Response to comments by reviewer 1

Thank you for your detailed comments. The responses to each comment are provided below.

Comment: My primary concern is the notable discrepancy between proxy records and transient model simulations in North America and North Africa. As the authors noted, while proxy records suggest divergent trends in these regions, the climate models indicate a common forced response. Since data assimilation depends on modelled covariance to distribute local information from assimilated variables across space and other variables, I question the effectiveness of these models for reconstructing global hydroclimate over the Holocene. The authors attempt to address this issue by applying a localization radius, but this adjustment may limit some of the advantages of data assimilation in understanding hydroclimate variability.

Response: As noted in the text, we agree with this concern. However, we argue that the effectiveness of models to capture covariance relationships is scale dependent so that global climate relationships are more difficult to accurately simulate than regional ones. This is likely more true of precipitation than temperature. We highlight the advantage of using data assimilation with a multi-model prior which provides more confidence when the two independent simulations agree (e.g., Line 402). Additionally, the use of covariance localization is a common feature in data assimilation (e.g., Osman et al., 2021) to ensure that local proxies are favored over remote proxies when the modeled covariance relationships may be imperfect. Data assimilation also has an advantage over proxy-only reconstructions as it provides a method for considering site specific relationships between lake status and precipitation so that in regions where the two variables covary strongly, the model prior is updated by a greater amount.

Comment: Regarding the mechanisms behind these trends, the authors mainly focus on the Northern Hemisphere, particularly North America. They also mention wetter conditions in southeastern Australia during the mid-Holocene. Many records in mid- and high-latitude regions often reflect changes in wind direction and intensity. The authors could enhance their analysis by exploring the potential drivers of the wetting trend. Specifically, do the authors identify a particular influence of atmospheric circulation variability in this wetting trend?

Response: We focus our discussion for this topic on western North America because it remains an important topic of open discussion in the community and there is substantial independent data for comparison. Sect. 3.7 explores this topic and we invoke SLP anomalies in the North Pacific and resulting changes to atmospheric circulation strength as a likely source of the wetting trend (Fig 10). For Australia we are limited in the available independent proxies to validate results for this region. However, we note that the pattern of the wetting in southern Australia (Fig 7d-f) is potentially consistent with a northward position of the southern westerly winds. This could represent the extratropical response to a northward shift in the position of the ITCZ (Cheng et al., 2012; van der Bilt et al., 2022). We did not consider these results robust enough to include in the manuscript. Future research to integrate isotope proxies from New Zealand with currently unavailable isotope enabled simulations would provide more detailed insight into the dynamics of this region.

References

van der Bilt, W.G.M., W.J., D'Andrea, L.T., Oppedal, Bakke, J., Bjune, A. E., Zwier, M.: Stable Southern Hemisphere westerly winds throughout the Holocene until intensification in the last two millennia, Commun Earth Environ, 3, 186, https://doi.org/10.1038/s43247-022-00512-8, 2022.

Cheng, H., Sinha, A., Wang, X., Cruz, F. W., and Edwards, R. L.: The global paleomonsoon as seen through speleothem records from Asia and the Americas, Clim. Dyn., 39, 1045–1062, https://doi.org/10.1007/s00382-012-1363-7, 2012.

Osman, M.B., Tierney, J. E., Zhu, J., Tardif, R., Hakim, G. J., King, J., C.J. Poulsen: Globally resolved surface temperatures since the Last Glacial Maximum. Nature, 599, 239–244, https://doi.org/10.1038/s41586-021-03984-4, 2021

Response to comments by reviewer 2

Thank you for these detailed comments. The responses to each comment are provided below.

Comment: The lake status data are mainly from the Oxford Lake Status (OLS) Databank (n = 98; Street-Perrott et al., 1989), which is really old. Why not the newer one, the Global Lake Status Data Base (GLSDB: Kohfeld and Harrison, 2000, Harrison et al., 2003)(https://pmip2.lsce.ipsl.fr/synth/lakestatus.shtml)? Even though, it is still a little old and there's some controversy on the chronology and proxy used. In Africa and Australia, new compilations have been completed (Cort et al., 2021 QSR; Clerke, 2022 Hydrological Regime of Australian Lakes Over the Late-Quaternary and Holocene). And in Americas and Asia, I'm sure there are more continuous lake level records (Lowry and Morrill, 2019; Li and Zhang, 2020).

Response: Thank you for your detailed list of relevant data sources. Some of these were previously considered but not included for a variety of reasons. One major constraint is that many of these papers (Harrison et al., 2003; Lowry and Morrill, 2019; Li and Zhang, 2019) focus on only the LGM and/or mid-Holocene and therefore only publish lake status values for these narrow time periods. The PSM used in this study requires transient data to compare relative (percentile) changes through time, and the limited published data are therefore insufficient. We also note that using time slice simplifies the designation for "higher" or "lower" as it requires less precise age control; within a single time slice, values can span multiple millennia. Additionally, many of the referenced publications are primarily based on pollen or geochemical interpretations (e.g., Lu et al., 2015; Moreno and León 2002; Sun et at al., 2013; Vélez et al., 2003) which we exclude (Line 125), or the data do not have sufficient duration or resolution to meet our criteria for inclusion in the dataset (Line 131). We evaluate each of the listed references to determine if the publications provide relevant data for our study.

Kohfeld and Harrison, 2000 and Harrison et al., 2003 data available at https://pmip2.lsce.ipsl.fr/synth/lakestatus.shtml reports time slice lake status values for the mid-Holocene and LGM. Although the "core of [GLSDB] is the Oxford Lake-level Data Base" (Qin et al., 1998), the GLSDB (677) nearly doubles the number of lakes within the OLS (360 lakes) mainly by "improvements to the coverage temperate and wet tropical regions…achieved through the incorporation of information from four regional data bases" (Qin et al., 1998). We tracked down three regional databases and found 74 records which match our stated criteria from the former Soviet Union and Mongolia, (FSUDB: Tarasov et al., 1996; n=38); European (ELSDB: Yu and Harrison, 1995; n=16), and Chinese (Yu, 2001; n=20) lake-level data bases. We could not identify data for the African Lake Status Data Base (Jolly et al. 1998), but many of these data would likely be superseded by the new De Cort et al. (2021) compilation.

De Cort et al. (2021) was overlooked because the composite records cannot be used in the data assimilation framework. However, by looking at the original data, we identify 19 sites in eastern and southern Africa which meet our criteria.

Clerke et al. (2022) was not published at the time of data collection. Thank you for bringing it to our attention. There appear to be 5 lakes which fit our criteria.

Lowry and Morrill (2019) and Li and Zhang (2020) report time slice lake status values for the LGM, but we searched the publications for each of the 44 and 37 sites respectively to identify records which we may have overlooked. We found 1 site (Uyuni Basin, Argentina) which we will add to the database.

In total, we identified 99 new records that we will add to the dataset, primarily from Eurasia. Some of these replace previously included data, so the updated dataset will include 216 total sites. A discussion of these sites is provided in Sect. 2.1.1 Lake status data (Lines 131-179). We have also updated the assimilation, created a new reconstruction, and revised the text accordingly (Sect. 3 Results and discussion). Upon evaluation of the results, we do not find any major changes that affect our conclusions.

Comment: Are TraCE-21k and HadCM3 the only available transient simulations? What is the difference between these two models, such as the temporal resolution and reconstruction skill. I think they should be compared before combined.

Response: Transient GCM simulations for our time-scale of interest are rare, and we are unaware of additional simulations run with different models. Precipitation values for the two simulations are compared by Hancock et al. (2023) which found that reconstruction skill, as assessed by agreement with proxy values, varies regionally. Results assimilating a model prior sampled from only one model would more closely resemble the model values shown in Fig. S6. Data for both models is available at annual temporal resolution, but we use multidecadal (50 year) mean values for the model prior (see line 243).

Comment: The conclusion is too long.

Response: Agreed. We have made conclusion more concise (lines 254-674).

Comment: Line 86: please add “(details in section 2.4)”.

Response: We have included this reference (now Line 87)

Comment: In figure 2(b), What is the purpose of removing the Holocene mean value? And how do you deal with the negative values?

Response: Fig 2 shows the preprocessing steps applied to the proxy values. For data assimilation, values are often converted to anomalies prior to assimilation (e.g. Erb et al., 2022), and we found that our multi-model prior necessitated this step. The lake status percentiles were assigned between 0-100 for each model, but the precipitation values do not have a bias correction. We found significant mean offsets between the TraCE-21ka and HadCM3 which impacted the covariance structure in the assimilation. Therefore, it was necessary to convert the data to anomalies to accurately reconstruct precipitation values. Negative values are not an issue at this step because both the proxy values and the model prior are both converted to anomalies. The difference between these is the innovation which may be positive or negative even without converting to anomalies.

Comment: Line 305: “then” should be “than”.

Response: We have fixed grammatical error.

Comment: In section 3.2, as Pearson’s correlation coefficient is really not high, the significance level should be mentioned.

Response: We have revised the text to mention the mean significance level (0.18 and 0.13 for the prior and reconstruction respectively). We will also mention the percentage of records significantly correlated to the reconstruction (60% of records with a p-value < 0.05 compared to 47% for the model prior; Lines 379-380).

Comment: In equation S1, no Xa.

Response: We have fixed this error.