an overestimate in tree cover (even more common at a higher resolution grid), which needs to be explained.

Firstly, we do believe that validations are an essential part of a data publication and would like to highlight that this is indeed a common practice in ESSD (e.g. https://essd.copernicus.org/articles/16/2917/2024/, https://essd.copernicus.org/articles/16/2465/2024/, https://essd.copernicus.org/articles/16/2449/2024/).

Secondly, we would like to address your concerns. We choose to include samples younger than 100 BP as determined by Age models used in the LegacyPollen dataset. We clearly state in our methods section which values are being compared. These are tree pollen percentages, tree REVEALS estimates (reconstructed tree cover), and remote sensing forest cover (a temporal average for the years 2000, 2005, 2010, and 2015). We do this for each grid cell that has a valid REVEALS tree cover estimate for the timeslice 100 BP - present. We define grid cells as valid, which include samples from at least one large lake or several smaller basins.
We chose to do this with the averaged gridded data since we recommend data to be used in this format. we provide additional validations with different grid cell sizes in the supplementary materials. In our revised discussion, we now mention the slight differences between validations results of differing spatial resolutions and highlight areas in which especially high errors occur (see tracked changes pasted below).

~~Reconstruction results are also~~ However, continental differences are evident in the quality of tree cover reconstruction, with Europe showing a significantly larger reduction in errors compared to other regions. North America and Asia exhibit larger reconstruction errors in the REVEALS estimates, though these are still lower than those derived from tree pollen percentages. Notably, regions such as the Great Lakes, the Labrador Peninsula, and the Pacific Northwest display particularly high errors in tree cover reconstruction. Asia, characterized by sparser coverage, presents fewer large errors increasing the overall continental reconstruction error. This highlights the need for improved vegetation reconstruction, especially in North America and Asia. The reason for this reduced performance could lie in a lack of RPP studies, especially in North America, or in a significantly higher regional variability of RPP values compared to Europe. While differences in validation outcomes across varying spatial resolutions are marginal (see S4), some variability is observed when different grids are employed, highlighting spatial heterogeneity in reconstruction success. Despite these caveats, overall trends in tree cover appear consistent, with acceptable correlation coefficients, though absolute values in certain regions remain challenging to interpret with confidence as tree cover continues to be overestimated in all continents.

Also the calculation of the source area for 80% of the pollen is not well discussed in terms of its correctness, usefulness and implication. Personally, I don't think that the theoretical 80% area is useful with respect to openness comparisons where the signal may be determined by the theoretical source area of 30 to 50 % of pollen (compare Matthias & Giesecke (2014) QSR, 87, 12-23.). I recommend removing these additional analysis, presenting them elsewhere.

Following another reviewer's suggestion, we decided to move the calculation and description of the 80% pollen source area to the supplementary materials (see S5 of the uploaded new manuscript version).

Instead the authors could provide examples of past vegetation composition or compare their results in more detail to existing quantitative reconstructions. They could elaborate how this data differs from previous continental studies, stress what this specific dataset is better at, and where the limitations are (see below under Data).

In our edited manuscript, we expand on methodological differences to previous quantitative reconstructions, compare our reconstructed tree cover to the most recent European REVEALS reconstruction by Serge et al. (2023), and expand on the limitations of our dataset in the discussion. Our rasterization follows a different methodology, which features a trade-off between increased flexibility in temporal and spatial resolution with slightly increased uncertainty. The comparison with Serge et al. highlights much lower reconstruction errors in our dataset than in Serge et al.'s, while their data coverage is slightly better.

Differences in rasterization:

This method of temporal and spatial averaging differs from several previous REVEALS applications. Pollen counts are often summed in temporal bins prior to running REVEALS to increase pollen counts and reduce uncertainty (Trondman et al., 2015; Githumbi et . However, temporally averaging after the REVEALs application, as implemented by us, increases the flexibility of the dataset with the trade-off of potentially increased uncertainty. Rasterization has previously been perfomred by using a weighted average taking into account the basin size of the original record (Trondman et al., 2015; Githumbi et al., 2022; Serge et al., 2023). However, the most recent REVEALS-based North American vegetation reconstruction uses the same arithmetic mean as described above (Dawson et al., 2024b). When comparing our method of temporal and spatial aggregation to that used by previous European reconstructions (e.g. Serge et al., 2023), we also found no significant differences in the validation of reconstructed tree cover (see S6).

Comparing with Serge et al. (methods):

~~Using REVEALS , radii of~~ Additionally, we compare our REVEALS reconstruction to the most recently published REVEALS reconstruction in Europe by Serge et al. (2023, version: RPPs.st1). We average our reconstruction in the same grid and temporal bins as used by Serge et al. to compare the reconstructed tree cover between both reconstructions. To get the total tree cover, we sum evergreen and summergreen tree cover values in Serge et al.'s dataset, while excluding broadleaved summergreen temperate warm shrubs (BSTWS) and broadleaved evergreen xeric shrubs (BEXS). We validate the previous reconstruction and our reconstruction in the most recent time slice available in Serge et al.'s reconstruction (-65 to 100 BP, https://doi.org/10.48579/PRO/J5 with the remote sensing forest cover and compare validations. Unfortunately, direct validation could only be performed with the most recent time slice available online, rather than the historical time slice used in the validation by Serge et al., which limits the ability to reproduce their validation results exactly. We do not apply any openness correction here as we do not have comparable 80~~% pollen source areas were calculated for large lakes(see Fig. ??) . The radii indicate in which area 80%~~

~~of the deposited pollen originated from (see Section 2.2.2) and yield an understanding of which area the pollen record is representative of, which is especially useful when individual time series from large lakes are being used for analyses. The 80~~% pollen source areas ~~are roughly a function of basin size (see Fig.??)and range between 155 km and 762 km. The median 80% pollen source radius is 225 km including all large lakes~~ available for the records used in Serge et al. (2023). The reconstruction by Serge et al. differs in the temporal as well as spatial aggregation routine, as described above. Definition of arboreal taxa varies, a different RPP-value set was used, and the amount of total records included is higher than in our reconstruction (Serge et al.: 1607, LegacyVegetation: 1287).

Comparison with Serge et al. (results):

A specific comparison with Serge et al. (2023) reveals that our reconstruction generally shows lower forest cover across Europe. Unfortunately, direct validation could only be performed with the most recent time slice available online, rather than the historical time slice used in the validation by Serge et al., which limits the ability to reproduce their validation results exactly. Despite this limitation, our reconstruction demonstrates a lower MAE, indicating improved accuracy. This is notable given that Serge et al. utilized a larger number of records in their study. One potential explanation for these differences could lie in the variations in RPP values and the selection of arboreal taxa used in the reconstruction, as we employ an arboreal tree treshold and include more taxa in our REVEALS reconstruction.

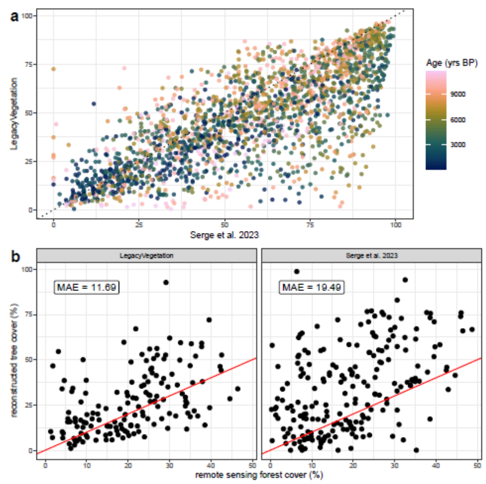


Figure 10. (a) Comparison between LegacyVegetation (this publication) and the tree cover from Serge et al. (2023) and (b) validations with modern, remote-sensing forest cover for both data sets.

Specific comments

I am using the line numbering of the track changes document.

Title: I don’t think the short title “LegacyVegetation” is informative and adding a version number already here is confusing. What if there will not be an update?

- We removed the version number, but would like to maintain the dataset title.

LegacyVegetation~~1.0~~: Northern Hemisphere reconstruction of ~~vegetation composition~~ past plant cover and ~~forest~~ total tree cover from pollen archives of the last 14 ka

11: Better “The source area of 80% of the deposited pollen ...”

- We moved calculations and descriptions of the 80% pollen source area to the supplementary materials (S5).

22: Better “compiling” rather than “collecting”. I thought that work was done by Wieczorek et al. 2020, so perhaps better “updating”.

- “collecting” here refers to conducting more RPP studies in North America to have a larger amount of taxa and/or regional variability represented. It does not refer to the compilation of RPP values in a synthesis.

70: This is not really related to the previous sentence, which outlines the need for reconstructing forest cover. Validation is only possible for the modern situation which is often different from the natural that we are aiming to reconstruct: e.g. trees are harvested when they are still young reducing the overall pollen production, alien taxa (Pseudotsuga) make up large proportions of the forest, N-fertilization may lead to higher pollen production ....

- We extended our discussion on remote sensing data characteristics that could influence validation.
  We also improved wording in this text passage in the introduction. (two screenshots)

With its importance for the assessment of biome stability, carbon storage, climatic feedbacks, and land-use-change, ~~forest~~ tree cover is an often reconstructed variable (e.g. Fyfe et al., 2015; Githumbi et al., 2022; Serge et al., 2023). Due to the global availability of remote sensing data on contemporary ~~forest cover, it also offers good opportunities for the validation of reconstructions~~ tree cover, reconstructions of tree cover in modern time slices may even be validated (Hjelle et al., 2015; Roberts et al., 2018). Yet, only Serge et al. (2023) and Pirzamanbein et al. (2014) use this opportunity for extensive validation and even improvement of reconstructions from European pollen records. No grid-cell based validations exist for the Northern Hemisphere.

Validating pollen-based tree cover estimates with remote sensing-derived forest cover also presents a challenge. One key issue is the inherent errors associated with remote sensing forest cover ~~poses as the best option for validation. Even with an accurate forest coverreconstruction, uncertainties persist regarding the abundance of individual taxadue to the aggregated nature of the forest cover measure. To address this, global syntheses of forest~~ data. While validation using other sensors is possible, only a limited subset of the available data is cross-validated with Lidar data, which itself is characterized by limited spatial coverage (Sexton et al., 2013). A critical limitation of surface reflectance methods, as used in the Landsat-based forest cover, is their reliance on a 2D perspective, primarily capturing the forest canopy. This means that the understory is often not detected, resulting in an incomplete representation of the forest structure. In contrast, pollen-based estimates provide a more comprehensive, stratified view of the vegetation, as they incorporate all contributing taxa, not just the tree canopy. Despite this broader scope, pollen data and REVEALS estimates tend to emphasize trees more than other vegetation types consistently as is evident in the validations. Furthermore, pollen-based estimates are derived from records that span a much longer timescale than the modern forest cover data available, even though modern timeslices are used for validation. Increased anthropogenic impact could exacerbate discrepancies between pollen-based and ~~other plant inventories or compositional remote sensing products could offer better validation.~~ remote-sensing estimates. This could contribute to the overestimation of forest cover, which persists in all continents. Additionally, these modern and arguably unnatural vegetation conditions may not correspond to past vegetation and may therefore have reduced significance for the reconstruction of past, natural landscapes.

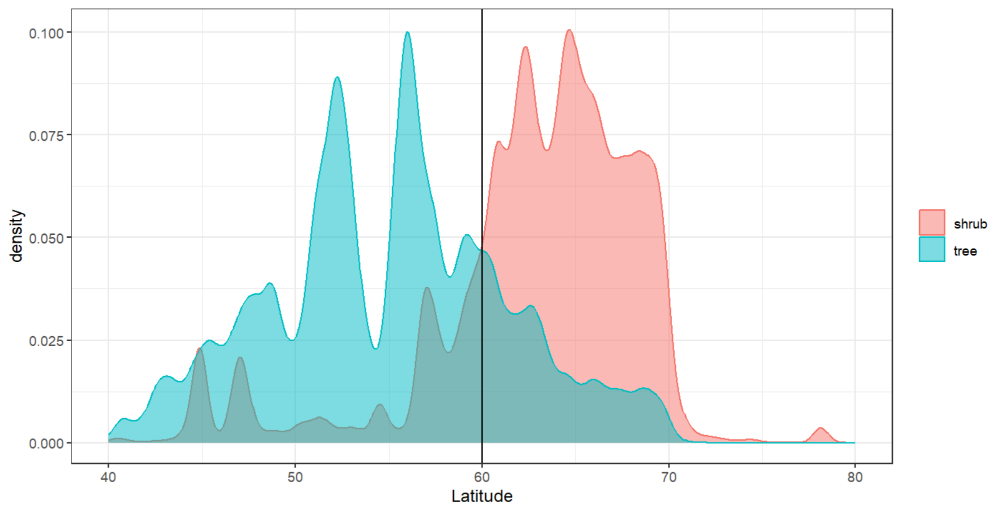
153: This cut off remains a weak point of the analysis, should be motivated and highlighted in what the data can be used for.

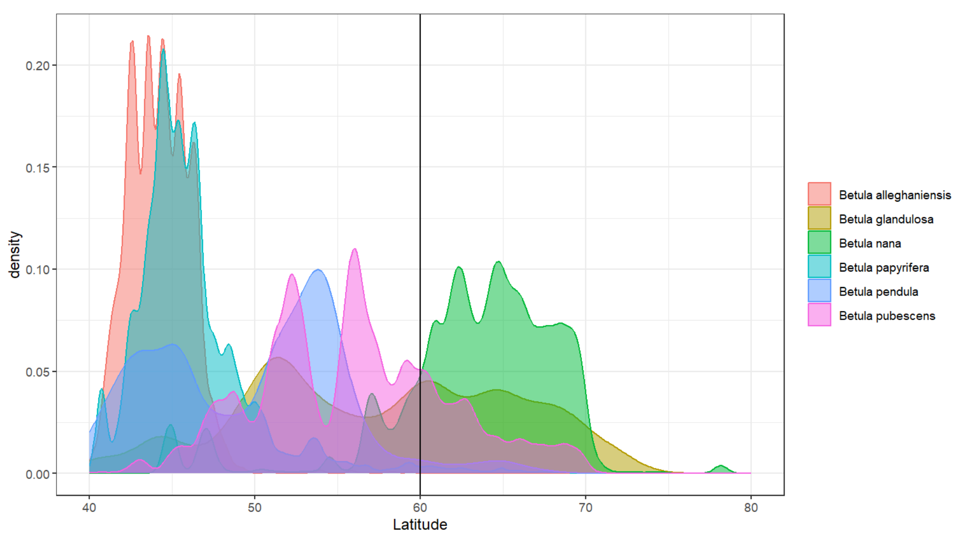
- We implemented this threshold (of 60°N, we apologize the typo in the manuscript) to reflect the present distribution of *Betula* plant functional types, as illustrated by GBIF occurrences in the two plots below. For this example we define both *Betula nana* and *Betula glandulosa* as shrub PFTs. We added an emphasis on the limitations of this static threshold in the manuscript's discussion (see tracked changes below). We also added the present Betula PFT distribution as illustrated below to the supplementary materials (S2).

tend to have higher certainty compared to taxon-specific reconstructions, as they are based on aggregation across taxa. However, the static latitudinal arboreal threshold for Betulaceae, *Betula*, and *Alnus* poses a limitation in our reconstruction. This could be improved by incorporating a dynamic, climate-dependent threshold in future work.

- GBIF references:
  - GBIF.org (01 October 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.2pw3qw
  - GBIF.org (01 October 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.vgchrb
  - GBIF.org (7 October 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.7fdwhs

- GBIF.org (7 October 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.achmvv
- GBIF.org (1 October 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.xyv4ge





161: How was the age of the samples determined?

- The ages of the samples were determined using standardized age modeling. The exact source of the age model is documented for each record in the LegacyPollen 2.0 dataset (please see Section 2.1, https://essd.copernicus.org/articles/14/3213/2022/essd-14-3213-2022-discussion.html, and https://doi.pangaea.de/10.1594/PANGAEA.965907)

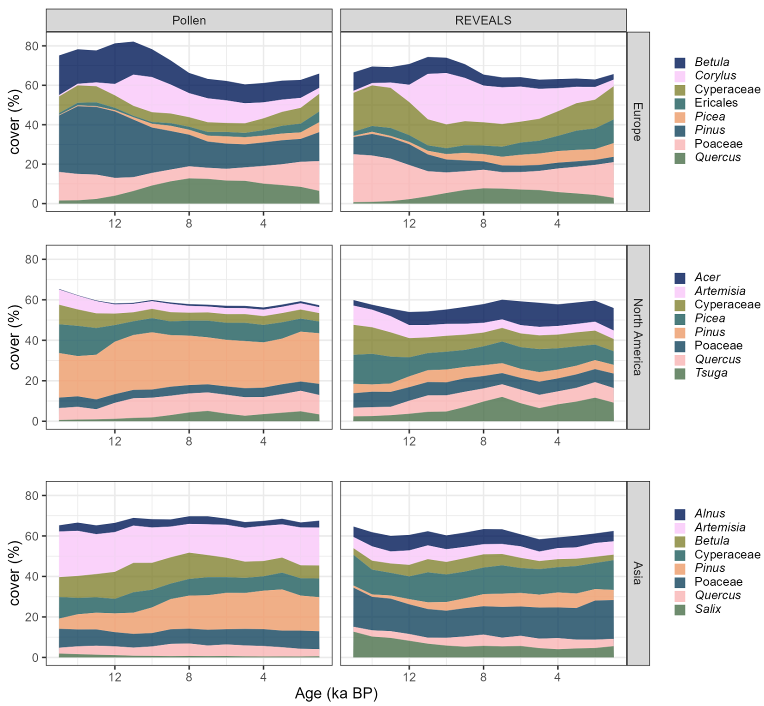
reconstruction only lake and peat records in the Northern Hemisphere were used (n = ~~2732~~2752) Analogous to the preceding LegacyPollen 1.0 dataset (Herzschuh et al., 2022), the data synthesis involved revising and standardizing age modeling and taxonomic harmonization for consistency of records. Reconstruction chronologies may, therefore, differ slightly from previous reconstructions due to this revised age modeling. Spatial data coverage of records in the reconstruction is dense in Europe

Figure caption to Fig 4: Smaller than what? Also the results of this analysis should be discussed more using the original paper (Sugita 1993). Are the results making sense for the reconstruction of tree cover. Would it not be more useful to look at the distance that 80% of the herbs come from? If the results are making sense, what are the consequences? Can result represent several grid cells? How do the result compare in cases were the area for one site is a subset of the area for another?

- Following another reviewer’s suggestion we decided to move the calculation and description of the 80% pollen source are to the supplementary materials. The results depend both on the choice of the dispersal model and the maximum vegetation extent (Zmax). Both basin size and its composition will impact the calculated 80% pollen source radius. And have to be assumed. The results have, therefore, an illustrative character, aiming to highlight the regional nature of pollen-based vegetation data to potential data users. For rasterization we follow previous reconstructions in that we average compositions from records located within a grid cell not source areas located within a grid cell. Norms on Zmax, dispersal models and percentage of source area considered would need to be established first, before using pollen source areas of any kind for spatial averaging rather than record location.

Fig 6: The new title of the manuscript indicates reconstructions for the last 14 ka but here 20 ka are presented. X-axis label is missing.

- See corrected figure below.



-

260: “lower” than what?

- Than tree pollen percentages.

coast and in the boreal biome. Rather open areas exist at the Tibetan Plateau and at very high latitudes. The ~~forest~~ tree cover derived from the REVEALS reconstruction is generally lower than tree pollen percentages. However, the difference between

265: Here you use “relevant”. Is that the 80% absolute? If yes, I think this may be too large for some lakes. Or are you using the grid squares as suggested in the rest of the sentence. Please explain.

- Grid cell remote sensing forest cover values were used. Please see corrected sentence below.

Remote sensing forest cover within ~~relevant pollen source areas~~ grid cells was used to validate the modern, reconstructed ~~forest~~ tree cover from the original pollen data and the REVEALS ~~run~~ estimates for each grid cell. Here we present validation of

307: This is a well-documented fact for areas with a dominance of wind pollinated trees - note not true in the subtropics and tropics.

- We remove this paragraph in favor of a more detailed discussion of remote sensing and pollen data differences in the discussion.

355: Global?

- We corrected this.

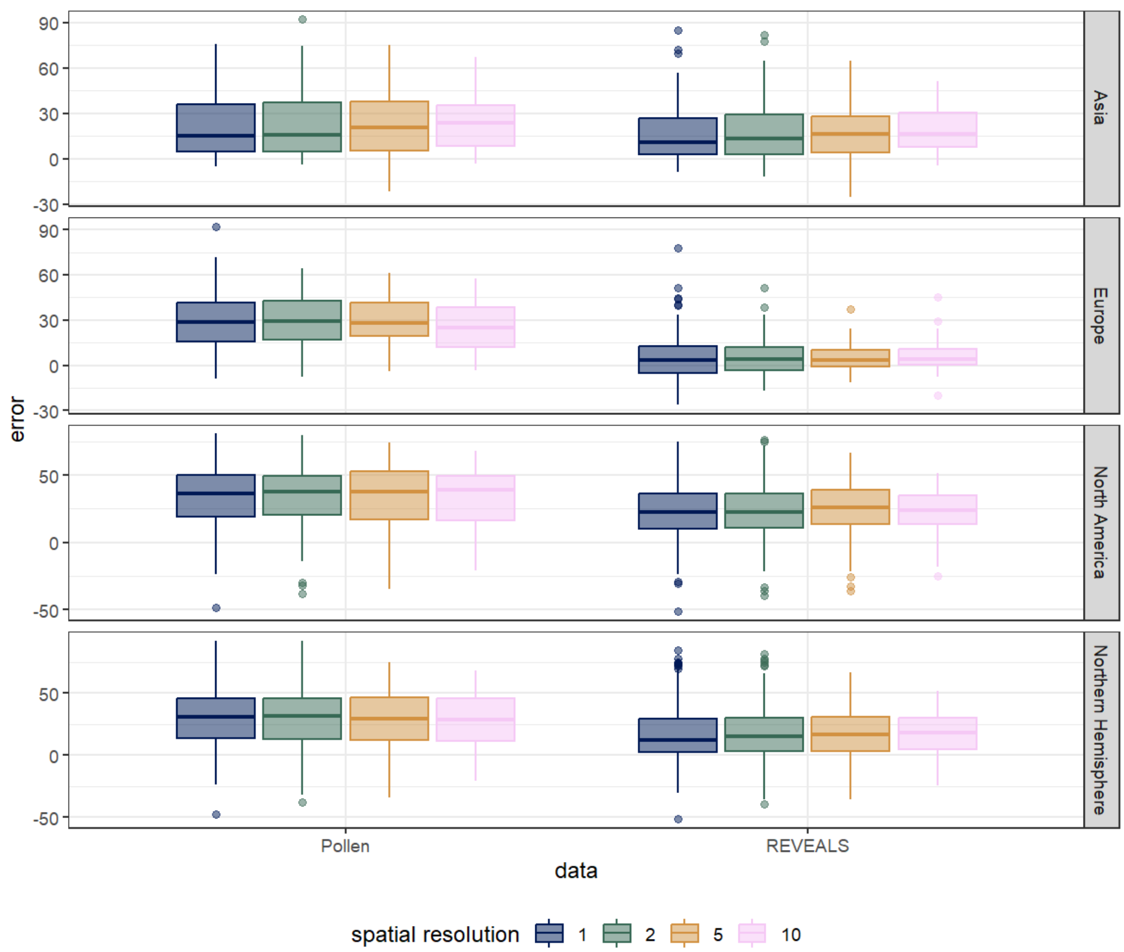
433: Did you use data from all these constituent databases?

- The LegacyPollen dataset uses data from all these constituent databases. We do not. We have removed the one we do not use. (screenshot)

Acknowledgements. We thank Thomas Böhmer for support with dataset curation and harmonization. The project was supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie through the German Climate Modeling Initiative PALMOD (grant no. 01LP1510C to UH), the European Union (ERC, GlacialLegacy grant no. 772852 to UH), and the China Scholarship Council (grant no. 201908130165 to CL). Data were partly obtained from the Neotoma Paleoecology Database (http://www.neotomadb.org) and its constituent databases (European Pollen Database, ~~QUAVIDA data base for Australasia, Latin American Pollen Database, African Pollen Database~~ and the North American Pollen database). The work of data contributors, data stewards, and the Neotoma community is gratefully acknowledged.

Supplement Fig 4: I see large differences between the different grid resolutions. Particularly interesting are the overestimations in the finest grid in Central Europe.

- We would like to disagree concerning large differences between the different spatial resolutions. While we do see a bit of variability concerning the median and mean error values (see Figure below), we cannot discern a clear trend towards better or worse reconstructions with a lower or higher spatial resolution or any large differences. A possible trend could be an increase in variability in REVEALS errors in Europe with higher spatial resolution, but this does not seem to be the case in the other continents, nor the Northern Hemisphere as a whole. We do however add a sentence in our discussion, highlighting this small variability between different spatial resolutions (see tracked changes below).



Data

Reading the manuscript I got the impression that the data is presented in a gridded format with binned time steps, while the data on PANGAEA is the sample based REVEALS estimate.
At second reading I discovered that a script for rasterization and binning is available from Zenodo. The script works well but it is using the data from Zenodo where duplicate files are available. As criticized before the data still contains marine cores e.g. MD84-629 which is wrongly labeled as peatland. REVEALS estimates are still provided for sites beyond 14,000 years ago as now indicated by the title in response to earlier criticism.

- It seems you had a look at the deprecated and not the reviewed dataset. Due to the length of data editor processing time, the data on PANGAEA has not been updated yet. Instead, we uploaded the revised dataset to Zenodo to make it available to reviewers before being finalized on PANGAEA. We outlined this interim solution in our reply to your last review and updated the dataset links in the manuscript. The dataset deposited on Zenodo does not include any records other than lakes and peat and does not extend to samples older than 14 ka BP.

Looking at the data for Europe I see supposed links to the data in the EPD/Neotoma: e.g. Event = Handle (in Neotoma), Site_ID and Dataset_ID identical to Neotoma. It is nice to see them but I did not see any documentation, stating that indeed these columns are links to the data in Neotoma. Moreover, the Event seems inconsistent e.g. “AGE_neotoma” versus AGE in Neotoma, while other EPD/Neotoma datasets don’t have the suffix.

- The input dataset (LegacyPollen2.0) used for this REVEALS reconstruction includes additional columns indicating Neotoma_DOIs for records that were originally obtained from Neotoma. A list of Neotoma records used in the reconstruction is additionally given in S1.
  The event inconsistencies are due to PANGAEA restriction in event uniqueness. Each event name can only be awarded once. The Event “AGE” already existed on PANGAEA, but evidently for a different location. This is why PANGAEA requires suffixes to be added to unique Event-Names. Again the original Neotoma event names can be acquired from the input dataset (LegacyPollen2.0, https://doi.pangaea.de/10.1594/PANGAEA.965907).

The basin diameter for peat bogs seems to have been set to 100 (Table 2) without further explanation. However the 100 m are also given as basin size in the site metadata for all peatlands, which is not correct! Estimated lake diameters and derived data are given with 12 digits after the comma suggesting a precision that is not there.

- The peatland basin size was only set to 100 m for peatlands where any other basin size is missing (n = 488 for all continents). We correct this inaccuracy in table 2. In the course of the revisions we had added basin areas from previous reconstructions where possible. We have now also added this updated metadata to the zenodo upload. We agree that the amount decimals given is unsuitable and have now updated to two decimals in the Zenodo upload.

Table 2. Static model parameters and model settings for REVEALS runs using REVEALSinR (Theuerkauf et al., 2016).

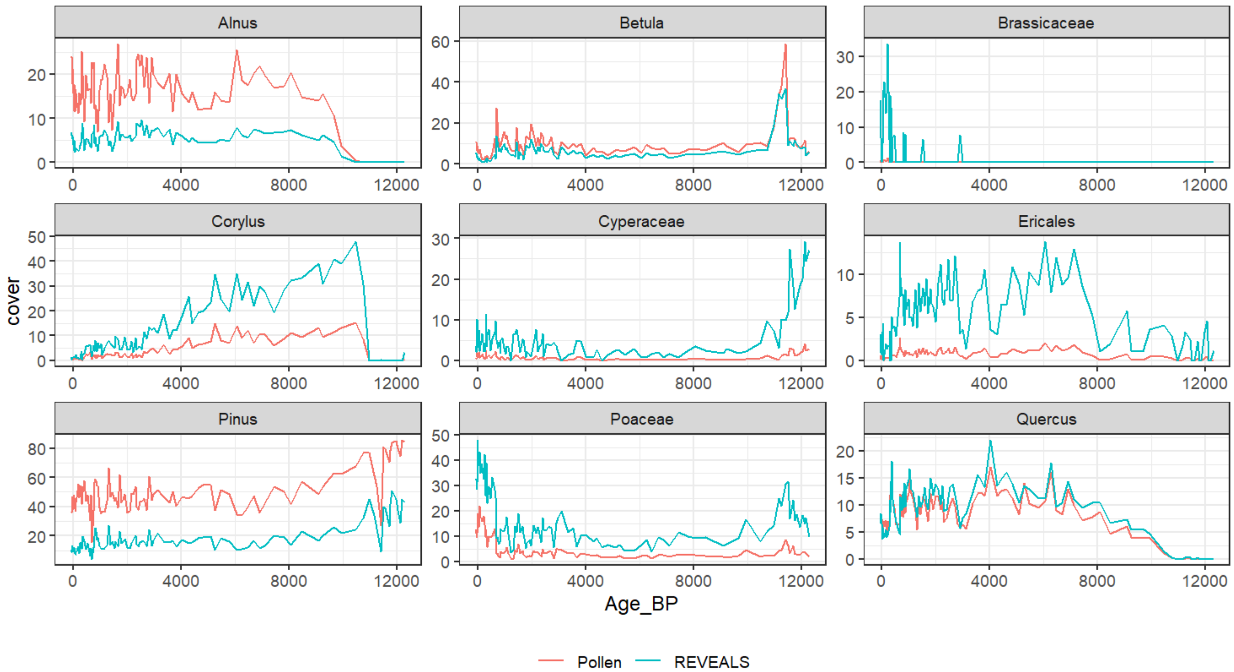
Parameter	Values and settings used in model run
atmospheric model	unstable atmosphere
dispersal model	gaussian plume
wind speed	$3m \times s^{-1}$
maximum extent of regional vegetation (~~region cutoff~~Z_{max})	1000 km
number of RPP ~~variations~~ and pollen count variations (n)	2000
peatland basin ~~radius~~ area (for missing sizes)	~~100 m~~31.41 ha
lake basin area (for missing sizes)	49 ha
function to randomize pollen counts	rmultinom_reveals

The data include for each taxon the 10_percentile, 90_percentile, mean, and median, as well as the sd of cover in %, while the manuscript does not mention these calculations and how they may be used. It needs to be better communicated what kind of data is presented and how the authors anticipate it to be used.

- We added a paragraph in the manuscript to describe the data structure and when and how it should be rasterized. Please see tracked changes below.

The published dataset includes vegetation reconstructions for individual records in Asia, Europe, and North America up until 14 ka BP. The reconstructed coverage values include mean, median, standard deviation, and 10% and 90% quantile values for each taxon. Mean values and standard deviations are given for tree cover. For each sample its validity as a site is given. Only reconstructions from large lakes are valid independently. To include all other records a spatial and temporal average is necessary (rasterization, https://doi.org/10.5281/zenodo.12800291).

I went on checking the REVEALS estimates for Grosser Treppelsee in Brandenburg, for which I counted the pollen. While the overall forest cover reconstruction through time looks reasonable the taxon specific estimates are unlikely for several taxa. The region is dominated by Pinus forest while the estimate of Pinus cover in the surface sample is 8.5% mean cover, which is much too low. REVEALS estimates for Brassicaceae cover 100 to 200 years ago is with around 20% way too high. Thus the data may be useful for continental scale forest cover reconstructions while regional studies would benefit from regionally estimated PPEs. The caveats of using continental scale RPPEs and particularly of setting RPPEs to 1 for some taxa need to be discussed in the publication.



- We show the pollen percentages and the REVEALS reconstruction for 10 common taxa at Großer Treppelsee above. The cover of *Pinus* is especially corrected to be lower due to its high pollen productivity. It should also be remembered that the pollen record from a lake as large as Großer Treppelsee (~59 ha) will have a regional signal rather than a local one, including the mosaic of open and closed vegetation in Brandenburg and likely even Poland and not just the rather closed forest surrounding the actual lake. The modern forest cover (landsat) in a circle with a 100 km radius surrounding Großer Treppelsee is 28.194% and the reconstructed modern value (with the openness correction accounting for urban areas) is ~ 29%. We believe that this shows that tree cover is reconstructed well in this area. Nevertheless, the accuracy of the REVEALS reconstruction depends on RPP values, the availability and quality of which varies for each taxon. We already show this in Figure 3 and our discussion. We did however expand on the variation between RPP value sources and the validity of reconstructions on a larger spatial scale and the aggregated tree cover scale as opposed the compostion in the discussion. (see tracked changes)

The ~~REVEALS forest cover reconstructions presented here offer valuable insight into past vegetation changes. The global dataset provides an opportunity to explore past vegetation dynamics, gaining a deeper understanding of responses, trajectories, and potential feedback mechanisms. Given the increasing discussions surrounding the possibility of tipping events in vegetation cover (Armstrong McKay et al., 2022; Lenton and Williams, 2013), this could be of considerable use. While a reconstruction of exact tree lines is not trivial with pollen data,~~ reliability of reconstructions also varies among different taxa due to the quality of RPP values, which is documented in detail in a supplementary file outlining the sources of RPP values (see Section "Code and Data Availability"). Reconstructions based on taxa with continental RPP values are the most reliable, followed by those based on hemispheric data, with standardized RPP values being the least reliable. This hierarchy should be considered when interpreting results. The use of continental RPP values could also make our reconstruction more reliable at larger spatial scales as opposed to local reconstructions. Additionally, uncertainties in RPP values themselves can affect reconstruction success and could be leading to the persistent overrepresentation of tree taxa despite the application of REVEALS~~and subsequent biomization improve treeline reconstructions as shown by Binney et al. (2011). Additionally, this dataset can address unanswered questions about Holocene vegetation dynamics, including the deglacial forest conundrum (Dallmeyer et al., 2022; Strandberg et al., 2022). It also serves as a valuable tool for validating models with coupled climate and vegetation, which rely on extensive time series and vegetation data for accurate predictions (Dallmeyer et al., 2023; Dawson et al., 2024a). Comparing modeled vegetation to reconstructed vegetation could help uncover missing dynamics in coupled climate-vegetation models. New insights gained from these applications could enhance our ability to predict future changes.~~ Tree cover reconstructions tend to have higher certainty compared to taxon-specific reconstructions, as they are based on aggregation across taxa. However, the static latitudinal arboreal threshold for Betulaceae, *Betula*, and *Alnus* poses a limitation in our reconstruction. This could be improved by incorporating a dynamic, climate-dependent threshold in future work.

Reply to Gaillard

General reply

Dear Marie-Jose Gaillard,

Thank you very much for your thorough review of our manuscript. We are pleased to inform you that we were able to implement most of your suggestions and have addressed your concerns accordingly.

Regarding your larger comments:

- **Aggregation differences from previous continental-scale reconstructions:** We now describe these differences more clearly in the text and provide supplementary materials demonstrating that the method of aggregation does not significantly impact the absolute reconstructed tree cover values.
- **REVEALSinR error:** We have clarified the error calculation in REVEALSinR and now explain in more detail how coverage distributions are calculated across repeated model runs while accounting for total pollen counts and RPP variability.
- **Move 80% PSA to supplementary materials:** This suggestion has been implemented, and the relevant section has been moved to the supplementary materials.

In addition, we have addressed several smaller phrasing issues throughout the manuscript. We also corrected an error in the validation and reconstruction figures, where a 2x4° rasterization was mistakenly used instead of the intended 2x2° format. Importantly, this adjustment does not impact dataset validity.

Please see our detailed responses to each of your comments below. We also replied to your in-line comments in the previous manuscript version and append this document as well. The current dataset version (https://doi.org/10.5281/zenodo.13902921), code (https://doi.org/10.5281/zenodo.10191859), and rasterization script (https://doi.org/10.5281/zenodo.13902976) can be found on Zenodo.

Once again, we would like to express our thanks for your input, which we believe has improved the clarity of the manuscript and emphasized its usability compared to other datasets. We are confident that we have addressed all of your suggestions.

Best regards,
Laura Schild and Ulrike Herzschuh

Specific replies

General comments

The authors have made substantial revisions that were necessary such as deleting the southern hemisphere from the reconstruction and producing REVEALS estimates based on pollen records from several sites within areas (grid cells) of various sizes and for time windows of various lengths. This leads to more acceptable results. I appreciate the hard work made to finalize this revision, but there are still misunderstandings that needs to be clarified in the paper.

1. One of my major concerns is the calculation of REVEALS mean estimates based on the REVEALS reconstructions for several sites within grid cells and several pollen counts within time windows, i.e. the step that the authors call “aggregation” in space and time. For the “aggregation” in space the authors calculate the mean of the individual site REVEALS estimates without any weighting by the K coefficient that is dependent of basin size (the larger the basin, the heavier the weighting should be for each taxon, and vice versa). Such a weighting is implemented in Sugita’s REVEALS computer program but not in REVEALSinR. In Sugita’s method, the REVEALS estimates from individual sites within a grid cell are weighted with the taxon-specific “pollen dispersal-deposition coefficient K” of all pollen taxa involved, se e.g. Li et al. (2017). This should be clarified under METHODS.
   For the “aggregation” in time the authors similarly calculate the mean of the individual counted level REVEALS estimates. The reliability of REVEALS estimates depends, among other things, on the size of the pollen count. In this context, the usual size of pollen counts (often around 1000, seldom more, quite often around 500 and sometimes less) is a low pollen count. This implies that all REVEALS estimates in the Schield et al. REVEALS dataset are of relatively low reliability and calculating the mean of these REVEALS estimates does not make them more reliable. All earlier continental Holocene REVEALS reconstructions have worked with time windows of such a length that it would maximize the size of the counts without using too long time windows (generally maximum 500 years). The compromise to make depends on the aim of the study. One has then to sum pollen counts within each time window and use this new pollen count for the REVEALS application to obtain the REVEALS estimates for the time window (see e.g. Githumbi et al., 2022). This procedure is very different from calculating mean REVEALS estimates and is statistically the correct way to do. I understand that it would be a huge work to redo the work in this way for this manuscript. But this should be listed as one of the many differences between this REVEALS dataset and earlier ones. I do not know whether the error on REVEALS estimates as calculated by REVEALSinR (see my point below) is sensitive to the size of pollen counts. I guess not, but I can’t find anything about this issue in the REVEALSinR original paper or elsewhere. In that case, this is also an aspect that makes REVEALS applications using REVEALSinR weaker if the size of pollen counts is not considered in the error estimate on REVEALS results.

Thank you for highlighting these differences. We added an explanation of these differences to previous reconstructions in our manuscript (see tracked changes below). We highlight the trade-off of reconstruction robustness with application flexibility with the changed temporal binning. While our spatial aggregation differs from previous European reconstructions, it is actually the same arithmetic mean used by Dawson et al. (2024, see lines 306-309,398 in https://github.com/andydawson/reveals-na/blob/master/r/reveals.r).

and Ord (1994), using the same equation as Wieczorek and Herzschuh (2020). We provide a script for rasterization with adjustable temporal and spatial resolution for users of the dataset on Zenodo (https://zenodo.org/doi/10.5281/zenodo.12800290). This method of temporal and spatial averaging differs from several previous REVEALS applications. Pollen counts are often summed in temporal bins prior to running REVEALS to increase pollen counts and reduce uncertainty (Trondman et al., 2015; Githumbi et . However, temporally averaging after the REVEALs application, as implemented by us, increases the flexibility of the dataset with the trade-off of potentially increased uncertainty. Rasterization has previously been perfomred by using a weighted average taking into account the basin size of the original record (Trondman et al., 2015; Githumbi et al., 2022; Serge et al., 2023). However, the most recent REVEALS-based North American vegetation reconstruction uses the same arithmetic mean as described above (Dawson et al., 2024b). When comparing our method of temporal and spatial aggregation to that used by previous European reconstructions (e.g. Serge et al., 2023), we also found no significant differences in the validation of reconstructed tree cover (see S6).

For validation, the reconstructed ~~forest~~ tree cover of the past 100 years was rasterized and compared to modern remote

We did implement a temporal binning of pollen counts prior to REVEALS and did a weighted mean (based on basin size) to compare the tree cover reconstruction of the modern time slice between our method of rasterization and previous European reconstructions' method (titled "Gaillard suggestion" in the Fig. R1 below). We find that absolute values differ minimally and that the reconstruction error is virtually the same. We, therefore, feel confident in stating that our method of rasterization does not impact reconstruction success significantly negatively.
We explain the error calculation in REVEALSinR in detail in our reply to the next general comment.

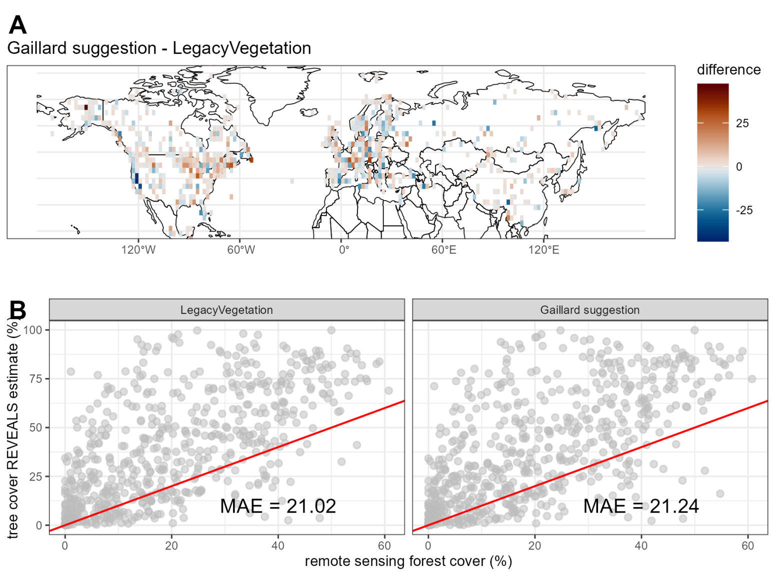


Fig. R1: Overview of differences between modern tree cover timeslices of LegacyVegetation rasterization and methods used in previous European REVEALS reconstructions ("Gaillard suggestion"). (A) Gridcell (2x2°) differences in tree cover between the two versions. (B) Validations with landsat forest cover for both versions.

2. Another major difference between implementation of the REVEALS model with the computer programs of Sugita and REVEALSinR of Theuerkauf et al. (2016) is the calculation of the uncertainties (errors) on the REVEALS estimates. The REVEALS standard error accounts for the

standard errors (or deviations) of the relative pollen productivities for the individual pollen taxa and on the number of pollen counted; i.e. the size of the pollen count matters. The error calculated in REVEALSinR does not consider the RPP errors. I do not mean that the errors from the REVEALSinR program are wrong, but it is a pity not to use the errors on RPPs as this parameter is very influential on the final REVEALS estimate of plant cover. This difference between the two applications should at least been mentioned.

REVEALSinR runs REVEALS on each sample multiple times (n = 2000 in our study) while altering pollen counts as well as RPP each time. Errors are added (or subtracted) on the pollen counts. The size of the error depends on the total pollen counts with smaller total counts resulting in larger errors being added (hence the larger standard deviations of the randomized pollen counts with total pollen counts are low). This is visualized in Fig. R2 below. Not summing pollen counts in time slices before running REVEALS does therefore lead to higher standard deviations, but does not significantly effect absolute values as highlighted in our reply to your general comment 1.
During repeated REVEALS runs in REVEALSinR, RPP are also generated randomly from a normal distribution (μ = RPP value, σ = RPP SD). The 2000 REVEALS results are then used to calculate statistics such as the mean and median REVEALS estimate, as well as quantiles and standard deviations. REVEALSinR therefore accounts both for total pollen count and RPP SDs in the calculation of REVEALS estimates. We include these statistics as REVEALS outputs in our description of the dataset in the manuscript (see tracked changes below).
Description of differences between Sugita's program and REVEALSinR:

We use REVEALSinR from the DISQOVER package in R to implement REVEALS (Theuerkauf et al., 2016, Version 0.9.13, https://github . It mainly differs from the original program by Sugita (2007) in the process of error calculation. REVEALSinR includes repeated model runs with random error added to RPP values and pollen counts (see Table 2 for the number of variations). The resulting distribution of REVEALS results allows for an estimation of the standard deviation of vegetation cover per taxon. The program by Sugita (2007), however, derives error estimates with a hybrid method from a variance-covariance matrix of PPE and Monte Carlo simulations. For further details on the REVEALS model see the original publication Sugita (2007) ~~or Githumbi et al. (2022).~~ and for previous REVEALS applications on continental scales see e.g Li et al. (2017), Githumbi et al. (2022) , Serge et al. (2023), and Dawson et al. (2024a).

Description of values included in dataset:

3.1 Dataset description

The published dataset includes vegetation reconstructions for individual records in Asia, Europe, and North America up until 14 ka BP. The reconstructed coverage values include mean, median, standard deviation, and 10% and 90% quantile values for each taxon. Mean values and standard deviations are given for tree cover. For each sample its validity as a site is given. Only reconstructions from large lakes are valid independently. To include all other records a spatial and temporal average is necessary (rasterization, https://doi.org/10.5281/zenodo.12800291).

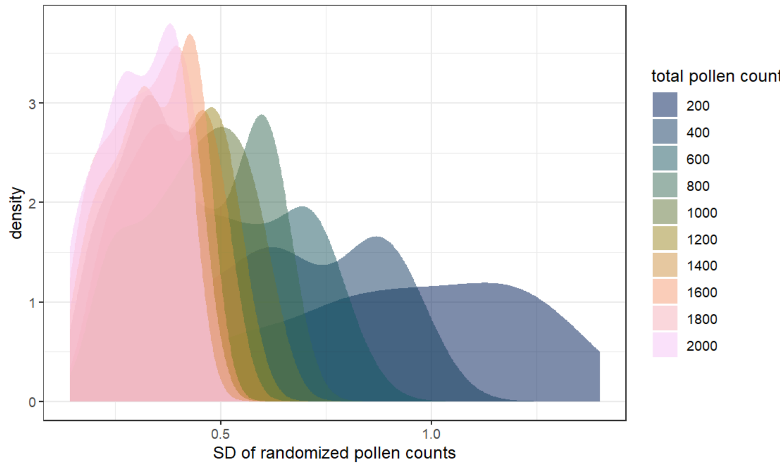


Fig. R2: Distribution of random pollen count standard deviations with changing total pollen counts in REVEALSinR.

3. 80% pollen source area: this information should be presented as an alternative to estimate the size of the region that is represented by REVEALS estimates of plant cover. Sugita (2007a) who developed the REVEALS model assumes that Zmax is the size of the region represented by REVEALS estimates (see also Li et al., 2017). Zmax can only be assumed (you assumed it to be 1000 km over the entire study region) and the region from which most of the pollen are coming (in your case 80%) can be estimated. See also Hellman et al., 2008b (in VHA) who assumed Zmax to be 400 km (distance from the pollen site) in S Sweden and the 90% source area (200 km) was considered to be the area from which most of the pollen came. One should therefore state that the assumed value for Zmax influences the estimate of x% pollen source area. Please, also specify what dispersal model you use, the Gaussian Plume Model or the Lagrangian Stochastic Model, for estimating your 80% pollen source area, which makes also a difference (see Theuerkauf et al., 2016).

Following the suggestion by two other reviewers, we decided to move the calculation and results of the 80% pollen source area into the supplementary materials (see S5). We add two additional sentences in the manuscript highlighting the assumption of a maximum spatial extent of regional vegetation (Zmax) and include a reference to Hellmann et al. 2008 (see tracked changes below). All parameters used in REVEALSinR are listed in Table 2, which is also referred to in the text (see below as well).

Zmax highlight:

The REVEALS model follows a set of assumptions. Firstly, neither directionality nor pollen transport through agents other than wind are considered in the model. The maximum spatial extent for this pollen transport (Z_{max}, see Table 2) has to be set to define the region in which most of the pollen originates. This value will always be an assumption and has only been tested empirically by Hellman et al. (2008b). Additionally, it is assumed that the basin is circular with no source of pollen within the basin radius. The peatland and bog sites used in our reconstructions inherently violate this assumption. Nevertheless, the

Parameter table:

Table 2. Static model parameters and model settings for REVEALS runs using REVEALSinR (Theuerkauf et al., 2016).

Parameter	Values and settings used in model run
atmospheric model	unstable atmosphere
dispersal model	gaussian plume
wind speed	$3m \times s^{-1}$
maximum extent of regional vegetation (~~region cutoff~~ Z_{max})	1000 km
number of RPP ~~variations~~ and pollen count variations (n)	2000
peatland basin ~~radius~~ area (for missing sizes)	~~100 m~~ 31.41 ha
lake basin area (for missing sizes)	49 ha
function to randomize pollen counts	rmultinom_reveals

Two additional comments, minor but still important:

4. Avoid the term reconstruction for pollen percentages or raw pollen data. These are simply data, pollen% are not a reconstruction of vegetation, they are proxy data of vegetation, while a traditional narrative interpreting the pollen percentages using various kind of information is a reconstruction, as REVEALS-based estimates of plant cover is a reconstruction of past plant cover. I advise you to revise this throughout the manuscript, text and Figures. I made comments in the manuscript about that, but not everywhere. Using “reconstruction” for pollen data is misleading, and makes the text difficult to understand in some places.

We removed “reconstructions” in the context of pollen and changed it to your suggested wording of tree pollen percentages. Please see an example text passage below.

reconstructions compared to the original pollen data. Both ~~reconstructions~~ pollen percentages and REVEALS estimates still overestimate ~~forest~~ tree cover.

5. I would use the terms “(total) tree pollen” and “(total) tree cover” instead of “forest cover” when it refers to pollen % and REVEALS-based estimates of tree cover. It is important to be clear in terms of what you are comparing the satellite vegetation (forest cover) with. If you choose to follow my advice, revise the manuscript consequently. I made comments in the manuscript about that, but not everywhere.

We changed it to “tree pollen” and “tree cover” at the applicable locations. Please see an example text passage below.

LegacyVegetation~~1.0~~: Northern Hemisphere reconstruction of ~~vegetation composition~~ past plant cover and ~~forest~~ total tree cover from pollen archives of the last 14 ka

In conclusion

I miss a description of your new REVEALS dataset for the N Hemisphere in comparison to the earlier continental REVEALS dataset for Europe, China and N America. What is different and what are the improvements.

1. In terms of what is different in the methodology, please see my major comments above, and specific comments in the revised manuscript. Do not forget that you use different chronologies than those used in earlier reconstructions. They might not be so different, but we do not know. The best solution is to describe all the differences in methodology already in the METHODS section, in the part describing REVEALSinR and in the part describing how you “aggregate” site-specific and level (time)-specific REVEALS estimates to mean REVEALS estimates (level-specific meaning using single analysed levels/samples to run REVEALS.

We now note in the manuscript that chronoligies may differ to previous reconstructions (see tracked changes below).

reconstruction only lake and peat records in the Northern Hemisphere were used (n = ~~2732~~2752) Analogous to the preceding LegacyPollen 1.0 dataset (Herzschuh et al., 2022), the data synthesis involved revising and standardizing age modeling and taxonomic harmonization for consistency of records. Reconstruction chronologies may, therefore, differ slightly from previous reconstructions due to this revised age modeling. Spatial data coverage of records in the reconstruction is dense in Europe

As described above we added a description of how temporal and spatial averaging/aggregation differ between our and previous reconstructions. Please see the tracked changes again below.

and Ord (1994), using the same equation as Wieczorek and Herzschuh (2020). We provide a script for rasterization with adjustable temporal and spatial resolution for users of the dataset on Zenodo (https://zenodo.org/doi/10.5281/zenodo.12800290). This method of temporal and spatial averaging differs from several previous REVEALS applications. Pollen counts are often summed in temporal bins prior to running REVEALS to increase pollen counts and reduce uncertainty (Trondman et al., 2015; Githumbi et . However, temporally averaging after the REVEALs application, as implemented by us, increases the flexibility of the dataset with the trade-off of potentially increased uncertainty. Rasterization has previously been perfomred by using a weighted average taking into account the basin size of the original record (Trondman et al., 2015; Githumbi et al., 2022; Serge et al., 2023). However, the most recent REVEALS-based North American vegetation reconstruction uses the same arithmetic mean as described above (Dawson et al., 2024b). When comparing our method of temporal and spatial aggregation to that used by previous European reconstructions (e.g. Serge et al., 2023), we also found no significant differences in the validation of reconstructed tree cover (see S6).

For validation, the reconstructed ~~forest~~ tree cover of the past 100 years was rasterized and compared to modern remote

2. In my view, the improvements in your REVEALS dataset are:
   - You have included in your synthesis the pollen records from the northern hemisphere between Europe and China, those sites that were included in Cao et al (2019) REVEALS reconstruction, and applied REVEALS on them in accordance with the methodology you use for the rest of the Northern Hemisphere.
   - Further, it would be informative to know how many pollen records you use overall and in specific continents (Europe, China, N America) for which earlier REVEALS reconstructions exist.

The number of records used for each continent were already present in Section 2.1. We correct small errors here and added the amount of records used in Serge et al. in Section 2.3 (see tracked changes below).

reconstructions due to this revised age modeling. Spatial data coverage of records in the reconstruction is dense in Europe (~~1275~~ 1287 records) and North America (~~1016 records~~1040) and sparsest in Asia (~~441~~446) (see Fig. 1). The records’ sample density decreases with age (see Fig. 2). Only samples dated to 14 ka BP or younger were used to ensure that the climatic conditions of recorded vegetation were similar to the modern climate.

For Europe, compare with Serge et al. (2023).

We added a comparison of modern forest cover between our and Serge et al.'s reconstruction. Please see the methods and the results of this comparison in the manuscript's tracked changes below. The tree cover in Serge et al. tends to be higher than our reconstructed tree cover, leading to higher reconstruction errors.

Serge comparison (method):

~~Using REVEALS , radii of~~ Additionally, we compare our REVEALS reconstruction to the most recently published REVEALS reconstruction in Europe by Serge et al. (2023, version: RPPs.st1). We average our reconstruction in the same grid and temporal bins as used by Serge et al. to compare the reconstructed tree cover between both reconstructions. To get the total tree cover, we sum evergreen and summergreen tree cover values in Serge et al.'s dataset, while excluding broadleaved summergreen temperate warm shrubs (BSTWS) and broadleaved evergreen xeric shrubs (BEXS). We validate the previous reconstruction and our reconstruction in the most recent time slice available in Serge et al.'s reconstruction (-65 to 100 BP, https://doi.org/10.48579/PRO/J5 with the remote sensing forest cover and compare validations. Unfortunately, direct validation could only be performed with the most recent time slice available online, rather than the historical time slice used in the validation by Serge et al., which limits the ability to reproduce their validation results exactly. We do not apply any openness correction here as we do not have comparable 80~~% pollen source areas were calculated for large lakes(see Fig. ??) . The radii indicate in which area 80%~~

~~of the deposited pollen originated from (see Section 2.2.2) and yield an understanding of which area the pollen record is representative of, which is especially useful when individual time series from large lakes are being used for analyses. The 80~~% pollen source areas ~~are roughly a function of basin size (see Fig.??)and range between 155 km and 762 km. The median 80% pollen source radius is 225 km including all large lakes~~ available for the records used in Serge et al. (2023). The reconstruction by Serge et al. differs in the temporal as well as spatial aggregation routine, as described above. Definition of arboreal taxa varies, a different RPP-value set was used, and the amount of total records included is higher than in our reconstruction (Serge et al.: 1607, LegacyVegetation: 1287).

Serge comparison (results):

A specific comparison with Serge et al. (2023) reveals that our reconstruction generally shows lower forest cover across Europe. Unfortunately, direct validation could only be performed with the most recent time slice available online, rather than the historical time slice used in the validation by Serge et al., which limits the ability to reproduce their validation results exactly. Despite this limitation, our reconstruction demonstrates a lower MAE, indicating improved accuracy. This is notable given that Serge et al. utilized a larger number of records in their study. One potential explanation for these differences could lie in the variations in RPP values and the selection of arboreal taxa used in the reconstruction, as we employ an arboreal tree treshold and include more taxa in our REVEALS reconstruction.

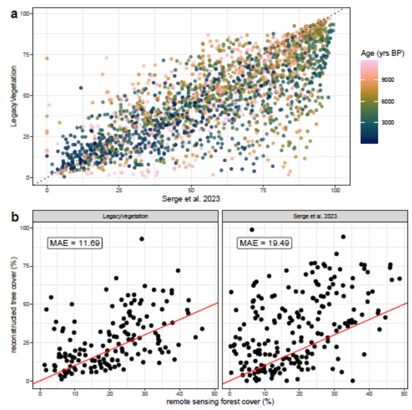


Figure 10. (a) Comparison between LegacyVegetation (this publication) and the tree cover from Serge et al. (2023) and (b) validations with modern, remote-sensing forest cover for both data sets.

In terms of RPP, you should also mention if you use more RPP values than in earlier studies and also clarify that your RPP synthesis is made in a different way (different rules) than those by Githumbi et al. (2022) for Europe and Li et al. (2018) for China. For China, the improvement is that you have added new recent RPP values from recent papers.

The same synthesis rules used by Githumbi et al. (2022) were used in our synthesis. We synthesized RPP values on different taxonomic levels to account for the harmonized pollen dataset used in this reconstruction. This is why more values are available. Please see the expanded explanations in the tracked changes below.

2021; Zhang et al., 2021a, b; Wan et al., 2020, 2023; Jiang et al., 2020). The methods ~~by Wieczorek and Herzschuh (2020) were followed fore~~ for study selection and calculation of synthesis values ~~.~~ follow Wieczorek and Herzschuh (2020) as well as Githumbi et al. (2022). We expanded the synthesis calculation of RPP to different taxonomic levels (genus, family, and order) to account for the taxonomic harmonization in the pollen dataset. An overview of original values and synthesized values can be found in Appendix A and B respectively. The amount of RPP values in Asia (59) and Europe (69) is higher than in previous RPP synthesis due to the inclusion of multiple taxonomic levels (Li et al., 2018; Githumbi et al., 2022).

Finally, your new REVEALS dataset should be presented as an alternative dataset that is more flexible that the earlier continental ones as it allows users to amalgamate the REVEALS estimates in space choosing various sizes of grid cells, and in time choosing various length of time windows. It should be stated, however, that mean REVEALS estimates over space do not weight the K coefficient according to lake/bog size, and that mean REVEALS estimates over time are not as reliable as REVEALS estimates based on the total pollen count in a time window (see my comment above). With flexibility you loose reliability. This should be clarified for the users.

In addition to the description of differences in the methods section (see our reply to your general comment 1), we point out the trade-off once more in the discussion. Please see the tracked changes below.

Although our reconstruction method is more flexible than previous efforts, the temporal and spatial aggregation used may reduce its reliability, due to smaller total pollen counts used in REVEALS runs and the use of an arithmetic as opposed to a weighted spatial mean.

LegacyVegetation 1.0: ~~Global~~ Northern Hemisphere reconstruction of vegetation composition and forest cover from pollen archives of the last ~~50~~ 14 ka

Laura Schild[1,2], Peter Ewald[1,2], Chenzhi Li[1,2], Raphaël Hébert[1], Thomas Laepple[1,3], and Ulrike Herzschuh[1,2,4]

[1]Helmholtz Centre for Polar and Marine Research, Research Unit Potsdam, Alfred Wegener Institute (AWI), Germany
[2]Institute of Environmental Sciences and Geography, University of Potsdam, Karl-Liebknecht-Straße 24-25, Potsdam, Germany
[3]MARUM-Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Germany
[4]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, Potsdam, Germany

Correspondence: Ulrike Herzschuh (ulrike.herzschuh@awi.de)

Abstract. With rapid anthropogenic climate change future vegetation trajectories are uncertain. Climate-vegetation models can be useful for predictions but need extensive data on past vegetation for validation and improving systemic understanding. Even though pollen data provide a great source of this information, the data is compositionally biased due to differences in taxon-specific relative pollen productivity (RPP) and dispersal.

Here we present a Northern Hemisphere reconstruction of quantitative regional vegetation cover from a ~~global~~ sedimentary pollen data set for the last ~~50~~ 14 ka using the REVEALS model to correct for taxon- and basin-specific biases. ~~In a first~~ For the reconstruction, we ~~used previously published ,~~ expanded on a previously published synthesis of continental RPP values. ~~For a second reconstruction, we statistically optimized RPP values for common taxa with the goal of improving the fit of reconstructed forest cover from modern pollen samples with remote sensing forest cover.~~

The data sets include taxonomic compositions as well as reconstructed forest cover for each original pollen sample. ~~Relative~~ 80% pollen sources areas were ~~also calculated~~ calculated for large lakes and are included in the data set~~of the original REVEALS run~~. Additional metadata includes modeled ages, age model sources, basin locations, types and sizes.

The improvements in forest cover reconstructions with the REVEALS reconstruction using ~~original/optimized parameters range from 1/0% (Australia and Oceania/Australia and Oceania) to 58/65~~continental RPP values range from 24% (North America) to 72% (Europe~~/North America~~) relative to the mean absolute error (MAE) ~~in~~ of the pollen-based reconstruction. ~~Optimizations were considerably more successful in reducing MAE when more records and RPP estimates were available. The optimizations were purely statistical and only partly ecologically informed and should, therefore, be used with caution depending on the study matter~~The dataset can be used as a grid with binned and aggregated samples (adjustable script provided on Zenodo; https://zenodo.org/doi/10.5281/zenodo.12800290) or as individual timeseries if the record's basin size exceeds 50 ha.

This improved quantitative reconstruction of vegetation cover is ~~invaluable~~ beneficial for the investigation of past vegetation dynamics and modern model validation. By collecting more RPP estimates ~~for taxa in the Southern Hemisphere~~ especially in

North America and adding more records to existing pollen data syntheses, reconstructions may be improved even further. ~~Both reconstructions are~~ The REVEALS reconstruction is freely available on PANGAEA (see Data availability section).

1 Introduction

Anthropogenic climate change is driving vegetation shifts that could lead to disruptions in ecosystem functions and services, and even trigger feedback effects with other earth system elements (IPCC, 2023; Armstrong McKay et al., 2022). Predicting these changes through modeling is challenging. A sufficient mechanistic understanding of vegetation dynamics and interactions with climate is needed, which requires validation and testing of model data with extensive vegetation data across climatic transitions ~~akin~~ comparable to those anticipated in the future (Dearing et al., 2012). Given the relatively brief duration of available instrumental climate and vegetation data, there is a clear need for long-term ~~environmental~~ vegetation records derived from paleoecological archives that cover broader climatic gradients than modern datasets (Dearing et al., 2010; Dallmeyer et al., 2023).

Pollen data as a direct proxy for paleo-vegetation is especially useful for comparisons with modeled data as it can be used to reconstruct land-use (Fyfe et al., 2015; Davis et al., 2015), biomes (Woodbridge et al., 2014; Prentice et al., 1996), and climate (Herzschuh et al., 2023a, b; Bartlein et al., 2011; Viau et al., 2012). The compilation of pollen data syntheses is essential to aid this purpose (Anderson et al., 2006; Gaillard et al., 2010; Strandberg et al., 2014). Several subcontinental and continental collections of pollen data already exist, spanning regions such as Europe, North America, Africa, Siberia, and China (Fyfe et al., 2009a; Whitmore et al., 2005; Vincens et al., 2007; Cao et al., 2014, 2020) and have been integrated into the global database Neotoma (Williams et al., 2018). To allow for a broader application of pollen data, LegacyPollen 2.0 (Li et al., 2024b) offers a global, harmonized pollen dataset that underwent taxonomic standardization, metadata verification and consistent age modeling (Li et al., 2022a, 2021; Herzschuh et al., 2022). This taxonomic harmonization trades off higher taxonomic resolution of some datasets for equivalence, resulting in overall comparability useful for analyses at large spatial scales. Despite advances in harmonization, the use of pollen data remains limited due to the fact that pollen compositions do not accurately reflect vegetation (Davis, 1963; Prentice, 1985; Prentice and Webb III, 1986). This limitation arises from variations in taxon-specific parameters ~~like~~ such as relative pollen productivity (RPP) and pollen dispersal characteristics, leading to discrepancies between the pollen record and ~~real~~ actual past vegetation. This hinders quantitative vegetation assessment as taxa with high pollen productivity and efficient pollen dispersal tend to be overrepresented in the pollen record, while those with low pollen productivity and less effective dispersal are underrepresented. These factors, together with the compositional nature of pollen data, result in a non-linear relationship between pollen and vegetation (Prentice and Webb III, 1986). Approaches such as the R-value model (Davis, 1963; Webb et al., 1981) and the extended R-value model (Parsons and Prentice, 1981) were created to address this issue and were refined with Sugita's (2007) model for "Regional Estimates of Vegetation Abundance from Large Sites" (REVEALS) . By accounting for taxon-specific RPP and fall speed values, as well as basin-specific parameters such as basin size and type, REVEALS models quantitative vegetation cover in ~~relevant pollen source areas~~ the region surrounding a basin from pollen compositions. The model has been applied in several regional-scale studies ~~(Nielsen et al., 2012; Mazier et al., 2015; Hellman et al., 2008; Nielsen and Odgaard, 2010)~~ (Nielsen et al., 2012; Mazier et al., 2015; Hellman et al., 2008) and multiple validations have demonstrated its ~~accuracy~~ ability in approximating actual vegetation (Sugita et al., 2010; Hellman et al., 2008; Soepboer et al., 2010; Mazier et al., 2012), even though the model's performance heavily relies on accurate taxon-specific parameters. While Wieczorek and Herzschuh (2020) and Githumbi et al. (2022)provide a comprehensive compilation of RPP and fall speed values for taxa of the Northern Hemisphere, the overall availability of RPP studies is still limited and regional variations in RPP values exist (Harris et al., 2020; Broström et al., 2008; Li et al., 2017; Mazier et al., 2012). This makes the application of REVEALS on larger scales particularly challenging. Only some (sub-) continental REVEALS reconstructions are available for Europe ~~(Trondman et al., 2015; Roberts et al., 2018;~~ (Trondman et al., 2015; Roberts et al., 2018; Githumbi et al., 2022; Serge et al., 2023), Asia (Cao et al., 2019; Li et al., 2022b, 2023, 2024a) and North America ~~(Dawson et al., 2018)~~(Dawson et al., 2024). Currently, no global or Northern Hemispheric quantitative vegetation cover reconstructions using REVEALS exist.

With its importance for the assessment of biome stability, carbon storage, climatic feedbacks, and land-use-change, forest cover is an often reconstructed variable ~~(e.g. Fyfe et al., 2015; Githumbi et al., 2021; Serge et al., 2023)~~(e.g. Fyfe et al., 2015; Githumbi et . Due to the global availability of remote sensing data on contemporary forest cover, it also offers good opportunities for the validation of reconstructions (Hjelle et al., 2015; Roberts et al., 2018). Yet, only Serge et al. (2023) and Pirzamanbein et al. (2014) use this opportunity for extensive validation and even improvement of reconstructions from European pollen records. No ~~site-wise validations or attempts at improvements of forest cover reconstructions by adjusting RPP values exist for other regions or on global scales~~grid-cell based validations exist for the Northern Hemisphere.

Here we present ~~global~~ reconstructed quantitative vegetation cover for the Northern Hemisphere from the LegacyPollen2.0 dataset - an updated global taxonomically and temporally standardized fossil pollen dataset of ~~3728~~ 3680 palynological records - using REVEALS spanning ~~primarily the last 50k years, with some records reaching back even further~~the last 14k years. The data sets were created using existing estimates of taxon-specific parameters~~and also applied an optimization approach to improve parameters. Using remote sensing forest cover we adjust RPP values for the ten most common taxa on each continent for better agreement of reconstructed with remote sensing forest cover. The REVEALS reconstructions with original and optimized parameters include~~ . The REVEALS reconstruction includes corrected vegetation compositions as well as reconstructed forest cover.

2 Methods

2.1 Pollen Data Set

The pollen data synthesis LegacyPollen2.0 (Li et al., 2024b) includes ~~3728~~ 3680 temporally resolved records (time-series) distributed globally. Data were collected from individual publications and the Neotoma Paleoecology Database which includes data from the European Pollen Database, the QUAVIDA data base for Australasia, the Latin American Pollen Database, the

African Pollen Database and the North American Pollen database (Flantua et al., 2015; Fyfe et al., 2009b; Giesecke et al., 2014; Lézine et a . An overview of Neotoma records included in LegacyPollen 2.0 and this reconstruction can be found in S1. Sediment and peat cores used for the creation of pollen data are of lacustrine, peat and marine origin. For the REVEALS reconstruction only lake and peat records in the Northern Hemisphere were used ($n = 2732$) Analogous to the preceding LegacyPollen 1.0 dataset (Herzschuh et al., 2022), the data synthesis involved revising age modeling and taxonomic harmonization for consistency of records. Spatial data coverage of records in the reconstruction is ~~densest in North America (1132~~ dense in Europe (1275 records) and ~~Europe (1451) , sparser~~ North America (1016 records) and sparsest in Asia (~~706) and very scattered in South America (191), Africa (164) and Australia and Oceania (84,~~ 441) (see Fig. 1). The records~~primarily span the last 50 ka with temporal coverage being a lot sparser before 20 ka BP~~' sample density decreases with age (see Fig. 2).

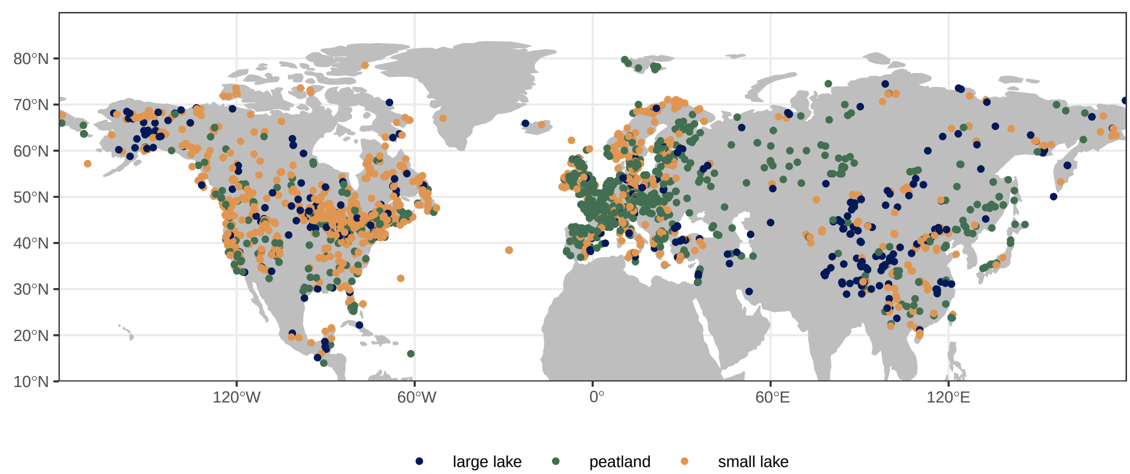


Figure 1. Pollen record locations in the LegacyVegetation dataset. Colors indicate record type (large lake $\geq$ 50 ha). Record density is ~~highes~~ highest in Europe and Eastern North America, and lowest in ~~Africa~~ Northern and ~~Australia and Oceania~~Central Asia.

2.2 Implementing REVEALS

The REVEALS model ("Regional Estimates of Vegetation Abundance from Large Sites") estimates quantitative vegetation coverage from pollen assemblages using site and taxon-specific parameters (Sugita, 2007). Based on wind speed and taxon-specific fall speed, pollen dispersal is modeled in ring sources around the basin and deposition over the basin is integrated to give pollen influx. Together with RPP this dispersal factor is used to correct original pollen counts to better represent ~~real~~ actual vegetation (see Equation 1 and Table 1). By running the model with variations of relative pollen productivity (RPP) values, a statistical distribution of results is calculated.

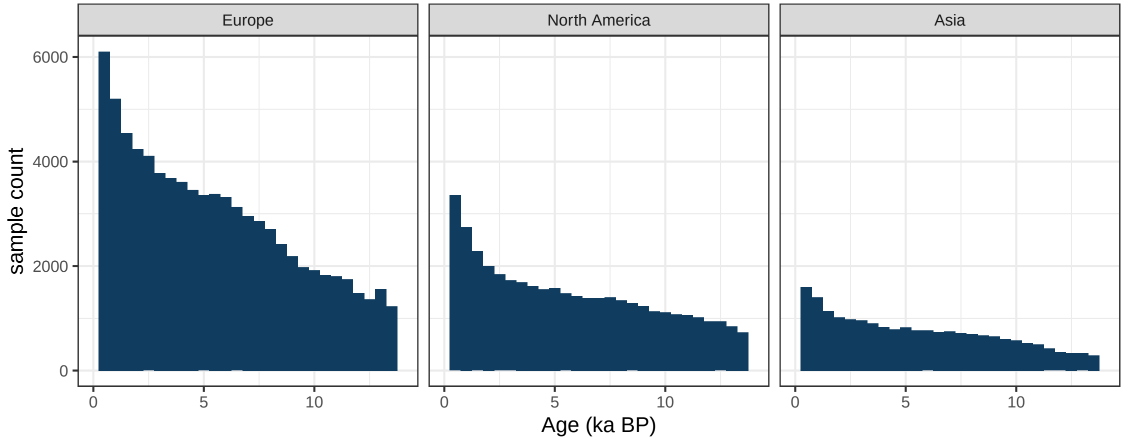


Figure 2. Temporal coverage of records in the LegacyVegetation dataset per continent. Bins are ~~1000~~ 500 years wide. Sample count decreases with age ~~with a noticeable drop in~~ and Europe has the most samples ~~at 20 ka BP~~overall.

$$\hat{V}_i = \frac{n_{i,k}/\hat{\alpha}_i \int_R^{Z_{max}} g_i(z)dz}{\sum_{j=1}^{m}(n_{j,k}/\hat{\alpha}_j \int_R^{Z_{max}} g_i(z)dz)} \tag{1}$$

The REVEALS model follows a set of assumptions. Firstly, neither directionality nor pollen transport through agents other than

Table 1. Algebraic terms in the REVEALS equation (see Equation 1)

Function term	~~explanation~~ definition
$\hat{V}_i$	vegetation estimate of taxon i
$n_{i,k}$	pollen counts of taxon i at site k
α_i	relative pollen productivity of taxon i
R	basin radius
Z_{max}	maximum extent of regional vegetation
z	distance from a point in the center of a basin
g_i	dispersal and deposition function for taxon i

wind are considered in the model. Additionally, it is assumed that the basin is circular with no source of pollen within the basin radius. The peatland and bog sites used in our reconstructions inherently violate this assumption. Nevertheless, the quantitative reconstruction of vegetation cover from peatland cores is possible by using Prentice's deposition model (Prentice, 1985, 1988) instead of Sugita's deposition model (Sugita, 1993) in the dispersal and deposition function (see Eq. 1; Sugita, 2007). Previous studies show that results from small bogs are still reliable when aggregated, while results from large bogs tend to deviate from those of large lakes ~~(Trondman et al., 2015; Mazier et al., 2012)~~(Trondman et al., 2015; Mazier et al., 2012; Trondman et al., 2016). Using peatland records for reconstructions is, therefore, appropriate ~~. All sitesthat were not classified as lakes were run with peatland settings~~when spatially averaging multiple sites. We use the implementation of REVEALS from the R package REVEALSinR (Theuerkauf et al., 2016).

2.2.1 ~~Parameters~~

~~For each site,~~ For further details on the REVEALS model ~~also requires information on basin type, basin size and original pollen counts, all of which were collected in the LegacyPollen 2.0 dataset (Li et al., 2024b).~~see the original publication Sugita (2007) or Githumbi et al. (2022).

2.2.1 Parameters and Model Settings

For each taxon, values for RPP (with uncertainties provided as standard deviation) and fall speeds are used. ~~When available, we use continent-specific values in our reconstruction following~~ We made use of the synthesis of Northern Hemisphere RPP and fall speed values by Wieczorek and Herzschuh (2020). Several RPP studies published since this synthesis were added to the compilation (Geng et al., 2022; Li et al., 2022b; Wang et al., 2021; Huang et al., 2021; Zhang et al., 2021a, b; Wan et al., 2020, 2023; Jian . The methods by Wieczorek and Herzschuh (2020) were followed fore study selection and calculation of synthesis values. An overview of original values and synthesized values can be found in Appendix A and B respectively.

When available, we use continent-specific values in our reconstruction. For taxa with no continental values present, we use ~~northern hemispheric~~ Northern Hemispheric values. If no values exist for a taxon, RPP is set to a constant (RPP = 1, σ=0.25) and fall speeds are filled with mean continental fall speeds~~(see Appendix A: Original RPP and fall speed values per continent~~. Continental RPP values are available for the majority of pollen counts in all three continents (see Fig. 3). The fraction of pollen counts for which ~~RPP estimates are available are much higher in the Northern Hemisphere than in the Southern Hemipshere (see Fig. 3).~~ standard RPP values were assumed is highest in North America but still < 10%. For each site, the REVEALS model also requires information on basin type, basin size and original pollen counts, all of which were collected in the LegacyPollen 2.0 dataset (Li et al., 2024b). Apart from taxon- and basin-specific parameters the REVEALS model requires several constant parameters to be set, which can be found in Table 2.

2.2.2 Modifications in REVEALSinR

We calculate the radius of ~~relevant~~ the 80% pollen source area by finding the radius in which the median influx of all taxa is 80% of the total influx (as defined by the total influx in the maximum extent of regional vegetation chosen). This is calculated by employing the lake deposition model in REVEALSinR (Theuerkauf et al., 2016). Starting from z_{max} the deposited pollen is calculated per taxon. This is assumed to be the total pollen each taxon deposits. In a step-wise process the radius around the basin is increased and the deposited pollen relative to the total influx at z_{max} is calculated for each taxon. We define our 80%

Table 2. Static model parameters and model settings for REVEALS runs using REVEALSinR (Theuerkauf et al., 2016).

Parameter	Values and settings used in model run
atmospheric model	unstable atmosphere
dispersal model	gaussian plume
wind speed	$3m \times s^{-1}$
maximum extent of regional vegetation (region cutoff)	1000 km
number of RPP variations	2000
peatland basin radius	100 m
function to randomize pollen counts	rmultinom_reveals

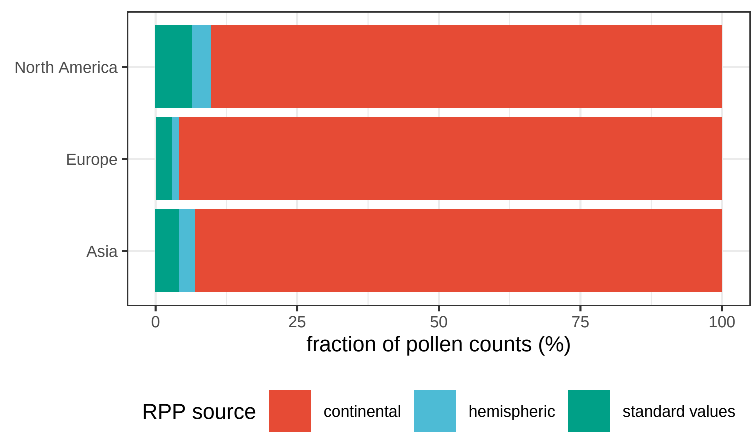


Figure 3. ~~Percentage~~ Regional source of RPP values for percentage of pollen counts per continent~~for which RPP estimates are available~~. A ~~higher~~ majority of pollen counts is covered by continental RPP values with the highest fraction in Europe. Only a small percentage of pollen counts has only hemispheric RPP ~~information in the Northern Hemisphere compared~~ values available. No available RPP values lead to the ~~continents~~ use of ~~the Southern Hemisphere~~a standardized RPP value of 1 ± 0.25.

pollen source radius as the radius where the median of the relative influx of all taxa reaches 80%. The primary objective of this calculation is to provide a clear understanding of the scale of the source area for users unfamiliar with pollen data. It highlights the regional nature of lacustrine pollen data and demonstrates the influence of lake size on this source area.

We also reduced computational effort in REVEALSinR by implementing a maximum number of steps in the lake model used to model mixing in the basin. The number of steps was set to 500 unless n falls below that maximum value for $n = basin\,radius/10$ for basins with a radius of at least 1000 m and $n = basin\,radius/2$ for basins with a radius smaller than 1000 m.

2.3 Reconstruction of forest cover and validation

Forest cover was reconstructed by summing up percentages of arboreal taxa (see ~~S1~~S2: List of arboreal taxa) with Betulaceae, *Betula*, and *Alnus* being classified as arboreal at sites below 70° N. The mean reconstructed compositional coverages from the REVEALS results were used for the forest cover reconstructions. REVEALS results were then rasterized to aggregate and include records from smaller basins as well. Reconstructed time series were averaged in 500 year bins and then rasterized in grids of differing spatial resolution. A grid cell was classified as having a valid reconstruction when it contained records from at least one large lake (>= 50 ha) or at least two small basins following Serge et al. (2023). Standard deviations of the REVEALS estimates were aggregated by applying the delta method by Stuart and Ord (1994), using the same equation as Wieczorek and Herzschuh (2020). We provide a script for rasterization with adjustable temporal and spatial resolution for users of the dataset on Zenodo (https://zenodo.org/doi/10.5281/zenodo.12800290). For validation, the reconstructed forest cover of the past ~~500 years was~~ 100 years was rasterized and compared to modern remote sensing forest cover. Only valid grid cells as defined above were used for validation. Average tree canopy cover ~~within pollen source areas of all sites~~ for all grid cells was extracted from the Landsat Global Forest Cover Change (GFCC) data set from the temporal average of the years 2000, 2005, 2010 and 2015 (Sexton et al., 2013; Townshend, 2016). An openness correction was applied to sites containing urban areas and paved surfaces within the 80% pollen source areas (PSA) to correct for areas without any pollen sources and thus ~~improve~~ ensure comparability to modern remote sensing forest cover (see Equations 2-4). For this, the percentage of unvegetated land cover classes for the year 2015 in the ESA CCI land cover data set was used (ESA, 2017, see Table 3). Areas covered by water or ice are already considered as missing values in the remote sensing forest cover data set and do not need to be corrected for. Forest cover was validated ~~site-wise~~ for each grid cell and mean absolute error (MAE) and correlation coefficients calculated for each continent. No openness correction was applied to the reconstruction values in the final dataset. Validation for a 2x2° grid is included in the results section. Further validations using 1°,5°, and 10° resolution are included in the supplementary material (S3: Validation results for different spatial resolutions).

Table 3. Unvegetated land cover classes in ESA CCI LC chosen for the openness correction.

Name	Code
Urban areas	190
Bare areas	200
Consolidated bare areas	201
Unconsolidated bare areas	202

$$unvegetated\ classes = \{190, 200, 201, 202\} \quad (2)$$

$$unvegetated\,(\%) = \cancel{\frac{\sum cells\ in\ PSA \in open\ classes}{\sum cells\ in\ PSA}} \frac{\sum cells\ in\ PSA \in unvegetated\ classes}{\sum cells\ in\ PSA} \tag{3}$$

$$corrected\,tree\,cover = reconstructed\,tree\,cover \times (1 - unvegetated) \tag{4}$$

2.4 ~~Optimization~~

~~In addition to the REVEALS approach, which is motivated by a biophysical model but also based on a large number of model choices and parameters, we also apply a statistical approach. Here, RPP values for common taxa are estimated by minimizing the misfit of reconstructed and remote sensing forest cover. For the optimization we rely on the "L-BFGS-B" method (Byrd et al., 1995), which allows for box constraints, and minimize the residual sum of squares (RSS) of reconstructed forest cover with remote sensing forest cover. RPP values were bound by upper and lower limits based on original RPP values (see Equation ??). Fall speeds and standard deviations of RPP were kept constant to the REVEALS approach.~~

$$\cancel{original\ RPP \times 0.25 < new\ RPP < original\ RPP \times 4}$$

~~The RPP values were optimized for the ten most common taxa in the REVEALS reconstruction for all sites on a continent, forest cover reconstructed, and the residual sum of squares (RSS) with remote sensing forest cover calculated. The results were validated using a spatial leave-one-out (SLOO) cross-validation. In this cross-validation one site and all sites within a predefined radius (exclusion buffer) were excluded from the optimization to account for spatial autocorrelation. The optimized RPP values were then applied to the forest cover reconstruction of the site left-out and the absolute error with remote sensing forest cover recorded. This was repeated with 20 sites to estimate the spread of MAE. The exclusion buffer around the validation site was set to 200 km. Due to computational limitations (roughly 3 hours for one continental SLOO fold using 20 threads with 1.2 GHz CPU each), the number of sites used per continental optimization during the cross-validation was limited to 100, leading to a rather conservative estimate of the true error.~~

3 Data summary

3.1 80% Pollen Source Areas

Using REVEALS~~and original RPP values~~, radii of ~~relevant~~ 80% pollen source areas were calculated for ~~all sites~~ large lakes(see Fig. 4). The ~~relevant pollen source areas~~ radii indicate in which area 80% of the deposited pollen originated from (see Section 2.2.2) and yield an understanding of which area the pollen record is representative of~~. The~~, which is especially useful when individual time series from large lakes are being used for analyses. The 80% pollen source areas are roughly a function of basin size (see Fig. 5) and range between ~~68 km and 729~~ 155 km and 762 km. The median 80% pollen source radius is ~~86~~ 225 km including all ~~basins and 138 km including only lakes.~~ large lakes.

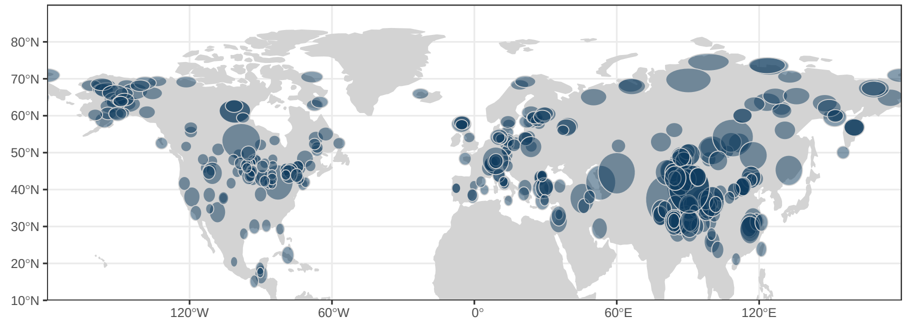


Figure 4. Map indicating the ~~size of~~ relevant pollen source areas for ~~all records~~large lakes. Many small basins in Europe lead to smaller 80% pollen source areas. Several large basins and correspondingly large 80% pollen source areas exist in Asia. In general the 80% pollen source areas highlight the regional nature of the pollen record signal.

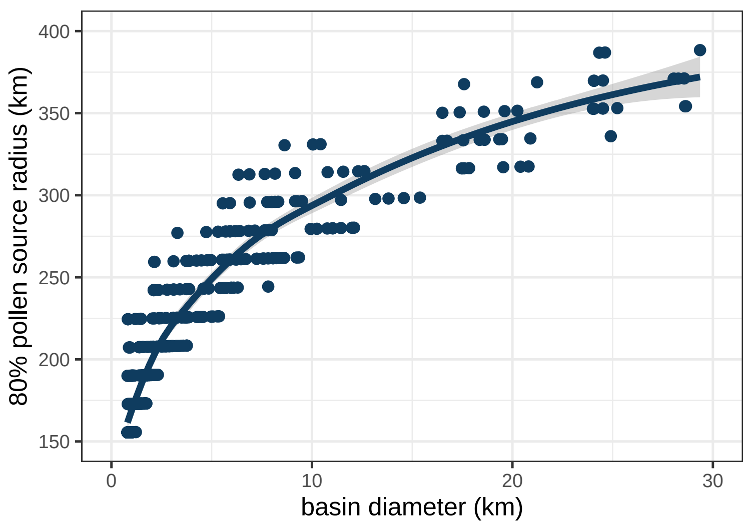


Figure 5. ~~Scatter plot~~ Scatterplot of basin diameter and 80% pollen source ~~radius~~ area of ~~a subset of~~ large lakes in the REVEALS ~~records with original RPP values~~data set. ~~Larger~~ In general, larger basins have larger pollen source areas with the relationship between ~~basin~~ diameter and 80% pollen source radius being roughly logarithmic.

3.2 ~~Comparison of original and optimized RPP values~~

~~The calculated pollen source areas (see section 3.1) were used to extract modern remote sensing forest cover per site. Within the optimization, RPP values were adjusted for the ten most common taxa per continent to improve the fit between reconstructed and remotely sensed modern forest cover. The RPP values are one of the main correction factors applied in REVEALS. Here we compare original and optimized RPP values for the relevant continental taxa.~~

~~The magnitude of adjustment from original to optimized RPP values differs between continents (see Fig. ??). The highest and lowest absolute change respectively occurred for *Quercus* (4.08) and Fabaceae (0.09) in Africa, for *Picea* (87.81) and *Ephedra* (0.43) in Asia, for *Pinus* (32.58) and Asteraceae (0.16) in Europe, for *Alnus* (1.79) and Amaranthaceae (in which we included Chenopodiaceae, 0.02) in Australia and Oceania, for Amaranthaceae (63.81) and *Tsuga* (0.43) in North America, and for Amaranthaceae (15.91) and Melastomataceae (0.74) in South America (see Appendix B). Relative change of RPP values is mostly positive with many taxa reaching an increase of three times the original RPP value. This is the maximum RPP value that can be reached, as the upper constraint for RPP optimization was set as 4 times the original RPP value (see Section 2.4). In most cases RPP values for arboreal taxa are increased. This increase represents reconstructed forest cover being regulated down as can be seen in the validations (see Fig. 9). Dumbbell graph illustrating original and optimized RPP values per continent and taxon. Arboreal taxa such as Pinus, Picea, Quercus have increases that are especially large.~~

3.2 Reconstructed compositions

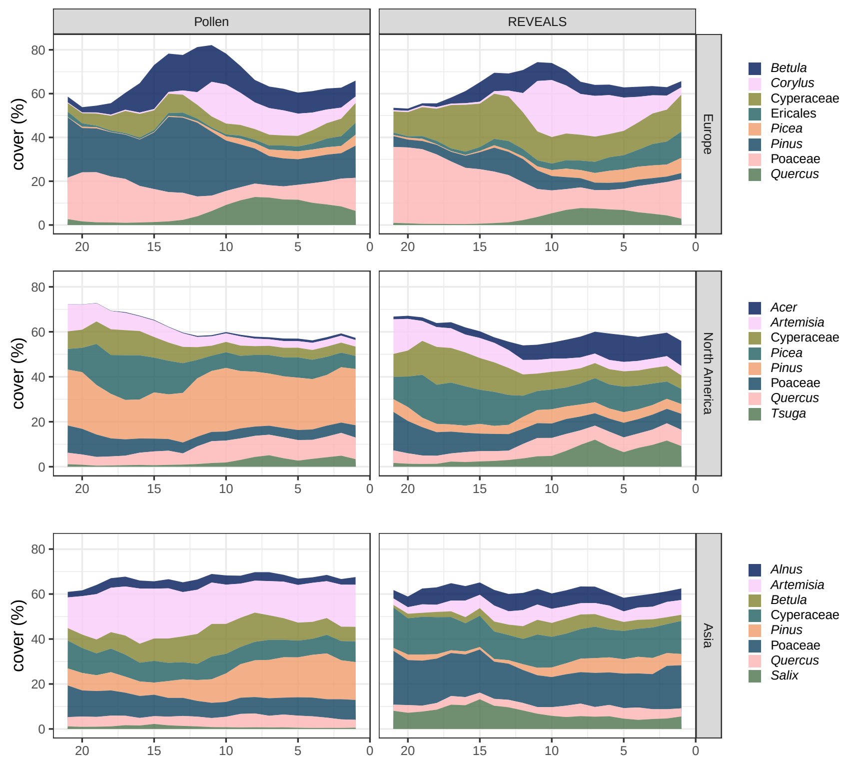


Figure 6. Average continental taxonomic coverages per reconstruction for the 8 most common taxa per continent. ~~Compositional differences~~ Differences are ~~more pronounced in~~ especially evident for Pinus, Artemisia, and Betula, which all have decreased coverages after the ~~Northern Hemisphere due to the availability~~ application of ~~more RPP values~~REVEALS, as well as Poaceae and Cyperaceae with increased coverages.

~~Both the original and optimized RPP values were used to run REVEALS and~~ REVEALS was used to reconstruct quantitative vegetation cover. ~~Due to the differences in RPP values the reconstructed compositions differ between both REVEALS runs.~~ Here we compared these reconstructed compositions ~~among each other and with~~ to the original pollen composition.

Differences in composition ~~are especially apparent for~~ between Pollen data and REVEALS are apparent for all continents of the Northern Hemisphere. ~~For example, compared to the original pollen composition REVEALS runs with the original and the optimized RPP values both increase *Larix* cover in Asia, Ericales cover in Europe, and decrease *Picea* cover in North America, although the version with optimized RPP values does so more strongly (see Fig. 6). The original and the optimized version also diverge in the adjustment of some taxa. *Artemisia* cover in Asia is reduced by the original version and increased by the optimized one. *Picea* cover stays roughly the same with original RPP values in North America and decreases with optimized ones and while Asteraceae cover in Europeis increased in the REVEALS version with original RPP values, it is considerably higher in the optimized one.~~

~~In the Southern Hemisphere the differences between reconstructions are much less pronounced (see Fig. 6). The REVEALS reconstruction with original RPP values is almost indistinguishable from the original pollen spectra and adjustments in the optimized version are also much smaller than in the Northern Hemisphere. An increase in Cyperaceae cover in Australia and Oceania, decreases of Asteraceae and Cyperaceae in South America, and~~ Some clear examples include: increases of Cyperaceae in all continents, decreases of ~~*Quercus* in Africa are evident in the REVEALS run with optimized RPP values.~~

~~The difference in reconstructions between the hemispheres is most likely due to the availability of regional RPP and fall speed values. For South American taxa many RPP values are unknown and for remaining taxa average values of Northern Hemispheric studies were used~~ Betula in Europe, decreases of Pinus in all continents, and increases of Acer in North America with the application of REVEALS and its intended correction of taxon-sepcific biases (see Fig. ~~3 and Appendix A).These are often close to 1 and, therefore, do not change the original compositions drastically. Improving reconstructions without more available RPP estimates for Southern Hemispheric taxa is unrealistic.~~ 6).

3.3 Reconstructed forest cover

Using the compositional data available from the original pollen data ~~, the REVEALS run with original RPP values,~~ and the REVEALS run~~with optimized RPP values (see section 3.3)~~, we reconstructed forest cover for all sites and samples and rasterized the result with different spatial resolutions. The temporal trend in Northern Hemisphere forest cover is the same for ~~all three~~ both reconstructions. Forest cover increases from ~~20~~ 14 ka BP until roughly 6 ka BP and decreases again towards the present (see Fig. 7). REVEALS reconstructed forest cover is generally lower than forest cover from original pollen compositions. On average forest cover values from the REVEALS run ~~with original/optimized RPP values are roughly 11/19~~are roughly 14.54% lower than values from original pollen compositions. The temporal trends in Asia and North America are positive, whereas forest cover in Europe has its maximum around 6 ka BP and has been decreasing since.

Forest cover is ~~higher in the Northern Hemisphere in all time slices and reconstructions with the exception of the Eurasian Steppe, which is always characterized by a low reconstructed forest cover~~ generally highest in Eastern North America. This

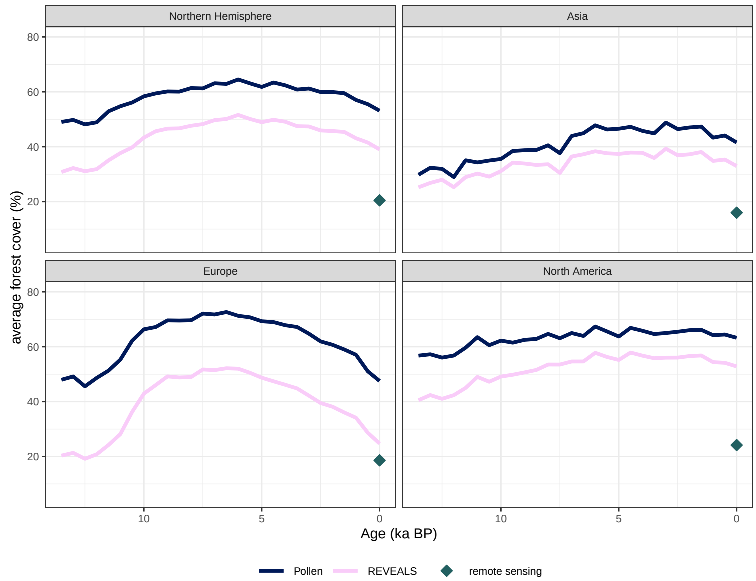


Figure 7. ~~Global~~ Northern Hemisphere and continental average forest cover from ~~10x10~~2x2° grid cell means for raw pollen data ~~, the REVEALS reconstruction with original RPP values,~~ and the REVEALS reconstruction ~~with optimized RPP values~~(Northern Hemisphere and continental averages from different grid cell resolutions are available in S2: Reconstruction results for different spatial resolutions). Remotely sensed global average forest ~~cover~~ clover for the grid cells with valid pollen ~~record locations~~ coverage is indicated with the diamond. Temporal trends are the same, but absolute forest cover is reduced in the REVEALS reconstructions compared to the original pollen data. ~~Forest cover from REVEALS~~ Both reconstructions ~~with optimized RPP is lowest~~still overestimate forest cover.

is also where data coverage is best in North America (see Fig. 8). ~~Within REVEALS reconstructions, forest cover is reduced more in~~ Density of valid grid cells is very high in Europe, where forest cover increases until roughly 6 ka BP and then decreases. Data coverage in Asia is sparse, but valid grid cells indicate higher forest cover on the Southeastern coast and in the boreal biome. Rather open areas exist at the Tibetan Plateau and at very high latitudes. The forest cover derived from the ~~Northern Hemisphere than in the Southern Hemisphere. A continuous band of highly forested boreal forest is visible in the REVEALS reconstructions using original RPP values. The intensity of this band is reduced in the~~ REVEALS reconstruction ~~using optimized RPP values~~is generally lower. However, ~~areas in northeastern Siberia, China, and eastern North America remain strongly forested~~the difference between Pollen and REVEALS forest cover is smaller in North America than in Europe and Asia.

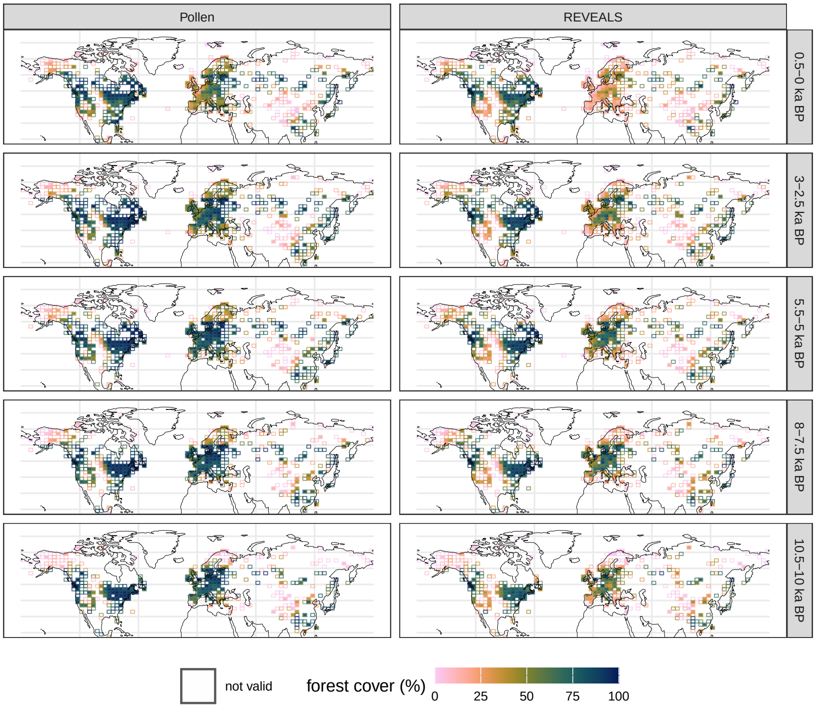


Figure 8. Reconstructed forest cover in ~~10x10~~2x2° grid cells from raw pollen data ~~, the REVEALS reconstruction with original RPP values,~~ and the REVEALS reconstruction for 5 example time slices (reconstructions with ~~optimized RPP values~~different grid cell sizes are available in the in S2: Reconstruction results for different spatial resolutions). Valid cells are filled and include reconstructions from at least one large lake ($\geq$ 50 ha) or several smaller basins. Forest cover in Eastern North America is ~~generally~~ higher than in ~~the Northern Hemisphere~~Europe and Asia. ~~Reductions of~~ REVEALS reconstructed forest cover ~~with the REVEALS~~ is generally lower than raw pollen reconstructions~~are higher in the Northern Hemisphere~~.

3.4 Validation with gridded data sets

3.4.1 ~~Validation with complete data sets~~

Remote sensing forest cover within relevant pollen source areas was used to validate the modern, reconstructed forest cover from the original pollen data and ~~both REVEALS runs for each site. As the true error for the optimization results will be underestimated here, we also present results from the SLOO validation is Section 3.5.2.~~ the REVEALS run for each grid cell. Here we present validation of gridded data with a 2° spatial resolution. Validations with additional spatial resolutions differ only marginally and are included in the supplementary materials (S3: Validation results for different spatial resolutions). Forest cover reconstructed from original pollen data is predominantly higher than remote sensing forest cover with a ~~global~~ mean absolute error (MAE) of ~~34.39%~~ 33.05% in the Northern Hemisphere (see Fig. 10a). As reconstructed forest cover is much lower for ~~both REVEALS runs~~ the REVEALS reconstruction (see Fig.7), ~~MAE values are reduced for both REVEALS~~

~~reconstructions. Using the original RPP values yields an MAE of 20.35% of reconstructed to remotely sensed forest cover. This is further reduced to 14.36% using the optimized RPP values~~ the MAE value is reduced significantly to 19.73% (see Fig. 9a).

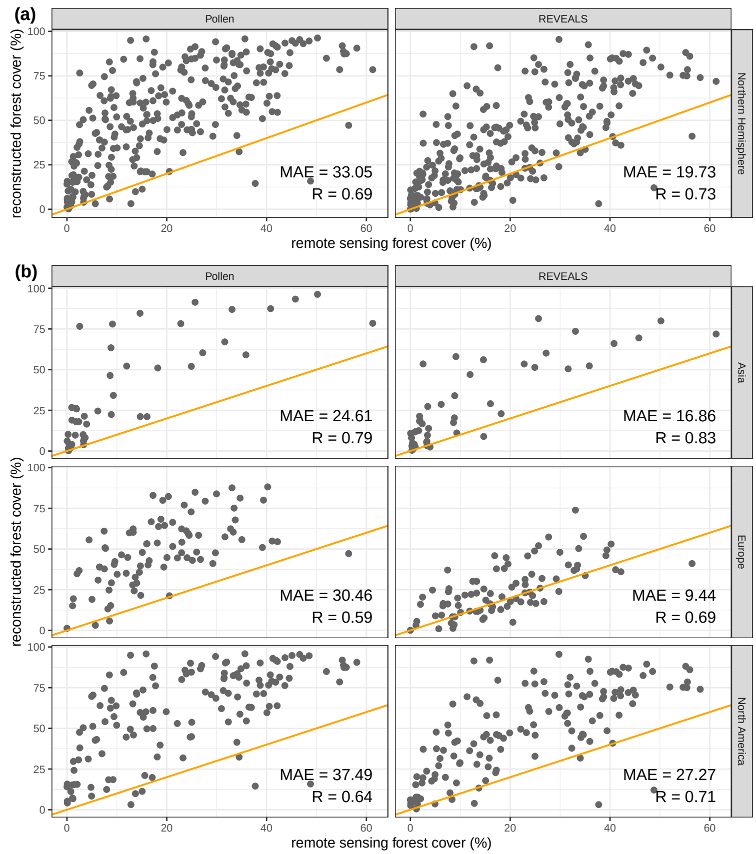


Figure 9. Remote sensing forest cover (LANDSAT) and modern reconstructed forest cover from Pollen ~~, REVEALS with original RPP values,~~ and REVEALS (< 100 years BP) in 2x2° grid cells with ~~optimized RPP values globally~~ mean absolute errors (~~a~~MAE) and ~~for all continents~~ correlation coefficient (~~b~~R) per group. Reconstructed forest cover from the original pollen data tends to overestimate observed (remote sensing) forest cover. ~~This is improved~~ Improvements with the REVEALS ~~run using original RPP values and even more so~~ reconstruction are especially high in Europe. Validations with different grid cell sizes are available in the ~~REVEALS run using optimized RPP values~~supplement (S3: Validation results for different spatial resolutions).

Continental mean absolute errors (MAE) in forest cover from original pollen data range from ~~12.44% (Africa) to 44.22~~24.61% (Asia) to 37.49% forest cover (North America, see Fig. 9b). All continental MAE values are lower for the REVEALS reconstruction ~~with original RPP values~~ and range from ~~12.33% (Africa) to 28.73~~9.44% (Europe) to 27.27% (North America). The improvement is largest in Europe (~~58~~72% relative to the initial MAE of the pollen-based reconstruction, see Fig. 9 and 10) and smallest in ~~Africa (1~~North America (24%). ~~Forest cover from the REVEALS reconstruction with optimized RPP values reduces continental MAE values even further with values ranging between 9.1% (Africa) and 21.08% forest cover (South America). MAE are generally improved more with optimized RPP values with the exception of records in Australia and Oceania. The largest improvement (relative to the pollen-based forest cover MAE) was achieved in North America (65%) but reconstructions in Europe (61%) and Asia (48%) also reduced the original MAE by more than or roughly half. The REVEALS runwith optimized RPP values~~REVEALS reconstructed forest cover also has higher correlation coefficients in all continents. The REVEALS run, therefore, produced ~~the~~ reconstructed forest cover that corresponds ~~best with~~ better remote sensing forest cover~~, with the exception of records from Australia and Oceania. Additionally, the reduction of forest cover MAE, and therefore the reconstruction improvement, was much larger in the continents of the Northern Hemisphere for both REVEALS runs~~. Nevertheless, forest cover still tends to be overestimated.

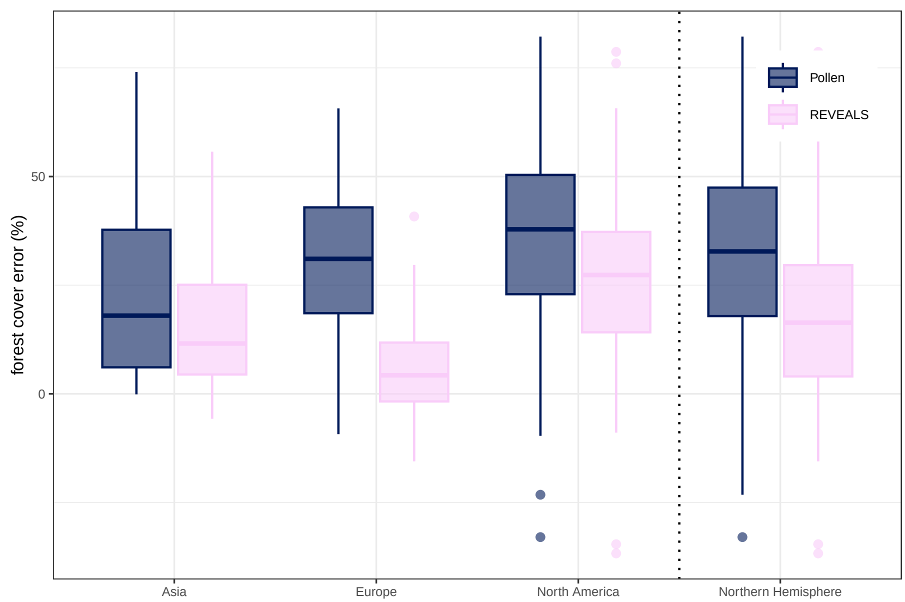


Figure 10. ~~Bar graph of MAE improvement relative to the MAE of the pollen-based~~ Forest cover reconstruction error per continent ~~and REVEALS~~ for a gridded 2x2° reconstruction. ~~The absolute MAE reduction is shown in~~ Mean errors decreased with the ~~text labels. Except for Australia and Oceania, the~~ REVEALS reconstruction ~~with optimized RPP values achieves higher improvements. Improvements~~ for all continents but are still generally ~~higher~~ > 0 (overestimation of forest cover). Lowest errors are present in ~~the Northern Hemisphere~~Europe.

Spatial patterns are present for the errors of ~~all three~~ both forest cover reconstructions (see Fig. 11). In ~~the Southern Hemisphere, especially western South America, forest cover is predominantly underestimated by the reconstructions. The highest errors in reconstructed forest cover occur in continents of the Northern Hemisphere where forest cover is predominantly overestimated by the pollen-based reconstruction. In~~ Europe the REVEALS ~~reconstructions manage~~ reconstruction manages to reduce errors extensively. In ~~eastern North Americasome records still tend~~ Eastern and coastal Northwestern North America, the

REVEALS reconstruction still tends to overestimate forest cover~~, even with the application of REVEALS and after optimizing~~. This could be due to a lack of continental RPP values. ~~The same is the case for several records in eastern Asia~~In North America, few RPP studies are available (see Appendix A) and more taxa are assigned hemispheric or standardized values than in the other continents.

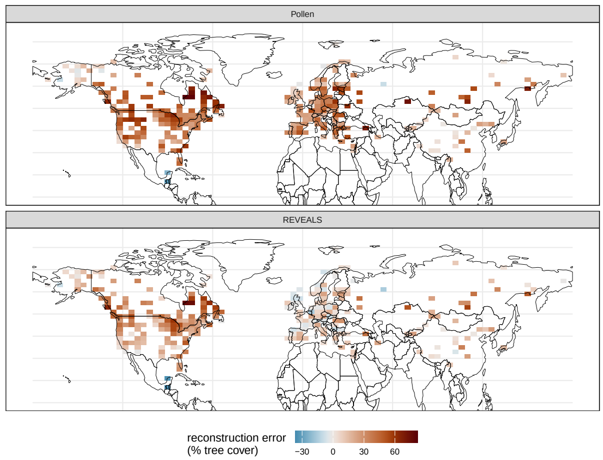


Figure 11. Map of the reconstruction error (in % forest cover) for forest cover reconstructed from Pollen ~~, REVEALS with original RPP values~~ and REVEALS data. Remaining errors with ~~optimized RPP values~~the overall better REVEALS reconstructions are especially high in North America (Northern West Coast, Labrador Peninsula).

The large difference between forest cover reconstructed from original pollen compositions and remote sensing forest cover could be due to the difference in the signal that is recorded. Remote sensing forest cover records the canopy, whereas pollen data also records the vegetation present below the tallest canopy. Several layers of trees could, therefore, increase the percentage of arboreal taxa recorded. Even though this comparison between these data sources may not be straightforward, it is still necessary for this large-scale validation of reconstruction as few other vegetation data is available globally. Additionally, it is more likely that the overestimation of forest cover in the initial pollen data is due to the higher production of pollen by trees than by non-arboreal taxa. This leads to an overrepresentation of arboreal taxa in the pollen record. By using REVEALS, the pollen productivity of taxa is taken into account and corrected for. The proportion of arboreal taxa is therefore strongly reduced in the vegetation compositions reconstructed using REVEALS.

~~The reasons for the difference in reconstruction improvements between the hemispheres could lie both in the smaller number of records available and the lack of regional RPP estimates for continents of the Southern Hemisphere. The latter play an important role as the optimization is based on the original RPP estimates and can only determine better values if these are in~~

~~the range of the original RPP values described in Equation **??** (see Sect. 2.6). An effective optimization of RPP values may, therefore, rely on some existing continental RPP estimates that can be refined with the optimization approach.~~

~~Optimizing more RPP could also solve the lack of regional improvements in eastern North America. This area is, amongst others, dominated by *Acer* which is not one of the ten most common taxa in the RPP optimization in North America. Optionally, this could also be solved by optimizing on subcontinental scales, though this requires a sufficient amount of regional records.~~

3.4.1 ~~SLOO Validation of Optimization~~

~~A spatial leave-one-out validation was conducted by excluding a subset of available records in the optimization (see Sect. 2.4). By separating testing and training sites, the true spread of forest cover error from the optimization of RPP values can be evaluated. This also indicates the potential error if the optimized parameters were to be applied to new records. The distribution of absolute error from the SLOO validation is comparable to that of the reconstruction utilizing the complete optimization for Africa, Asia, Europe and South America (see Fig. **??**). In North America, the absolute error spread and media are larger in the SLOO validation than in both REVEALS reconstructions. As errors in North America were comparably large to begin with (see Fig. 10 and 12), this could be due to the small number of folds conducted in the SLOO validation (n = 20) as well as the small number of records used (n = 100). The same could be the case for Australia and Oceania. Additionally, the spatial buffer in the SLOO validation leads to even fewer records being available for optimization. This could further decrease improvements in Australia and Oceania optimization. Overall the SLOO validation results indicate that the optimization success is relatively stable in Africa, Asia, Europe and South America. In North America, the spatial variability leads to higher uncertainty and in Australia and Oceania the optimization is not able to decrease absolute errors considerably. Boxplot of absolute errors from continental SLOO validations (20 folds) and from validations with complete Pollen, REVEALS (original RPP) and REVEALS (optimized RPP) data sets. The SLOO validation shows how reliable the optimized parameters are when testing sites were not included in the optimization. Variance and averages of absolute errors are comparable to the entire optimization dataset for Africa, Europe, Asia and South America. Errors are larger in Australia and Oceania and North America.~~

4 Dataset applications and limitations

Our reconstructed quantitative vegetation cover datasets using REVEALS provide ~~global coverage~~ reconstructions of taxonomic compositions as well as forest cover ~~and extend to 50 ka BPand beyond~~in Europe, Asia, and North America and extend to 14 ka BP. The reconstructions made use of taxon-specific parameters and were, thus, able to correct some of the compositional biases present in pollen compositions. Notably, the error in modern reconstructed forest cover was reduced compared to pollen-based reconstructions on all continents which shows that improvements in forest cover reconstructions from ~~both~~ REVEALS applications are considerable.

Reconstruction results are also similar to available large-scale pollen-based vegetation reconstructions. Increases in forest cover in northern and eastern Asia up until the Holocene thermal maximum as seen in our results are consistent with reconstructions by Cao et al. (2019) and Tian et al. (2016). The reconstructed spatial patterns of forest cover in China with low forest cover in the North China plain and the Tibetan Plateau and a higher forest cover along the east coast and the south agree with previous reconstructions as well (Li et al., 2023, 2022b, 2024a). Results for European forest cover also roughly correspond with previous REVEALS applications and show an increase of forest cover after the last glacial maximum until roughly ~~4 ka BP (Githumbi et al., 2021; Fyfe et al., 2015; Serge et al., 2023)~~6 ka BP (Githumbi et al., 2022; Fyfe et al., 2015; Serge et al., 2023; Strandberg . The gridded reconstruction by Serge et al. (2023) was even validated with modern remote sensing forest cover and showed a good fit.

The REVEALS forest cover reconstructions presented here offer valuable insight into past vegetation changes. The global dataset provides an opportunity to explore past vegetation dynamics, gaining a deeper understanding of responses, trajectories, and potential feedback mechanisms. Given the increasing discussions surrounding the possibility of tipping events in vegetation cover (Armstrong McKay et al., 2022; Lenton and Williams, 2013), this could be of considerable use. While a reconstruction of exact tree lines is not trivial with pollen data, the application of REVEALS and subsequent biomization improve treeline reconstructions as shown by Binney et al. (2011). Additionally, this dataset can address unanswered questions about Holocene vegetation dynamics, including the deglacial forest conundrum ~~(Dallmeyer et al., 2022)~~(Dallmeyer et al., 2022; Strandberg et al., 2022) . It also serves as a valuable tool for validating models with coupled climate and vegetation, ~~relying~~ which rely on extensive time series and vegetation data for accurate predictions ~~(Dallmeyer et al., 2023).~~ (Dallmeyer et al., 2023; Dawson et al., 2024). Comparing modeled vegetation to reconstructed vegetation could help uncover missing dynamics in coupled climate-vegetation models. New insights gained from these applications could enhance our ability to predict future changes.

However, the reconstructions are associated with some of the limitations of sedimentary pollen data. This includes age uncertainty, temporal mixing, and irregular spatial and temporal resolution of records. Age uncertainty is already treated as best as possible through consistent age modeling of the pollen dataset (Li et al., 2022a, 2021). Nevertheless, in general, replicating sediment and peat cores could provide more accurate estimates.

Moreover, there is uncertainty surrounding the success of the compositional reconstructions. As global compositional vegetation data is not readily available, using remote sensing forest cover poses as the best option for validation. Even with an accurate forest cover reconstruction, uncertainties persist regarding the abundance of individual taxa due to the aggregated nature of the forest cover measure. To address this, global syntheses of forest and other plant inventories or compositional remote sensing products could offer better validation. ~~The optimized RPP set can produce very unrealistic compositions, for example regarding Asteraceae in Europe. The optimization was conducted purely statistically and limited ecological information was provided as input. The use of original RPP values, originating from physical studies, is, therefore, the more conservative approach for compositional reconstructions and the optimized data set should be used with caution for compositional applications. Although, many missing RPP and fall speed values, especially for taxa in the Southern Hemisphere, result in uncertainties in the original REVEALS reconstruction as well. A higher number of RPP estimates could help increase not only the confidence~~

~~in compositional reconstructions, but also the optimization success in continents of the Southern Hemisphere, where the small amount of information led to lower improvements in forest cover reconstruction.~~

Another challenge lies in validating the results with past vegetation data. It is uncertain whether RPP values have remained stable over time, and historical compositional data are not only scarce but likely too recent to test this assumption (Baker et al., 2016). Vegetational compositions from sedimentary ancient DNA could provide a solution. Local aDNA vegetation signals could be averaged across multiple records within a pollen source area to generate a comparable reconstructed vegetation composition using a different proxy and to compare to pollen-based results (Niemeyer et al., 2017).

To ensure the correct utilization of the dataset and to obtain reliable analysis results, several key considerations should be followed. Firstly, rasterization mitigates individual errors by temporal and spatial averaging. This process is particularly useful in reducing the variance that might arise from individual measurements, providing a more reliable representation of the underlying signal.The reliability of reconstructions varies among different taxa due to the quality of RPP values, and this is explicitly documented in a supplementary file that outlines the sources of RPP values (see Section Code and Data availability). Reconstructions of taxa with continental RPP values are the most reliable, followed by those based on hemispheric data, with standardized RPP values being the least reliable. This hierarchy should be taken into account when interpreting the results. Higher certainty is associated with forest cover reconstruction, as it is based on aggregation among taxa. Reconstructions of temporal forest cover trends are reliable, as evidenced by high correlation coefficients, despite a tendency for absolute values to be overestimated, particularly in North America. For individual time series, the reliability of data varies with the size of the lakes from which samples were taken. Only data derived from large lakes ($\geq$ 50 ha) are reliable for site-wise analyses. This distinction is clearly indicated with validity flags in the dataset. Reconstructions from smaller basins should not be used alone.

5 Conclusions

We present data sets of reconstructed compositional vegetation and forest cover ~~from a globally distributed~~ in the Northern Hemisphere from a sedimentary pollen data set using the REVEALS model. We used ~~published (original), continental~~ synthesized RPP values for ~~one reconstruction , while in a second reconstruction, we optimized continental RPP values for common taxa by incorporating remote sensing forest cover data~~reconstruction and made use of hemispheric or standardized values, when continental ones were not available. This approach allowed us to address some of the inherent biases in pollen compositions~~and suggests a method for enhancing taxon-specific RPP estimates~~. Considerable improvement in the reconstruction of forest cover is ~~especially achieved in the continentsof the Northern Hemisphere. Even though improvements of reconstructions in the Southern Hemisphere were largely possible as well, the collection of more regional RPP values is indispensable for better reconstructions~~achieved in all continents. Improvements were smallest in North America, which suggest a need for further RPP studies.

Accurate data on past vegetation is invaluable for the validation of coupled climate-vegetation models and the testing of hypotheses on feedback effects and vegetation dynamics. This knowledge is essential for modeling and predicting vegetation trajectories under anthropogenic climate change.

6 Code and data availability

The produced datasets are freely available from ~~PANGAEA (, , Herzschuh et al. 2023c; Schild et al. 2023~~Zenodo (https://doi.org/10.5281/zenodo.12800159).

Input data from LegacyPollen 2.0 is available on PANGAEA ~~as well~~ (https://doi.pangaea.de/10.1594/PANGAEA.965907, Li et al. 2024b).

The code used to produce the datasets ~~is~~ and adjustable rasterization code are freely available from Zenodo (https://doi.org/10.5281/zenodo.10191859, https://doi.org/10.5281/zenodo.12800291, Schild and Ewald 2023).