Answer to Oliver Bothe

We'd like to thank Oliver Bothe for his effort and his very constructive comments, which helped us a lot to improve the manuscript. In the following, we provide point-by-point replies. The reviewer comments are in blue and bold, answers are in black, cited text in italic and new or changed text is marked in red.

Major

That is: I do not believe that the manuscript does what the authors claim in the title. While I appreciate a catchy title, I think overselling is a net-minus. My understanding of the title is that the authors claim to show why 1809 remains a cold case. The title plays with the cold-case terminology from criminal investigations and procedural crime series on TV and streaming. However, from my point of view the manuscript does not show enough in that respect to claim this title. A positive reading is, the title already announces the failure to provide major new insights into location, seasonality, or strength of the eruption, a negative reading is that the title suggests large new insights, why we won't have much success in becoming more certain about this eruption. I don't think these large new insights are provided. The manuscript is in a sense incremental while also presenting some very valuable - and I think new - simulations and in addition supporting previous understanding about the eruption in 1809.

We agree with the reviewer that we do not provide enough new material in our study to justify a "why it remains" in the title. Nevertheless, our findings contribute to the recent knowledge of the unidentified 1809 eruption and support previous understanding of its climate impact. Hence, we will revise our title to: ***"The unidentified eruption of 1809: still a climatic cold case".***

Minor

Line 19: The claim that the early 19th century is the coldest period of the last 500 years is based on a reference to Jungclaus et al. (2017). While I am not entirely sure whether this is meant to be a global statement or a hemispheric one, there have been a number of recent reconstruction efforts spanning the period of the last 500 years and beyond, which may or may not require to reassess or qualify the statement. That is, such an absolute statement requires assessing the newest evidence, which in this case may include, for example, Büntgen et al. (2020), PAGES2k Consortium (2019), and more.

The reference to Jungclaus et al (2017) was related to the 2nd part of the sentence that the early 19th century is a period of special scientific interest for PMIP4/past2K. We have revised the sentence and included now references to the PAGES2k Consortium (2019) and the original Wilson et al (2016) N-TREND study. Büntgen et al. (2020) uses only ring-width data and may therefore not always truly be a strong representative record of volcanic forcing on summer temperatures.

----" T*he early 19th century (~1800-1830 CE), at the tail end of the Little Ice Age, marks one of the coldest period of the last millennium (e.g. Wilson et al, 2016; PAGES2k Consortium, 2019) and is therefore of special interest in the study of inter-decadal climate variability (Jungclaus et al., 2017).*"

A note on the paragraph starting in line 78: the authors detail changes in their atmospheric component but they only mention in passing that there were also changes in their land component. That is fine, but the authors probably also can be briefer in describing the atmospheric model.

We have revised the following sentences in the model description to:

"*It consists of four components: the atmospheric general circulation model ECHAM6 (Stevens et al., 2013), the ocean-sea ice model MPIOM (Jungclaus et al., 2013), the land component JSBACH (Reick et al., 2013) and the marine biogeochemistry model HAMOCC (Ilyina et al., 2013). JSBACH is directly coupled to the ECHAM6 model and includes dynamic vegetation, whereas HAMOCC is directly coupled to the MPIOM. ECHAM6 and MPIOM are in turn coupled through the OASIS3-MCT coupler software. In MPI-ESM1.2, ECHAM6.3 is used, which is run with a horizontal resolution in the spectral space of T63 (~200 km) and with 47 vertical levels up to 0.01 hPa with 13 model levels above 100 hPa. In ECHAM6.3 aerosol microphysical processes are not included. The radiative forcing of the volcanic aerosol is prescribed by monthly and zonal mean optical parameters, which are generated with the EVA forcing generator, see section 2.1.2. The MPIOM, which is run in its GR15 configuration with a nominal resolution of 1.5◦around the equator and 40 vertical levels, has remained largely unchanged with respect to the CMIP5 version. Several revisions with respect to the MPI-ESM CMIP5 version have been made for the atmospheric model incl. a new representation of radiation transfer, land physics and biogeochemistry components and the ocean carbon cycle. A detailed description of all updates is given in Mauritsen et al. (2019).*"

Figures: some of the Figures have a strangely low resolution (on my screen). While this will be caught by the technical editing at Copernicus anyway, I wanted to mention it.

Thank you very much for this hint, we will carefully go over the figures again and will revise them if necessary.

Data: EEIC. I have to admit I am not really up to date and may express my ignorance but how does the EEIC data differ or improve on the most recent ICOADS data?

The EEIC data are included in the ICOADS reconstruction, and appear to represent the primary data source for the 1790-1830 period (Freeman et al., 2016), especially over the ocean areas presented in our study. Our analysis could well have used the ICOADS gridded data, and we expect the results would not depend on the choice of data set.

Data: Stations. I was thinking there should exist tropical stations for the period of interest. Do they have too little temporal coverage, or were they not of interest?

There is only one tropical station with a long series: Chennai ,India (13.5°N, 80°E) but this station has an unfortunate gap right after the eruption, so we consider NH extratropical staton only.

At line 284 I was wondering whether a very short comparison between reconstructed and observed results would be of value there.

We have revised Fig.2 by including the reconstructed spatial distribution of NH summer 1811 to cover a three year post volcanic period. The NH summer 1816 panel is therefore shifted to the supplementary material. As we compare reconstruction, station data and model simulation already in the discussion session, we have only slightly revised the text here.

"*The cooling after the 1809 eruption is not so pronounced as after the Mt. Tambora eruption in 1815. In general the station data support the spatial distribution of the reconstructed near surface temperature anomalies derived from tree-ring data. They show a local minimum over Northern, Eastern and Southern Europe in NH summer 1810, which does not appear over Western and Central Europe and New England. The warm anomalies in the order of 2°C, which are found in summer 1811 over Eastern Europe are not however captured by the N-TREND spatial reconstruction although some slight warming is seen in the data over East Poland, Belarus and the Baltic States.*"

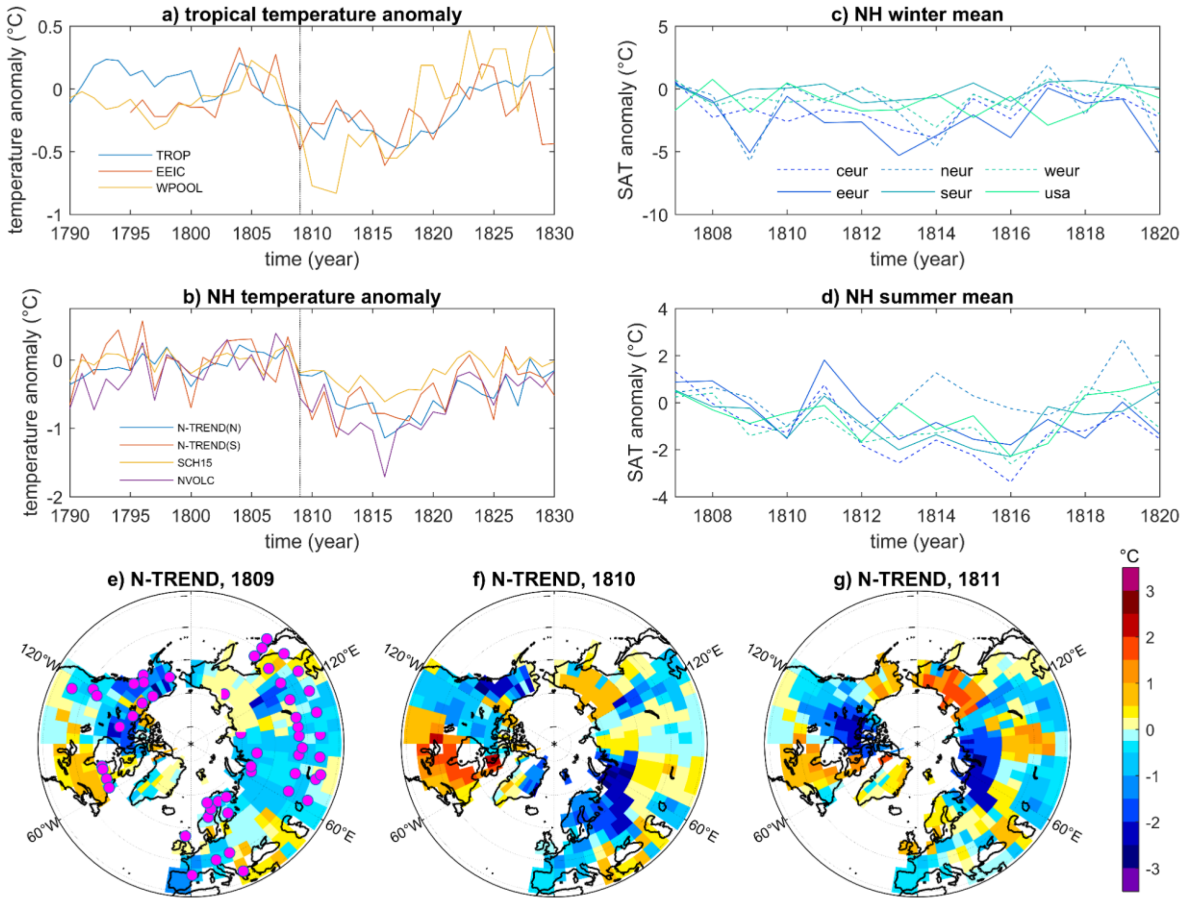

Figure 2: *Observed and reconstructed temperature anomalies around the 1809 volcanic eruption.* ***a***) *Reconstructed tropical (30°N–30°S, 34°E-70°W) sea surface temperature (TROP, D'Arrigo et al., 2009), measured tropical marine surface air temperatures from ship logs (EEIC, Brohan et al., 2012) and warm pool data (WPOOL, D'Arrigo et al., 2006).* ***b***) *NH summer land temperatures from four tree-ring based reconstructions (Wilson et al,, 2016 (N-TREND (N)), Anchukaitis et al., 2017 (N-TREND (S)), Guillet et al., 2017 (NVOLC), Schneider et al., 2015 (SCH15)).* ***c***) *Monthly mean NH winter and summer temperature anomalies (°C) from 53 station data averaged over different European regions (Central Europe (CEUR: 46.1-52.5°N, 6-17.8°E), Eastern Europe (EEUR: 47-57°N, 18-32°E), Northern Europe (NEUR: 55-66°N, 10 -31° E), Southern Europe (38-46°N, 7-13.5°E), Western Europe (WEUR: 48.5-56°N, 6°W- 6°O) and New England (NENG: 41-44°N, 73- 69°W).* ***d-f***) *Mean surface temperature anomalies (°C) for boreal summers of 1809 (d), 1810 (e) and 1811 (f) in NH TR data N-TREND (S) (Anchukaitis et al, 2017). Pink dots in panel d illustrate the location of the tree-ring proxies used in the N-TREND reconstruction. All anomalies are with respect to the 1800-1808 climatology.*

Index comparisons: Line 347ff. I fail to see the relevance of the circulation results for the manuscript. They feel unrelated to the rest of the manuscript. They are again referred to in the discussions but to me it also remains unclear there why they are relevant for the argument of the manuscript. The point in the discussion does not depend on the analysis, does it? This is not necessarily a problem but skipping the relevant parts may make the manuscript more concise.

The analysis of indices associated with known dominant modes of large-scale atmospheric circulation anomalies relevant for the Northern Hemisphere was motivated by the idea that the variable response patterns identified between the ensembles but also within each ensemble reflected a lack of robust imprint of volcanic forcing on such modes. Our analysis overall seems to confirm this, especially for the most relevant NAO and PNA. We will tone down the results on the climatic indices to better reflect the specific aim of this analysis. Specifically, we have moved appendix 1 to the supplement and revised the text as follows:

- lines 233-237 of the original manuscript are moved to the supplement
- the paragraph (lines 347-367) of the original manuscript is moved to the supplement, and replaced with a shortened version along the following lines: *"The substantial differences found in the post-eruption evolution of continental and subcontinental climates that can be produced by internal climate variability and forcing structure reflect substantial differences in the post-eruption evolution. Specifically, post-eruption anomalies of selected dominant modes of large-scale atmospheric circulation in the Northern Hemisphere and the tropics, including the Pacific/North American pattern, the North Atlantic Oscillation, the North Pacific Index and the Southern Oscillation, yield a spread of responses within individual ensembles that is often as large as the range of pre-eruption variability. Further, response distributions generated by different forcings in some cases do not overlap (see Supplementary Figure S1)."*

Discussions of central Asia: The authors give good reasons for potentially weak reconstructions in Asia and particularly central Asia. Has there been a general evaluation of how well the MPI-ESM performs in these regions?

Unfortunately, there exist no extensive and detailed model evaluation of the MPI-ESM1.2 over central Asia. Recently, an assessment of 30 CMIP6 models in simulating precipitation over arid Central Asia was published (Guo et al., 2021) but it does not include the MPI-ESM1.2. There exists some PMIP mid holocene studies which concentrate more on Central Asia but they are more general and partially related to previous versions of the model (e.g. Zheng et al., 2013). However, a comparison of the climatological mean state of near-surface air temperature with ERA-Interim shows that the MPI-ESM-LR model overestimates the mean temperature over the Central Asian highlands about a few degrees (1-5 °C), slightly underestimates it between 50-65°N and overestimates it again over East Siberia, north of the Arctic circle (Müller et al., 2018; Figure 7a). So we cannot rule out that part of the weak correlation is related to the model biases, which we already briefly mentioned in our conclusion: " ….*In contrast, the reconstructed temperature pattern over Asia is not produced by any model simulation, forced or control. Possible explanations for this model-data disagreement include deficiencies and uncertainties regarding both tools. In particular, the dendrochronological network remains ….*".

To emphasize this point a bit more, we have added now a sentence about the model performance in the description of Figure 12:

"T*his is most likely related to a substantial data quality issue as the tree-ring data, especially for central Asia, is solely based on TRW data (Wilson et al., 2016)*. *However, we can also not rule out the possibility of a model biases as the climatological mean state of near-surface air temperature in the MPI-ESM-LR over Central Asia deviates from ERA-Interim about a few degrees (Müller et al., 2018)."*

TR: I invite the authors to skip the abbreviation TR. "Tree ring" is not too long to be spelled out everywhere.

We use tree-ring now throughout the text.

NA: similarly to TR, is it really necessary to abbreviate North America?

Certainly it is not necessary, we spell North America now out.

There are a small number of grammar/spelling/typos etc that I assume are artefacts from tracking changes in a document. I only mention Line 655: proxi -> proxy

We went carefully over the text again and corrected grammar and spelling typos.

Is a "comparative assessment" a comparison?

It is similar but it is not the same: a "comparative assessment" includes an evaluating component. We agree that it might be more appropriate to use the more general term "comparison" and have revised the text accordingly.

I wonder whether it makes sense to spell out S in line 638.

For clarity, we refer now to sulfur isotopes.

References:

Anchukaitis, K.J., R. Wilson, K. Briffa, U. Büntgen, E.R. Cook, R.D. D'Arrigo, N. Davi, J. Esper, D. Frank, B. Gunnarson, G. Hegerl, S. Helama, S. Klesse, P.J. Krusic, H. Linderholm, V. Myglan, T. J. Osborn, Z. Peng, M. Rydval, L. Schneider, A. Schurer, G. Wiles and E. Zorita, Last millennium Northern Hemisphere summer temperatures from tree rings: Part II: spatially resolved reconstructions, Quaternary Science Reviews, 163, 1-22, doi: 10.1016/j.quascirev.2017.02.020, 2017.

Brohan, P., Allan, R., Freeman, E., Wheeler, D., Wilkinson, C., and Williamson, F.: Constraining the temperature history of the past millennium using early instrumental observations, Clim. Past, 8, 1551–1563, https://doi.org/10.5194/cp-8-1551-2012, 2012.

Büntgen, Ulf, et al. "Prominent role of volcanism in Common Era climate variability and human history." Dendrochronologia 64 (2020): 125757.

Chenoweth, M. Two major volcanic cooling episodes derived from global marine air temperature, AD 1807–1827 https://doi.org/10.1029/2000GL012648, 2001.

D'Arrigo, R., Wilson, R., Palmer, J., Krusic, P., Curtis, A., Sakulich, J., Bijaksana, S., Zulaikah, S., Ngkoimani, L. O., and Tudhope, A.: The reconstructed Indonesian warm pool sea surface temper-atures from tree rings and corals: Linkages to Asian monsoon drought and El Niño–Southern Oscillation, Paleoceanography,21, PA3005, doi:10.1029/2005pa001256, 2006.

D'Arrigo, R., Wilson R., and Tudhope, A.: The impact of volcanic forcing on tropical temperatures during the past four centuries, Nat. Geosci., 2, 51–56, doi:10.1038/ngeo393, 2009.

Freeman, E., et al.: ICOADS Release 3.0: a major update to the historical marine climate record, Int. J. Climatol., 37(5), 2211–2232, doi:10.1002/joc.4775, 2017.

Guillet, S., et al.: Climateresponse to the Samalas volcanic eruption in 1257 revealed by proxy records, Nature geoscience, 10, 123–128, 2017.

Guo, Hao, et al. Assessment of CMIP6 in simulating precipitation over arid Central Asia. Atmospheric Research, 2021, 252. Jg., S. 105451.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Nunez-Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, Journal of Adva--nces in Modeling Earth Systems, 5, 287-315, 10.1029/2012ms000178, 2013.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K.,Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J. S.: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model: Mpiom CMIP5 Ocean Simulations, J. Adv. Model. Earth Sy., 5, 422–446, https://doi.org/10.1002/jame.20023, 2013.

Jungclaus, J. H. et al: The PMIP4 contribution to CMIP6 – Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 past1000 simulations, Geosci. Model Dev., 10, 4005–4033, https://doi.org/10.5194/gmd-10-4005-2017, 2017.

Mauritsen, T., et al..: A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM 1.2-HR), J. Adv. Model. Earth Sy., 10, 1383–1413, https://doi.org/10.1029/2017MS001217, 2018.

PAGES2K "Consistent multi-decadal variability in global temperature reconstructions and simulations over the Common Era." Nature geoscience 12.8 (2019): 643.

Reick, C. H., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, Journal of Advances in Modeling Earth Systems, 5, 459-482, 10.1002/jame.20022, 2013.
Stevens, B., et al. .: Atmospheric component of the MPI-M Earth System Model: ECHAM6: ECHAM6, J. Adv. Model. Earth Syst., 5(2), 146–172, doi:10.1002/jame.20015, 2013.
Schneider, L., Smerdon, J. E., Büntgen, U., Wilson, R. J. S.,Myglan, V. S., Kirdyanov, A. V., and Esper, J.: Revising mid-latitude summer temperatures back to A.D. 600 based on a latewood density network, Geophys. Res. Lett., 42, GL063956, https://doi.org/10.1002/2015gl063956, 2015.
Toohey, M., Stevens, B., Schmidt, H., and Timmreck, C.: Easy Volcanic Aerosol (EVA v1.0): an idealized forcing generator for climate simulations, Geosci. Model Dev., 9, 4049 – 4070, doi:10.5194/gmd-9-4049-2016, 2016.
Wilson, R., et al.: Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context, Quat. Sci. Rev., 134, 1–18, doi:10.1016/j.quascirev.2015.12.005, 2016.
Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P. et al.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6, Geosci. Model Dev., 9, 2701 – 2719, doi:10.5194/gmd-9-2701-2016, 2016.
Zheng, W., Wu, B., He, J., and Yu, Y.: The East Asian Summer Monsoon at mid-Holocene: results from PMIP3 simulations, Clim. Past, 9, 453–466, https://doi.org/10.5194/cp-9-453-2013, 2013.