Comments to the author:
Dear colleagues, thank you very much for addressing the points by the 4 reviewer. Can I please ask you to revise the paper accordingly. Just a few points related to your answers to the reviewers:

Dear Editor Jürg Luterbacher,

Thank you for your comments.

We have carefully checked the suggested proxy data sets and have chosen to include a comparison with the PAGES 2k (2019) and the Luterbacher et al. (2016) data sets. We have added those figures to the appendix. Our response to the comments is given in further detail below, in blue and bold text.

1. On the use of the Buentgen al paper which is certainly the latest published large scale reconstruction and suggested by different reviewers. Another option would have been Neukom et al. that was presented in the latest IPCC AR6. Using and showing only reconstruction is critical as it will not allow a complete and unbiased picture. Also Buentgen et al. has limitations/challenges, is annual with a summer and tree ring only biased I would like to see a more broader comparison including other hemispheric reconstructions (as suggesting by reviewer 4, point 2)

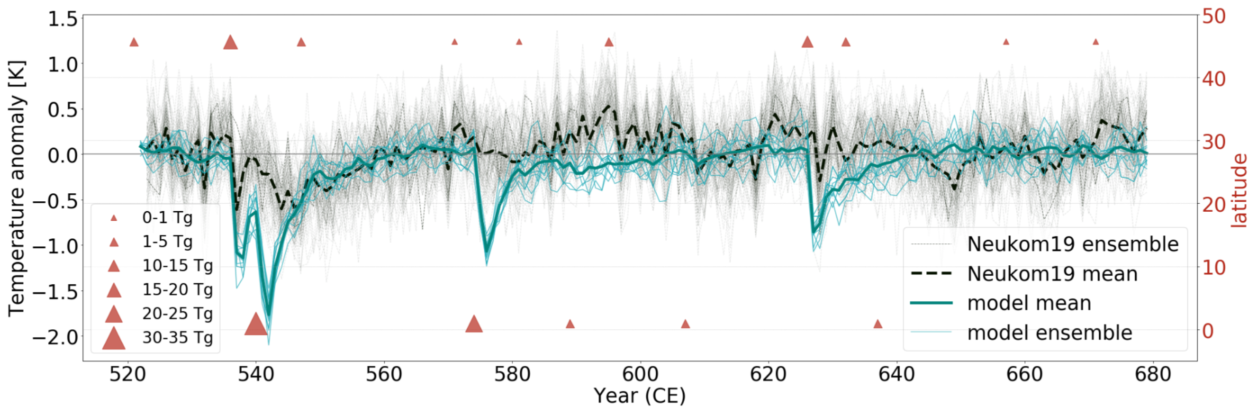
R4 C2. Along the same line, starting from line 32ff the authors state that cooling might have exceeded that of the LIA and focus on two site chronologies that were presented in 2016. While the authors rightly present the results of this study, and add the reply provided by Helama and colleagues from 2017, they ignore a vast body of proxy studies that have been published on the topic and where the chronologies cover many sites of the NH. By focusing only on the LALIA study, they ignore a large body of spatial and temporal reconstructions covering the period of interest. The authors should therefore present a more balanced assessment of the existing data by including e.g., Schneider et al. (2015, ERL), NTREND (spatial and temporal; Wilson et al., 2016 QSR; Anchukaitis et al., 2017 QSR), Guillet et al. (2017 NGEO) or the most recent TRW-based paper from the tree-ring community published lately by Buntgen et al (2021) in NCOMM.

As stated in our previous reply we compare our simulations now with the ensemble reconstruction from Buentgen et al. (2021), which encompasses 15 different NH tree ring reconstructions.

The explicit data sets suggested by Reviewer 4 do not cover our period of interest: Schneider et al. (2015) covers 600-2000 CE, NTREND (Wilson et al., 2016; Anchukaitis et al., 2017) covers 918-2004 CE and Guillet et al. (2017) covers the 1257 Samalas eruption. On top, the three latter reconstruction methods are included in the Büntgen et al. (2021) study.

According to your suggestions, we have extended our comparison to two different reconstructions. We compare our model results now with the hemispheric scale proxy reconstructions from the PAGES2k consortium (Neukom et al. 2019) for the NH annual mean, as well as for Europe JJA with the reconstruction by Luterbacher et al. (2016), see our answers below. For both reconstructions, we used the composite plus scaling (CPS)

method data set. We have added these two figures to the appendix as Fig. A3 and A2 d) and elaborated on them in the discussion and appendix.

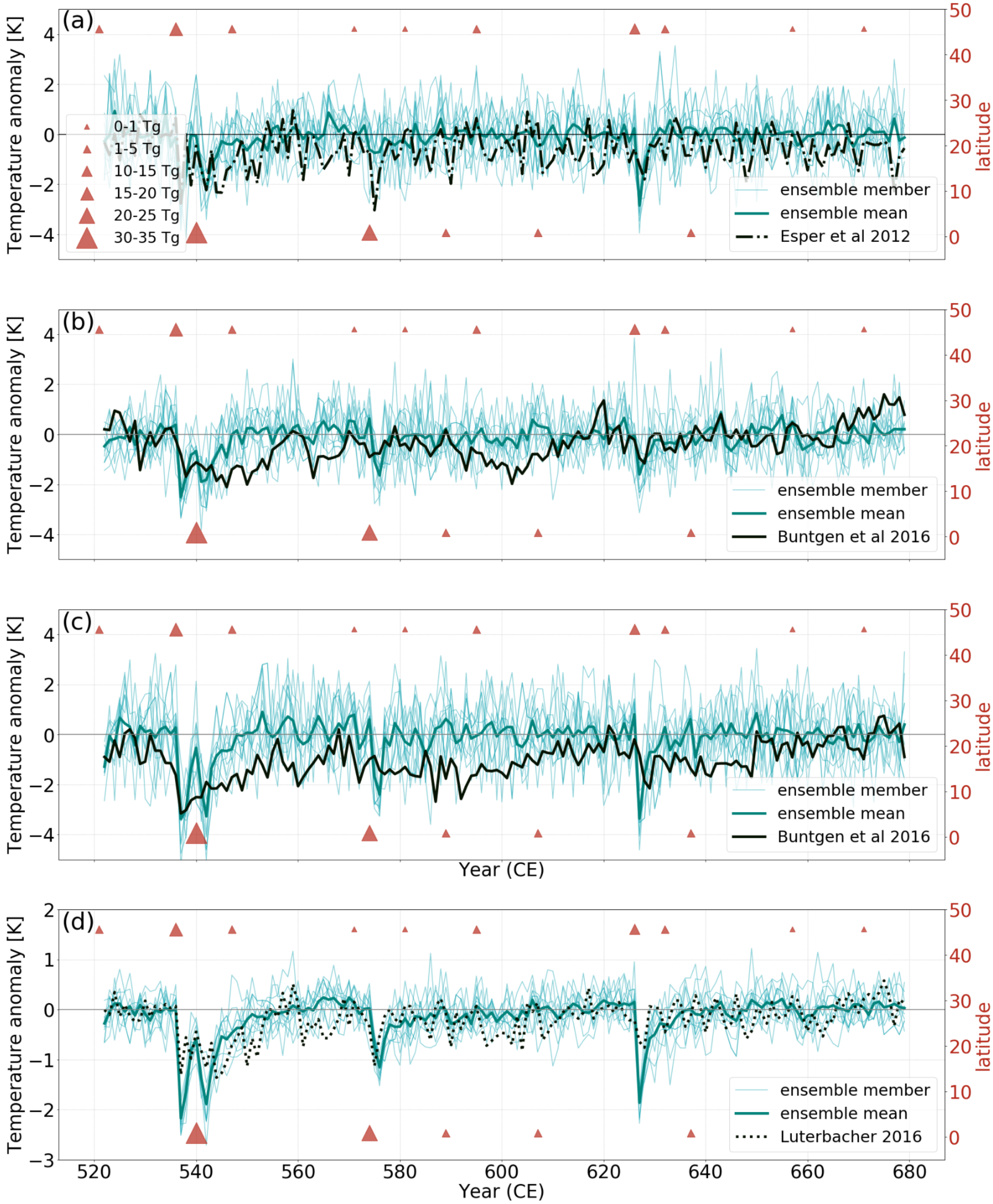


New **Figure A3: Model - proxy comparison for the entire NH annual mean. The proxy data is the PAGES2k consortium (Neukom 2019). Anomalies are calculated wrt 0-1850 CE.**

2. as well as a comparison at continental scale as suggested by reviewer 2, point 11 at least for some of the parts of the manuscript.

R2C11. A more appropriate comparison could be made with continental reconstructions for Europe (Luterbacher et al. 2016, Env. Res. Lett) and for Asia (Zhang et al. 2018 Nat. Sci. Reports, 8, 7702).

When it comes to continental reconstructions, we have now added Luterbacher et al. (2016) to figure A2 in the Appendix (see figure A2 d) below). The data used by Zhang et al. (2018) consist of decadal-mean data and thus does not include the annual peak cooling.



New Figure A2 Model - tree-ring comparison for a) Northern Scandinavia (N-Scan) (Esper et al., 2012b), b) the Alps and c) the Altai (Büntgen et al., 2016) and d) Europe (Luterbacher et al., 2016). Near-surface air temperatures from the model are taken for the corresponding area and for JJA and land only. Anomalies are calculated wrt 0-1850 CE.

On the Buentgen/model comparison.: It seems that there is a large difference around 540, around 575 and 630, also obvious is the general overestimation from around 545 to 575 and underestimation until 640. Indeed, it would be interesting if this is a feature for this reconstruction only or also shared with independent evidence from other reconstructions.

In figure A3, we compare the entire NH annual mean from the model simulations with the PAGES 2k data. The PAGES 2k (2019) data show different signals and responses following large eruptions compared to the model simulations. The temperature response following the 536/540 CE and 626 CE eruptions are weaker in the proxy data. The recovery time after 540 CE however, agrees well with the model simulations. Whereas Büntgen et al. (2021) shows a weaker signal to the 574 CE eruption, the PAGES 2k data reveal no response at all. In addition, the PAGES 2k proxy data is up to 0.5 degrees warmer between 580 and 610 CE.

Wrt to the new Figs. A2 and A3, we have added the following text to the ms:

4 Discussion

Model - tree ring comparison

…

“Reconstruction records are becoming more uncertain further back in time due to the sparseness of available proxy records especially before 1200 CE (Masson-Delmotte, 2013; Esper et al., 2018; PAGES 2k, 2019). Comparing the model results with temperature reconstructions for the entire NH annual mean (PAGES 2k, 2019) reveals a less good agreement (see Appendix A, Figure A3).”

…

“Previous proxy-based studies (Larsen et al., 2008; Esper et al., 2012b; Luterbacher et al., 2016; Helama et al., 2017; PAGES 2k, 2019) found a cooling up to 570 CE for Scandinavia, Europe, and the NH, based on tree-ring, ice-core, lake sediment, and documentary records (Figures A2a and d, and A3).”

Appendix A - Model - tree ring, and multi-proxy reconstruction comparison:

“Most temperature reconstruction data sets go back to about 1200 CE, and the further back in time, the fewer proxy records remain, and the more uncertainties they contain (Masson-Delmotte, 2013; Esper et al., 2018; Neukom et al., 2019). The main proxy type that remains to reconstruct the temperatures in the Northern Hemisphere (and especially mid-high latitudes, Europe) are tree rings (Neukom et al., 2019), and they are often used to reconstruct the temperature in especially Europe (Luterbacher et al., 2016). Other reconstructions available consist of a mix of proxies with a more limited dating precision, which leads to a reduction of the amplitude of the signals (Sigl et al., 2015; Büntgen et al., 2020; Plunkett et al., 2021). Only 25% of the proxies available for our study period have annual dating precision (Sigl et al., 2015).

The data sets used for the individual tree-ring sites are from Northern Scandinavia (N-Scan) (Esper et al., 2012b, a), from the Alps and Altai (Büntgen et al., 2016) and from Europe (Luterbacher et al., 2016) (Table 2).”

Table 2. Overview of tree-ring/proxy locations and type (MXD: maximum latewood density; TRW: tree ring width) used in this individual comparison. For the proxy reconstructions from Luterbacher et al. (2016) and PAGES 2k the method used is composite plus scaling (CPS).

Location name	Coordinates	Type of data	Reference
N-Scan	66°- 70°N, 19°- 29°E	MXD	Esper et al. (2012b, a)
Alps	46°N, 12.5°E	TRW	Büntgen et al. (2016)
Altai	50°N, 87.5°E	TRW	Büntgen et al. (2016)
Europe	35°- 70°N, -25°- 40°E	TRW and MXD	Luterbacher et al. (2016)
NH	0°- 90°N, -180°- 180°E	mixed proxies	(PAGES 2k, 2019)

...

"For Europe, the model and proxy data (Luterbacher et al., 2016) agree well. The proxy reconstruction falls within the spread of the model ensemble (Fig. A2d). As for the comparison with Büntgen et al. (2021), the peak cooling is less for the reconstructions, and there is an lag in the proxy data after 540 CE.

To illustrate the reduction in amplitude of the signals, we compared the PAGES 2k multi-proxy reconstruction (Neukom et al., 2019) to the simulated NH annual mean near surface temperature (Fig. A3). The proxy data agrees well for the recovery period after the 536/540 CE double eruption, and from ~650 CE - 675 CE, but the reconstructions show a weaker cooling following the volcanic eruptions, and for the 574 CE eruption the reconstruction does not show a signal at all."

Review 2:

point 11: please see comment above

point 17: I don't agree on the following answer: The 0-1850 CE mean was taken because this is a long enough period for the volcanic response to be negligible.

Within this period there are clearly times with strong volcanic activity. I suggest to show the difference with respect to a shorter period without any large volcanic events.

In the figure below, we have calculated the temperature anomalies with respect to a) the control run for 0 CE conditions, b) the 0-1850 CE past2k mean, as we have used in our anomaly calculations in the paper, and c) the 525-535 CE mean (a period before without volcanic eruptions) from the past2k run. As you can see, this gives the same result (within +/- 0.15 K) as when using the entire 0-1850 period for the anomaly calculation from the past2k run. We therefore argue that using the 0-1850 CE mean does not affect the results of our study. We have added this information to the Methods section and we have revised the sentence to: "*We use the average over 0-1850 CE to have a reference climate that is representative for the entire pre-industrial Common Era.*"

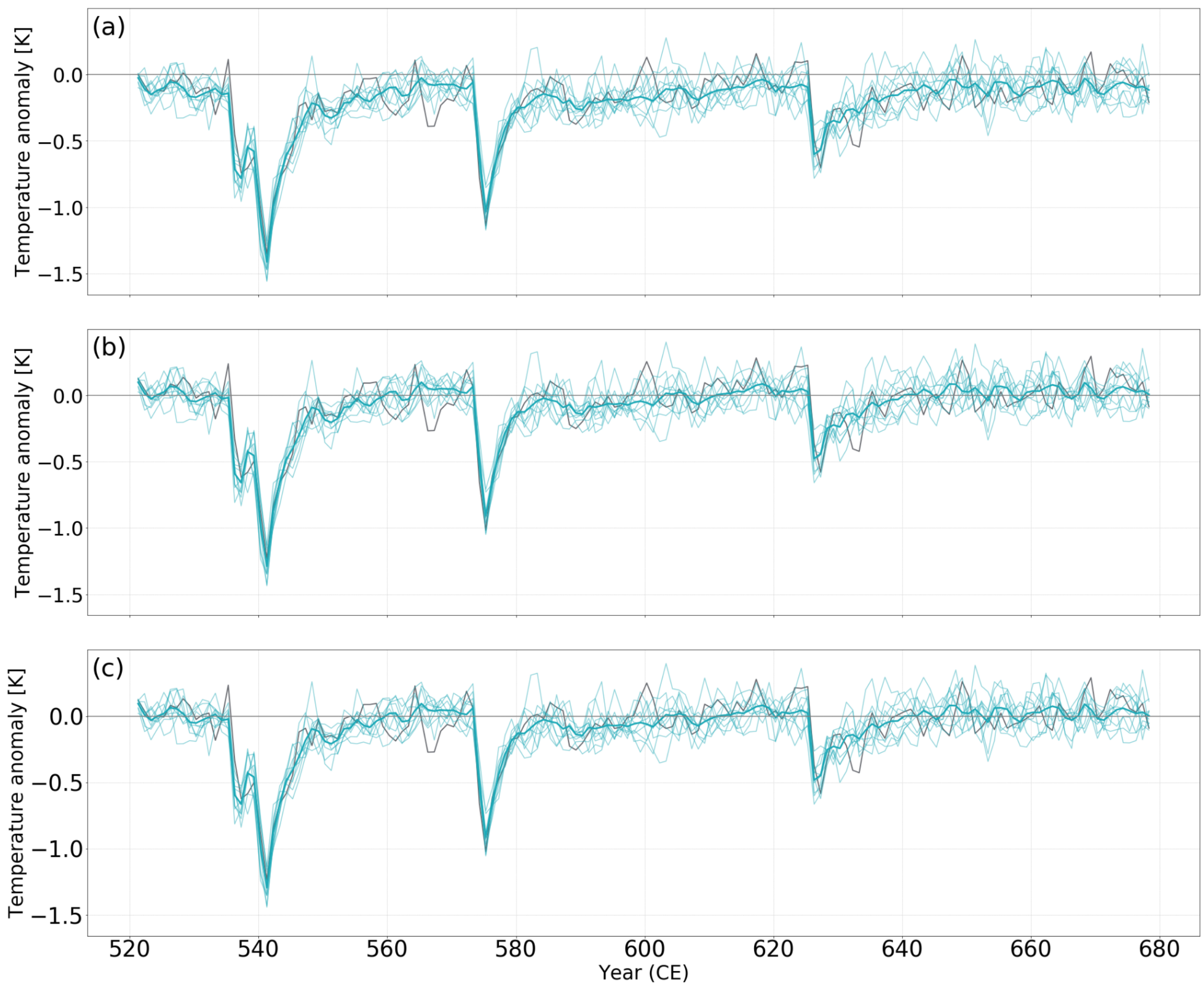


Figure. Ensemble temperature anomalies wrt a) CTR (0 CE), b) 0-1850 CE, and c) 520-535 CE mean of the study period.

point 18, NAO comparison. This indeed is not clear to me why the model shows the opposite of the reconstructions and what we know from recent strong volcanoes with strong positive NAO during winter within the first few years.

A positive NAO and a surface winter warming pattern has been observed after large tropical eruptions in the past (Robock and Mao, 1992; Robock, 2000 and following work). However, not all IPCC models show this signal (Stenchikov et al., 2002; Driscoll et al., 2012 and following papers). The cause for this is still highly debated, f.e., discussing model deficiencies (low top versus high top models; Charlton-Perez et al., 2013), volcanic aerosol forcing details (Toohey et al., 2014), strength of the volcanic eruption and forcing (Bittner et al., 2016a), tropical vs high latitude eruption impacts (Schneider et al., 2009), high internal variability and number of model ensemble members (Bittner et al., 2016b), the role of the ENSO state during the eruption (Coupe and Robock, 2021), up to if the observed signal is due to volcanic eruptions at all (Polvani et al., 2019).

Here, we show that there is a positive NAO in the first winter after the eruptions visible after three out of the four large eruptions during 520-680 CE (old Figure A3). The maps in Figure 3 are a mean of two winters as well as four eruptions times 12 ensemble members, and so the signal is smoothed. We have added these details to the discussion.

point 19: this might be misleading to mention and discuss the non-significant areas. I propose to concentrate on the stat. sign. Differences.

In some cases, like for the NAO and the atmospheric circulation pattern, we would like to show and stress these results due to the ongoing debate in this research field.

References:

Anchukaitis, K.J., Wilson, R., Briffa, K.R., Büntgen, U., Cook, E.R., D'Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B.E. and Hegerl, G., 2017. Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews*, *163*, pp.1-22.

Bittner, M., Timmreck, C., Schmidt, H., Toohey, M. and Krüger, K., 2016a. The impact of wave‐mean flow interaction on the Northern Hemisphere polar vortex after tropical volcanic eruptions. *Journal of Geophysical Research: Atmospheres*, *121*(10), pp.5281-5297.

Bittner, M., Schmidt, H., Timmreck, C. and Sienz, F., 2016b. Using a large ensemble of simulations to assess the Northern Hemisphere stratospheric dynamical response to tropical volcanic eruptions and its uncertainty. *Geophysical Research Letters*, *43*(17), pp.9324-9332.

Büntgen, U., Allen, K., Anchukaitis, K.J., Arseneault, D., Boucher, É., Bräuning, A., Chatterjee, S., Cherubini, P., Churakova, O.V., Corona, C. and Gennaretti, F., 2021. The influence of decision-making in tree ring-based climate reconstructions. *Nature communications*, *12*(1), pp.1-10.

Charlton‐Perez, A.J., Baldwin, M.P., Birner, T., Black, R.X., Butler, A.H., Calvo, N., Davis, N.A., Gerber, E.P., Gillett, N., Hardiman, S. and Kim, J., 2013. On the lack of stratospheric dynamical variability in low‐top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres*, *118*(6), pp.2494-2505.

Coupe, J. and Robock, A., 2021. The Influence of Stratospheric Soot and Sulfate Aerosols on the Northern Hemisphere Wintertime Atmospheric Circulation. *Journal of Geophysical Research: Atmospheres*, p.e2020JD034513.

Driscoll, S., Bozzo, A., Gray, L.J., Robock, A. and Stenchikov, G., 2012. Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. *Journal of Geophysical Research: Atmospheres*, *117*(D17).

Guillet, S., Corona, C., Stoffel, M., Khodri, M., Lavigne, F., Ortega, P., Eckert, N., Sielenou, P.D., Daux, V., Churakova, O.V. and Davi, N., 2017. Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. *Nature geoscience*, *10*(2), pp.123-128.

Luterbacher, J., Werner, J.P., Smerdon, J.E., Fernández-Donado, L., González-Rouco, F.J., Barriopedro, D., Ljungqvist, F.C., Büntgen, U., Zorita, E., Wagner, S. and Esper, J., 2016. European summer temperatures since Roman times. *Environmental research letters*, *11*(2), p.024001.

Neukom, R., Barboza, L.A., Erb, M.P., Shi, F., Emile-Geay, J., Evans, M.N., Franke, J., Kaufman, D.S., Lücke, L., Rehfeld, K. and Schurer, A., 2019. Consistent multi-decadal variability in global temperature reconstructions and simulations over the Common Era. *Nature geoscience*, *12*(8), p.643.

Polvani, L.M., Banerjee, A. and Schmidt, A., 2019. Northern Hemisphere continental winter warming following the 1991 Mt. Pinatubo eruption: reconciling models and observations. *Atmospheric Chemistry and Physics*, *19*(9), pp.6351-6366.

Robock, A. and Mao, J., 1992. Winter warming from large volcanic eruptions. *Geophysical Research Letters*, *19*(24), pp.2405-2408.

Robock, A., 2000. Volcanic eruptions and climate. *Reviews of geophysics*, *38*(2), pp.191-219.

Schneider, D.P., Ammann, C.M., Otto‐Bliesner, B.L. and Kaufman, D.S., 2009. Climate response to large, high‐latitude and low‐latitude volcanic eruptions in the Community Climate System Model. *Journal of Geophysical Research: Atmospheres*, *114*(D15).

Schneider, L., Smerdon, J.E., Büntgen, U., Wilson, R.J., Myglan, V.S., Kirdyanov, A.V. and Esper, J., 2015. Revising midlatitude summer temperatures back to AD 600 based on a wood density network. *Geophysical Research Letters*, *42*(11), pp.4556-4562.

Stenchikov, G., Robock, A., Ramaswamy, V., Schwarzkopf, M.D., Hamilton, K. and Ramachandran, S., 2002. Arctic Oscillation response to the 1991 Mount Pinatubo eruption: Effects of volcanic aerosols and ozone depletion. *Journal of Geophysical Research: Atmospheres*, *107*(D24), pp.ACL-28.

Toohey, M., Krüger, K., Bittner, M., Timmreck, C. and Schmidt, H., 2014. The impact of volcanic aerosol on the Northern Hemisphere stratospheric polar vortex: mechanisms and sensitivity to forcing structure. *Atmospheric Chemistry and Physics*, *14*(23), pp.13063-13079.

Wilson, R., Anchukaitis, K., Briffa, K.R., Büntgen, U., Cook, E., D'arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B. and Hegerl, G., 2016. Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, *134*, pp.1-18.

Zhang, H., Werner, J.P., García-Bustamante, E., González-Rouco, F., Wagner, S., Zorita, E., Fraedrich, K., Jungclaus, J.H., Ljungqvist, F.C., Zhu, X. and Xoplaki, E., 2018. East Asian warm season temperature variations over the past two millennia. *Scientific Reports*, *8*(1), pp.1-11.