Answer to the reviewer 1

We'd like to thank reviewer 1 for his/her effort and very constructive comments, which helped us a lot to improve the manuscript. In the following, we provide point-by-point replies. Reviewer comments are in blue and bold, answers are in black, cited text in italic and new or changed text is marked in red.

Reviewer 1

Section 2.1.1: It might be worth introducing the parameterization scheme for the aerosol microphysical processes in MPI-ESM1.2-LR. As suggested in recent studies (e.g. LeGrande et al., 2016, Nat. Geosci.), some CMIP5-era climate models can produce overly strong volcanic cooling due to unrealistic aerosol microphysics. How is the scheme in MPI-ESM1.2-LR different?

The simulation of aerosol microphysical processes is computationally very expensive therefore only a very few CMIP type models treat aerosol microphysical processes interactively. In the MPI-ESM1.2 model we also do not consider aerosol microphysical processes explicitly. The radiative volcanic forcing is prescribed by monthly mean optical parameters (extinction, asymmetry factor, single scattering albedo) which are considered in the models radiative scheme. The optical parameters are precalculated with the EVA volcanic forcing generator (Toohey et al., 2016) which we explain in detail in section 2.1.2. A multimodel comparison for the Tambora eruption of different aerosol microphysical models (Zanchettin et al., 2016; Clyne et al., 2020) revealed large differences between the different aerosol models and that the more idealized forcing approach with the EVA tool is within the multi model range.

For clarification, we included now in the model description (2.1.1) the following sentences:

"*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 (Toohey et al., 2016), see section 2.1.2.*"

To address the uncertainties in the microphysical processes we have add in the forcing description (2.1.2) the following sentences:

"*The applied volcanic forcing is compiled with the Easy Volcanic Aerosol (EVA) forcing generator (Toohey et al., 2016). EVA provides an analytic representation of volcanic stratospheric aerosol forcing, prescribing the aerosol's radiative properties and primary modes of their spatial and temporal variability. Although Eva represents an idealized forcing approach, its forcing estimates lie within the multi model range of global aerosol simulations for the Tambora eruption (e.g. Zanchettin et al., 2016; Clyne et al., 2020).*

Eruption timing: In L107 it is mentioned that the 1809 eruption is set to occur on Jan 1st of 1809. Recent studies (e.g., Predybaylo et al., 2020, Commun. Earth & Environ) have suggested that the eruption timing may also affect the climate response, especially the ENSO response, due to different circulation conditions and ENSO phases. Since the SOI response is also assessed in Fig. 6, it might make this study more complete to add one extra experiment testing the sensitivity to the eruption timing.

The reviewer raised an important point about the influence of the eruption season on the climate response to large volcanic eruptions. This is of course of particular interest with respect to tropical hydroclimate and ENSO which is discussed widely in the recent literature (e.g., Stevenson et al., 2017; Predybalo et al., 2017, 2020; Zhuo et al., 2021) but is beyond the central scope of this paper and we therefore refrain from doing an extra sensitivity study. It could be of course a topic of a separate study with the focus on a comparison of model data and tropical SST reconstructions for large historic eruptions. For such a study not only additional simulations for different eruption seasons are of interest but also simulations for different initial oceanic background states.

According to reviewer 2 comments we will shift Figure 6 (which we have changed according to your suggestions) to the appendix (Figure S1) and shorten the text about the circulation results, see comment below. We have now included also the “eruption timing“ as a critical point in our discussion:

“ ...with strong tropical forcing but relatively weak (and/or short-lived) NH extratropical forcing, appears to be a possible scenario for the 1809 eruption. A factor which might have an influence on the mismatch between model results and proxy data is the season of the eruption, which has an impact on the climate response to volcanic eruptions (e.g., Stevenson et al., 2017; Predybalo et al., 2020). We have not discussed this uncertainty for the 1809 eruption here as we chose January 1809 as the starting date following most studies which suggest that the eruption probably happened in boreal winter months across 1808 and 1809. Nevertheless the timing of the eruption remains uncertain. Chenoweth (2001) for example dated the 1809 eruption back to March-June 1808 based on a sudden cooling in Malaysian land surface temperature data and a peak cooling of marine air temperature in 1809. To take into account the event season may be necessary to better describe the climate response to volcanic eruptions and should be addressed in further studies. “

Fig. 2: It might be better to use green triangles or other readable colors and symbols to denote the location of the tree-ring proxies. Red can be misleading given that red also represents a high temperature anomaly. The colorbar may also be adjusted to drop the white color to differentiate the missing values.

Thank you very much for your suggestion, we have revised Fig. 2. and use now pink dots for the tree ring locations. According to a comment from the second reviewer we also included a panel of the NH summer temperature anomalies of 1811 instead of 1816 in Fig. 2, and shifted the 1816 NH summer panel to the supplements (Fig. S3).

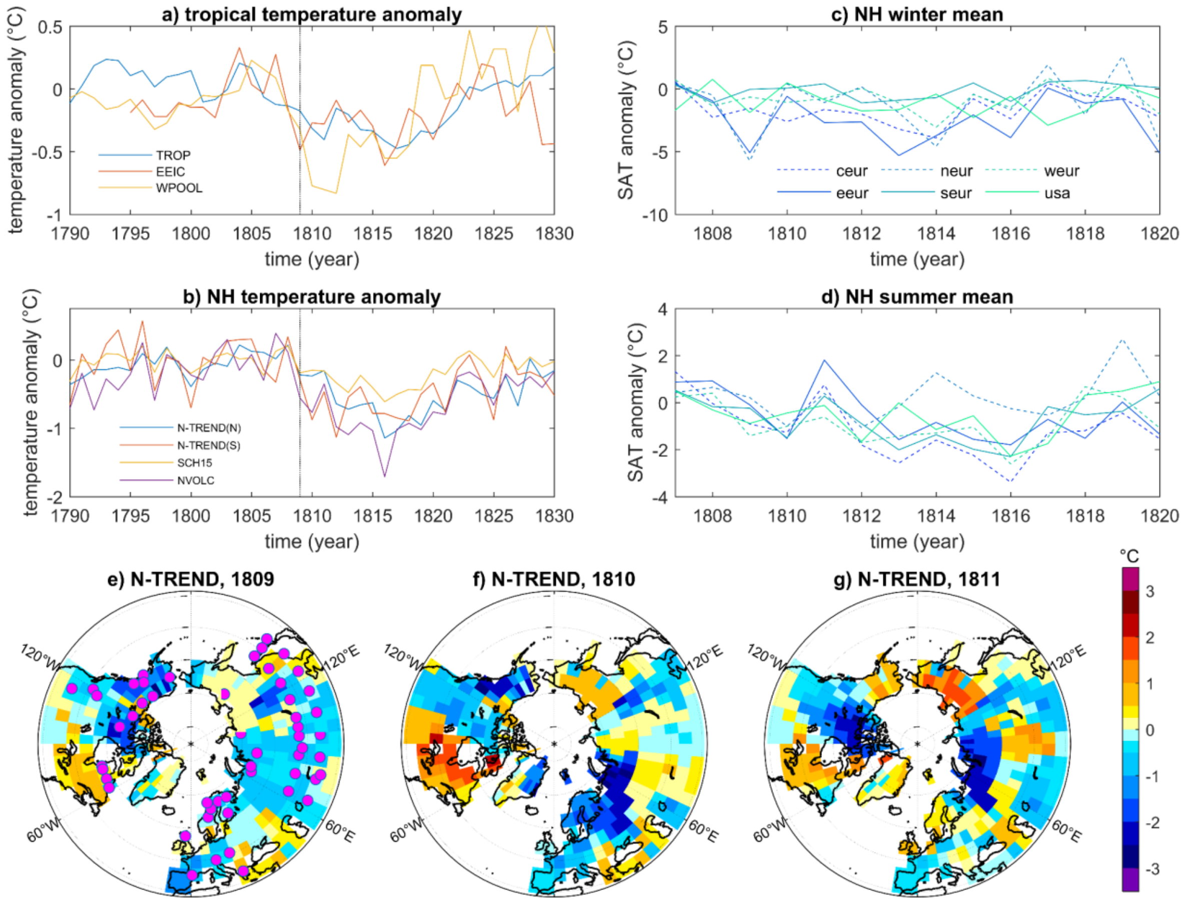

Figure 2: *Observed and reconstructed temperature anomalies around the 1809 volcanic eruption.* ***a****) Reconstructed tropical (30°N–30°S, 34°E-70°W) sea surface temperature (TROP, D'Arrigo et al., 2009), measured tropical marine surface air temperatures from ship logs (EEIC, Brohan et al., 2012) and 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.*

Fig. 5 & 6d: It might be worth showing the relative sea surface temperature (RSST) (Khodri et al., 2017, Nat. Commun.) to highlight the impact of volcanic forcing on ENSO relative to tropical mean cooling in the supplementary information. It may or may not affect the results significantly, but either way it is a valuable assessment.

We have now calculated the relative sea surface temperature according to Khodri et al. (2017) and included an additional plot showing the relative SST in the tropics in the supplements (Figure S4).

We have slightly revised the discussion of Figure 5:

" High is the only experiment where a significant El Nino type anomaly is seen over the Pacific Ocean in boreal summer 1810, while in the other three experiments a slight but non-significant warming appears off the coast of South-American. Looking to the relative SST anomalies as calculated after Khodri et al. (2017) an El Nino type anomaly is seen for all 4 scenarios in boreal summer 1810, while in winter 1809/1810 a significant warming anomaly appears in the central tropical Pacific in all experiments except Best. "

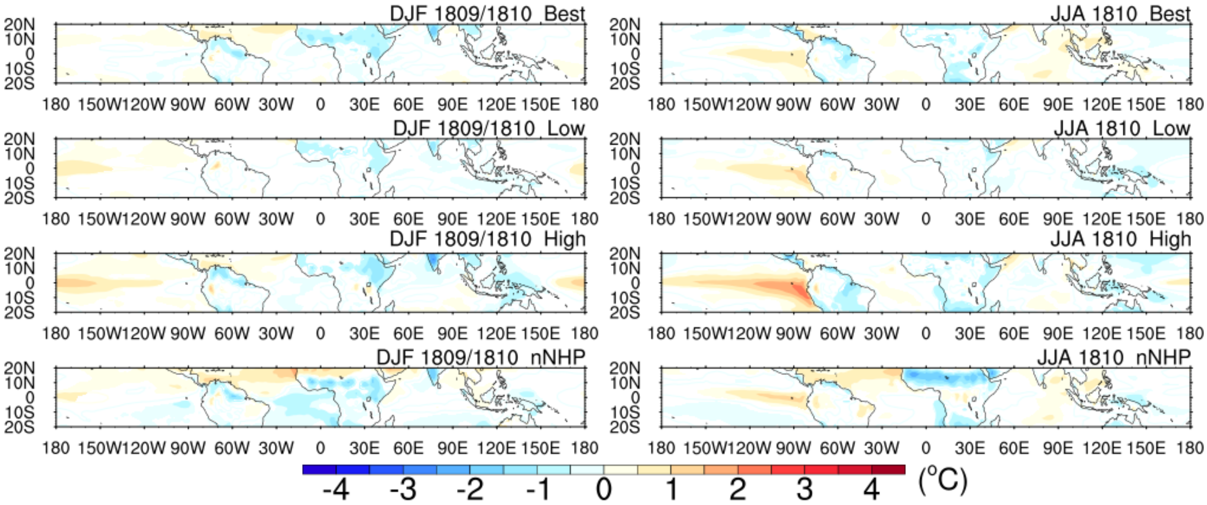

Figure S4: *Seasonal mean relative sea surface temperature anomalies in the tropical regions. Simulated ensemble-mean relative sea surface temperature anomalies for the 1st winter (1809/1810) and the second summer (1810) after the 1809 eruption for the four different MPI-ESM simulations. Shaded regions are significant at the 95% confidence level according to a t-test. Anomalies are calculated with respect to the period 1800-1808.*

Fig. 6: If I am not misunderstanding, the volcanic forcing magnitude in Best, Low, and High experiments can be ranked as High > Best > Low according to Fig. 1, and there's no difference in their meridional structure. I am curious about the reason that the SOI response in the Low experiment seems to lie between that of the Best and High experiments during both winter and summer. Particularly, does the fact that the High and Best experiments show opposite signs during the summer indicate that the SOI response is actually more internally driven than externally forced? Similar doubts exist for other indices that the impact seems not following the same monotonic order of the magnitude of the forcing. It might be good to reorder the bars in the figures as pre, Low, Best, High to highlight the potential impact of magnitude, no matter exists or not.

We have changed the order of the experiments in Figure 6. The revised figure is shown below. Indeed, the revised order better illustrates a possible link between the response of some of the indices and the magnitude of the applied forcing. This seems more evident in NAO, PNA, with noticeably opposite tendencies seen in the two seasons for both indices. The relation is less clear for NPI and SOI, which we interpret as being related to the different ENSO-related dynamics triggered by the different volcanic forcing. Following a comment by reviewer #2, we have toned down the discussion about the circulation indices and put the amended Figure 6 as supplement in the revised manuscript, as a detailed analysis and discussion of the underlying dynamical aspects are beyond the scope of this paper and deserve a dedicated study. 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 in 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). "*

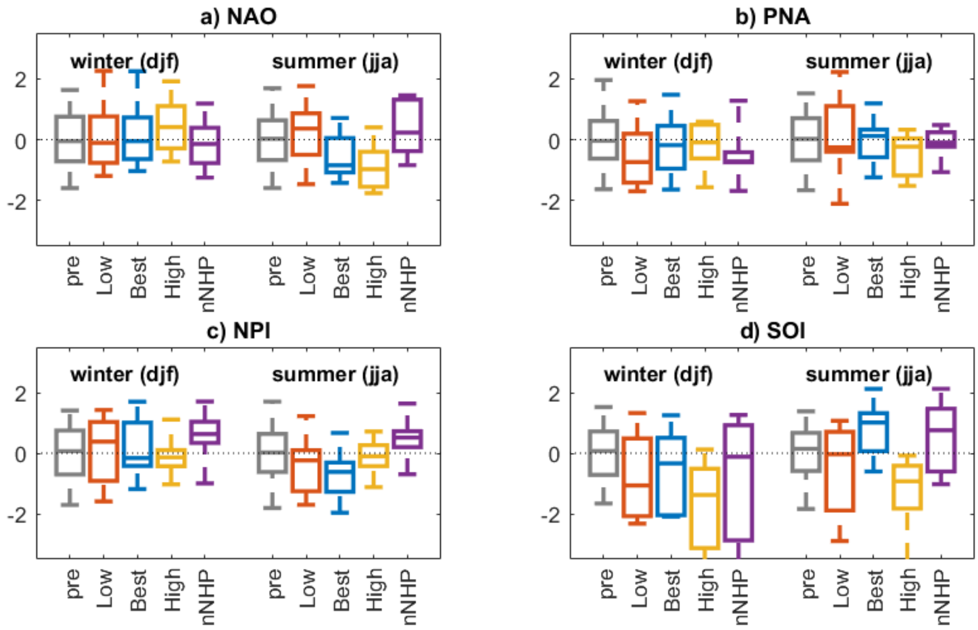

Figure 6 (revised, now Figure S1). *Atmospheric circulation indices. Box-Whisker plots (median, 25th-75th and 5th-95th percentile ranges) of seasonal anomalies of circulation indices: a) North Atlantic Oscillation (NAO), b) Pacific/North American pattern (PNA), c) the North Pacific Index (NPI) and d) the Southern Oscillation Index (SOI) from the model experiments for the first post-eruption winter (1809/1810, DJF) and second post-eruption summer (1810, JJA) following the 1809 eruption. Pre-eruption (1800-1808) data, shown as grey Box-Whisker plots, are used to standardize the indices.*

Fig. 7a and L200-204: In L200-204, it is mentioned that the authors accounted for the sparsity and irregularity in spatial and temporal sampling of the EEIC data, but it is unclear how good the performance of the processing is, and the authors still see overall dampened tropical SST anomalies in EEIC compared to model simulations in Fig. 7a. I was wondering what the comparison would look like if compare the mean of the model simulated SST anomalies over grid cells nearest to the locales of the EEIC logs to the mean of the original EEIC data. Similar strategy might be worth taking for other comparisons if the observations/reconstructions are available over multiple sites instead of a processed regional mean.

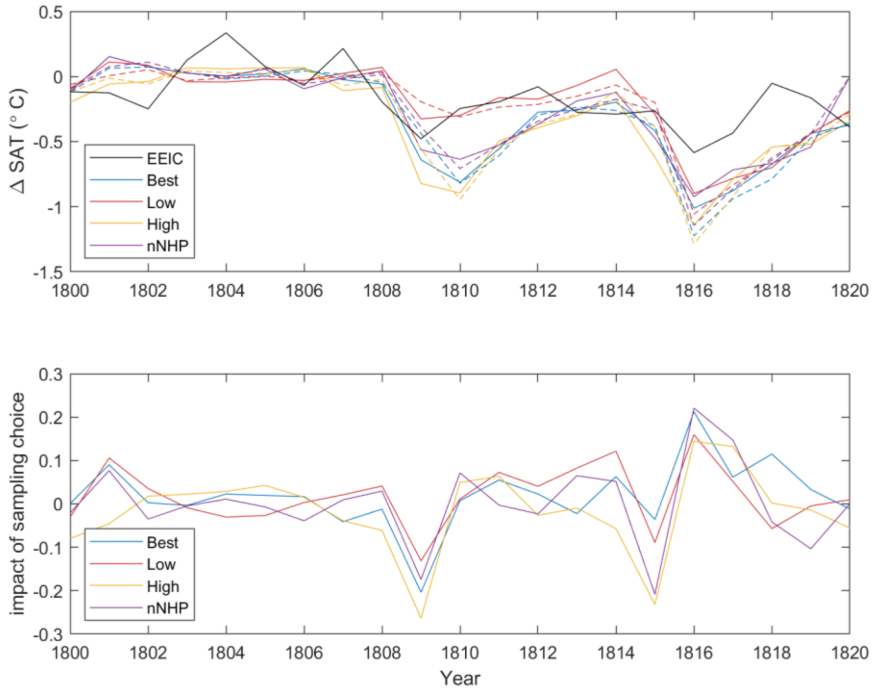

__Figure S5:__ Impact of limited sampling on the average tropical surface air temperature (SAT) anomalies in the MPI-ESM simulations. (top), EEIC and MPI-ESM tropical SAT anomalies, the solid colored lines show the overall simulated tropical (30°S-30°N) average values, while the dashed lines show the average of model points corresponding to the EEIC measurements. (bottom). Difference between the model values (Full sampling - EEIC sampling).

Indeed, as suggested by the reviewer, the sampling of the ship based measurements does introduce some uncertainty into the average values. Figure S5 (above) shows the results of two sampling strategies used on the model data: the solid lines show the average tropical SAT anomaly, while the dashed lines show the average of the model sampled at the point and time of each EEIC measurement. The difference between the two is shown in the bottom panel. This analysis shows that the sampling acts to diminish the temperature anomaly in 1809 compared to the full tropical average. The magnitude of this impact is around 10-30% in 1809. The impact in 1810, at the peak of the tropical cooling (in the model) is quite small. For the Tambora eruption, sampling impacts the mean differently in 1815 and 1816.

This comparison highlights, and provides a rough estimate of the impact that limited sampling in the EEIC measurements has on estimating the tropical mean temperature anomaly. We have included the figure above in the supplemental information (Figure S5), and refer to it in the text. This uncertainty in the measurements does not however significantly impact the model overestimation of the tropical cooling in 1810.

We have added text (line 376): "When the model results are sampled at the locations and times of the EEIC measurements (Fig. S5), the mean negative temperature anomalies in 1809 are 10-30% smaller, with Best, High and nNHP experiments all producing anomalies similar to that of the EEIC measurements. For the 1810-1812 period, the sampling makes little difference compared to the full tropical average, with Best, High and nNHP experiments all showing larger negative temperature anomalies than the EEIC measurements.

Fig. 9: It seems that the strategy mentioned above is taken here in Fig. 9 as the model simulations are "similarly sampled". Perhaps can add an extra column for the visualization similar to Fig. 7a, comparing EEIC to ensemble means, but for two separated regions. Is it overall a better agreement

in Fig. 9 than in Fig. 7a? If so, does it mean the sparsity of the EEIC logs is not well accounted for as mentioned in L200-204?

The reviewer is correct that the EEIC sampling of the model is used in Fig. 9. The inclusion of each ensemble member in this plot allows for a better appreciation of the variability in each ensemble of simulations, of which the sampling plays a part. As suggested, we propose to include a figure displaying the impact of sampling on the tropical average of Figure 7a. in the supplements. As discussed above, while the sampling clearly makes an impact on the calculated averages, it does not appear to significantly change the interpretation of the comparison between the EEIC measurements and the model simulations.

Fig. 13 & 14: What is the rationale that the anomalies are calculated with respect to the years 1806-1820 here instead of 1800-1808 as in previous figures? The decision will largely affect the model-data comparison on the response to volcanic forcing.

We agree that this would be more consistent but unfortunately the station data are sparse and irregular and often only partly available between 1800- 1808. Hence, we decided to take into account the full period as reference period.

L527: a typesetting issue (delta-18-O)

The typo is corrected.

References:

Anchukaitis, K.J., et al: 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.

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

Clyne, M., et al.: Model physics and chemistry causing intermodel disagreement within the VolMIP-Tambora Interactive Stratospheric Aerosol ensemble, Atmos. Chem. Phys., 21, 3317–3343, 2021.

D'Arrigo, R. et al..: The reconstructed Indonesian warm pool sea surface temperatures 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.

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

Khodri, M., et al.: Tropical explosive volcanic eruptions can trigger El Nino by cooling tropical Africa, Nat. Commun., 8(1), 778, doi:10.103 8/s41467-017-00755-6, 2017.

LeGrande, A. N., Tsigaridis, K., and Bauer, S. E.: Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections, Nat. Geosci., 9, 652–655,https://doi.org/10.1038/ngeo2771, 2016.

Predybaylo, E., G. L. Stenchikov, A. T. Wittenberg, and F. Zeng, 2017: Impacts of a pinatubo-size volcanic eruption on enso. J. Geophys. Res., 122, 925-947, https:/doi.org/10.1002/2016jd025796.

Predybaylo, E., Stenchikov, G., Wittenberg, A. T., and Osipov, S.: El Niño/Southern Oscillation response to low-latitude volcanic eruptions depends on ocean pre-conditions and eruption timing. Communications Earth & Environment, 1(1), 1-13, 2020.

Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. L., Tomas, R. A., and Gao, C.: Role of eruption season in reconciling model and proxy responses to tropical volcanism, P. Natl. Acad. Sci. USA, 114(8), 1822-1826, doi:10.1073/pnas.1612505114, 2017.
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., etLL: 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.
Zhuo, Z., Kirchner, I., Pfahl, S., and Cubasch, U.: Climate impact of volcanic eruptions: the sensitivity to eruption season and latitude in MPI-ESM ensemble experiments, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-260, in review, 2021.