Paleoceanography and Paleoclimatology

RESEARCH ARTICLE
10.1029/2020PA004128

Key Points:

- Paleo data assimilation products capture the magnitude and persistence of volcanic cooling estimates in MXD tree-ring reconstructions
- The magnitude and persistence of regional cooling varies significantly between data assimilation products, requiring further explanation
- The persistence of the cooling in CESM-LME is shorter than in the data assimilation products, but the magnitude estimates are similar

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. Tejedor,
etejedor@albany.edu

Citation:

Tejedor, E., Steiger, N., Smerdon, J. E., Serrano-Notivoli, R., & Vuille, M. (2021). Global temperature responses to large tropical volcanic eruptions in paleo data assimilation products and climate model simulations over the last millennium. *Paleoceanography and Paleoclimatology*, *36*, e2020PA004128. https://doi.org/10.1029/2020PA004128

Received 29 SEP 2020
Accepted 26 MAR 2021

Author Contributions:

Conceptualization: E. Tejedor, J. E. Smerdon, M. Vuille
Data curation: E. Tejedor
Formal analysis: E. Tejedor, N. Steiger, R. Serrano-Notivoli
Funding acquisition: J. E. Smerdon, M. Vuille
Investigation: E. Tejedor, N. Steiger, R. Serrano-Notivoli

Global Temperature Responses to Large Tropical Volcanic Eruptions in Paleo Data Assimilation Products and Climate Model Simulations Over the Last Millennium

E. Tejedor[1], N. Steiger[2,3], J. E. Smerdon[2], R. Serrano-Notivoli[4], and M. Vuille[1]

[1]Department of Atmospheric and Environmental Sciences, University at Albany (SUNY), Albany, NY, USA, [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA, [3]Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel, [4]Department of Geography, Autonomous University of Madrid, Madrid, Spain

Abstract Large volcanic eruptions are one of the dominant perturbations to global and regional atmospheric temperatures on timescales of years to decades. Discrepancies remain, however, in the estimated magnitude and persistence of the surface temperature cooling caused by volcanic eruptions, as characterized by paleoclimatic proxies and climate models. We investigate these discrepancies in the context of large tropical eruptions over the Last Millennium using two state-of-the-art data assimilation products, the Paleo Hydrodynamics Data Assimilation product (PHYDA) and the Last Millennium Reanalysis (LMR), and simulations from the National Center for Atmospheric Research Community Earth System Model-Last Millennium Ensemble (NCAR CESM-LME). We find that PHYDA and LMR estimate mean global and hemispheric cooling that is similar in magnitude and persistence once effects from eruptions occurring in short succession are removed. The estimates also compare well to Northern-Hemisphere reconstructions based solely or partially on tree-ring density, which have been proposed as the most accurate proxy estimates of surface cooling due to volcanism. All proxy-based estimates also agree well with the magnitude of the mean cooling simulated by the CESM-LME. Differences remain, however, in the spatial patterns of the temperature responses in the PHYDA, LMR, and the CESM-LME. The duration of cooling anomalies also persists for several years longer in the PHYDA and LMR relative to the CESM-LME. Our results demonstrate progress in resolving discrepancies between proxy- and model-based estimates of temperature responses to volcanism, but also indicate these estimates must be further reconciled to better characterize the risks of future volcanic eruptions.

1. Introduction

Volcanic eruptions are one of the main natural causes of radiatively forced climatic changes over the Common Era, producing fluctuations on interannual to multi-year timescales. The analysis and interpretation of such changes has resulted in a wide body of research, especially following the influential work of H.H. Lamb (Lamb & Stanley, 1970), and expanding into many efforts to characterize and model the climate response to volcanism (e.g., Robock, 2000; Sigl et al., 2015; Zanchettin et al., 2016). It is now well understood that explosive volcanic eruptions can alter climate by injecting large amounts of sulfur-containing gases into the stratosphere, such as SO_2 and H_2S, leading to the formation of liquid sulfate aerosols (Toohey & Sigl, 2017). These aerosols consequently scatter incoming solar radiation and absorb infrared radiation, which in turn leads to a net decrease in radiation reaching the Earth's surface and associated global cooling (Robock, 2000). The introduction of satellite observations led to rapid advancements in our understanding of how stratospheric aerosols affect global temperature, particularly since the El Chichon eruption in 1983 (e.g., Robock, 1983; Robock & Matson, 1983). This event served as a natural experiment to test and develop monitoring instruments and improved representations of atmospheric chemistry in climate models, which in turn improved simulations of the climate effects due to volcanic aerosols in the stratosphere (Rampino & Self, 1984). With the exception of the eruption of Mt. Pinatubo in 1991, however, the volcanic events monitored during the 20th century are of a much smaller magnitude than many of those that have occurred over the Common Era. Further improvement in our understanding of the impacts and dynamics of large volcanic eruptions is therefore dependent on high-resolution proxy reconstructions and model simulations of the larger volcanic events that occurred before the widespread availability of instrumental records. The

Methodology: E. Tejedor, N. Steiger, J. E. Smerdon, R. Serrano-Notivoli, M. Vuille
Project Administration: M. Vuille
Supervision: J. E. Smerdon, M. Vuille
Validation: E. Tejedor, J. E. Smerdon, M. Vuille
Visualization: E. Tejedor
Writing – original draft: E. Tejedor
Writing – review & editing: E. Tejedor, N. Steiger, J. E. Smerdon, R. Serrano-Notivoli, M. Vuille

significance of such work is that despite the quiescence of large volcanic eruptions over the last century, they may occur at any time in the future and could significantly disrupt global society (Papale, 2018; Toohey et al., 2019) through modifications to hydroclimate (Tejedor et al., 2021) and the impacts on agriculture due to short-term cooling (Engvild, 2003). It is therefore critical that proxy records of past climate and climate model simulations are broadly interrogated and used to derive a cohesive picture of the impacts and dynamics associated with past large eruptions. Such investigations are essential for anticipating future volcanic hazards and risks.

Recent proxy-based studies of the temperature impacts of volcanic eruptions over the last millennium have focused primarily on paleoclimatic records from the Northern Hemisphere (Anchukaitis et al., 2017, 2012; D'Arrigo et al., 2009; Esper, et al., 2015; Stoffel et al., 2015) and are predominantly limited to reconstructions of climate during the growing season (June, July, and August or JJA) (Schneider et al., 2017). Proxy-based characterizations of the temperature response to volcanic events over the Last Millennium are therefore limited in the Southern Hemisphere and to a lesser extent in the tropics. Multiple studies have used the collection of models from the fifth phase of the Coupled Modeling Intercomparison Project (CMIP5, Taylor et al., 2012), such as the National Center for Atmospheric Research-Community Earth System Model-Last Millennium Ensemble (NCAR CESM-LME) to partially address these limitations by analyzing the modeled temperature response to volcanism globally (e.g., Colose et al., 2016; Stevenson et al., 2016). Despite these advances, there are still unresolved questions about the robustness of large-scale proxy-estimated cooling, particularly in tree-ring estimates, and the magnitude of cooling simulated by climate models (e.g., Anchukaitis et al., 2012; Esper et al., 2013, 2015; Mann et al., 2012). It is indeed still unclear whether differences between the temperature response seen in proxy data, such as in high-resolution maximum latewood density (MXD) reconstructions, and that which is simulated by climate models like CESM, arise from biases in the proxy signal, inaccuracies in the volcanic forcing reconstructions, or flaws in the model simulations of the climate response (Lücke et al., 2019).

In addition to studies of large-scale surface temperature responses to volcanic eruptions, volcanic impacts on ocean dynamics have also been analyzed. These studies have primarily focused on changes in interannual to multidecadal climate modes, such as the Atlantic Multidecadal Oscillation (AMO, e.g., Knudsen et al., 2014; Swingedouw et al., 2017; Zanchettin et al., 2012), or Indo-Pacific variability (e.g., Maher et al., 2015 and references herein). Characterizations of enhanced probabilities of El Niño or La Niña occurrences following large volcanic eruptions have been a recurring and important topic of debate, based mostly on studies that primarily analyze either model experiments (Stevenson et al., 2016) or proxy reconstructions (Dee et al., 2020; Li et al., 2013; Wahl et al., 2014). These studies have generally produced contradictory results, with some finding that volcanic eruptions enhance the likelihood of El Niños, La Niñas, or have no discernable impact on the state of the tropical Pacific. There is nevertheless some evidence that the state of the tropical Pacific prior to the eruption may be an important control on its response to volcanic eruptions, which has not always been a focus of previous studies. Accounting for this preconditioning may therefore help reconcile the currently contrasting conclusions about how the tropical Pacific responds to eruptions (e.g., Predybaylo et al., 2017; Wahl et al., 2014). Reconciling these findings, and the more general study of the dynamical ocean responses to volcanism, is important for characterizing how volcanic impacts, such as droughts or flooding, may be driven by changes in these ocean modes. But these studies are also important because alterations in the ocean dynamics may be one mechanism by which volcanic impacts persist in the climate system after the direct effects of volcanic aerosols have dissipated (e.g., Tejedor et al., 2021).

Given the above-described state of the science, it is crucial to better characterize the magnitude, persistence and spatial extent of climatic changes due to large volcanic eruptions. These efforts will improve our understanding of the dynamics that cause these changes and their representation in climate models. Such improvements will thus enhance risk assessment and preparedness for future volcanic eruptions. These motivations drive the work that we present herein, which uses recently derived paleoclimate data assimilation products that combine high-resolution proxy records with information from last-millennium climate model simulations. We specifically employ two new data assimilation products to evaluate the global temperature response to large tropical volcanic eruptions. The Paleo Hydrodynamics Data Assimilation product (PHYDA; Steiger et al., 2018) is a new, publicly available, and global data set that provides reconstructions of hydroclimate, temperature and associated dynamic climate variables for the past two thousand years. We

compare the PHYDA to another new and publicly available data assimilation product, the Last Millennium Reanalysis product (LMR) (Tardif et al., 2019), which was recently used to investigate disagreements between model and proxy estimates of global temperature responses to volcanism (Zhu et al., 2020). These two state-of-the-art data assimilation products, which provide global representations of volcanic responses, are compared to simulations from the NCAR CESM-LME, which served as the prior for the PHYDA. We also employ the most recent volcanic stratospheric sulfur injection reconstruction ("eVolv2k_v2"; Toohey & Sigl, 2017), which considerably refines the dating of certain eruptions and improves the estimates of eruption magnitudes over the last millennium.

Equipped with the new and global paleo data assimilation products, the CESM-LME and the most recent volcanic reconstruction, this study addresses the following questions: 1) What are the large-scale temperature responses to volcanic events in the two data assimilation products and how do they compare to other proxy-derived estimates? 2) What do the data assimilation products tell us about past responses in ocean dynamics to volcanism? and 3) How do the estimated temperature and oceanic responses in the data assimilation products compare to the modeled volcanic responses?

2. Data and Methods

2.1. Data Assimilation Products and Model Data

We use the PHYDA (Steiger et al., 2018), which provides global reconstructions of hydroclimate and temperature along with associated dynamic climate variables over the past two thousand years (Steiger et al., 2018). PHYDA combines 2,978 global high-resolution paleoclimatic time series with the physical constraints of an atmosphere-ocean climate model (the CESM-LME simulation number 10 Otto-Bliesner et al., 2015). We use a 100-member PHYDA ensemble, randomly selected from the full 1,000-member PHYDA ensemble, to provide robust uncertainty estimates for the derived reconstruction. We limit our analysis to the 1000–1850 CE interval, as the uncertainties in PHYDA significantly increase prior to 1000 CE in many regions. We select the PHYDA-reconstructed 2-m surface temperature data during boreal summer (JJA), which is reconstructed on approximately a 2° latitude-longitude grid. Anomalies of temperature (in °C) were calculated with respect to 1000–1850 CE. Additionally, we selected two reconstructed sea surface temperature indices from the boreal winter season (December, January, February or DJF): The El Niño-Southern Oscillation (ENSO) index and the AMO index as potential indicators of both rapid (inter-annual) and long-term (decadal) climate system responses due to volcanic impacts. We also note that the PHYDA includes a filtered selection of proxies by season, in which only proxies with strong annual, DJF, or JJA signals were used to reconstruct each season.

We additionally use LMR (Hakim et al., 2016; Tardif et al., 2019), which is also a data assimilation-based global climate reconstruction of the past two thousand years. Similar to PHYDA, LMR also reconstructs multiple climate variables, including near-surface air temperature (as used herein and processed the same as PHYDA). The employed data assimilation methodologies used to derive the LMR and PHYDA products are similar, both based on the variant of the Kalman filter and "offline" approach described by Hakim et al. (2016). Specific methodological differences do exist, however, including the formulation of the transfer functions that were used to relate the proxies and climate model output. LMR additionally used the CCSM4 last-millennium climate simulation as its prior (Landrum et al., 2013), while the PHYDA used the CESM-LME. The two products both assimilated proxies from the PAGES2k database (Emile-Geay & Consortium, 2017), but LMR did not include the many additional dendroclimatic records that were assimilated in PHYDA. The selection criteria for which proxies were assimilated and whether they were used to inform temperature, hydroclimate, or both, were also different between the two datasets. Finally, the CESM-LME prior was bias-corrected based on instrumental climatology before the assimilation process in PHDYA, while the LMR did not employ bias correction. These details are important with regard to any differences in the estimated responses that we characterize for the PHYDA and LMR. To the degree that differences exist, there are multiple potential reasons tied to methodological and input data choices. While understanding the impact of these choices is critical for vetting the two data assimilation products, a comprehensive diagnosis of the causal differences between the two products is beyond the scope of this study.

Our model analysis employs the CESM-LME (Otto-Bliesner et al., 2015), from which we select the same climate variables with the same resolution as we use from PHYDA. Ensemble member 10 was used for the prior in PHYDA, ensuring that any differences in the estimated volcanic responses between CESM and PHYDA are a result of the proxy information that is assimilated in PHYDA. To assess the robustness of the CESM-LME results, we additionally analyze all of the LME simulations. While the focus herein is on the temperature responses in PHYDA and the CESM-LME, the hydroclimate response to volcanism in the two datasets has also been investigated (Tejedor et al., 2021).

Data assimilation and model products were compared to two state-of-the-art Northern Hemisphere tree-ring reconstructions that target the boreal summer: 1) the N-TREND reconstruction, which combines data from tree-ring width (TRW) measurements and Maximum Latewood Density (MXD) (Anchukaitis et al., 2017); and 2) the NH-MXD, which relies exclusively on MXD data and removes the temporal persistence associated with the biological memory effect (Schneider et al., 2015). The N-Trend reconstruction is a spatially resolved warm season (May-August) reconstruction that spans 750–1988 CE, which characterizes a consistent imprint of volcanism in NH temperature anomalies (Anchukaitis et al., 2017). Because it includes TRW records, it also likely includes a biological memory effect (Esper et al., 2015) in the initial temperature response to volcanic events (Zhu et al., 2020). The NH-MXD reconstruction is derived exclusively from regional MXD chronologies that were used to target mean boreal summer Northern Hemisphere temperatures from 600 to 2000 CE (Schneider et al., 2015). This reconstruction has been argued to include less of a biological memory effect relative to tree-ring width reconstructions (Esper et al., 2015). Additionally, the MXD chronologies have been used to develop an independent estimate of large volcanic events over the Last Millennium (Schneider et al., 2017).

We use the previously developed dendroclimatic products as a ground truth for our analyses, which allows for a more impartial comparison between the surface temperature responses to volcanism in different assimilation products and model simulations. We then choose to focus primarily on the boreal summer season in the other datasets because it is the target reconstructed by the N-TREND and NH-MXD datasets, which have been argued to be the most accurate tree-ring reconstructions for assessments of volcanic responses over the last millennium (Esper et al., 2018). We also note that LMR only targets annual mean climate fields. The focus on boreal summer thus requires that we compare annual LMR results to the JJA variables in PHYDA and CESM. Comparisons between the global average time series in the PHYDA and LMR products nevertheless suggest that LMR (Annual) and PHYDA (JJA) (1000–1900 CE; $r = 0.75$ $p < 0.01$) actually agree better than LMR (Annual) and PHYDA (Annual) (1000–1900 CE; $r = 0.58$ $p < 0.01$). Additionally, the results from Zhu et al. (2020), in which the authors specifically reconstruct the JJA season, show a very similar temperature response to volcanic events when computing the annual LMR or the JJA LMR. Despite these arguments for choosing JJA as the principal target throughout our study, we additionally analyze the annual responses in PHYDA and CESM.

A final note is necessary regarding the two data assimilation products. Both the PHYDA and LMR were constructed as offline data assimilation products. One implication of this approach is that while the model simulations that are incorporated as the prior have explicit temporal histories tied to their forcing data, the temporal information is not incorporated in the data assimilation methodology, that is, the timing of climate events such as volcanic eruptions or trends over specific periods (e.g., 20th-century warming) are not dictated by the prior. Consequently, all temporal structure in the PHYDA or LMR products is tied specifically to information contained in the assimilated proxy networks (Steiger et al., 2018; Tardif et al., 2019).

2.2. Volcanic Forcing Index

To identify volcanic events in the PHYDA and LMR, we use the most recent global volcanic forcing reconstruction, named "eVolv2k_v2" (Toohey & Sigl, 2017), which is based on a suite of ice-core records from Greenland and Antarctica. This database contains estimates of the magnitudes and approximate source latitudes of major volcanic stratospheric sulfur injection (VSSI) events from 500 BCE to 1900 CE. The eVolv2k_v2 reconstruction is a significant improvement over prior estimates, which relied on volcanic forcing reconstructions with greater chronological uncertainty (Crowley & Unterman, 2013; Gao et al., 2008; Sigl et al., 2015). Accurately selected eruption dates are crucial for resolving sub-annual to annual volcanic responses in the proxy data and to obtain a precise state of the global climate response to large volcanic

events, especially when considering both proxy and model data. This is in contrast to stand-alone model simulations, which, by construction, use the dates of the eruptions as imposed by the adopted forcing history.

We select those volcanic events located within the tropics (25°N to 25°S) with magnitudes greater than the Mt. Pinatubo eruption (VSSI > 8.78 Tg S; Toohey & Sigl, 2017) from 1000 CE to 1850 CE. A total of 19 eruptions during the LM satisfy this criterion in the eVolv2k_v2 reconstruction. Because we select these events exclusively using the eVolv2k_v2 data, they will be referred to herein as the eVolv2k_v2 Tropical Volcanic Events or TVE_{ev}. The focus on tropical eruptions is motivated by the known climatic signal of large tropical volcanic events, related to seasonal variations in the ITCZ, which facilitates the transfer of aerosols between hemispheres (Kravitz & Robock, 2011; Oman et al., 2006). To assess the potential bias induced by "double-event" occurrences, we also repeat our analyses for a 13-event subset of TVE_{ev} that only selects the second event when two events occur within 10 years (e.g., we included the event in 1458 but not the one in 1453, see Table S1 in the supporting information). To perform the same analyses using CESM, we select the 19 or 13 equivalent volcanic events in the Gao et al. (2008) reconstruction because it was used to force the CESM LME. For those cases in which the dates do not align between the two volcanic forcing reconstructions, we select the Gao et al. (2008) event that is closest to the selected Toohey and Sigl (2017) event in the data assimilation products. To differentiate between the events selected based exclusively on the eVolv2k_v2 reconstruction, the collection of events selected from the Gao et al. (2008) reconstruction that approximate the occurrences in the eVolv2k_v2 reconstruction will be noted as TVE_{Gao}. The 19 and 13 selected events for the TVE_{ev} and TVE_{Gao} collections are listed in Table S1 together with some of the event characteristics.

2.3. Analysis Methods

We use a Superposed Epoch Analysis (SEA; Hegerl et al., 2003) approach to analyze the global temperature response to tropical volcanic events. SEA is a common approach used to quantify the climatic response to specific volcanic events (Gao & Gao, 2017; Schneider et al., 2017). Annual and seasonal temperature data are sorted into categories dependent on the date of volcanic events (Table S1 in the supporting information). The year of a volcanic eruption is assigned as year 0, and then the temperature data are extracted for the five years prior to the eruption and for 20 years post-eruption, to obtain a SEA matrix (number of volcanic events × 26 years). The temperature anomalies for the 20 post-eruption years are calculated as the departures from the pre-eruption mean (reference period averaged over years −5 to −1) and then averaged across all 19 or 13 events, thereby obtaining a 26-years composite of temperature anomalies averaged over all of the events. The reference period of 5 years ensures that our results are not affected by changes in the mean background state as a result of low-frequency changes during the LM.

Statistical significance testing of the composite anomalies is performed by means of a Monte-Carlo simulation, based on the null hypothesis that there is no association between the eruptions and the temperature response (i.e., no significant difference in temperature during the years following eruptions, when compared to non-eruption years). Pseudo-event years are randomly chosen from all years in which eruptions did not occur for each approach (TVE_{ev} and TVE_{Gao}, see Table S1 in the supporting information), and temperature anomalies for years −5 to +20 are calculated in the same way as was done for the years before and after real eruptions, thus creating 10,000 randomly generated composite matrices. Finally, we create a random composite distribution for each column in the matrix (i.e., for each year from year −5 to year +20, relative to the eruption year 0). We used these distributions to test the significance of the actual composites at the 95% confidence level. To assess the global temperature response in space, this method is applied separately for each grid cell included in the 100-member PHYDA ensemble, that is, each ∼2° grid cell is treated as an independent reconstruction, yielding a total of 13,824 grid cells. The same process is applied to the CESM simulation, including the same number of grid cells, and the LMR.

The largest volcanic eruptions of the 20th century (Agung 1963, El Chichon 1982, and Mt. Pinatubo 1991) coincided with prevailing El Niño conditions (Ménégoz et al., 2018). Some studies analyzing the impacts of volcanic events on climate have thus suggested the need to remove the ENSO-signal prior to the analysis, by either detrending the data (D'Arrigo et al., 2009) or considering the residuals after regressing the data against ENSO (Gao & Gao, 2018; Lehner et al., 2016). We do not remove the ENSO signal from the variables

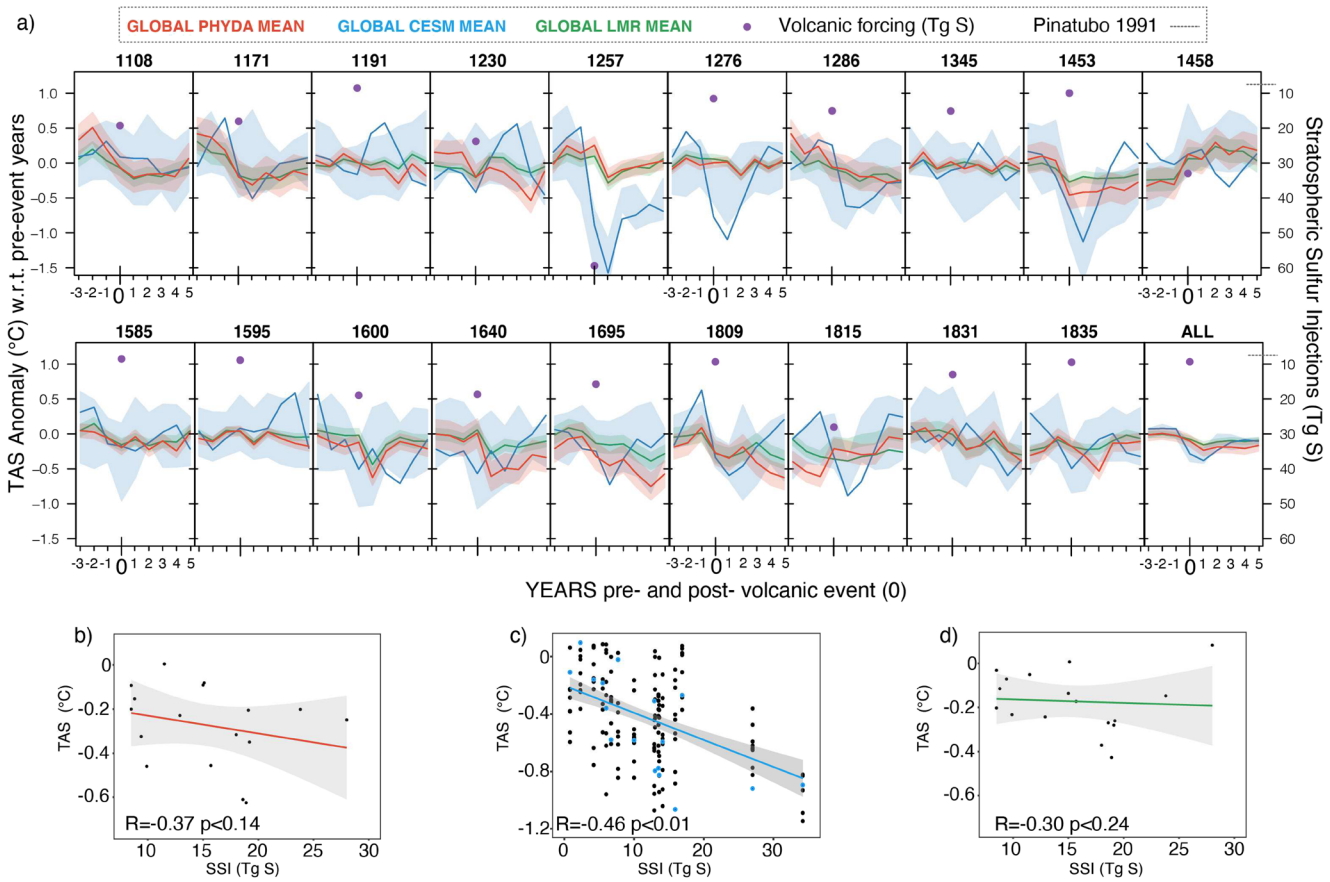

Figure 1. (a) Global temperature response, with respect to the 5-years pre-event mean, to Last Millennium large tropical volcanic events in PHYDA (JJA season), CESM (ensemble member number 10) (JJA season) and LMR (Annual). Bold lines represent the mean and the envelopes represent the uncertainty of the reconstruction ($\pm 1\sigma$) based on 100 PHYDA ensemble members (see text for details) and 20 LMR ensemble members. Bold lines for CESM represent the ensemble member 10 response and the blue shading represents the spread in the additional 9 CESM ensemble members. Purple dots represent the eruption magnitude, expressed by the stratospheric sulfur injection (in Tg S), of each volcanic event as estimated by Toohey and Sigl (2017). For CESM, the volcanic events from Gao et al. (2008) have been temporally aligned to match eVolv2k_v2 dating from Toohey and Sigl (2017). The magnitude of the stratospheric sulfur injections estimated by Gao et al. (2008) are not shown. (b) Spearman correlation between the globally averaged near-surface air temperature response in PHYDA and eVolv2k_v2 magnitude estimates. The events 1458 and 1257 are not included in the analysis, the first because its magnitude is likely shifted by the event in 1453, and the latter because it is normally considered an outlier. (c) Same as in (b) except using CESM and Gao et al. (2008). 1257 is not included in the analysis. (d) Same as in (b) except using LMR. Note different scales of *y*- and *x*-axis in (c) as compared to (b) and (d). CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE, Tropical Volcanic Events.

studied, as the median value of the PHYDA Niño 3.4 index indicates neutral ENSO conditions during the DJF seasons prior to the selected eruption events (Figure S1 in the supporting information).

3. Results

3.1. Temporal Evolution of the Global Temperature Response to TVE_{ev} and TVE_{Gao} Over the Last Millennium

The global temperature response to each year within the TVE_{ev} and TVE_{Gao} collections for the three analyzed datasets is shown in Figure 1 and Table 1. Overall, a global cooling response begins in the year of the volcanic event, with the most pronounced signal occurring in the year following the eruption (year +1). In general, the cooling following the eruption is coherent across all events, although inconsistencies in the peak year of cooling are evident in the 1600, 1640, and 1835 events. In the data assimilation products (PHYDA and LMR), the largest global cooling (mean) with respect to the five previous years is observed after the Huaynaputina eruption in 1600 CE (−0.63°C and −0.44°C, respectively). Additionally, the Parker eruption in 1630 CE (−0.61°C and −0.28°C), the unidentified event of 1695 CE (−0.46°C, −0.16°C), and

Table 1
Peak Global Mean Temperature Anomaly Following TVE_{ev} or TVE_{Gao} in PHYDA, CESM (Ensemble Member Number 10) (Both JJA), and LMR (Annual) With Respect to the 5-Years Pre-Event Mean

Year of volcanic eruption	Anomaly in PHYDA (in °C)	Anomaly in CESM (in °C)	Anomaly in LMR (in °C)
1108	−0.20*	−0.11*	−0.23*
1171	−0.31*	−0.80*	−0.24*
1191	−0.10*	−0.18	−0.03*
1230	−0.21	−0.27	−0.20
1257	−0.21*	−1.57*	−0.29*
1276	0.0	−1.06*	−0.05
1286	−0.10*	−0.82*	−0.13*
1345	−0.08*	−0.02	−0.03
1453	−0.46	−0.89*	−0.27
1458	−0.05*	0.0	−0.06
1585	−0.20	−0.35	−0.15
1595	−0.15*	0.09*	−0.11*
1600	−0.63*	−0.60	−0.44*
1640	−0.61*	−0.31	−0.28*
1695	−0.46*	−0.58*	−0.16*
1809	−0.35*	−0.78*	−0.34*
1815	−0.25*	−0.91*	−0.15*
1831	−0.23*	−0.16	−0.22*
1835	−0.32*	−0.58	−0.22*

Note. Anomalies with an asterisk refer to the year following the volcanic event. Anomalies that refer to the year of the event have no asterisk.

CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE_{ev}, eVolv2k_v2 Tropical Volcanic Events.

the Kuwae event in 1453 CE (−0.46°C and −0.27°C) are also associated with large cooling magnitudes. It is worth noting, however, that as a result of fewer proxy records available in the earlier part of the PHYDA database, the uncertainty of the response is larger in the early part of the Last Millennium, especially before the volcanic eruption of 1453 CE after which the uncertainty of the 100-member PHYDA ensemble is considerably reduced. To a lesser extent, this is also true for the LMR. In the CESM-LME, the largest global JJA cooling (mean) is found during the summer following the Samalas event (−1.57°C) in 1257 CE, the largest known event of the Last Millennium (Vidal et al., 2016). In general, the magnitudes of the post-eruption cooling are larger in CESM, as in the case of the well-known Tambora event in 1815, which led to a global temperature drop of −0.91°C in CESM, while the mean global cooling is only −0.25°C in PHYDA and −0.15°C in LMR. The CESM shows a significant correlation (Figure 1c) between the magnitude of the volcanic event (expressed by the stratospheric sulfur injection) and the induced post-eruption cooling, with a negative slope that is significantly different from zero (F-test; $p < 0.01$). The correlation between the ice-core stratospheric sulfur injection estimated magnitude and the proxy estimated cooling, however, is not significant in either PHYDA or LMR (Figures 1b and 1d) and the slope of the regression is not significantly different from zero (F-test; $p = 0.33$ and $p = 0.79$, respectively).

The SEAs using the 19 TVE_{ev} or TVE_{Gao} (Figures 2a–2d) highlight the abrupt and significant temperature drop in the three datasets, especially in year +1 (−0.24°C, −0.28°C, and −0.17°C, in PHYDA, CESM, and LMR, respectively). In contrast, the duration of significant cooling differs widely: CESM only shows significant cooling for 2 years after the volcanic event (Figure 2b), while in PHYDA the cooling remains significant until year +8 (Figure 2a), and in LMR the significance of the cooling extends for 15 years (Figure 2c). The magnitude of the cooling is not affected by the occurrence of double-events, but once these are removed, the persistence of the cooling is reduced (Figures 2e and 2g) in both PHYDA (until year +6) and LMR (until year +10). CESM, on the other hand, indicates a similar temperature response regardless of whether we control for the double events (Figure 2f), which is perhaps expected given that the cooling only remains significant for 2 years in the model.

3.2. The Double-Event Effect and the Tree-Ring Density Response

In order to test for the extended persistence of the data assimilation products, we further explore the volcanically induced cooling over the Northern Hemisphere, including the double-event effect, by comparing our results to the N-TREND and NH-MXD tree-ring reconstructions. Our results (Figures 3a and 3b) highlight that the cooling is more persistent when 19 events are used for the SEA, as compared to only the 13 non-overlapping events in all data sets. When all 19 events are used, N-TREND and LMR show a persistent cooling of more than a decade (13 and 16 years, respectively) in the composite using all events. PHYDA portrays a weaker, albeit still significant cooling until year +10. In NH-MXD the persistent cooling lasts until year +6, while CESM only shows significant cooling until year +2. The magnitude of the cooling in year +1 is very similar in all products, being slightly larger in the N-TREND reconstruction.

When the effect of double-events is removed, the persistent significant cooling is reduced to less than 10 years in all reconstructions (Figure 3b). N-TREND and LMR indicate the largest persistence (10 years), followed by PHYDA (6 years), while 4 years of significant cooling are shown in NH-MXD and only 2 years in CESM. N-TREND and NH-MXD estimate similar magnitudes of post-eruption cooling in the Northern

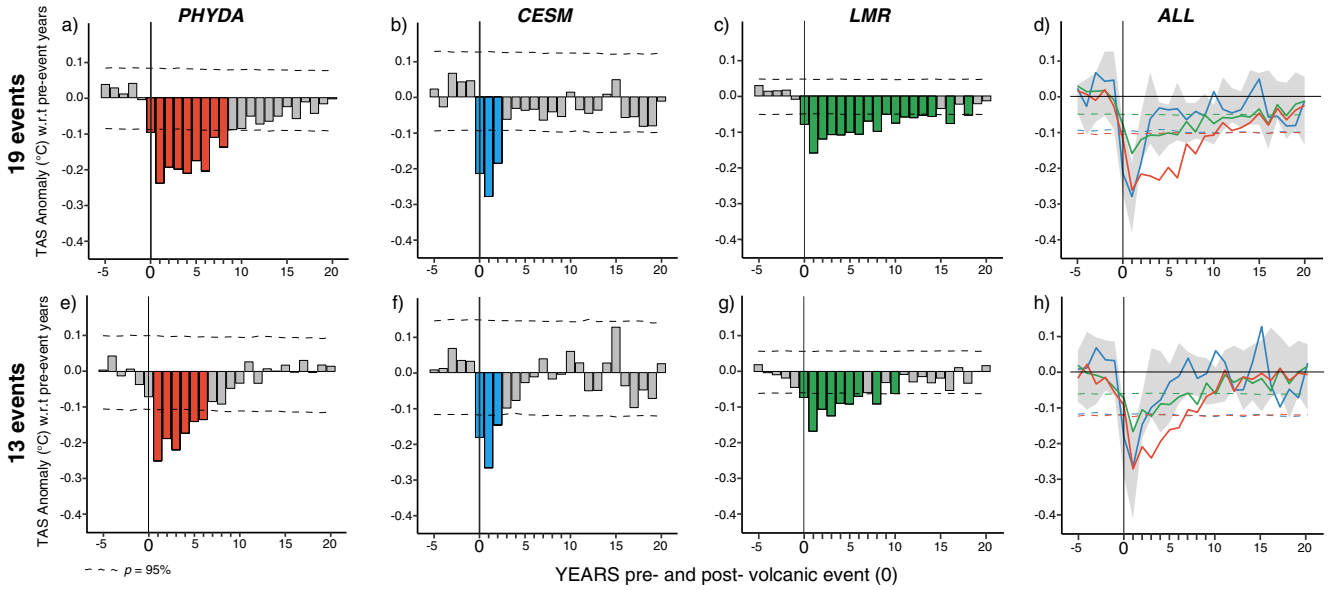

Figure 2. (a)–(d) Superposed Epoch Analysis of the globally averaged temperature anomaly (near-surface air temperature, TAS) following 19 TVE_{ev} (in JJA-PHYDA-ensemble mean and in Annual-LMR-ensemble mean) and TVE_{Gao} (in JJA-CESM) over the Last Millennium. Results shown in (b) and the bold blue lines in (d) represent the response of the CESM ensemble member number 10. The spread of the additional 9 CESM ensemble members' response is shown in gray in (d). (e)–(h) As in (a)–(d) except using only 13 events (not including the double-events). Color bars represent values that are significant at the 95% confidence-level. CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE, Tropical Volcanic Events.

Hemisphere (~−0.5°C) followed by PHYDA (−0.4°C) and CESM (−0.35°C), with LMR estimating the smallest magnitude (−0.32°C).

When the first of the double-events is selected (Figure 3c), the post-eruption cooling magnitude is larger in year 4 in N-TREND and PHYDA. LMR, NH-MXD, and CESM still show the larger magnitude of cooling in year +1, although persistence is increased to 16 years after the volcanic event in the LMR. N-TREND also displays an extended persistence of 13 years, while PHYDA, NH-MXD and CESM show the same persistence as when the second events are selected.

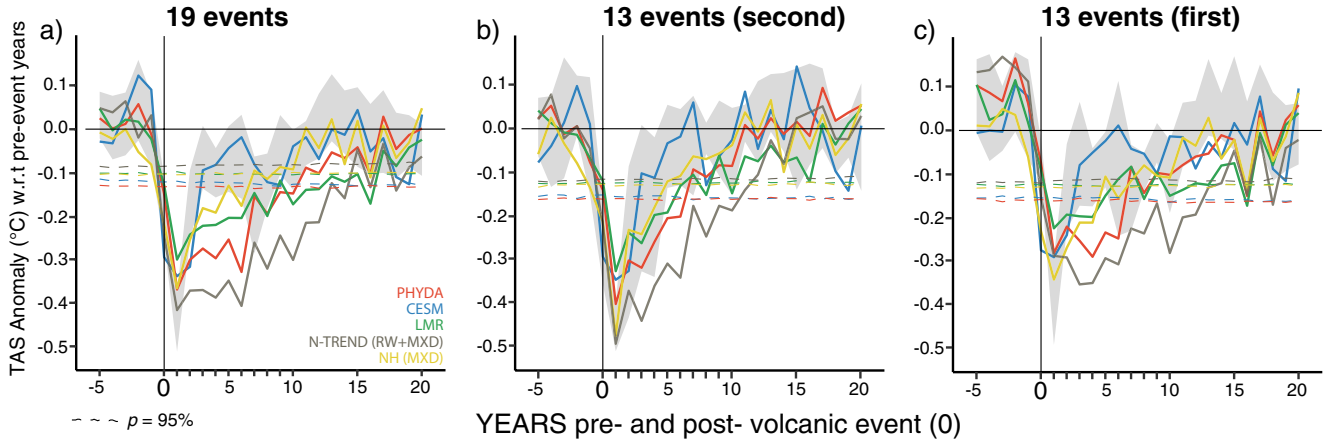

Figure 3. NH (>25°N) JJA temperature (near-surface air temperature, TAS) response to the TVE_{ev} and TVE_{Gao} in PHYDA, CESM (ensemble member number 10), N-TREND (Anchukaitis et al., 2017) and NH-MXD (Schneider et al., 2015); LMR (Annual) responses over the same domain are also shown. The spread of the additional 9 CESM ensemble members' responses is shown in gray. (a) SEA including 19 events. (b) SEA based on only 13 events, including only the second event whenever double-events occurred. (c) as in (b) but including only the first event from double-events. Horizontal lines represent the significance at the 95% confidence-level for each estimate. CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; SEA, Superposed Epoch Analysis; TVE, Tropical Volcanic Events.

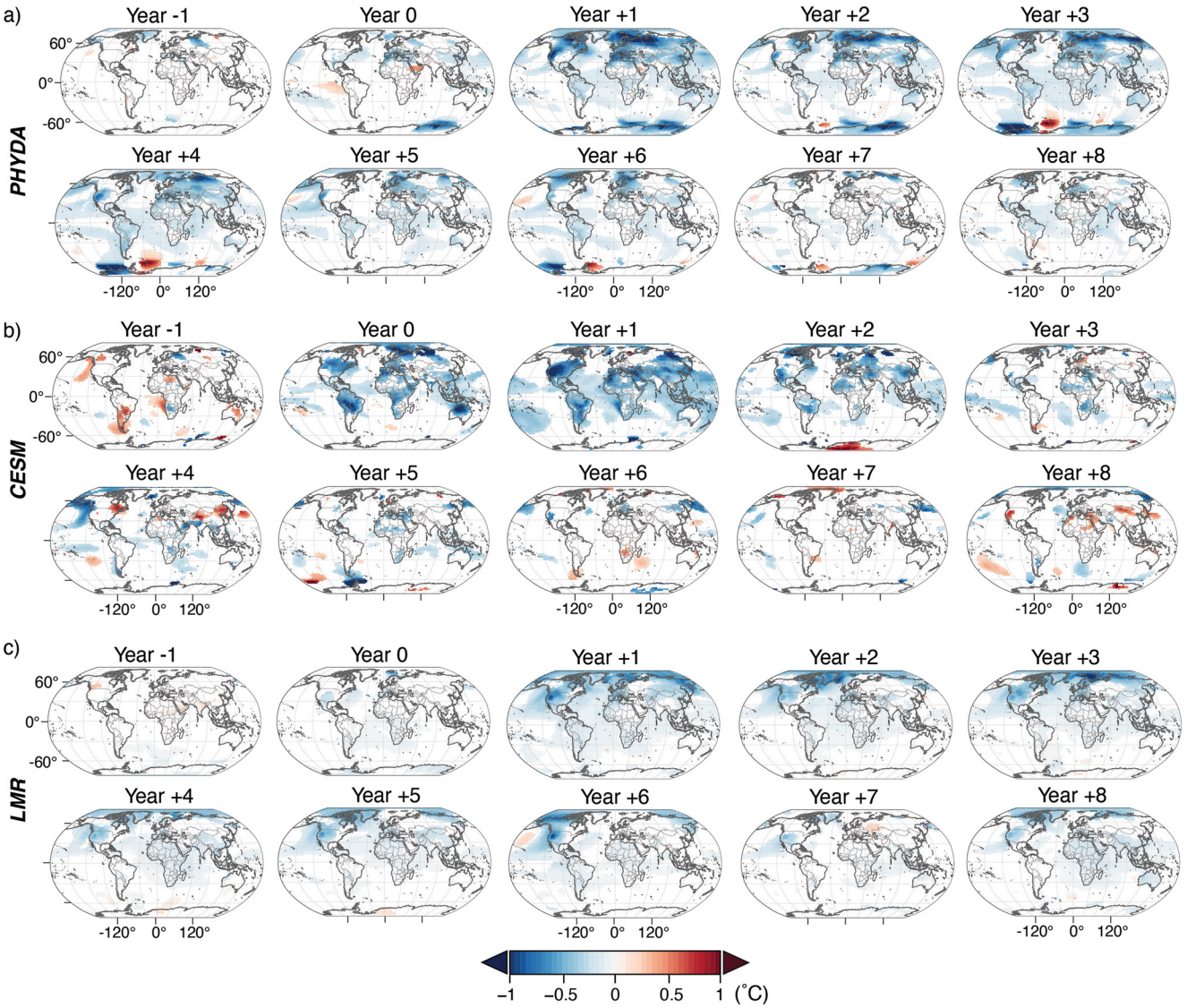

Figure 4. (a) Superposed Epoch Analysis of the spatial temperature (near-surface air temperature, TAS) anomaly response to the 13-event subset in the TVE_{ev} in the JJA-PHYDA-ensemble mean from year −1 to +8 after the volcanic event (0). (b) As in (a) except for JJA-CESM (ensemble member number 10) over the 13-event subset in the TVE_{Gao}. (c) as in (a) except for Annual-LMR-ensemble mean over the 13-event subset in the TVE_{ev}. Only values significant at the 95%-confidence level are shown. CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE, Tropical Volcanic Events.

3.3. Spatial Analysis of the Global Temperature Response to TVE_{ev}

Based on the above results, we perform the subsequent analysis using the 13-event subsets of TVE_{ev} or TVE_{Gao}, assuming that double-events in the SEA induce an extended cooling signal in PHYDA and LMR, but not CESM. In PHYDA, substantial cooling occurs globally in year +1 following tropical volcanic eruptions (Figure 4a, and Figure S2 in the supporting information), including a temperature drop larger than 1°C overs some areas of the Northern Hemisphere and Antarctica. The tropics and the Southern Hemisphere show a slightly smaller cooling (~−0.5°C) although cold anomalies still persist until year +6 after the volcanic event. We compute a weighted zonal-mean temperature response (Figures 5a and 5d), in which grid cells that contain non-significant temperature anomalies ($p < 0.05$) are not included in the mean estimate. These results confirm the larger magnitude of the cooling in the Northern Hemisphere and from 70°S to 90°S (Antarctica). This analysis also highlights that the persistence of the cooling is reduced to 3–4 years once the double-events are removed (Figure 5d).

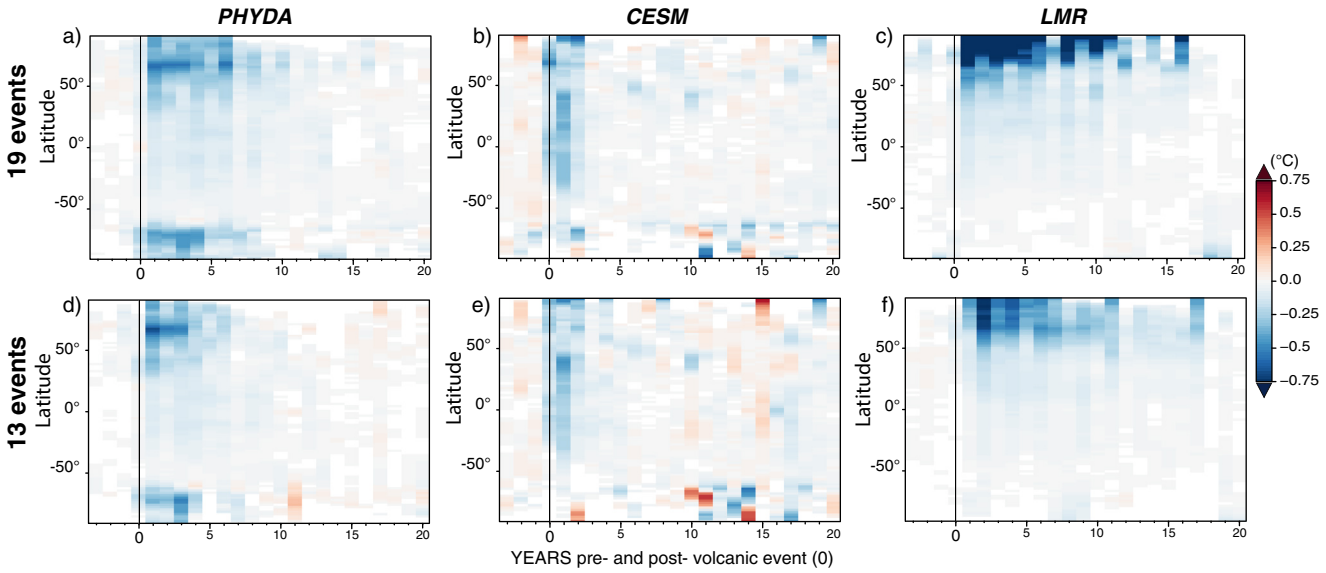

Figure 5. (a–c) Hovmöller diagram showing the weighted zonal-mean temperature (near-surface air temperature, TAS) response to 19 TVE_{ev} in PHYDA and LMR and 19 TVE_{Gao} in CESM (ensemble member number 10). (d–f) as in (a–c) except using only 13 TVE_{ev} and TVE_{Gao} (without the double-events). Non-significant values ($p < 0.05$) are given a weight of zero. Vertical black line indicates the onset of the volcanic event. CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE, Tropical Volcanic Events.

In CESM, extensive post-eruption cooling is already evident globally during the year of the eruption (year 0) (Figures 4b and S3 in the supporting information), including a temperature drop of more than 1°C overs South America, Australia, and the circumpolar areas of the Northern Hemisphere. The significant cooling is more evident over North America and the tropics in year +1, with few regions of cooling over Eurasia that are significant. Some areas still indicate a significant cooling in year +2, but this cooling dissipates by year +3. These characteristics become more apparent when the weighted zonal-mean temperature is computed (Figures 5b and 5d), showing that the largest cooling occurs over the tropics in year +1 for both the 13 and 19 TVE_{ev} selections.

In LMR, a widespread post-eruption cooling is evident in the Northern Hemisphere from years +1 to +8 (Figures 4c and S4 in the supporting information), with some cooling also apparent over the tropics and South America. The Hovmöller plot (Figures 5c and 5f) indicates a significant and persistent cooling north of 60°N, especially when using the full 19 TVE_{ev} selection, and a rather muted cooling response for the Southern Hemisphere.

3.4. Impacts on ENSO and AMO

The impact of the 13 TVE_{ev} on oceanic modes is further explored by analyzing the response of two sea surface temperature indices (Niño 3.4 and AMO). In the case of Niño 3.4 (Figure 6a) results vary widely, with only PHYDA revealing a significant La Niña occurring in year +4 of the SEA. The SEA for the AMO (Figure 6b) indicates more consistent results across datasets, highlighting a significant post-eruption cooling in all products. CESM shows the largest cooling, which extends from years 0 to +4. PHYDA also reveals significant cooling in years 1, 2, and 4, while LMR displays a significant cooling starting in year +1 and lasting over a decade. Results using the annual PHYDA and CESM AMO indices are very similar to those derived for the boreal winter means (see Figure S5 in the supporting information), while annual results for Niño 3.4 indicate no significant response in any product.

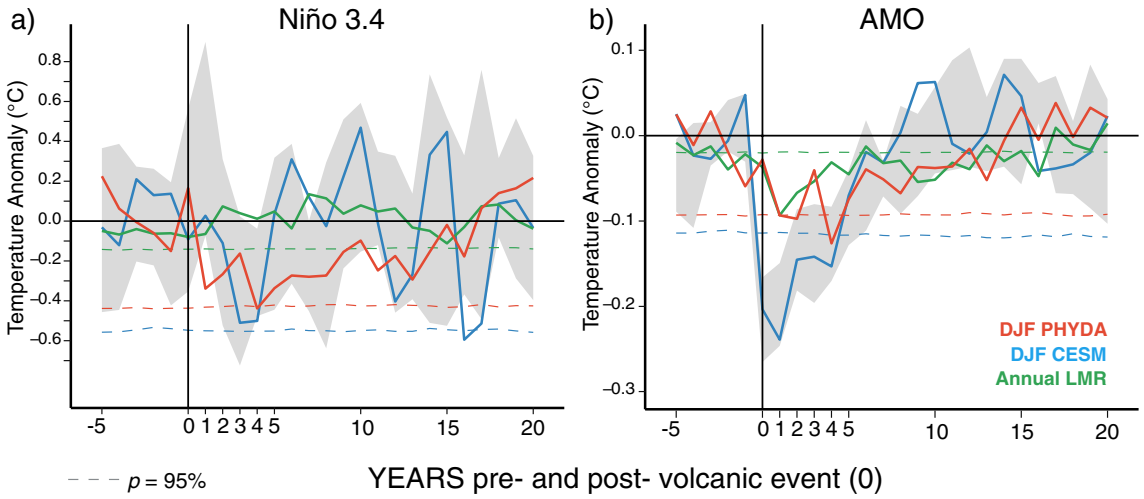

Figure 6. (a) Temperature response in the Niño 3.4 region (5°N–5°S, 170°W–120°W) to 13 TVE_{ev} in DJF PHYDA and Annual LMR, and TVE_{Gao} in DJF CESM (ensemble member number 10). The spread of the 9 additional CESM ensemble member's responses is shown in gray. (b) Same as in (a) except for the Atlantic Multidecadal Oscillation region (AMO, 0°N–60°N, 25°W–80°W). Horizontal lines represent the significance at the 95%-confidence level. CESM, Community Earth System Model; JJA, June, July, and August; LMR, Last Millennium Reanalysis; PHYDA, Paleo Hydrodynamics Data Assimilation product; TVE, Tropical Volcanic Events.

4. Discussion

4.1. Magnitude of the Cooling

The global impact of volcanic eruptions on temperature is confirmed and further analyzed in this study using two paleoclimate data assimilation products (PHYDA and LMR) and the CESM LME climate model simulations. We note that the NCAR models (CESM LME and CCSM4) form the priors for PHYDA and LMR, implying that the proxies are adding information beyond what the prior is alone providing (this is explicitly the case for PHYDA given its use of the CESM LME member 10 as its prior).

The global temperature response post-eruption (year +1) in these datasets is consistent with previous studies showing significant global (land and ocean) cooling in response to large tropical eruptions (e.g., Robock, 2000). There are, however, evident differences between the estimated responses in the PHYDA, LMR, and CESM datasets, such as the magnitude of the induced-cooling and the persistence of the temperature response in general.

In the case of magnitude, the two paleoclimate data assimilation products are consistent with the decreases in temperature estimated from MXD-based temperature reconstructions for the Northern Hemisphere (e.g., Esper et al., 2015; Schneider et al., 2017). There are, however, differences in the response in the tropics and the Southern Hemisphere between PHYDA and LMR; PHYDA estimates a response that is more similar to CESM, while the LMR response is more muted. On the other hand, the CESM results show a significant relationship between the magnitude of the volcanic forcing and the induced cooling, which is in contrast to the two paleo data assimilation products (Figure 1c). In the PHYDA and LMR, the magnitude of the cooling does not depend linearly on the magnitude of the eruption as estimated by the eVolv2k_v2 reconstruction. There are three potential reasons for this discrepancy: 1) The uncertainty in both PHYDA and LMR increases back in time due to decreasing proxy availability. Hence, the gradual reduction of variance in the ensemble mean over time (Figure 1a) should not be interpreted as a reduction in the variance of the historical climate, but as an indication of the decreasing amount of information from proxies. The limited proxy availability during the first half of the Last Millennium could affect the sensitivity of the relationship between volcanic forcing and the magnitude of the cooling. In the case of LMR, the targeted annual season could also contribute to the lower magnitude of the cooling when compared to PHYDA and CESM.

2) CESM may overestimate the climate sensitivity to volcanic forcing (e.g., Chylek et al., 2020; LeGrande et al., 2016; Stoffel et al., 2015). The fact that most of the eruptions in CESM are imposed in January would explain the larger magnitude and the significance of the cooling in year 0 (see discussion in Colose

et al., 2016). Additionally, the imposed aerosol optical depth in the model, which determines the spatial distribution of the radiative forcing, is based on an aerosol transport model (Gao et al., 2008), which is likely not a realistic representation of the actual dispersion and transport of individual events. Although this is not an error of the model itself, it may contribute to the existing differences between models, proxy records, and data assimilation products.

3) Uncertainties in the volcanic forcing reconstruction. The use of the most recent volcanic reconstruction "eVolv2k_v2" (Toohey & Sigl, 2017) increases the quality of our SEA because of its refined estimates for event dates and the magnitude of the stratospheric sulfur injections, relative to previous reconstructions (Crowley & Unterman, 2013; Gao et al., 2008; Sigl et al., 2015). There are still uncertainties, however, which are associated with both the dating of certain events, such as 1108 CE (±1), 1171 CE (±1), 1230 CE (±1), 1286 CE (±1), 1453 CE (±1), or 1458 CE (±1), and the stratospheric sulfur injection amounts, such as in the case of Mt Samalas (1257 CE, 59.42 Tg/s ± 10.86) or Huaynaputina (1600 CE, 18.95 Tg/s ± 4.03). Based on our results, and consistent with suggestions by Schneider et al. (2017), it is likely that the magnitude estimates for the 1453 and 1458 events might have been shifted in eVolv2k_v2, namely the event in 1453 or 1452 was larger than 1458, despite the opposite estimate in eVolv2k_v2. Additionally, further analyses of the Huaynaputina event in 1600 CE are required, which has recently been upgraded to a stronger eruption than previously estimated (Prival et al., 2019). Consistent with the findings of Prival et al. (2019), the Huaynaputina eruption induced the largest global cooling anomaly over the Last Millennium in both PHYDA (−0.63°C) and LMR (−0.44°C).

Overall, estimates of the global temperature response magnitude agree well in PHYDA and CESM. LMR, on the other hand, displays a more muted temperature response in the tropics and Southern Hemisphere that is reflected in a more muted estimate of the global temperature response. Despite these uncertainties, the robustness of the analysis is documented across the three datasets by the abrupt climatic response in the year of the event and especially in the first year after the eruption (Figures 1 and 2)—a response that includes significant temperature anomalies over a vast portion of the globe (Figures 4 and 5).

4.2. Persistence of the Cooling

Several studies have suggested that TRW temperature reconstructions underestimate and temporally extend the response to volcanic eruptions when compared to MXD reconstructions (e.g., Anchukaitis et al., 2012; D'Arrigo et al., 2013; Esper et al., 2015). PHYDA and LMR, while capturing a cooling magnitude that is consistent with MXD reconstructions, also show a more persistent response to volcanic forcing than the MXD reconstruction (+2 years). CESM on the other hand simulates a response that is more short-lived than what is reconstructed from MXD data. We also note that even though we compare PHYDA to the same ensemble member used as the data assimilation prior, the results of our comparisons are not dependent on the employed CESM LME member. The temperature response to volcanic eruptions is equally short in the nine other available ensemble members of the CESM LME (see Figure S6 in the supporting information). Our results therefore are not dependent on a single ensemble member of CESM.

The life cycle of the effective radiative forcing after large tropical volcanic events (Toohey et al., 2019), which persists for ∼24 months, determines the abrupt climate response seen in our study, lasting until approximately year +3. CESM seems to be capturing only these direct radiative effects (Figures 4b, 5b, and 5e), which cause direct cooling of the Earth's surface. But the radiative forcing effects of stratospheric aerosols also invoke multiple feedbacks that further modify the atmospheric and oceanic response, thereby inducing significant decadal-scale climate variations (Church et al., 2005; Miller et al., 2012; Zanchettin et al., 2012). These feedbacks likely serve as mechanisms to extend the responses to volcanic forcing as seen in the data assimilation products, although analyzing the true nature of these feedbacks and how they are embedded in the reanalysis products is beyond the scope of this study.

Whether the differences between the proxy estimates of persistence and the CESM response arise from a bias in the proxy data, or indicate a problem in the CESM model, is an open question (Lücke et al., 2019). Zhu et al. (2020) recently stated that discrepancies between volcanic forcing signals in iCESM (the prior used in Zhu et al., 2020) and tree-ring reconstructions (or data assimilation products including tree-ring records) since 1600 CE can largely be resolved by assimilating tree-ring density records only (MXD-based),

targeting the growing season instead of annual temperature, and by performing the comparison only at locations with sufficient proxy records. They also hypothesize that discrepancies during the early part of the Last Millennium may reflect uncertainties in forcing and modeled aerosol microphysics. In fact, they chose CESM as their reconstruction target for capturing the temperature response to large volcanic eruptions. Chylek et al. (2020) have nevertheless noted that CMIP5 models (including CESM) generally overestimate the volcanically induced cooling, and that this overestimation is caused by the model's response to volcanic forcing, not by the error in the forcing itself. They further hypothesize that the model's parameterization of aerosol-cloud interactions within ice and mixed-phase clouds is a likely source of such a discrepancy (Chylek et al., 2020). Furthermore, Esper et al. (2015) showed that MXD reconstructions over the NH correctly capture the post-eruption cooling and its persistence through comparisons to JJA instrumental observations using 13 volcanic events since 1850 (smaller in magnitude than the TVEP used in this study, with the exception of Krakatoa in 1883). Based on our results and the existing literature, we therefore suggest that MXD reconstructions are closer to the true temperature impact of TVE_{ev} than CESM. Particularly when the double-event effect is taken into account, data assimilation products, and especially PHYDA, show a response similar to the MXD reconstructions and thus are able to reproduce both the immediate temperature effect through direct radiative forcing, as well as the longer-term persistence of the signal mediated through feedbacks.

The AMO and ENSO could both potentially serve as oceanic feedbacks, capable of prolonging the initial radiatively forced cooling following volcanic eruptions (Knudsen et al., 2014; Swingedouw et al., 2017; Zanchettin et al., 2012). A significant cooling is observed in the region that defines the AMO starting in year +0 (in CESM), and in year +1 (in PHYDA and LMR). This result is in agreement with Birkel et al. (2018), suggesting a persistent multidecadal cooling in this region following large volcanic eruptions, although our results indicate that cooling is only significant until year +5 in CESM and PHYDA. LMR shows a very low-magnitude, albeit significant, AMO cooling that lasts until year +16.

Whether or not volcanic eruptions trigger ENSO events remains a continued topic of debate, although a recent study by Dee et al. (2020) found no evidence for a consistent ENSO response in the year following volcanic eruptions over the Last Millennium. Some studies, however, have suggested that an El Niño-like response is more likely to happen 1–2 years after large volcanic events (e.g., Adams et al., 2003; Emile-Geay et al., 2008; Mann et al., 2005). Regardless, no significant El Niño develops in either PHYDA or CESM in the DJF season, which is when El Niño typically reaches its peak phase. Instead, results from PHYDA suggest that a volcanically induced La Niña response is more likely to develop in DJF of year +4, characterized by a cooling in the eastern equatorial Pacific, in agreement with McGregor and Timmermann (2011), Sun et al. (2019), Wahl et al. (2014) (although the latter indicating a La Niña-like response in year +2), and contrary to some studies using only proxies (Adams et al., 2003; Mann et al., 2005) or those using CESM with a different selection of volcanic events (Stevenson et al., 2016). Additional implications of a La Niña-like state would include a La Niña-like pattern of the Antarctic Dipole, shown in our study (Figure 3a), which is a part of the Southern Hemisphere ENSO wave train response and represents the largest ENSO-related temperature anomaly outside of the tropical Pacific (Yuan & Martinson, 2001). Nevertheless, further studies analyzing the individual background state of ENSO for each particular event are needed to better understand the volcanically induced effects.

Finally, the mismatch that was shown by the IPCC AR5 (Intergovernmental Panel on Climate Change, 2014) between the response in CESM and tree-ring width (similar to Figure 3c) reconstructions can to a large extent be reconciled by avoiding overlapping events (i.e., removing the first event when there is a second event within the next 10 years) (Figure 3b).

4.3. Spatial Cooling Response

The global PHYDA temperature response also shows significant and widespread cooling, consistent with the CESM initial response (years 0 and + 1) and with other studies using model simulations (CESM-LME, e.g., Stevenson et al., 2017). Such cooling is attributed to more efficient backscatter of shortwave radiation due to the enhanced sulfate aerosol load, increasing the planetary albedo for the lifetime of the aerosols (Kravitz & Robock, 2011; Robock, 2000). On the other hand, the response in both PHYDA and LMR is to some extent proxy-dependent. Therefore, both data assimilation products might be underestimating the

true impact on temperature over land in the tropics, where there is an evident lack of annually resolved and temperature-sensitive proxy records. This is especially evident in the case of LMR, where the signal is rather muted away from the high northern latitudes (Figure 5f).

The spatial response in both PHYDA and LMR should therefore be interpreted with caution where the number of proxy records is limited. In PHYDA, this is the case for tropical and West Africa, and eastern South America. In contrast, the Arctic and Antarctica, the Northern Hemisphere in general, as well as the tropical oceans have a fairly dense proxy representation in space and time, hence results tend to be more robust and reliable over those regions. In LMR, results over the Southern Hemisphere should be interpreted with caution until more proxies are included in future versions. Differences in the persistence and the spatial cooling response between PHYDA and LMR may also arise from differences in the adopted model prior (CESM LME and CCSM4).

5. Conclusions

By analyzing the temperature response to large tropical volcanic eruptions over the Last Millennium in two state-of-the art data assimilation products, PHYDA and LMR, and simulations from the CESM-LME, we have shown:

- There is a robust agreement in the magnitude of the cooling responses to TVE_{ev} or TVE_{Gao} between the data assimilation products (PHYDA and LMR) and the CESM-LME
- The magnitude estimates of PHYDA and LMR are consistent with other proxy-informed responses, that is, they capture the magnitude of the post-eruption cooling in the NH, even though they assimilate TRW proxies
- The spatial response of the cooling seen in PHYDA and LMR should be treated with caution over certain regions where proxy availability is reduced back in time. One place where there is disagreement is in the Southern Hemisphere, where the LMR signal is very muted, suggesting the need to further isolate the causes of the differences
- Data assimilation products, such as PHYDA and LMR, effectively reduce the TRW cooling persistence by 4 years (Figure 3b), although there seems to remain a small extended persistence that may be associated with the potential memory effect from TRW
- The agreement in the persistence and magnitude of the cooling seen in both the PHYDA and the NH-MXD is particularly encouraging, given the fact that the PHYDA assimilates many TRW records
- Large tropical volcanic events induce an extended persistence of the temperature signal in reconstructions that include TRW measurements (Figure 3c) because of the so-called “double-event” effect. The biological persistence inherent to the TRW and MXD data or the potential failure of CESM to capture the full extent of ocean-atmosphere feedbacks might be potential causes of such model-proxy disagreements. If the origins of these differences are associated with modeling choices, our findings would suggest that adaptation measures to volcanic eruptions that are based solely on climate models may be subject to biases. However, we stress that a more comprehensive analysis over an ensemble of models, particularly with the CMIP6 generation of climate models, is required before larger conclusions can be drawn
- Finally, the estimated response of the AMO to volcanic eruptions is consistent in PHYDA and CESM, and suggests a shift toward the negative phase of the AMO for at least five years after an eruption. The estimates with respect to the Niño 3.4 index agree in the sign of the response (inducing negative anomalies), although the magnitude and the persistence vary across products

Collectively, the above conclusions demonstrate that paleoclimate data assimilation products are an emerging and important tool for assessing the impacts of volcanism. We nevertheless have also shown that large spatial inconsistencies exist in the estimated temperature response to volcanism in the PHYDA and LMR. Multiple differences exist in the methods and input data used to construct the two data assimilation products, indicating that a larger and more controlled assessment of the differences and their causes is necessary in future work. Further refining the estimates from the paleo data assimilation products, in addition to fully characterizing the uncertainties associated with the model estimates, is ultimately critical for accurately characterizing the risks associated with future volcanic eruptions. If historical evidence is indeed a guide to the future, large tropical eruptions are very likely to occur this century. Accurately constraining the

likely impacts of such eruptions, particularly in terms of how they will combine with those of persistent greenhouse gas forcing in the future, is thus a fundamental step toward preparing for the agricultural and socioeconomic impacts of these unavoidable events.

Data Availability Statement

The PHYDA data set are available via Zenodo web repository; https://zenodo.org/record/1198817. The LMR data set are available via web repository; https://atmos.washington.edu/~hakim/lmr/LMRv2/. The "Reconstructed volcanic stratospheric sulfur injections and aerosol optical depth, 500 BCE to 1900 CE-eVolv2k_v2" are available via DKRZ web repository; https://cera-www.dkrz.de/WDCC/ui/cerasearch/entry?acronym=eVolv2k_v2

Acknowledgments
The authors thank Cody Routson and one anonymous reviewer for their helpful and constructive comments. E. Tejedor, N. Steiger, J. Smerdon, and M. Vuille were partially supported by NSF-PIRE (OISE-1743738). E. Tejedor and M. Vuille were partially supported by NSF-P2C2 (AGS-1702439). M. Vuille was partially supported by NSF-CNH2-L (DEB-1923957). J. Smerdon and NS were partially supported by NSF AGS-1805490. RSN is funded by a "Juan de la Cierva" postdoctoral grant FJCI-2017-31595. E. Tejedor and R. Serrano-Notivoli are partially supported by the Government of Aragón through the "Program of research groups" (group H38, "Clima, Agua, Cambio Global y Sistemas Naturales"). Lamont contribution number 8489.

References

Adams, J. B., Mann, M. E., & Ammann, C. M. (2003). Proxy evidence for an El Niño-like response to volcanic forcing. *Nature*, *426*(6964), 274–278. https://doi.org/10.1038/nature02101

Anchukaitis, K. J., Breitenmoser, P., Briffa, K. R., Buchwal, A., Büntgen, U., Cook, E. R., et al. (2012). Tree rings and volcanic cooling. *Nature Geosciences*, *5*(12), 836–837. https://doi.org/10.1038/ngeo1645

Anchukaitis, K. J., Wilson, R., Briffa, K. R., Büntgen, U., Cook, E. R., D'Arrigo, R., et al. (2017). Last millennium northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews*, *163*, 1–22. https://doi.org/10.1016/j.quascirev.2017.02.020

Birkel, S. D., Mayewski, P. A., Maasch, K. A., Kurbatov, A. V., & Lyon, B. (2018). Evidence for a volcanic underpinning of the Atlantic multidecadal oscillation. *NPJ Climate and Atmospheric Science*, *1*(1), 24. https://doi.org/10.1038/s41612-018-0036-6

Church, J. A., White, N. J., & Arblaster, J. M. (2005). Significant decadal-scale impact of volcanic eruptions on sea level and ocean heat content. *Nature*, *438*(7064), 74–77. https://doi.org/10.1038/nature04237

Chylek, P., Folland, C., Klett, J. D., & Dubey, M. K. (2020). CMIP5 climate models overestimate cooling by volcanic aerosols. *Geophysical Research Letters*, *47*(3), e2020GL087047. https://doi.org/10.1029/2020GL087047

Colose, C. M., LeGrande, A. N., & Vuille, M. (2016). Hemispherically asymmetric volcanic forcing of tropical hydroclimate during the last millennium. *Earth System Dynamics*, *7*(3), 681–696. https://doi.org/10.5194/esd-7-681-2016

Crowley, T. J., & Unterman, M. B. (2013). Technical details concerning development of a 1,200 yr proxy index for global volcanism. *Earth System Science Data*, *5*(1), 187–197. https://doi.org/10.5194/essd-5-187-2013

D'Arrigo, R., Wilson, R., & Anchukaitis, K. J. (2013). Volcanic cooling signal in tree ring temperature records for the past millennium. *Journal of Geophysical Research: Atmospheres*, *118*(16), 9000–9010. https://doi.org/10.1002/jgrd.50692

D'Arrigo, R., Wilson, R., & Tudhope, A. (2009). The impact of volcanic forcing on tropical temperatures during the past four centuries. *Nature Geoscience*, *2*(1), 51–56. https://doi.org/10.1038/ngeo393

Dee, S. G., Cobb, K. M., Emile-Geay, J., Ault, T. R., Edwards, R. L., Cheng, H., & Charles, C. D. (2020). No consistent ENSO response to volcanic forcing over the last millennium. *Science*, *367*(6485), 1477–1481. https://doi.org/10.1126/science.aax2000

Emile-Geay, J., & Consortium, P. 2k. (2017). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, *4*(1), 170088. https://doi.org/10.1038/sdata.2017.88

Emile-Geay, J., Seager, R., Cane, M. A., Cook, E. R., Haug, G. H., & Haug, G. H. (2008). Volcanoes and ENSO over the past millennium. *Journal of Climate*, *21*(13), 3134–3148. https://doi.org/10.1175/2007JCLI1884.1

Engvild, K. C. (2003). A review of the risks of sudden global cooling and its effects on agriculture. *Agricultural and Forest Meteorology*, *115*(3), 127–137. https://doi.org/10.1016/S0168-1923(02)00253-8

Esper, J., Büntgen, U., Luterbacher, J., & Krusic, P. J. (2013). Testing the hypothesis of post-volcanic missing rings in temperature sensitive dendrochronological data. *Dendrochronologia*, *31*(3), 216–222. https://doi.org/10.1016/j.dendro.2012.11.002

Esper, J., George, S. S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F. C., Luterbacher, J., et al. (2018). Large-scale, millennial-length temperature reconstructions from tree-rings. *Dendrochronologia*, *50*, 81–90. https://doi.org/10.1016/j.dendro.2018.06.001

Esper, J., Schneider, L., Smerdon, J. E., Schöne, B. R., & Büntgen, U. (2015). Signals and memory in tree-ring width and density data. *Dendrochronologia*, *35*, 62–70. https://doi.org/10.1016/j.dendro.2015.07.001

Gao, C. C., & Gao, Y. J. (2018). Revisited Asian monsoon hydroclimate response to volcanic eruptions. *Journal of Geophysical Research: Atmosphere*, *123*(15), 7883–7896. https://doi.org/10.1029/2017JD027907

Gao, C., Robock, A., & Ammann, C. (2008). Volcanic forcing of climate over the past 1,500 years: An improved ice core-based index for climate models. *Journal of Geophysical Research*, *113*(D23), D23111. https://doi.org/10.1029/2008JD010239

Gao, Y., & Gao, C. (2017). European hydroclimate response to volcanic eruptions over the past nine centuries. *International Journal of Climatology*, *37*(11), 4146–4157. https://doi.org/10.1002/joc.5054

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., et al. (2016). The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres*, *121*(12), 6745–6764. https://doi.org/10.1002/2016JD024751

Hegerl, G. C., Crowley, T. J., Baum, S. K., Kim, K.-Y., & Hyde, W. T. (2003). Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophysical Research Letters*, *30*(5). https://doi.org/10.1029/2002GL016635

Intergovernmental Panel on Climate Change (2014). *Climate Change 2013—The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324

Knudsen, M. F., Jacobsen, B. H., Seidenkrantz, M.-S., & Olsen, J. (2014). Evidence for external forcing of the Atlantic multidecadal oscillation since termination of the Little Ice Age. *Nature Communications*, *5*, 3323. https://doi.org/10.1038/ncomms4323

Kravitz, B., & Robock, A. (2011). Climate effects of high-latitude volcanic eruptions: Role of the time of year. *Journal of Geophysical Research*, *116*(D1). https://doi.org/10.1029/2010JD014448

Lamb, H., & Stanley, S. J. (1970). Volcanic dust in the atmosphere; with a chronology and assessment of its meteorological significance. *Philosophical Transactions of the Royal Society of London: Series A. Mathematical and Physical Sciences*, *266*(1178), 425–533. https://doi.org/10.1098/rsta.1970.0010

Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., & Teng, H. (2013). Last millennium climate and its variability in CCSM4. *Journal of Climate*, *26*(4), 1085–1111. https://doi.org/10.1175/JCLI-D-11-00326.1

LeGrande, A. N., Tsigaridis, K., & Bauer, S. E. (2016). Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections. *Nature Geoscience*, *9*(9), 652–655. https://doi.org/10.1038/ngeo2771

Lehner, F., Schurer, A. P., Hegerl, G. C., Deser, C., & Frölicher, T. L. (2016). The importance of ENSO phase during volcanic eruptions for detection and attribution. *Geophysical Research Letters*, *43*(6), 2851–2858. https://doi.org/10.1002/2016GL067935

Li, J., Xie, S.-P., Cook, E. R., Morales, M. S., Christie, D. A., Johnson, N. C., et al. (2013). El Niño modulations over the past seven centuries. *Nature Climate Change*, *3*(9), 822–826. https://doi.org/10.1038/nclimate1936

Lücke, L. J., Hegerl, G. C., Schurer, A. P., & Wilson, R. (2019). Effects of memory biases on variability of temperature reconstructions. *Journal of Climate*, *32*(24), 8713–8731. https://doi.org/10.1175/JCLI-D-19-0184.1

Maher, N., McGregor, S., England, M. H., & Sen, G. , A. S. (2015). Effects of volcanism on tropical variability. *Geophysical Research Letters*, *42*(14), 6024–6033. https://doi.org/10.1002/2015GL064751

Mann, M. E., Cane, M. A., Zebiak, S. E., & Clement, A. (2005). Volcanic and solar forcing of the tropical pacific over the past 1,000 years. *Journal of Climate*, *18*(3), 447–456. https://doi.org/10.1175/JCLI-3276.1

Mann, M. E., Fuentes, J. D., & Rutherford, S. (2012). Underestimation of volcanic cooling in tree-ring-based reconstructions of hemispheric temperatures. *Nature Geoscience*, *5*, 202. https://doi.org/10.1038/ngeo1394

McGregor, S., & Timmermann, A. (2011). The effect of explosive tropical volcanism on ENSO. *Journal of Climate*, *24*(8), 2178–2191. https://doi.org/10.1175/2010JCLI3990.1

Ménégoz, M., Bilbao, R., Bellprat, O., Guemas, V., & Doblas-Reyes, F. J. (2018). Forecasting the climate response to volcanic eruptions: Prediction skill related to stratospheric aerosol forcing. *Environmental Research Letters*, *13*(6), 064022. https://doi.org/10.1088/1748-9326/aac4db

Miller, G. H., Geirsdóttir, Á., Zhong, Y., Larsen, D. J., Otto-Bliesner, B. L., Holland, M. M., et al. (2012). Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophysical Research Letters*, *39*(2). https://doi.org/10.1029/2011GL050168

Oman, L., Robock, A., Stenchikov, G. L., Thordarson, T., Koch, D., Shindell, D. T., & Gao, C. (2006). Modeling the distribution of the volcanic aerosol cloud from the 1783–1784 Laki eruption. *Journal of Geophysical Research*, *111*(D12). https://doi.org/10.1029/2005JD006899

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., et al. (2015). Climate variability and change since 850 CE: An ensemble approach with the community earth system model. *Bulletin of the American Meteorological Society*, *97*(5), 735–754. https://doi.org/10.1175/BAMS-D-14-00233.1

Papale, P. (2018). Global time-size distribution of volcanic eruptions on Earth. *Scientific Reports*, *8*(1), 6838. https://doi.org/10.1038/s41598-018-25286-y

Predybaylo, E., Stenchikov, G. L., Wittenberg, A. T., & Zeng, F. (2017). Impacts of a Pinatubo-size volcanic eruption on ENSO. *Journal of Geophysical Research: Atmospheres*, *122*(2), 925–947. https://doi.org/10.1002/2016JD025796

Prival, J.-M., Thouret, J.-C., Japura, S., Gurioli, L., Bonadonna, C., Mariño, J., & Cueva, K. (2019). New insights into eruption source parameters of the 1600 CE Huaynaputina Plinian eruption, Peru. *Bulletin of Volcanology*, *82*(1), 7. https://doi.org/10.1007/s00445-019-1340-7

Rampino, M. R., & Self, S. (1984). The Atmospheric Effects of El Chichón. *Scientific American*, *250*(1), 48–57. https://doi.org/10.1038/scientificamerican0184-48

Robock, A. (1983). El Chinchón eruption: The dust cloud of the century. *Nature*, *301*, 373. https://doi.org/10.1038/301373a0

Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics*, *38*(2), 191–219. https://doi.org/10.1029/1998RG000054

Robock, A., & Matson, M. (1983). Circumglobal transport of the El Chichon volcanic dust cloud. *Science*, *221*, 195–197. https://doi.org/10.1126/science.221.4606.195

Schneider, L., Smerdon, J. E., Büntgen, U., Wilson, R. J. S., Myglan, V. S., Kirdyanov, A. V., & Esper, J. (2015). Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. *Geophysical Research Letters*, *42*(11), 4556–4562. https://doi.org/10.1002/2015GL063956

Schneider, L., Smerdon, J. E., Pretis, F., Hartl-Meier, C., & Esper, J. (2017). A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures. *Environmental Research Letters*, *12*(9), 094005. https://doi.org/10.1088/1748-9326/aa7a1b

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., et al. (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*, 543. https://doi.org/10.1038/nature14565

Steiger, N. J., Smerdon, J. E., Cook, E. R., & Cook, B. I. (2018). A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Scientific Data*, *5*, 1–15. https://doi.org/10.1038/sdata.2018.86

Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. L., Tomas, R. A., & Gao, C. (2017). Role of eruption season in reconciling model and proxy responses to tropical volcanism. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(8), 1822–1826. https://doi.org/10.1073/pnas.1612505114

Stevenson, S., Otto-Bliesner, B., Fasullo, J., & Brady, E. (2016). "El Niño Like" hydroclimate responses to last millennium volcanic eruptions. *Journal of Climate*, *29*(8), 2907–2921. https://doi.org/10.1175/JCLI-D-15-0239.1

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., et al. (2015). Estimates of volcanic-induced cooling in the northern Hemisphere over the past 1,500 years. *Nature Geoscience*, *8*(10), 784–788. https://doi.org/10.1038/ngeo2526

Sun, W., Liu, J., Wang, B., Chen, D., Liu, F., Wang, Z., et al. (2019). A "La Niña-like" state occurring in the second year after large tropical volcanic eruptions during the past 1,500 years. *Climate Dynamics*, *52*(12), 7495–7509. https://doi.org/10.1007/s00382-018-4163-x

Swingedouw, D., Mignot, J., Ortega, P., Khodri, M., Menegoz, M., Cassou, C., & Hanquiez, V. (2017). Impact of explosive volcanic eruptions on the main climate variability modes. *Global and Planetary Change*, *150*, 24–45. https://doi.org/10.1016/j.gloplacha.2017.01.006

Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., et al. (2019). Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate of the Past*, *15*(4), 1251–1273. https://doi.org/10.5194/cp-15-1251-2019

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Tejedor, E., Steiger, N. J., Smerdon, J. E., Serrano-Notivoli, R., & Vuille, M. (2021). Global hydroclimatic response to tropical volcanic eruptions over the last millennium. *Proceedings of the National Academy of Sciences*, *118*(12), e2019145118. https://doi.org/10.1073/pnas.2019145118

Toohey, M., Krüger, K., Schmidt, H., Timmreck, C., Sigl, M., Stoffel, M., & Wilson, R. (2019). Disproportionately strong climate forcing from extratropical explosive volcanic eruptions. *Nature Geoscience*, *12*(2), 100–107. https://doi.org/10.1038/s41561-018-0286-2

Toohey, M., & Sigl, M. (2017). Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE. *Earth System Science Data*, *9*(2), 809–831. https://doi.org/10.5194/essd-9-809-2017

Vidal, C. M., Métrich, N., Komorowski, J.-C., Pratomo, I., Michel, A., Kartadinata, N., et al. (2016). The 1257 Samalas eruption (Lombok, Indonesia): The single greatest stratospheric gas release of the Common Era. *Scientific Reports*, *6*, 34868. https://doi.org/10.1038/srep34868

Wahl, E. R., Diaz, H. F., Smerdon, J. E., & Ammann, C. M. (2014). Late winter temperature response to large tropical volcanic eruptions in temperate western North America: Relationship to ENSO phases. *Global and Planetary Change*, *122*, 238–250. https://doi.org/10.1016/j.gloplacha.2014.08.005

Yuan, X., & Martinson, D. G. (2001). The Antarctic dipole and its predictability. *Geophysical Research Letters*, *28*(18), 3609–3612. https://doi.org/10.1029/2001GL012969

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., et al. (2016). The model intercomparison project on the climatic response to volcanic forcing (VolMIP): Experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, *9*(8), 2701–2719. https://doi.org/10.5194/gmd-9-2701-2016

Zanchettin, D., Timmreck, C., Graf, H.-F., Rubino, A., Lorenz, S., Lohmann, K., et al. (2012). Bi-decadal variability excited in the coupled ocean-atmosphere system by strong tropical volcanic eruptions. *Climate Dynamics*, *39*(1–2), 419–444. https://doi.org/10.1007/s00382-011-1167-1

Zhu, F., Emile-Geay, J., Hakim, G. J., King, J., & Anchukaitis, K. J. (2020). Resolving the differences in the simulated and reconstructed temperature response to volcanism. *Geophysical Research Letters*, *47*(8), e2019GL086908. https://doi.org/10.1029/2019GL086908