Reviewer 1, 2nd review

In my opinion, the revised version of the manuscript "Climate field reconstructions for the North Atlantic region of annual, seasonal and monthly resolution spanning CE 1241-1970" by Jesper Sjolte and Qin Tao, did not sufficiently adress the concerns of the first round of reviews and therefore requires additional corrections and improvements before it can be published.

Thank you for again taking the time to review our manuscript. We hope that our revisions will meet your expectations this time. In particularly, we have expanded the comparison to climate field reconstructions to include the Last Millennium Reanalysis in seasonal and annual resolution, as well as expanding (and correcting) the comparison to the ModE-RA reanalysis.

Main points:

I still believe that the monthly reconstruction is of low quality as the authors mention themselves in the comparison with the instrumental time series. Therefore it should be removed from the title, mentioned in the text much less prominently and be discussed more critically.

We have removed “monthly” from the title and only use the monthly reconstruction as an investigation of the seasonality of the reconstruction and an option to compare to monthly observations and reconstructions.

One reason for this are probably misunderstandings about existing data sets:

The authors write: "As the ice core records are dated using a combination of layer counting and volcanic horizons an entire ice core record is never shifted by one year"
This is a little overconfident as long as eruptions dates and locations are still partly unknown and constantly updated, e.g. https://doi.org/10.1016/j.jvolgeores.2025.108346

We have softened up this formulation from ”never shifted” to ”unlikely to be shifted”. The eruptions used for synchronization in both the GICC05 chronology (Vinther et al., 2006) and the GICC21 chronology (Sinnl et al., 2022) are historically identified with known sources and timing of eruptions, and Sinnl et al. (2022) showed that GICC05 (used for the Vinther et al. (2010) data) is still valid back to CE ~1200. We have now explained this in the text L90-95.

The authors seem to have missed that ModE-RA includes annually resolved proxy data as input, too: The statement in line 287 and 318ff: "This might be due to fewer data constraining ModE-RA" is incorrect. ModE-RA assimilates more records than this study in any time period. Especially in case of winter over continental Europe ModE-RA assimilates monthly historical data and early instrumental observations while there is little seasonal proxy information used in this study. The authors only mention their poor monthly skill in the comparison with instrumental data (line 280ff).

Unfortunately, there was an error in our comparison with the ModE-RA reanalysis, due to a coding typo in the re-gridding of the ModE-RA data for the comparison. After correcting this error, the comparison shows better correlation between our reconstruction and ModE-RA

(Figure 10). Although this doesn't change the main conclusions of the comparison, we hope that the results are more convincing now.
The ModE-RA reanalysis uses the Iso2k database which features a subset of the Vinther et al. (2010) ice core data. For example, the Iso2k entry for the Dye-3 data only has *annual* data back to 1778 (we checked the Iso2k meta data) (also indicated in Figure 3a, Valler et al., 2024). Iso2k features two seasonal resolution Greenland ice cores spanning the whole ModE-RA reanalysis (GRIP and Crete, apparently representing DJF, which doesn't match the seasonality defined by Vinther et al. (2010)). We use 8 ice core records in seasonal resolution spanning the entire time frame of our reconstruction, including records from two Dye-3 cores (Table 1). In our experience it is the ∂^{18}O gradient between ice core sites which provide the best constraints on a climate field reconstruction (Sjolte et al., 2018) (L159-160), not local correlation between ice core ∂^{18}O and temperature.
The sites of long-term documentation data used in ModE-RA is concentrated in Central Europe (Figure 3b, Valler et al., 2024), which does not provide optimal constraints on NAO-type variability for the winter season, during which we see the largest discrepancy between ModE-RA and SAT25. In the expanded analysis of correlation to climate field reconstructions (section 3.5) we see a quite stable relation to the Last Millennium Reanalysis through time (annually and seasonally), while the correlation to the ModE-RA reanalysis clearly drops going back in time, especially for winter. This could indicate that the ModE-RA reanalysis does not capture the extra-tropical winter variability in the North Atlantic region in the time period before CE ~1800. Furthermore, the long-term variability of ModE-RA is determined by the model prior due to 71-year assimilation window (Valler et al., 2024), which could also be a factor.
We now discuss the points above in more detail in our revision (L372-388).

Hence, I already suggest in the first review a validation against gridded monthly instrumental data sets for temperature and precipitation in the 20th century. The authors use 20CR for validation, which is a reanalysis only constraint by SLP and SSTs. Therefore, outputs like precipitation are not the ideal validation data sets.

We did compare the 20CRv3 to gridded precipitation and found them comparable (as we noted in the data section). However, to accommodate this comment we have now included a comparison to seasonal gridded HadCRUT5 temperature and GPCC precipitation data (Figure S13) as well as monthly gridded HadCRUT5 temperature (Figure S15). Our reconstruction has higher correlation to the gridded HadCRUT5 temperature data than to 20CRv3 temperature.

Although I missed to mention this in the last round of reviews, I am surprised that no other proxy-based reconstructions are shown or discussed, e.g. PAGES2k or the Last Millennium Reanalysis, which also exists in a seasonal version.

As written above, we have now included a comparison to the Last Millenium Reanalysis in seasonal and annual resolution. See section 3.5 in the revision.

Finally, the entire paper is very much focused on the reconstruction data set. However, I miss some discussion of the results. Why has the Qaqo timeseries hardly any variability in the reconstruction in contrast to observations? What happened between 1770 and 1790 in Fig.

5? This looks suspicious. Is it some reconstruction artefact or do you have a climatological explanation? Please add some discussion.

There is a model bias in JJA variance (see supplementary Figure S11) causing the variability in Southern Greenland to be underestimated. We now explain this in the text (L288-290).
We are not sure exactly what the reviewer is referring to with regards to the interval 1770-1790. There is an oscillation in temperature out of phase between Greenland and Europe, which points to NAO-type variability. The reconstruction follows the temperature observations fairly well in this interval, so it is not an artefact. The Laki eruption was 1783. High latitude eruptions tend to be followed by negative NAO (Sjolte et al., 2021; Tao et al., 2025), which could explain the specific variability in this interval. However, attribution of external forcing goes beyond the scope of this study.

Minor suggestions:

I would also prefer marking/hatching the significant grid boxes in the map in an IPCC style rather then just writing the number of significant grid boxes on top.
The reviewer may have missed the contour line for p = 0.01 in our correlation maps. Hatching can make it more difficult to read the results and we prefer the contour.

I would also add the time period for which the correlation is calculated in the captions of fig. 2 and 3.
Corrected.

I agree with the second reviewer that it is difficult to understand the methodology of how the different versions of the reconstruction are generated. I suggest adding a figure with the workflow.
We have added a new Figure 2 illustrating the workflow of the reconstruction method.

Fig. 5: I miss information in the caption if this time series shows monthly values, seasonal means or annual means.
It does say DJF in the figure caption. We have changed the captions to make it more clear what data we are showing.

Line 233: This should only reach the same level of skill throughout time if dating and proxy uncertainties would not increase further back in time. I guess that is what you mean by "in theory".
Our reconstruction is specifically designed to be consistent through time as we explain in the data section. We now provide additional information on the ice core chronology and sample replication for tree-ring data in the method and data section.

line 344: "the any long-term information" ?
Corrected.

Check references for consistency, e.g. all letters are capitalized in Johnsen, S.J. and is it really required to have an doi and an _eprint?

We have checked the reference list for consistency. We use the citation style in the Latex package provided by Climate of the Past.

Reviewer 3

This study by Sjolte and Tao presents a North Atlantic climate field reconstruction (CFR) of 2m temperature (T2m), sea surface temperature (SST), sea level pressure (SLP), and precipitation at annual, seasonal, and monthly scales over CE 1241–1970. Their approach uses an analogue reconstruction method to assimilate a small but carefully curated proxy network consisting of Greenland ice-core $\delta^{18}O$, tree-ring width (TRW), maximum latewood density (MXD), blue intensity (BI), and tree-ring cellulose $\delta^{18}O$, combined with an isotope-enabled climate model prior. The authors validate the reconstruction against reanalysis products and observations, and claim that their method shows promise when the proxy network is small but of high quality.

This work appears to be primarily methodological, as it does not present major climate science conclusions; therefore, my comments focus on the methodological aspects. The current mainstream of CFR relies on Bayesian data assimilation (DA) frameworks (e.g., EnKF), which explicitly account for uncertainties in both proxy observations and model priors, yield mathematically optimal reconstructions, and enforce a degree of dynamical consistency. Such approaches have been shown to be quite successful. Given this context, it is important for the authors to more clearly justify their choice of method. The approach adopted in this paper has at least two limitations: (i) proxy errors are not treated probabilistically, and (ii) the weighting function used to combine analogues is heuristic rather than statistically optimal. The authors therefore need to demonstrate that the analogue method exhibits at least competitive performance relative to Bayesian DA approaches. While the comparison with ModE-RA is helpful, the current presentation is very brief, lacks detailed discussion, and does not constitute a clean methodological comparison. Similar points have been made by previous reviewers but the revision is still weak in addressing such comments.

We thank the reviewer for the insightful comments, which has helped us improve our manuscript.

Justification of method.

The analogue method has been tested with pseudo-proxy experiments and shown to be suitable for reconstructions of the common era (Goméz-Navarro et al., 2017). In previous work the NAO extracted from climate field reconstructions based on the analogue method, has been shown to have similar performance compared to NAO reconstructions based on statistical methods (Sjolte et al., 2018), while providing the benefits of a consistent climate field reconstruction with several reconstructed climate variables, which has applications to, for example, forcing attribution studies (Sjolte et al., 2018; 2021; Tao et al., 2025). The choice of method is now discussed in the revision L399-407.

Extended comparison to climate field reconstructions.

As also written in the reply to Reviewer 1, there was an error in our comparison with the ModE-RA reanalysis, due to a coding typo in the re-gridding of the ModE-RA data for the comparison. After correcting this error, the comparison shows better correlation between

our reconstruction and ModE-RA. Although this doesn't change the main conclusions of the comparison, we hope that the results are now more convincing.
We have expanded the comparison to climate field reconstructions to include the Last Millennium Reanalysis in seasonal and annual resolution, as well as expanding the comparison to the ModE-RA reanalysis (Section 3.5).
We also discuss the different properties of each of the reconstructions, as well as comparing them in terms of performance, selection of proxy data, spatial coverage and temporal range (L359-407).
A true method comparison would require testing a range of reconstruction methods with the same proxy data collection and model prior. While such a study would be very interesting, it is beyond the scope of this work.

As a methodological paper, I believe the manuscript would benefit from a clearer justification of the chosen approach and a more thorough methodological evaluation (e.g., through a pseudoproxy experiment), comparing to the mainstream methods, to meet the standards for publication.

We hope that the added explanations on our choice of method and new results of the expanded comparison to climate field reconstructions have improved our manuscript, and that it now meets the expectations of the reviewer.

References.
Gómez-Navarro, J. J., Zorita, E., Raible, C. C., and Neukom, R.: Pseudo-proxy tests of the analogue method to reconstruct spatially resolved global temperature during the Common Era, Clim. Past, 13, 629–648, https://doi.org/10.5194/cp-13-629-2017, 2017.

Sjolte, J., Adolphi, F., Guðlaugsdòttir, H., & Muscheler, R. (2021). Major differences in regional climate impact between high- and low-latitude volcanic eruptions. *Geophysical Research Letters*, 48, e2020GL092017. https://doi.org/10.1029/2020GL092017

Tao, Q., Shen, C., Muscheler, R., and Sjolte, J.: Distinct winter North Atlantic climate responses to tropical and extratropical eruptions over the last millennium in PMIP simulations and reconstructions, Clim. Past, 21, 2561–2578, https://doi.org/10.5194/cp-21-2561-2025, 2025.