# Review of ”Holocene land cover change in North America: continental trends, regional drivers, and implications for vegetation-atmosphere feedbacks”

The manuscript by Dawson et al. presents new gridded reconstructions of land cover changes in North America, combining pollen-based vegetation cover reconstructions and a Bayesian spatial interpolation model. The new reconstructions are a valuable community effort and will be of great use for future studies of large-scale vegetation changes, land-atmosphere feedbacks, and anthropogenic land use during the Holocene. The maps can serve as boundary conditions for climate simulations and for evaluating Earth system model simulations with dynamic vegetation. Especially the high number of collected records covering the early Holocene is an impressive feature. The paper is well-written and my comments are mostly minor.
We thank Reviewer 1 for their positive and constructive review, and for their thoughtful and detailed comments.

## General comments

1. The Bayesian interpolation methodology is sound and has been established over several studies. Nevertheless, it was originally developed for individual time slices (spatial reconstructions) while the new LandCover6k efforts and hopefully further data compilations in the future aim at spatio-temporal reconstructions. While I don’t think that any adjustments of the interpolation strategy are needed for this study which does not aim at progressing the statistical interpolation methodology, I would appreciate discussing not just limitations of REVEALS but also of the interpolation methodology in Sect. 4.3. Moving from time slice to spatio-temporal reconstructions offers new statistical and data science challenges and opportunities which would be worthwhile discussing. In particular, can you comment on the potential for handling age uncertainties in the reconstruction algorithm, how uncertainties from REVEALS are propagated to the interpolation algorithm, using an actual spatio-temporal interpolation algorithm instead of reconstructing a set of time slices (see my comment below), and testing the impact of the non-uniform distribution of site locations through, e.g., bootstrapping. Do you think that cross-validation experiments in which some portion of the REVEALS reconstructions is left out from the interpolation could be a way to evaluate the spatial (or spatio-temporal) reconstructions?

   The REVEALS-GMRF interpolation approach used here is consistent with the approach used in the land cover reconstructions for Europe and China (e.g., Githumbi et al., 2022a and Li et al., 2023, respectively) . The process operates on individual time slices, and the interpolation approach itself does not have a temporal component. We will clarify this in the text as needed. We are in the process of developing an approach that includes a temporal component; the challenge is adding this complexity in a way that allows the approach to still be computationally tractable. We recognize that sample age uncertainty may influence results. The methods used in this work do not account for this age uncertainty. Given the temporal grain of the time bins (500 years throughout the Holocene, except for the last 700 years, 100-350 years), we expect that within uncertainty the majority of sample ages will not shift among time bins. When a temporal component is added to the REVEALS-GMRF approach, we will consider how to account for this uncertainty.

REVEALS does quantify standard errors associated with relative abundance estimates. These standard errors are not considered by the REVEALS-GMRF approach. This means that uncertainty estimates from the REVEALS-GMRF approach quantify the uncertainty determined by the variability of REVEALS fractional land cover around the estimated spatial field.

We agree that validation experiments would be useful to understand the impacts of sample unevenness. Validation experiments have been done using the REVEALS-GMRF approach for Europe (Pizamanbein et al., 2018). In that effort a 6-fold cross validation technique where 10, 3 and 1 random selections of the block were left out. This work generally showed little change in the predictions when data was withheld.

The challenge in repeating these experiments for North America lies in the computational burden associated with the REVEALS-GMRF approach. Future work aims to improve computational efficiency of the approach, at which point such experiments will be possible.

We will add further discussion of uncertainty handling in the main text.

2. The GMRF method provides reconstruction uncertainties for all grid boxes. However, so far, the uncertainties are only visualized for the continental and regional mean curves. To understand the statistical significance of the spatial land cover variations, it would be very helpful to also plot maps of the reconstruction uncertainties, for example together with Fig. 2. For the change maps (Fig. 4a, 5a, 6a, 7a), hatching areas with statistically significant changes would be valuable to assess the importance of the temporal changes.
   To address this comment, we will add maps of the reconstruction uncertainties to the Supplementary Information.

3. There are three other aspects related to the applications and improvement of the datasets that I kindly ask the authors to discuss or enhance the respective discussion. The authors mention the under-representation of arid regions due to a lack of pollen records. Do you see prospects for including other proxies, either for vegetation or for (hydro-)climate to improve the reconstruction for arid regions (e.g., biomarkers, isotopes from speleothems, or lake levels)? Currently, only three land cover types are separated. Is there the potential in terms of data availability to also separate boreal, temperate, and subtropical forest / grassland types in addition to evergreen, summergreen, and open land? Finally, the current separation into time slices of 1kyr (and potentially further

smoothing from 1 age uncertainties) precludes the analysis of sub-millennial trends. Do you see the possibility to also identify multi-decadal to multi-centennial variations on the regional and continental scale with the existing data coverage, potentially using an improved spatio-temporal interpolation methodology?
These are all excellent points and we plan to add a short review of them to the discussion, when we discuss future work and next steps. First, we agree that other proxies could be used in arid regions to improve vegetation reconstructions. Second, we agree that it would be possible to separate the land cover types into boreal, temperate, and subtropical components, given what is known about the individual pollen types and their climatic affinities. However, statistically modeling more finely resolved land cover groupings or taxa is complicated by the many 0 count observations in more finely resolved groupings, especially for such a large spatio-temporal domain. Third, whether centennial-scale variations can be confidently interpolated to a continental extent is a more open question, given the relative scarcity of sufficiently well-sampled and well-dated records. However, it should be possible to study multi-decadal to multi-centennial variations in vegetation cover at regional to continental scales, by carefully selecting records with the highest sampling resolution and age constraints.

4. The authors discuss implications for biophysical atmosphere-vegetation feedbacks very well (e.g., l.49-53, Sect. 4). In this context, it would be suitable in my opinion to also mention biogeochemical feedback mechanisms as the new land cover reconstructions should also be a useful tool for studying these ones, in particular since more and more Earth system models having capabilities for prognostic carbon and nutrient cycles.
We plan to add a short section reviewing biogeochemical feedback mechanisms, with an emphasis on carbon cycle implications.

5. Data availability: The posterior mean reconstructions have been made available as csv files through a github repository. In the interest of maximing reusability and making it easy to cite the dataset, it would be very valuable to (i) publish the data sets also in a FAIR repository with a permanent identifier, and (ii) publish the reconstructions as netCDF files which are more suitable for gridded data and allow for a better interoperability with climate and vegetation simulations. Additionally, I would recommend to make not just the posterior means but also the uncertainties available in a suitable data format, either as marginal (point-wise) uncertainties or, better, by publishing MCMC samples which allow quantification of spatially correlated uncertainties.
We will make the reconstructions publicly available in a NetCDF format in a FAIR-compliant repository such as PANGAEA or DRYAD, for both the means and uncertainties. Reviewer 2 raised the same point.

## Specific comments:

Regarding all maps in the manuscript, please consider using a different projection that displays the size of regions better since in the current projection the high latitudes are heavily overrepresented compared to, e.g., Mexico.

We will change the projection to Albers equal area with standard parallels 33.333N and 66.667N and center point of 57N, 100W.

l. 33: I kindly ask you to use the term ”Holocene temperature conundrum” instead of just ”Holocene conundrum” given the number of other conundrums that have appeared in the literature over the last decade(s).
Will do.

l. 57: Would it be suitable to mention not just land cover dynamics in the last sentence of the abstract but also Holocene climate dynamics?
Agreed, will do.

l. 63: Should it be ”net-negative” instead of ”negative-net”?
Will do.

l. 114-117: Is there a specific reason why there was no prior continental-scale reconstruction for North America?
We will add a bit more historical context.

l. 163: The authors use Bchron for the age modeling which is an appropriate choice in my opinion. I'm just curious if there is a specific reason to not use BACON which seems to be used more often in recent studies? While both model would be justifiable choices from my outsider view, it could be of interest to the community if the authors see specific advantages of Bchron for the vegetation reconstructions at hand.
Both BACON and Bchron are widely used and follow very similar Bayesian modeling frameworks. We have used both in our prior work. We have found that Bchron establishes increasing uncertainty in age estimates between control points and requires fewer assumptions about sedimentary processes than Bacon. We will add some text (and possibly a supplementary figure) to support this decision.

l. 168-169: Given that the taxa selection seems very important for the vegetation cover reconstructions, can you provide some more information on the selection criteria and the representativity of these taxa for the vegetation at the different locations? In addition, I would strongly suggest to move Supplementary Table 2 to the main manuscript (potentially with some additional information on the importance of the taxa like the average pollen percentage of those taxa).
We will add information about why we chose the selected taxa. Much of the pollen-parameter information in Supplementary Table 2 has been published elsewhere, so we tend to think that it better belongs in the Supplementary Information, but we can promote it to the main text if the Editor so requests.

l. 192-207: Can you provide the number of lakes used in the different workflow steps?
Yes, we will add this.

l. 218: Can you provide some support, e.g., in the supplement, for the very strong statement that ”these vegetation reconstructions indisputably overrepresented larch”?
We will remove the word ‘indisputably’ from the revised ms. Larch is known to be a challenging taxon to model accurately in pollen-vegetation models, because it prefers wetland settings (at least towards the south of its range) and its large pollen grain does not disperse far and preserve well. Hence, there is a general challenge of differentiating the local populations of larch growing at the site where a core was retrieved from those occurring across the broader source area. And the scarcity of larch pollen grains means that the REVEALS reconstructions for larch are highly sensitive to the pollen-parameter settings for larch (i.e. PPEs and fall speed). When revising the text, we will add these points to either the main text or supplement.

l. 226: Please remove the gray shading in Trondman et al. (2015). 2
Will do.

l. 260-261: I don’t understand the rationale behind the ”mean relative cover”. Excluding ice covered areas and areas without reliable reconstructions is reasonable, but why would you not use the area weighted mean of all grid boxes with reconstructed vegetation instead of just the average over those grid boxes?
There are two different effects that we are trying to capture with these two metrics. If we only do the area-weighted mean in which the unglaciated land area for a given time period is the denominator (and this is one of our two metrics), then this metric helpfully describes the proportional mix for any given time period, but will not capture the total increase in vegetated land area across time intervals, as North America deglaciates. In the second vegetation metric, we set the denominator to the total deglaciated land area at 0.25 ka. This second metric, by using a constant denominator that represents late-Holocene deglaciated land area, captures the general increase in vegetated land areas and of individual components.

The manuscript text on this topic could be clearer and we will work to revise it; we will also develop better and more precise terminology for these two metrics.

Fig. 1: To understand the spatio-temporal coverage properties better, can you provide maps similar to Fig. 1a for the individual time slices in the supplement?
Yes, we will add maps of site coverage for individual time slices to the supplement.

l. 332: I struggle to connect the mentioned increase from 56% to 91% with the results presented in Fig. 2 where the increase looks much smaller. Can you please clarify where these numbers come from?
This is an error in the text, caused by a failure to update results from an earlier version. Thank you for catching the mistake. We will fix the text.

Fig. 4c: Spruce and pine are prominently mentioned in the text (l. 374-378) but are not included in the figure with the coverages of important taxa. Is there a reason for this exclusion?
We will revise Fig. 4c to add spruce and pine, but may need to make further adjustments to preserve figure readability and information density.

Sect. 4.1 and 4.2.1 are fairly long considering that they are mostly a literature review while being relative unconnected to results from the new reconstructions. Therefore, I'd ask you to either state more explicitly how the new results are in agreement with / contradicting previous studies or consider shortening this part given that the paper is already rather long. If the goal is mainly to state potential applications of the new reconstructions, I don't think this long discussion of the previous literature is needed.
We will look for ways to shorten this section and better connect to results.  We may move some of this to Discussion.

Sect. 4.1.1 - 4.1.3 should be 4.2.1 - 4.2.3.
We will fix this.

l. 784: It is stated that the GMRF creates a spatio-temporally complete vegetation reconstruction. Is this an appropriate characterization? From my understanding of the methods section and previous studies using the GMRF method, it creates spatially complete reconstructions for a set of predefined time slices but without considering temporal dependences between the time slices. If this is an misunderstanding on my site, it would be helpful to state more explicitly in the methods section how the time dimension is handled in the GMRF since the referenced studies only apply it for spatial reconstructions.
You are correct with your understanding of our methods. We will amend the text to clarify this point.

l. 843-845: Maybe consider simplifying or splitting this sentence.
Will do.

l. 861: Is data assimilation the appropriate word here? From my understanding, the reconstructions would either be used as boundary conditions in simulations or for comparison with dynamically simulated vegetation, whereas data assimilation would refer to a dynamic simulation in which the simulated values would be relaxed towards the reconstructions. The latter is something that hasn't be done so far with vegetation reconstructions as far as I know.
The word choice of data assimilation was intentional and the goal was to point to future research directions.  Agreed that no data assimilation efforts have yet been done with pollen-based vegetation reconstructions and Earth system models, but we view this as a future potential  application of this vegetation reconstruction and identical/similar ones from other continents (existing and to come).

In many instances (e.g., l. 129, 196, 214, 216, 252), the parentheses in citations are set inconsistently. I kindly ask you to check the citations throughout the paper for typesetting errors.
Will do and apologies for these formatting errors.

I kindly ask you to check the reference list again and harmonize the used reference style, in particular by providing DOIs consistently wherever available. Additionally, Dawson et al. 2019a

and Dawson et al. 2019b seem to be the same reference, and Githumbi et al. 2022b and Githumbi et al. 2022c also seem to be the same reference. 3
Will do.  We'll both track down DOIs and remove these duplicates.