Reply to Mariani

Laura Schild & Ulrike Herzschuh

General reply

Dear Michela Mariani,

We thank you for your careful review or our manuscript. We appreciate the valid points you have made and find them to be easily remediable. Regarding other matters you have mentioned, we see a good opportunity to provide clarification on how we follow common procedures and highlight the usability and validity of our dataset.

Firstly, the use of continental syntheses of RPP values is widely accepted in continental-scale reconstructions and has been applied in several previously published reconstructions in Europe, North America and Asia. The use of even hemispheric values, when continental ones are missing, maximizes the utility of available data while still producing improved reconstructions as our validations show. As we cannot achieve small-scale perfect reconstructions yet, we advocate for these broader and necessarily coarser insights into vegetation dynamics by generalizing reconstructions.

Furthermore, our REVEALS application using previously published synthesized RPP values is completely independent of remote sensing data. This means using remote sensing data allows for reliable validation. This shows us a clear improvement of forest cover reconstructions compared to raw pollen-data. While our optimization approach does use remote sensing data as training and testing data, we separate both in a spatial leave one out validation. This way we are effectively avoiding circularity while also considering potential spatial autocorrelation.

We do concur that uncertainties with Southern Hemispheric reconstructions are high and will exclude those from our evaluation along with samples prior to the 14 ka BP. Moreover, the few non-lake and non-peat records will be removed from the data set as they are unfit for reconstruction with REVEALS. Their removal has no effect on our spatial coverage and general reconstruction results. These adjustments are not only feasible without any problems but will improve the clarity and the quality of the manuscript.

We have added replies to your specific comments below.

Best
Laura Schild and Ulrike Herzschuh

Specific replies

Major issues:

Original Comment

Inadequate regional calibrations: The generalization of RPPEs across broad geographical scales (hemispheres) ignores crucial ecological and bioclimatic regional variations. This approach most likely leads to significant inaccuracies in the vegetation reconstructions, in spite of what the presumed ‘validation’ approach suggests (see below).

Reply

While we agree that a reconstruction using synthesized values will not reflect reality exactly, we still argue for the usability and informativeness of the result of this generalized approach. Continental syntheses of RPP values are standard practice in large-scale reconstructions as they allow an approximation of past vegetation dynamics on a large scale. Notable examples of previously used continental-scale syntheses in Europe include Serge et al. (2023), Trondman et al. (2015) and Pirzamanbein et al. (2014). Githumbi et al. (2022) also synthesize values for Northern and Central Europe and treat only mediterranean records differently. Reconstructions in North America (Dawson et al. 2024) and Northern Asia (Cao et al. 2019) synthesize values on large or continental scales as well.

While we recognize the variability of relative pollen productivity (RPP), we advocate for the use of even hemispheric averages when continental values are lacking. The direction of taxon-specific correction (over- or underproduction of pollen) will generally be correct and provide a vast improvement of REVEALS estimates to using pollen percentages alone while being able to make the most of the data currently available. We highlight that compositional reconstructions using this method come with uncertainties, but are confident that aggregates, such as reconstructed forest cover, are much closer to reality than previous pollen-based estimates. By employing this methodology, our overarching goal is to generate reconstructions that facilitate comparisons across the northern hemisphere while shedding light on general vegetation dynamics. This approach mirrors the methodology utilized in large-scale climate models, where local nuances are necessarily sacrificed for broader insights.

Importantly, this is underlined by our validation which uses independent remote sensing data and demonstrates notable improvements in reconstruction accuracy compared to reconstructions based on raw pollen data. We will expand on this in our reply to an issue below.

Githumbi, Esther, Ralph Fyfe, Marie-Jose Gaillard, Anna-Kari Trondman, Florence Mazier, Anne-Birgitte Nielsen, Anneli Poska, u. a. „European Pollen-Based REVEALS Land-Cover Reconstructions for the Holocene: Methodology, Mapping and Potentials“.

Earth System Science Data 14, Nr. 4 (8. April 2022): 1581–1619. https://doi.org/10.5194/essd-14-1581-2022.

Serge, M. A., F. Mazier, R. Fyfe, M.-J. Gaillard, T. Klein, A. Lagnoux, D. Galop, u. a. „Testing the Effect of Relative Pollen Productivity on the REVEALS Model: A Validated Reconstruction of Europe-Wide Holocene Vegetation“. *Land* 12, Nr. 5 (Mai 2023): 986. https://doi.org/10.3390/land12050986.

Dawson, Andria, John W. Williams, Marie-José Gaillard, Simon J. Goring, Behnaz Pirzamanbein, Johan Lindstrom, R. Scott Anderson, u. a. „Holocene Land Cover Change in North America: Continental Trends, Regional Drivers, and Implications for Vegetation-Atmosphere Feedbacks“. *Climate of the Past Discussions*, 20. Februar 2024, 1–52. https://doi.org/10.5194/cp-2024-6.

Trondman, A.-K., M.-J. Gaillard, F. Mazier, S. Sugita, R. Fyfe, A. B. Nielsen, C. Twiddle, u. a. „Pollen-Based Quantitative Reconstructions of Holocene Regional Vegetation Cover (Plant-Functional Types and Land-Cover Types) in Europe Suitable for Climate Modelling“. *Global Change Biology* 21, Nr. 2 (2015): 676–97. https://doi.org/10.1111/gcb.12737.

Pirzamanbein, Behnaz, Johan Lindström, Anneli Poska, Shinya Sugita, Anna-Kari Trondman, Ralph Fyfe, Florence Mazier, u. a. „Creating Spatially Continuous Maps of Past Land Cover from Point Estimates: A New Statistical Approach Applied to Pollen Data“. *Ecological Complexity* 20 (1. Dezember 2014): 127–41. https://doi.org/10.1016/j.ecocom.2014.09.005.

Cao, Xianyong, Fang Tian, Furong Li, Marie-José Gaillard, Natalia Rudaya, Qinghai Xu, und Ulrike Herzschuh. „Pollen-Based Quantitative Land-Cover Reconstruction for Northern Asia Covering the Last 40 Ka Cal BP“. *Climate of the Past* 15, Nr. 4 (8. August 2019): 1503–36. https://doi.org/10.5194/cp-15-1503-2019.

Original comment

Questionable data assumptions and methodological gaps: The use of northern hemisphere RPPE values for taxa not natively present in the southern hemisphere, such as *Alnus* in Australia, introduces substantial and confusing biases. Presumably, the authors have not consulted the relevant scholars who worked within this field and the geographical areas mentioned. Similarly, defaulting RPP to 1 for taxa without specific data oversimplifies pollen-vegetation relationships. The paper does not adequately address the absence of data for the Southern Hemisphere, leading to a misleading portrayal of global vegetation.

It is suggested >50% of RPPEs are missing for Australia and Oceanic pollen records. So, in this work a decision was made to run these records using the Northern Hemispheric RPPEs, despite very different bioclimatic and ecological contexts. This extrapolation of Northern Hemisphere RPPEs to southern locations missing PPEs without considering ecological or bioclimatic differences is particularly problematic. RPPEs empirically produced using ground truthing work (field surveys and surface pollen collection) were ignored, especially across the Southern Hemisphere (see some references below).

Duffin, K. I., & Bunting, M. J. (2008). Relative pollen productivity and fall speed estimates for southern African savanna taxa. Vegetation History and Archaeobotany, 17, 507-525.

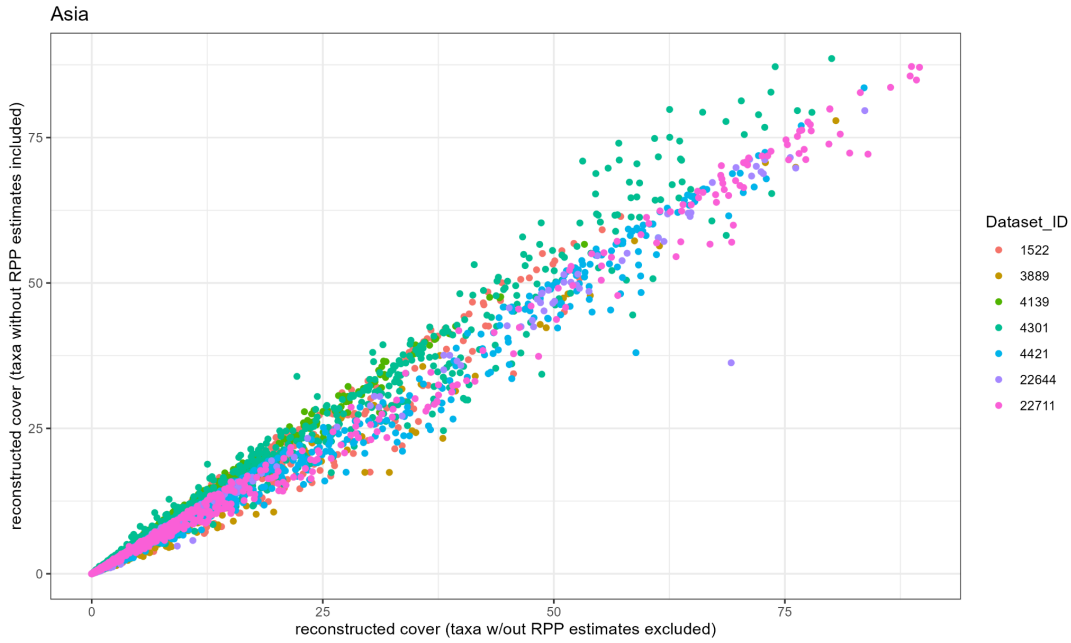
Mariani, M., Connor, S. E., Theuerkauf, M., Kuneš, P., & Fletcher, M. S. (2016). Testing quantitative pollen dispersal models in animal-pollinated vegetation mosaics: An example from temperate Tasmania, Australia. Quaternary Science Reviews, 154, 214-225.

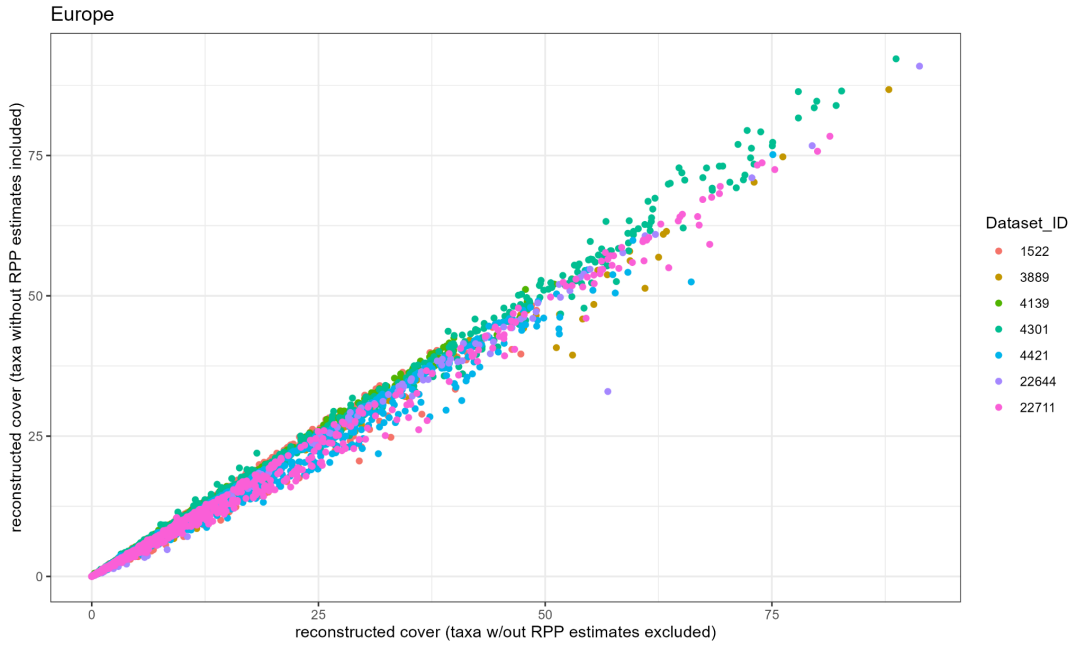
Mariani, M., Connor, S. E., Fletcher, M. S., Theuerkauf, M., Kuneš, P., Jacobsen, G., ... & Zawadzki, A. (2017). How old is the Tasmanian cultural landscape? A test of landscape openness using quantitative land‐cover reconstructions. Journal of Biogeography, 44(10), 2410-2420.

Mariani, M., Connor, S. E., Theuerkauf, M., Herbert, A., Kuneš, P., Bowman, D., ... & Briles, C. (2022). Disruption of cultural burning promotes shrub encroachment and unprecedented wildfires. Frontiers in Ecology and the Environment, 20(5), 292-300.

Reply

We do concur that uncertainties with Southern Hemispheric reconstructions are high due to a lack of regional RPP values and will exclude those from our data set.
We believe that including other observed taxa in the model and setting their RPP to 1, will still result in better estimates of aggregate values such as forest cover than the raw pollen data and therefore apply this standard value to include as much data as possible. Including missing RPP values by setting them to 1 or excluding taxa without RPP estimates leads to relatively similar coverage estimates as indicated by the figures below. Excluding any taxa tends to result in an overestimation of the remaining taxa as the total pollen count is reduced. By including all taxa we aim to account for this.


Asia
reconstructed cover (taxa without RPP estimates included)
reconstructed cover (taxa w/out RPP estimates excluded)
0
25
50
75
Dataset_ID
1522
3889
4139
4301
4421
22644
22711


Europe
reconstructed cover (taxa without RPP estimates included)
reconstructed cover (taxa w/out RPP estimates excluded)
0
25
50
75
Dataset_ID
1522
3889
4139
4301
4421
22644
22711

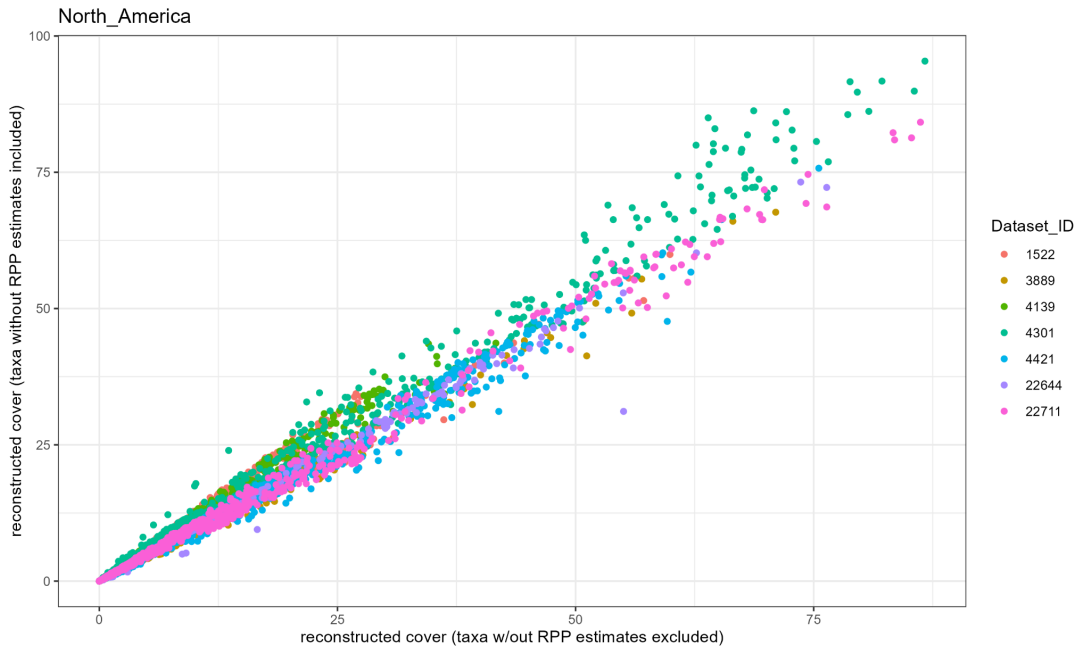


Fig 1: *(newly created)* Comparison of reconstructed cover values for taxa in a reconstruction excluding taxa, for which no RPP estimates are available, and in a reconstruction where all taxa are included and unknown RPP set to 1. **The results are highly correlated, showing that including taxa with unknown RPP does not impact the reconstruction of the taxa for which RPP were already known.**

Original Comment

Oversimplified and incorrect spatial and temporal settings: The inclusion of incorrect basin types in the model without appropriate adjustments is very concerning. Why are marine records included for a model explicitly designed to work for large lakes of closed basins with wind dispersal as the only mechanism for pollen deposition?

The manuscripts states that 'all sites that were not classified as lakes were run with peatland settings' = can we consider the ocean a peatland? REVEALS cannot work with marine records and it definitely does not make sense to apply the 'peatland' settings for marine records with some random arbitrary basin radius (100m?). Further, using a deep temporal scope (50ka) without any consideration for massive climatic shifts (likely larger than the effect of regional RPPEs values vs regional bioclimate variations) are concerning oversights, making any pre-Holocene glacial REVEALS reconstructions unrealistic with current interglacial PPEs.

Reply

We agree with the unsuitability of non-lake and non-peat records and apologize for their inclusion. **We will remove them from our data set.** We realize that many of the peat sites used do not have basin sizes assigned to them. However, peatlands tend to be relatively small and

therefore similar in size, with the mean size of peatlands used in Trondman et al. (2016) being lower than 100m and the average peatland size in the data used by Githumbi et al. (2022) being 716 m (with a rather large standard deviation of 1901 m due to few unrealistically large peatlands). These differences of several hundred meters at most do not influence the reconstruction of REVEALS estimates considerably, which is why a standardization of peatland sizes is appropriate here. Please see Figure 2 below for an example peatland reconstruction using different basin diameters.

Trondman, Anna-Kari, Marie-José Gaillard, Shinya Sugita, Leif Björkman, Annica Greisman, Tove Hultberg, Per Lagerås, Matts Lindbladh, und Florence Mazier. „Are Pollen Records from Small Sites Appropriate for REVEALS Model-Based Quantitative Reconstructions of Past Regional Vegetation? An Empirical Test in Southern Sweden“. *Vegetation History and Archaeobotany* 25, Nr. 2 (1. März 2016): 131–51. https://doi.org/10.1007/s00334-015-0536-9.

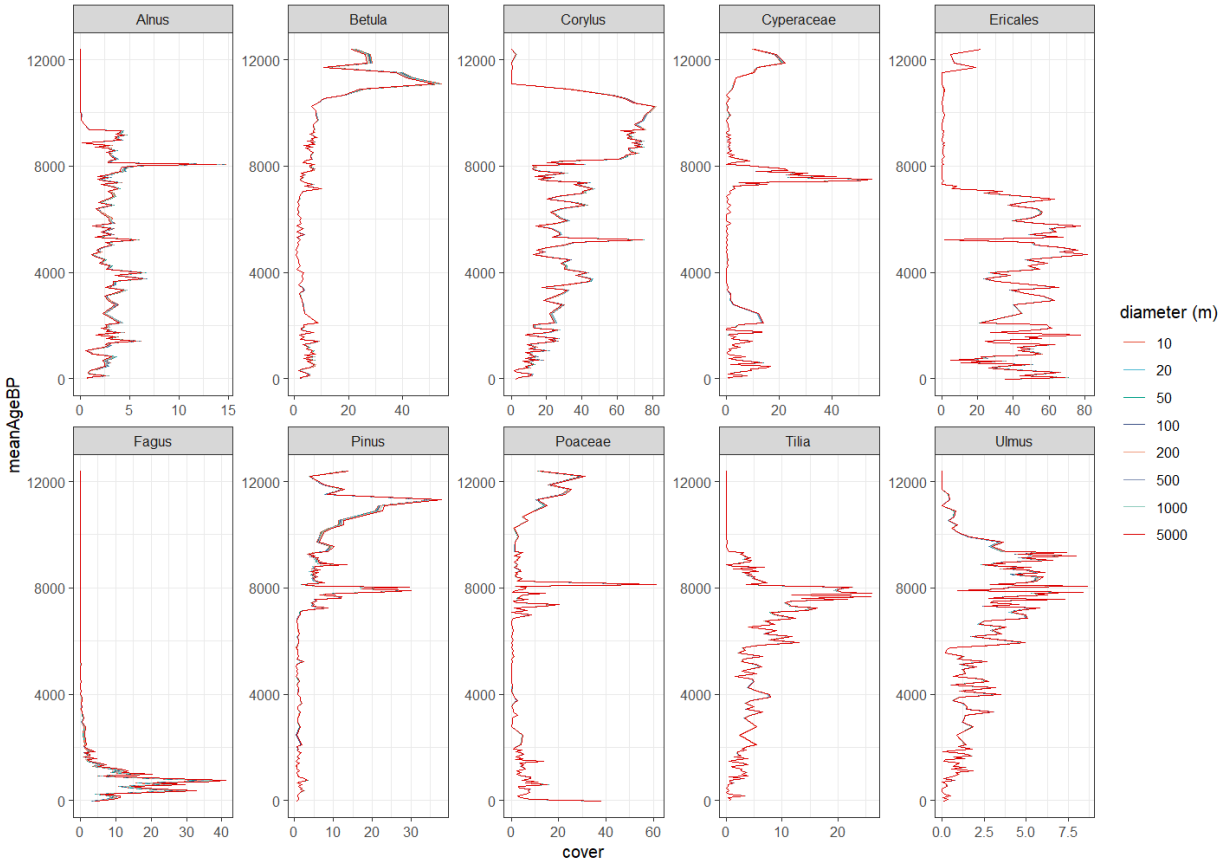


Fig 2: Example peatland site (Ageröds mosse, 13.42774W 55.93448N) with reconstructed vegetation at different set basin sizes (diameter in m). **The basin size has minimal impact on the reconstructed result when using peatlands and can therefore be standardized.**

Original Comment

Dubious optimization and validation: Optimizing PPEs to match remote sensing data risks validating the model based on its own assumptions rather than providing an unbiased estimation of past vegetation, which REVEALS is designed to do. This circular reasoning undermines the scientific integrity of the model's outputs. While an interesting concept this needs to be validated separately on a much smaller spatially and higher resolution scale before such a widespread application. This cannot really be called a ‘validation’.

Reply

We apologize for any confusion that may have arisen here, but **there is no circularity associated with the validation of the REVEALS reconstruction** making use of published syntheses (titled “REVEALS (original RPP)” in our manuscript). The remotely sensed forest cover is independent of the REVEALS reconstruction and has previously been used to validate large-scale reconstruction by Serge et al. (2023) and Pirzamanbein et al. (2014). **Our validations show a clear improvement in forest cover reconstruction compared to pure pollen data.**

We do use remote sensing data both as input data and validation data in the optimization approach. The spatial leave one out validation shown in the manuscript allows us to evade any circularity and spatial autocorrelation here. The optimization is repeated several times and each time one site is left out of the optimization entirely to check the result on this site (“leave on out”). As the sites in the vicinity may be relatively similar, we decide to exclude those as well (spatial buffer) from the optimization, but do not check the result on them. As this is computationally very expensive, we limited our repetitions of this procedure (“folds”) to 100 per continent, which will likely lead to a more conservative estimate of our error. Additionally, our focus lies more on the potential of this optimization method rather than the actual application of these “optimized” RPP values, as we describe in the section on data usability. We believe that this could be useful at potentially smaller scales and when more RPP values are available and are being optimized. As it stands it is more of a proof of concept and not the main contribution of our manuscript.

Original comment

The reconstructed forest cover for the past 500 years was compared to modern remote sensed cover. Why not a smaller and more recent age bin was considered? In the past 500 years many areas of the world have been colonised by Europeans and have experienced major shifts in vegetation structure, as management transferred from Indigenous to colonial regimes (e.g. the Americas and Australia). This means that forest cover over the whole 500 years bin is not comparable to modern remote sensing data. This highlights a Eurocentric view of the global vegetation patterns.

An example of validation of RPPEs using modern vegetation data (with surveys) has been done in the following papers:

Mariani, M., Connor, S. E., Fletcher, M. S., Theuerkauf, M., Kuneš, P., Jacobsen, G., ... & Zawadzki, A. (2017). How old is the Tasmanian cultural landscape? A test of landscape openness using quantitative land-cover reconstructions. Journal of Biogeography, 44(10), 2410-2420.

Mariani, M., Connor, S. E., Theuerkauf, M., Herbert, A., Kuneš, P., Bowman, D., ... & Briles, C. (2022). Disruption of cultural burning promotes shrub encroachment and unprecedented wildfires. Frontiers in Ecology and the Environment, 20(5), 292-300.

Reply

Our choice of 500 year age bins was founded in the aim to include as many records as possible, since not all have samples as young as 100 years BP. We have, however, **tested a smaller age bin for the REVEALS reconstruction** using published synthesis values and found similar validation results. Mean absolute errors will sometimes be slightly larger or slightly smaller but generally similar (Fig 3-5 below) **highlighting the validity of the data set.**

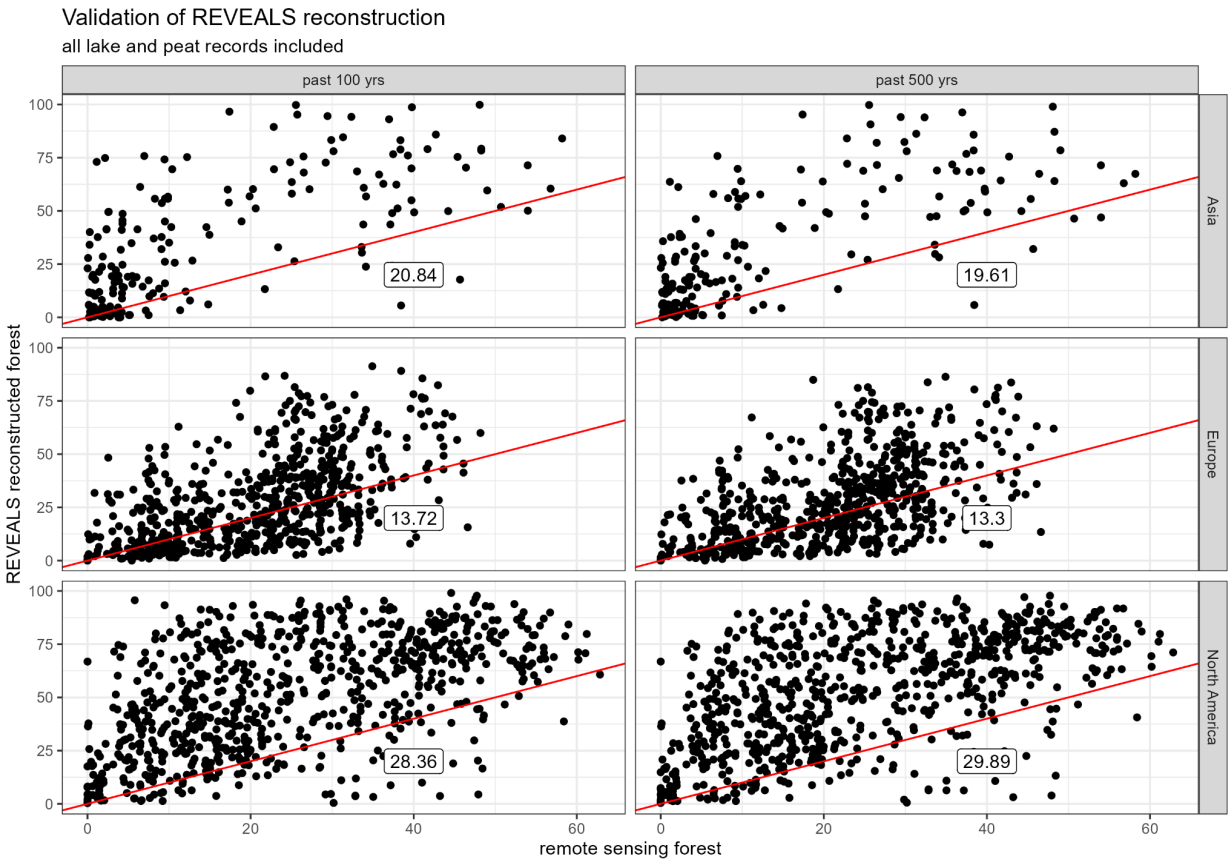


Fig 3: (newly created) Site-wise validation of reconstructed forest cover at all lake and peat records in Asia, Europe, and North America. Two age intervals were used for the calculation of

a “modern” site average forest cover. The label shows the mean absolute error for each subset. **The smaller modern interval changes error values only marginally. These values can be compared to the following two validations, which follow better practice, to see how the calculated MAE was potentially impacted.**

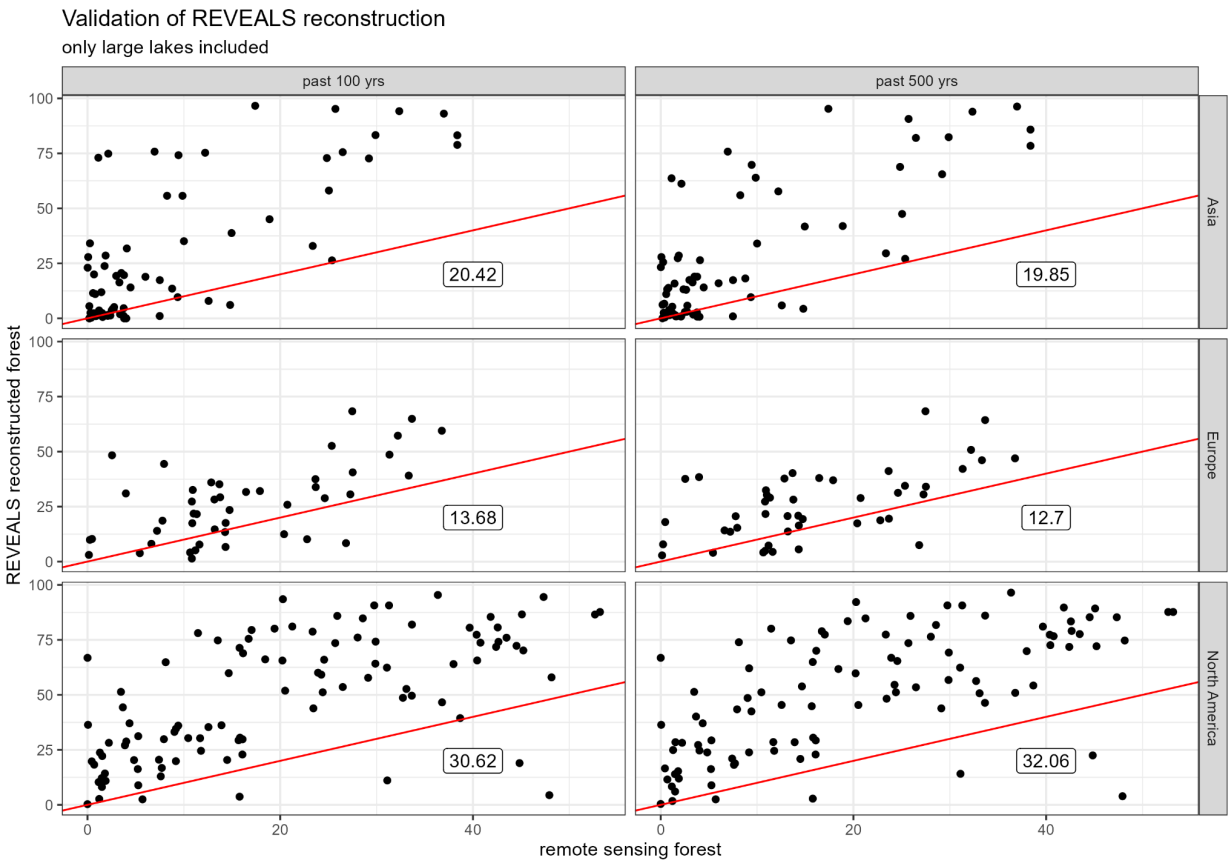


Fig 4: Site-wise validation of reconstructed forest cover at all large lake sites (>= 50ha) in Asia, Europe, and North America. Two age intervals were used for the calculation of a “modern” site average forest cover. The label shows the mean absolute error for each subset. **This site-wise validation even shows slightly smaller errors for European sites. MAE values in Asia and North America are slightly larger. The now smaller sample size could also impact this. Error values are slightly larger in the shorter modern interval with the exception of North America. Removing small sites from the site-wise validation does not impact it significantly.**

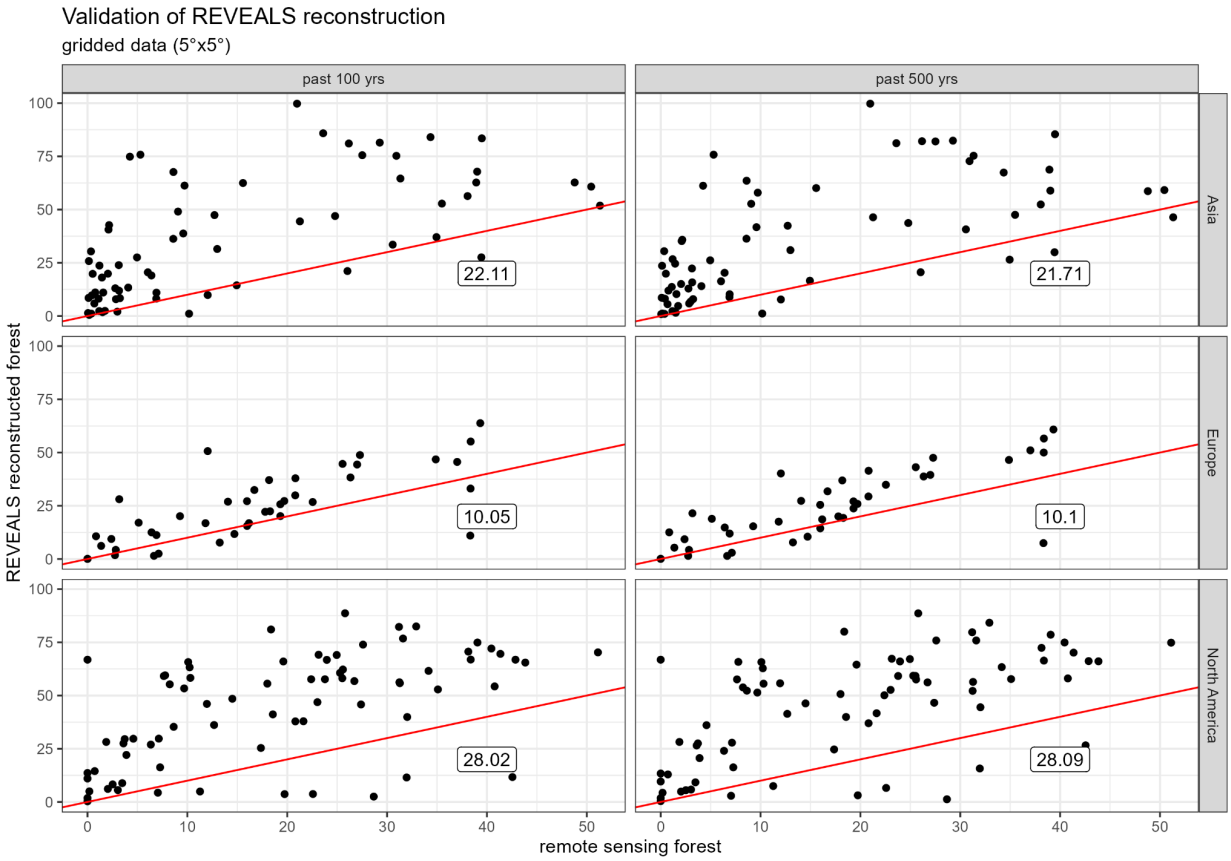


Fig 5: Grid-cell wise validation of reconstructed forest cover in Asia, Europe, and North America. Two age intervals were used for the calculation of a “modern” cell average forest cover. The label shows the mean absolute error for each subset. **This gridded validation shows smaller errors compared to the original site-wise validation (Fig 1) and the corrected site-wise validation (Fig 2). Only MAE values in Asia are slightly larger. Spatial averages of our data set, therefore, do not impact the quality.**