Dear Reviewer,

We thank you for your thorough and constructive feedback. This file provides a complete documentation of the changes made in response to each of your comments. Reviewer's comments are shown in normal text, author responses are shown in bold, italic, blue text.

Reviewer 2

Cuesta-Valero et al. provide a new estimate of continental heat storage including ground, inland waters and permafrost thawing. For continental heat storage, an update to the previous estimate (Cuesta-Valero et al. 2021) is provided. For inland waters and permafrost thawing, models are used to derive the estimates. I have some major reservations about their methodologies, listed below.

(1). The observation-based estimate for ground heat storage and model-based estimates for inland waters and permafrost thawing are merged together to provide the continental heat storage. I doubt if they can be put together, and then eventually be used in von Schuckmann et al. GCOS assessment (the other components are all observation-based).

We respectfully disagree with the reviewer on this point. Many relevant studies have combined raw observations, data assimilation (i.e., reanalysis and satellite products), as well as numerical simulations (i.e, global and regional climate simulations) in order to assess the state and evolution of a certain variable of interest. The different Assessment Reports (ARs) of the Intergovernmental Panel on Climate Change (IPCC) are the most important examples of this practice in the climate community. A particular example could be the combination of paleoreconstructions and paleosimulations to obtain a better picture of the last millennium in both IPCC-AR5 (Masson-Delmotte et al., 2013) and IPCC-AR6 (Arias et al., 2021).

Furthermore, the rest of components of the GCOS assessment of Earth heat inventory do not include only observational data. For example, the atmosphere heat storage is estimated from reanalysis data, while the heat uptake by glacier melting is estimated from indirect gravimetric observations retrieved from satellites. These two estimates are produced by using numerical techniques and models to assimilate and interpret raw observations. Similarly, we use an observational-based driver to force a numerical model to estimate permafrost heat storage. Concretely, we use ERA-Interim data as upper boundary condition for our permafrost model, which allows us to estimate the ground ice melting that is coherent with the evolution of surface conditions in the last decades. Please note

that this is not very different to the use of different reanalyses to estimate the heat storage by the atmosphere in the GCOS paper mentioned by the reviewer.

Finally, we must also note that the use of models in our estimates is mostly the result of a lack of adequate data to characterize heat storage in permafrost and inland waters systems. For lakes, for example, existing in-situ data sets with long-term temperature profiles contain only very few lakes relative to the total number of lakes worldwide, and existing repositories are spatially biased towards Europe and North America. For rivers, the data availability is even worse. As we make clear in the manuscript, we incorporate model results because, unfortunately, there are no adequate measurements to derive global estimates of permafrost heat storage and inland waters heat storage.

We have also added a new paragraph in the Conclusions stating that although this is not an ideal estimate of continental heat storage, we have identified a clear path toward complementing the use of models with more observational-based estimates.

(2). Uncertainty estimates for ground heat storage. In this study the uncertainty of the ground heat storage has been reduced by an order compared to their earlier estimate (for example line 200-205). The new estimate suggests a global land heat storage of 84.8 +/- 0.8 mWm-2 (previous estimate is 97+/-6). I found it hard to believe such a small error range, it is simply not possible. Remember you are using only 1000 station data to represent the entire land, even previous error range of 6 is a likely underestimation. I can't understand this small number and I don't understand how this small number is derived given the dataset is basically the same with the previous version.

Please, note that ground heat storage is estimated from subsurface (borehole) temperature profiles, not from meteorological stations. These temperature-depth profiles record the propagation of alterations in the surface energy balance though the ground, but due to the nature of heat diffusion, borehole profiles are able to retrieve only long-term past changes in surface conditions. That is, decadal to centennial changes in ground heat flux. This reduces greatly the variability in the global average of ground heat storage in our manuscript, even considering the variability at 1079 different locations, thus uncertainty ranges are always going to be narrower than those of estimates based on meteorological stations.

Furthermore, we have used a new bootstrap technique to estimate uncertainties from these geothermal data. Previous estimates of global ground heat flux from subsurface temperature profiles provided uncertainty estimates that were biased from a statistical point of view, as they were markedly conservative and included a much wider range than the 95 % *confidence interval that is typically* *provided with the global average. Please, check Cuesta-Valero et al. (2022) for a detailed comparison of previous uncertainty estimates in comparison with the new bootstrap method, as well as a prove that the uncertainty in previous inversions converge to the one reported here when appropriate error propagation methods are considered.*

(3). Uncertainty estimates for permafrost thawing. Only the uncertainty related to the soil thickness and ice saturation are taken into account. However, I think another major error come from the model and climate forcing. For example, the use of Mk3L and ERA-Interim, the errors/biases will definitely propagate into the estimate of this study. I have no idea how to resolve this, as it is related to the fundamental choices of this study: using models and reanalysis to drive the their estimates.

Indeed, the use of numerical simulations in the estimates of permafrost heat storage adds uncertainties to the results. As discussed in the manuscript, ERA-Interim is now superseded by the ERA5 reanalysis, and this new reanalysis should be used in a future iteration of this work. Regarding the Mk3L paleosimulation, new models contributing to the PMIP4 project may be more suitable. In any case, please note that the Mk3L simulation is used only to initialize the permafrost model, as finding an equilibrium state for ground ice under preindustrial conditions requires several centuries, and then a transitional period between preindustrial conditions and conditions at the starting date of ERA-Interim (1979 CE) should be provided. Also note that the surface boundary conditions for ~ 60 % *of the period of interest are obtained from the observation-based ERA-Interim reanalysis. Therefore, the effect of using a more advanced paleosimulation should be small.*

The largest uncertainties regarding the permafrost heat uptake are expected to be related to the effect of snow on the ground thermal regime and the distribution of ground ice. These first-order effects were addressed by our parameter ensemble simulations. Please refer to Nitzbon et al. (2022) for an extended discussion of the uncertainties and limitations of the permafrost heat uptake.

(4). Uncertainty estimates for inland waters. Is the ensemble spread used to estimate the uncertainty of heat storage in inland waters? If so, it is fundamentally different from the other two components, i.e. the assumption underlying this method is: model difference (whatever caused the difference) can fully represent the uncertainty. Such assumption is likely wrong as there are always common model biases. And such assumption is clearly different from the assumption for your permafrost thawing and ground heating uncertainty estimate, so they can not be simply added up, simply physically meaningless.

In our analysis, we explore and analyse each component separately, considering both spatial and temporal variability. Later, a common estimate for the entire continental system is derived. As pointed by the reviewer, using standard error propagation methods to derive the total uncertainty

in continental heat storage is excessively simplistic and omits critical differences in the methodology used to derive the heat storage within each subsystem. In the new version of the manuscript, we still provide with an uncertainty estimate for the continental heat storage but explaining the limitations of each method and the differences among the uncertainty estimated for each individual component. We also keep the uncertainty estimates for each individual component; thus the reader can reach their own conclusions about the trustworthiness of the reported uncertainties, and compute its own estimates.

(5). How the final estimate of land heat storage uncertainty been derived? Are you assuming independency of the three components? Are they independent?

The final continental heat storage series results from adding the global estimates for ground heat storage, permafrost heat storage, and inland waters heat storage using standard error propagation methods (lines 258-262 in the original manuscript). Here we should inform the reviewer of an error in the processing of the uncertainty estimates that lead to and underestimation of the total uncertainty in continental heat storage. Please, check the new version of the manuscript for an updated estimate.

Regarding the independence of the estimates, the three estimates are considered independent, as ground heat storage estimates do not include the heat uptake by permafrost thawing, and no ground heat storage nor permafrost thawing is possible in lakes or reservoirs. Nevertheless, the uncertainty estimate for the total continental heat storage is probably not robust, thus we have added a paragraph discussing the limitations in our estimate (see also our answer to the previous comment).

(6). Line 219: Please explain why “this large interannual variability is explained by the smaller surface of global lakes and reservoirs in comparison with the global land and permafrost areas”?

What we wanted to indicate here is that since inland water bodies cover a surface that is two orders of magnitude smaller than the land surface, and one order of magnitude smaller than the total permafrost area, we can expect a larger inter-annual variability in the estimated inland waters heat flux than in the other two components. We have rewritten this sentence in the new version of the manuscript.

(7). Line 257. The total land heat storage is 23.9+/-0.4 ZJ. The error range is too small to believe. Look at Fig. 1a, there are only several places with observations, and the spatial variability is large (that means you need more data to resolve these variability), so I don’t think the uncertainty can be so small. The uncertainty estimate should be better documented in this study, and any revision should be carefully assessed and validated.

Please, check our answer to the second comment. Cuesta-Valero et al. (2022) explains in detail the reason for this new smaller uncertainty in global estimates of ground heat flux, and it also includes a comprehensive comparison with previous techniques to retrieve uncertainty from inversions of subsurface temperature profiles. In a nutshell, previous estimates converge to the new uncertainty results when individual inversions from subsurface profiles are aggregated correctly.

To proceed (avoid rejection of this paper), I recommend the authors not putting the the estimates for the three estimates toghether, just presenting them separately, making a point that permafrost and lakes might be important in EEI, which is the best the authors' can do.. I disagree to put them together because some are model-based estimates, and the uncertainty etsimates are apparant very weak.

We disagree with the reviewer. We are already providing with the individual estimates for each component of the continental heat storage, analysing their temporal and spatial variability. Nevertheless, we recognize that we underestimated the differences between the sources of uncertainty considered in each continental subsystem, and we have included a discussion about the different uncertainties in each subsystem. Thereby, the readers of the manuscript have access to the individual estimates for ground, permafrost, and inland waters heat storage, information about how to interpret these estimates, the result of applying common error propagation methods, and a warning about the limitations in this uncertainty analysis.

Regarding the combination of measurements and models, we refer the reviewer to our answer to the first comment: combination of observation-based results with reanalysis and modelling estimates is a common practice in the climate community when observations for a relevant variable or component of the Earth system are not available. We consider that not using reanalysis or simulations to try to better understand the behaviour of the climate system is a mistake. Nevertheless, we agree that these lines of evidence cannot replace observations, and that they are very different between each other. Therefore, we clearly identify the source of data for our estimates, we clearly indicate in our manuscript that reanalyses and model simulations include additional uncertainties not present in subsurface temperature profiles, and we indicate that observations should be incorporated into the analysis where and when possible. We have also added a paragraph to the Conclusions section clearly indicating that more observation based data should be included in the new version of this analysis.

Unfortunately, it is not within our immediate reach to fill the observational gaps appearing in this

analysis, but we can still use other sources of information for mitigating those gaps as much as possible.

References

Arias, P., Bellouin, N., Coppola, E., Jones, R., Krinner, G., Marotzke, J., Naik, V., Palmer, M., Plattner, G.-K., Rogelj, J., Rojas, M., Sillmann, J., Storelvmo, T., Thorne, P., Trewin, B., Achuta Rao, K., Adhikary, B., Allan, R., Armour, K., Bala, G., Barimalala, R., Berger, S., Canadell, J., Cassou, C., Cherchi, A., Collins, W., Collins, W., Connors, S., Corti, S., Cruz, F., Dentener, F., Dereczynski, C., Di Luca, A., Diongue Niang, A., Doblas-Reyes, F., Dosio, A., Douville, H., Engelbrecht, F., Eyring, V., Fischer, E., Forster, P., Fox-Kemper, B., Fuglestvedt, J., Fyfe, J., Gillett, N., Goldfarb, L., Gorodetskaya, I., Gutierrez, J., Hamdi, R., Hawkins, E., Hewitt, H., Hope, P., Islam, A., Jones, C., Kaufman, D., Kopp, R., Kosaka, Y., Kossin, J., Krakovska, S., Lee, J.-Y., Li, J., Mauritsen, T., Maycock, T., Meinshausen, M., Min, S.-K., Monteiro, P., Ngo-Duc, T., Otto, F., Pinto, I., Pirani, A., Raghavan, K., Ranasinghe, R., Ruane, A., Ruiz, L., Sallée, J.-B., Samset, B., Sathyendranath, S., Seneviratne, S., Sörensson, A., Szopa, S., Takayabu, I., Tréguier, A.-M., van den Hurk, B., Vautard, R., von Schuckmann, K., Zaehle, S., Zhang, X., and Zickfeld, K. (2021). "Technical Summary". In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 33–144. DOI: `10.1017/9781009157896.002`.

Cuesta-Valero, F. J., Beltrami, H., Gruber, S., García-García, A., and González-Rouco, J. F. (2022). A new bootstrap technique to quantify uncertainty in estimates of ground surface temperature and ground heat flux histories from geothermal data. *Geoscientific Model Development*, **15**(20), 7913–7932. DOI: `10.5194/gmd-15-7913-2022`.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., and Timmermann, A. (2013). "Information from Paleoclimate Archives". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Chap. 5, pp. 383–464. DOI: `10.1017/CBO9781107415324.013`.

Nitzbon, J., Krinner, G., von Deimling, T. S., Werner, M., and Langer, M. (2022). *Quantifying the Permafrost Heat Sinkin Earth's Climate System*. Submitted to Geophysical Research Letters. DOI: `10.1002/essoar.10511600.1`.