

# Geophysical Research Letters

## RESEARCH LETTER

10.1029/2018GL081018

### Key Points:

- The 1815 Tambora eruption is used to demonstrate how initial conditions affect the simulation of posteruption climate anomalies
- Forcing uncertainties can overwhelm initial-condition spread in boreal summer, while the effect of initial conditions predominate in winter
- Current initial-condition uncertainties hamper inferences on magnitude of the Tambora eruption in relation to the “year without summer”

### Supporting Information:

- Data Set S1

### Correspondence to:

D. Zanchettin,  
davidoff@unive.it

### Citation:

Zanchettin, D., Timmreck, C., Toohey, M., Jungclaus, J. H., Bittner, M., Lorenz, S. J., & Rubino, A. (2019). Clarifying the relative role of forcing uncertainties and initial-condition unknowns in spreading the climate response to volcanic eruptions. *Geophysical Research Letters*, 46, 1602–1611. <https://doi.org/10.1029/2018GL081018>

Received 23 OCT 2018

Accepted 10 JAN 2019

Accepted article online 15 JAN 2019

Published online 2 FEB 2019

## Clarifying the Relative Role of Forcing Uncertainties and Initial-Condition Unknowns in Spreading the Climate Response to Volcanic Eruptions

**Daive Zanchettin<sup>1</sup>** , **Claudia Timmreck<sup>2</sup>** , **Matthew Toohey<sup>3</sup>** , **Johann H. Jungclaus<sup>2</sup>** , **Matthias Bittner<sup>2</sup>** , **Stephan J. Lorenz<sup>2</sup>**, and **Angelo Rubino<sup>1</sup>**

<sup>1</sup>Department of Environmental Sciences, Informatics and Statistics, University Ca' Foscari of Venice, Mestre, Italy, <sup>2</sup>Max Planck Institute for Meteorology, Hamburg, Germany, <sup>3</sup>GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany

**Abstract** Radiative forcing from volcanic aerosol impacts surface temperatures; however, the background climate state also affects the response. A key question thus concerns whether constraining forcing estimates is more important than constraining initial conditions for accurate simulation and attribution of posteruption climate anomalies. Here we test whether different realistic volcanic forcing magnitudes for the 1815 Tambora eruption yield distinguishable ensemble surface temperature responses. We perform a cluster analysis on a superensemble of climate simulations including three 30-member ensembles using the same set of initial conditions but different volcanic forcings based on uncertainty estimates. Results clarify how forcing uncertainties can overwhelm initial-condition spread in boreal summer due to strong direct radiative impact, while the effect of initial conditions predominate in winter, when dynamics contribute to large ensemble spread. In our setup, current uncertainties affecting reconstruction-simulation comparisons prevent conclusions about the magnitude of the Tambora eruption and its relation to the “year without summer.”

**Plain Language Summary** Strong volcanic eruptions are a major natural forcing of climate, and there is increasing awareness of their importance for understanding and prediction of climate. Much of our current understanding is based on past eruptions for which there is no direct observation and that are therefore subject to substantial uncertainties regarding the eruption characteristics, the associated volcanic forcing, and the climatic response. With these premises, our study tackles the question of whether, for a past or analogously future eruption, it is more important to constrain the uncertainty in the volcanic forcing it generates rather than the initial climate conditions upon which it occurs, in order to obtain robust information about the associated climate responses. We use the 1815 eruption of Mount Tambora as a test case. We demonstrate that the climatic effects of initial condition uncertainties can overwhelm those of forcing uncertainties especially in winter and at the regional scale. We show how the European “year without a summer” seen in climate reconstructions is compatible with very different combinations of forcing magnitude and initial conditions.

### 1. Introduction

Stratospheric volcanic aerosols are a major natural forcing factor influencing climate variability on subseasonal-to-multicentennial time scales (e.g., Robock, 2000; Timmreck, 2012; Zanchettin, 2017). Yet our understanding of volcanically forced climate variability remains limited by several gaps of knowledge regarding the forcing as well as the climate response.

Our understanding of volcanically forced climate variability is largely based on the study of past events: Reconstructions of volcanic forcing for such events is based on proxy data such as ice core sulfate records showing close agreement with variations in large-scale temperature records (e.g., Sigl et al., 2015). However, there are significant uncertainties associated with the forcing produced by any specific eruption, which for tropical eruptions are primarily controlled by the amount of sulfur injected into the stratosphere (Toohey & Sigl, 2017). Further uncertainties in volcanic radiative forcing may arise due to the impact of aerosol size distributions and spatial and temporal evolution due to, for example, influences of stratospheric circulation and chemistry (e.g., LeGrande et al., 2016; Timmreck, 2012). Recent studies have also demonstrated that the climate response to a specific eruption can critically depend on the climatic conditions at the time of

the eruption (e.g., Gagné et al., 2017; Moreno-Chamarro et al., 2015; Zanchettin et al., 2012; Zhong et al., 2011). These initial conditions are the expression of the interplay between background mean climate state, which is determined by the cumulative effect of external forcings, and ongoing intrinsic climate variability. Initial conditions are not merely a source of additive noise for posteruption climate variability but can actively influence the mechanisms involved in the posteruption climate evolution (Pausata et al., 2016; Swingedouw et al., 2015, 2017; Zanchettin et al., 2013).

A key, yet unaddressed, scientific question is whether constraining the magnitude of forcing estimates is more important than constraining initial conditions to accurately simulate the climate response to a specific volcanic eruption or attribute anomalous weather/climate to volcanic forcing. This question is relevant from a paleoclimate perspective given the still poorly constrained current ranges of reconstructed and simulated climate anomalies during preindustrial periods of strong volcanic activity (e.g., PAGES-Hydro2k-Consortium, 2017; Stoffel et al., 2015) and especially the large discrepancies between simulations and reconstructions over certain regions (e.g., PAGES-Hydro2k-Consortium, 2017; Stevenson et al., 2017; Zanchettin et al., 2015). The same question is also relevant from a climate prediction perspective, when the potential climatic implications of a future eruption of unknown characteristics and timing are explored (e.g., Boer et al., 2016; Illing et al., 2018).

This study tackles this question by systematically investigating the relative role of forcing and initial-condition uncertainties on the spread of posteruption surface climate anomalies in a large ensemble of coupled climate simulations of the 1815 Tambora eruption in Indonesia. The Tambora eruption is the strongest volcanic event of the last 500 years, with very clear volcanic sulfate signals observed in polar ice cores (Marshall et al., 2018; Sigl et al., 2015); its volcanic stratospheric sulfur injection is comparatively well constrained compared to other eruptions of the past millennium (Toohey & Sigl, 2017), whereas its radiative forcing remains affected by substantial uncertainty (Raible et al., 2016; Zanchettin et al., 2016); it is regarded as a test case for high-impact volcanic eruptions, with the 1816 “year without a summer” over Europe in its aftermath being iconic of posteruption regional cooling (Raible et al., 2016); it was preceded by a VEI-6 tropical eruption of unknown location in 1809 (Cole-Dai et al., 2009; Guevara-Murua et al., 2014); it is included in the protocol of the CMIP6-endorsed “Climate Model Intercomparison on the climatic response to volcanic forcing” (VolMIP, Zanchettin et al., 2016).

To quantify the relative contribution of forcing uncertainties and initial-condition spread to posteruption climate anomalies, we test how three climate simulation ensembles using the same set of initial conditions but different realistic volcanic forcing are distinguished by a blind cluster analysis. We also use climate reconstructions as a target to determine how both uncertainties affect simulation-reconstruction comparative assessments.

## 2. Data and Methods

### 2.1. Volcanic Forcing and Ensemble Simulations

The eVolv2k volcanic forcing reconstruction provides volcanic stratospheric sulfur injection estimates and uncertainties based on polar ice cores (Toohey & Sigl, 2017). Based on eVolv2k, we use three forcing time series including the 1809 and Tambora eruptions: a central estimate (hereafter Best), consistent with that used in VolMIP and PMIP4-past1000 (Jungclauss et al., 2017); a high-end estimate, corresponding to the best estimate plus 2 times the ( $1\sigma$ ) sulfur emission uncertainty (hereafter High); a low-end estimate, corresponding to the best estimate minus 2 times the ( $1\sigma$ ) sulfur emission uncertainty (hereafter Low). Based on the sulfur injection estimates, the Easy Volcanic Aerosol v1.0 (Toohey et al., 2016) module is used to produce volcanic aerosol forcing sets, consisting of aerosol extinction, single scattering albedo, and asymmetry factor as a function of latitude, height, time, and wavelength.

The superensemble is constructed following a two-step approach that allows to account for the effects of uncertainties in the magnitude of the 1809 eruption on surface climates at the onset of the Tambora eruption (e.g., Zanchettin et al., 2013). In the first step, three volcanic-forcing-only climate simulation ensembles covering the period 1800–1814 CE are generated using the Best, High, and Low volcanic forcing estimates for the 1809 eruption. Each ensemble consists of 10 simulations initialized on 1 January 1800 from different states sampled at centennial-scale intervals from a 1,200-year-long unperturbed (constant forcing,

excluding background volcanic aerosols) preindustrial control simulation. Therefore, we obtain 30 restart files for 1 January 1815 that describe a range of climate conditions accounting for uncertainty in the forcing of the 1809 eruption as well as the effect of intrinsic climate variability.

In the second step, the 30 restart files are used to initialize ensemble simulations covering the period 1815–1819 CE using the Best, High, and Low volcanic forcings for the Tambora eruption in April 1815 (some simulations are integrated to 1829). The so obtained three 30-member ensembles yield a superensemble of  $30 * 3 = 90$  simulations of the Tambora eruption.

Simulations are conducted with the low-resolution version of the Max-Planck-Institute Earth-System-Model (MPI-ESM-LR). MPI-ESM-LR couples the atmospheric general circulation model ECHAM6 run in its T63L47 configuration, which resolves the stratosphere up to 0.01 hPa with 21 model levels above 100 hPa (Stevens et al., 2013), with the ocean-sea ice model MPIOM in its GR15 configuration (Jungclaus et al., 2013). This configuration is the reference version of MPI-ESM for paleo-applications and has been widely tested in the context of the climate of the last millennium (e.g., Jungclaus et al., 2014; Moreno-Chamorro, Zanchettin, Lohman, & Jungclaus, 2017; Moreno-Chamorro, Zanchettin, Lohmann, Luterbacher, & Jungclaus, 2017; Zanchettin et al., 2015).

## 2.2. Cluster Analysis

To determine the relative roles of forcing and initial conditions on the simulated climate response to the Tambora eruption, we compare the three original ensembles (Best, High, and Low) with three clusters of simulations identified by performing a  $k$ -means cluster analysis (Lloyd, 1982) on the 90-member superensemble output. In practice, the  $k$ -means clustering partitions the 90 simulations into three clusters by minimizing the sum, over all clusters, of the within-cluster sums of squared Euclidean distances between the individual surface temperature anomalies and the ensemble-mean anomaly. So whereas the original ensembles differ by construction only for the different volcanic forcing by the Tambora eruption, the clusters are generated based on feature similarity and reflect the combined effect of forcing and initial-condition uncertainties on the spread of simulated post-Tambora climate anomalies.

If the original ensembles are mostly preserved, that is, most realizations of each ensemble fall into just one distinct cluster, we infer that the posteruption state is controlled by differences in the applied forcing compared to differences in the initial conditions. Otherwise, we infer that initial conditions significantly contribute to determine the climatic response compared to the forcing uncertainty, and they dominate when there is an almost equal spread of the original ensembles across the three clusters.

We use scatterplots of response ( $y$  axis) versus initial conditions ( $x$  axis) to further elucidate the importance of the initial conditions in determining the response. A sharp horizontal stratification between clusters indicates that the different forcings dominate over the initial-conditions spread. Considered responses encompass regional, hemispheric, and global average near-surface air temperature. Unless specified otherwise, we use global-mean near-surface air temperature in January–March 1815 to illustrate the initial conditions.

We also define a simple metric,  $F$ , that quantifies the degree of fragmentation of each original ensemble by the  $k$ -means clustering:  $F_i = (n_i - 1 + \prod_{k=1..3} C_{k,i})/h$ , where  $n_i$  is the number of clusters containing data from ensemble  $i$  obtained by the  $k$ -means procedure,  $C_{k,i}$  is the number of simulations from ensemble  $i$  that fall in cluster  $C_k$ , and  $h = 1,002$  is a normalizing constant.  $F$  ranges between 0 (no fragmentation, the original ensemble is completely preserved) and 1 (maximum fragmentation, with equal repartition of the original ensemble in the three clusters).

Unless reported otherwise, analysis is performed based on seasonal averages calculated from monthly mean anomalies from the associated climatological mean in the control simulation.

The cluster analysis is also performed using the absolute value of the difference between our simulations and a state-of-the-art reconstruction of summer European-mean near-surface air temperature (Luterbacher et al., 2016). The absolute value highlights the accuracy of the simulations independent of whether they overestimate rather than underestimate the post-Tambora cooling. In this case, initial condition is chosen as the last preeruption summer available in the reconstruction: This is summer 1814 since there is a cooling signal in summer 1815. A bivariate extension of cluster analysis is performed, where distances are calculated between points in the two-dimensional space of initial conditions and posteruption anomalies.

### 3. Results

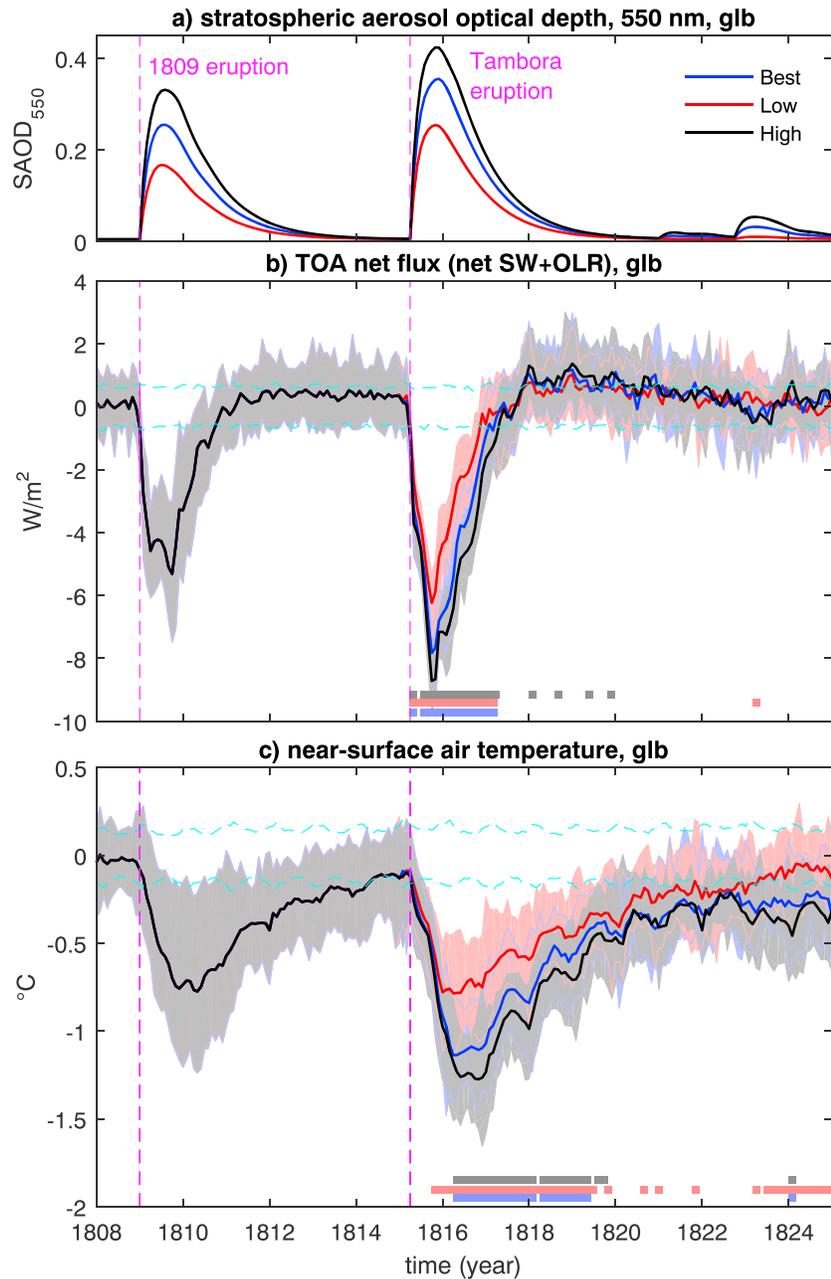
Figure 1 provides a classic representation of the volcanic forcing (panel a), the associated top-of-atmosphere global radiative flux anomaly (b), and global-mean near-surface air temperature evolution (c) simulated by the three ensembles. The mean evolutions and spreads overlap before 1815, as each of the Best, High, and Low ensembles contains all 30 parent simulations and therefore spans all three different forcing estimates of the 1809 eruption (see section 2.1). Ensemble-mean global radiative flux anomalies are within the range of internal variability at the onset of the Tambora eruption, whereas the global surface climate spans colder conditions than those determined by internal variability alone, indicating persistence of cooling initiated by the 1809 eruption. Distributional differences between ensembles (marked at the bottom of panels b and c) are significant in both variables at the peak of the Tambora forcing, despite ensemble ranges overlapping especially for temperature. More in detail, the Low forcing produces peak radiative flux anomalies that are about one-third weaker than the High forcing, whereas the Best forcing produces radiative flux anomalies that are only about 10–13% weaker than the High forcing. This asymmetry is similarly apparent in the global surface responses and does not emerge as clearly in the imposed aerosol optical properties.

Figure 2 illustrates scatterplots of climate response versus initial conditions for boreal summer (June–July–August) and winter (January–February–March) near-surface air temperatures spatially averaged over the whole globe (GST), the Northern Hemisphere (NHT), and two example continental-scale regions: Europe (EuT, 10°W:40°E; 35:70°N) and North America (NAT, 190:300°E; 25:70°N). At the global and hemispheric scales, summer responses are clearly stratified, with colder anomalies for the High ensemble and smaller negative anomalies for the Low ensemble; there is anyway at least some overlap between all ensembles. All simulations are outside the range of internal variability, in both GST and NHT (Figures 2a and 2c). In winter, stratification of original ensemble responses is less clear; hence, ensembles more strongly overlap particularly for NHT; a few realizations lie within the range of internal variability (Figures 2b and 2d). In all cases, the clusters differ from the original ensembles. For summer, this especially occurs because of the overlaps between the Best and High ensembles, as almost all simulations in the Low ensemble fall instead into the same cluster, whereas the High ensemble is almost split in two. In winter, simulations in the original ensembles are always distributed across all three clusters.

At the continental scale, we again observe that posteruption anomalies are larger and more clearly outside the range of internal variability in summer than in winter (Figures 2e–2h). In fact, winter responses exceed the range of internal variability only in a minority or even just a few of realizations, belonging to all the original ensembles. In winter, the coldest posteruption NAT anomalies are observed in the Low and Best ensembles, whereas some of the simulations with warmest posteruption anomalies belong to the High ensemble. A seasonal asymmetry is clear in the results of the clustering of both regions: In summer, original forcing-based ensembles remain rather preserved; in winter, they typically spread almost equally among all three clusters.

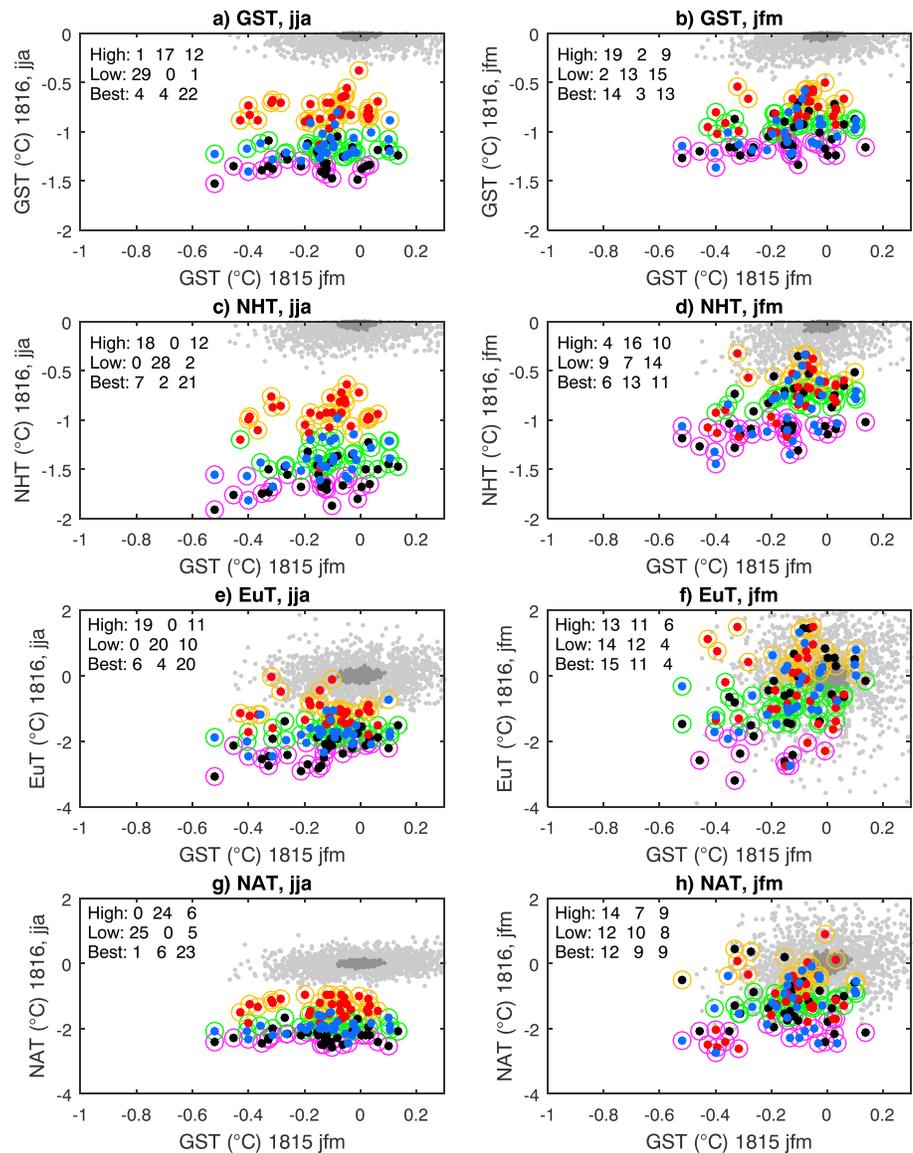
Overall, the imposed volcanic forcing largely determines the summer temperature response at global to continental scales. Nonetheless, initial conditions influence the response in winter already at the global scale, causing substantial overlapping of the ensemble distributions and thereby producing, occasionally, indistinguishable responses. Similar conclusions can be drawn at the hemispheric and continental scales, where EuT and NAT show a tendency toward a stronger spread of original ensembles into clusters by means of initial conditions compared to GST and NHT, and posteruption winter anomalies more closely superpose on the range of internal variability.

The importance of initial conditions increases at the grid point level. Figure 3 illustrates the fragmentation of original ensembles across the different clusters by the *k*-means algorithm applied on each summer and winter grid-point near-surface air temperature anomaly. The predominance of greenish and yellowish over blueish tones indicates overall substantial fragmentation of the original ensembles. This is particularly evident for the Best ensemble in both summer and winter, whereas the High and especially the Low ensembles present extensive regions of both ocean and land characterized by small fragmentation (i.e., blueish tones). Such regions differ in both ensembles and seasons and generally trace back to geographical features and features linked to the general circulation. They highlight how the specific forcing leads to distinguishable feedbacks. Weak fragmentation is observed, for instance, over the Arctic region in the Low ensemble in summer, likely reflecting a weaker sea-ice buildup in this ensemble, and over the West Pacific warm pool region and more



**Figure 1.** (a) Global-average monthly stratospheric aerosol optical depth at wavelength 550 nm produced by Easy Volcanic Aerosol v1.0; (b, c) ensemble mean (lines) and 5th–95th percentile intervals (shading) of global-mean top-of-atmosphere net radiative flux (positive means downward) and near-surface air temperature anomalies. Monthly anomalies are deviations from the respective climatology over the period 1800–1808. The three ensembles overlap by construction before 1815. Turquoise dashed lines are the 5th–95th percentile intervals for signal occurrence in the control run. Bottom rectangles indicate periods when an ensemble (color code as for the time series plots) is significantly different ( $p = 0.05$ ) from both the other two ensembles according to the Mann-Whitney  $U$  test. SAOD = stratospheric aerosol optical depth; TOA = top-of-atmosphere; SW= shortwave radiation; OLR= outgoing longwave radiation.

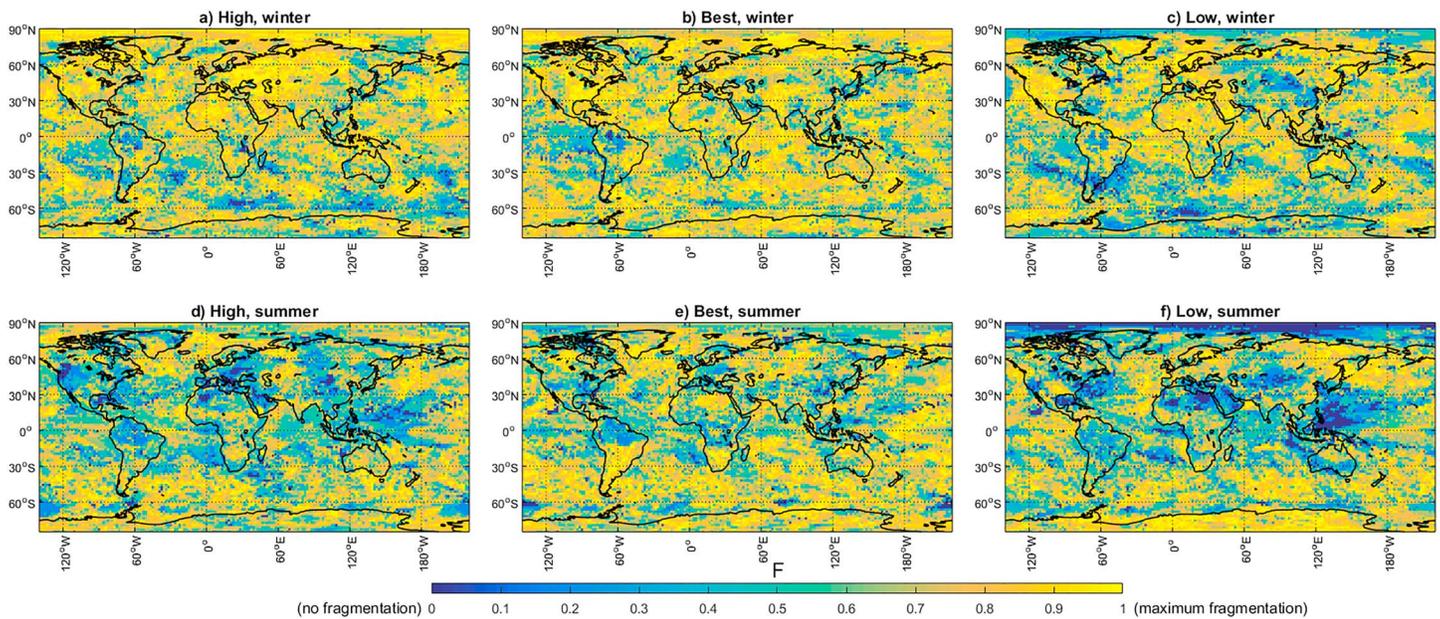
generally over monsoon regions, in both High and Low ensembles (Figures 3d and 3f). The Amazon region is bluish in all ensembles and both seasons, suggesting a strong dependency of the vegetation response on the magnitude of the applied forcing. Conversely, the extensive yellowish tone over Europe in winter suggests that forcing uncertainties are negligible for understanding posteruption anomalies in this area. Therefore, if a dynamical response mechanism exists, its activation and strength are strongly determined by initial



**Figure 2.** Scatterplot of surface temperature response (y axis) versus initial conditions (x axis) for (a, b) GST, (c, d) NHT, (e, f) EuT, and (g, h) NAT, for the summer 1816 (left column) and winter 1815/1816 (right column) responses. Small filled circles are the original ensembles (color code as in Figure 1); large empty circles are the clusters. The top three text lines in each panel report how the members of each of the original ensembles are redistributed across the clusters. The clouds of small gray dots show the range of variations that can be associated to internal variability through sampling along the control simulation (light: single simulation; dark: 30-member ensemble mean). GST = near-surface air temperatures spatially averaged over the whole globe; NHT = near-surface air temperatures spatially averaged over the Northern Hemisphere; EuT = near-surface air temperatures spatially averaged over Europe; NAT = near-surface air temperatures spatially averaged over North America.

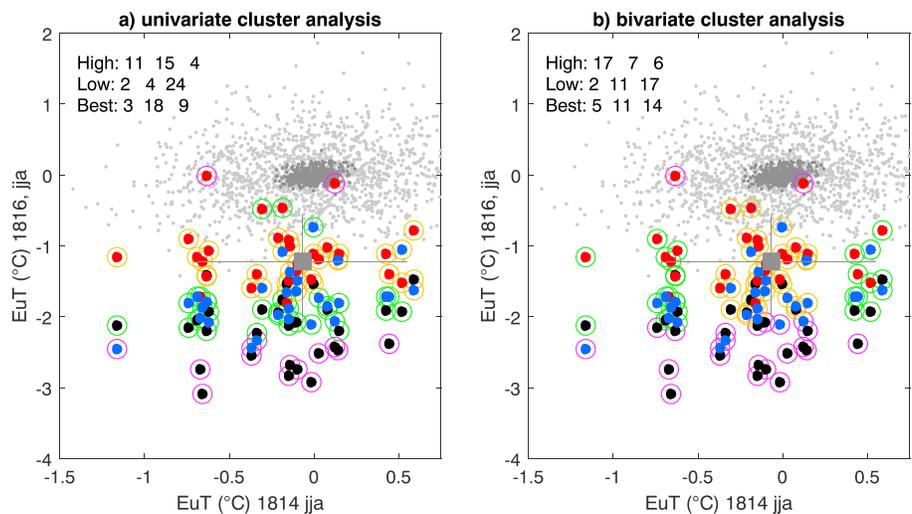
conditions and its prediction can be hampered by internal variability (see also the large ensemble spreads in Figure 2f).

From a paleoclimate perspective, a key issue to address is how the simulated responses compare with reconstructions. We therefore repeat the 1816 summer EuT analysis feeding the cluster algorithm with the absolute difference between simulated and reconstructed values. Considering only posteruption anomalies, there is substantial scatter of the Best and High ensembles across the three clusters, whereas most realizations in the Low ensemble fall in the cluster surrounding the reconstructed estimate (Figure 4a). Nonetheless, this “best matching” cluster contains realizations from all original ensembles, as does the “worst matching”



**Figure 3.** Fragmentation of grid point NH winter 1815/1816 (a–c) and summer 1816 (d–f) near-surface air temperature anomalies from the preeruption (1810–1814) climatology in the (a, d) High, (b, e) Best, and (c, f) Low ensembles.

cluster although it is mostly populated by simulations from the High ensemble. Including also the simulation-reconstruction absolute difference in the preeruption conditions in the analysis (Figure 4b), the clustering algorithm fragments the Low ensemble into two clusters, yielding less Low simulations to fall into the best matching cluster. The latter now includes more simulations from the High ensemble, which produces a stronger 1816 summer cooling than the reconstruction but compares better with the reconstructed preeruption conditions. Most realizations from the High ensemble remain nonetheless part of the worst matching cluster. The scatterplot clarifies that all ensembles contain realizations that lie within the uncertainty range of the reconstruction, although most members of the High ensemble



**Figure 4.** Scatterplot of climate response versus initial conditions for the 1816 EuT. Initial conditions are based on summer 1814 EuT anomalies. Clustering is performed on the difference between the simulated anomalies and the EuT reconstruction by Luterbacher et al. (2016), indicated by the gray square with associated uncertainty range (bars). (a) Cluster analysis on summer 1816 EuT anomalies alone (univariate); (b) cluster analysis on both summer 1816 EuT and summer 1814 EuT anomalies (bivariate). Representation as in Figure 2. Note that the data are the same in both panels, which differ only for the results of the clustering. EuT = near-surface air temperatures spatially averaged over Europe.

typically overestimate the reconstructed response. Most realizations lie outside the range of natural variability, yet their large scattering indicates a substantial impact of initial conditions.

Overall, the Low ensemble yields the most compatible results with reconstructed features of the European year without a summer, possibly reflecting the overestimation of posteruption surface cooling often observed in climate models (Marotzke & Forster, 2015); however, its substantial superposition with realizations of the Best ensemble in the surrounding of the reconstructed value within its range of uncertainty prevents concluding that the forcing difference matters. Further, the High ensemble overall compares poorly with the reconstruction, suggesting an overestimation of the forcing; yet it contains realizations that under favorable initial conditions are consistent with the reconstructed estimate and indistinguishable from realizations from the Best and even Low ensemble.

#### 4. Summarizing Discussion

Our results demonstrate that uncertainties in initial conditions can prevail or even dominate the impact of (realistic) choices on the eruption's magnitude on the surface temperature response to such a large eruption. Especially for winter, the interensemble overlap of posteruption temperature anomalies hinders conclusive assessments about the forcing magnitude and response pathways compatible with a certain volcanic signature on certain regional surface climates. Otherwise stated, different realistic combinations of volcanic forcing and initial conditions lead to indistinguishable temperature responses. This implies that improved constraints on the Tambora forcing would not allow for better understanding of the temperature response to this eruption within current modeling frameworks, unless this is accompanied by substantial progress in the constraint of initial conditions. Accounting for volcanic forcing uncertainty seems nonetheless necessary, as the use of just a current best estimate can bias the interpretation of reconstructed responses. This was exemplified for the European year without a summer as the cluster yielding the best correspondence between reconstructed and simulated features mainly contains realizations from an ensemble using the low-end estimate of volcanic forcing. In this sense, the classical truth-centered approach for ensemble analysis, where the ensemble mean and spread are regarded as forced response and uncertainty due to internal climate variability, respectively, may bring to misleading conclusions in reconstruction-simulations comparisons. On the other hand, certain continental and subcontinental regions appear to be particularly sensitive to the magnitude of volcanic forcing especially in boreal summer (e.g., North America in MPI-ESM-LR), which may provide guidance on identifying locations where climate proxies are most sensitive to the direct radiative impacts of volcanic eruptions.

The general validity of these conclusions stands beyond the single climate model used here, as different models currently used in paleoclimate applications share similar ranges of internal variability and climate sensitivity (e.g., PAGES2k-PMIP3 group, 2015). Generalization of our conclusions must consider that the forcing uncertainty used here accounts only for uncertainty in the volcanic stratospheric sulfur injection, not uncertainties related to the translation of sulfur injection into aerosol and radiative properties as performed here by the Easy Volcanic Aerosol v1.0 module. Such uncertainties result from, for example, the poorly constrained aerosol size distribution for eruptions larger than those recently observed (Toohey et al., 2016) and variations in the spatiotemporal evolution of the forcing due to differences in atmospheric circulation and sulfur injection height. Furthermore, the estimate of uncertainty in volcanic stratospheric sulfur injection for Tambora is smaller than most eruptions of the past 2,500 years (Toohey & Sigl, 2017): Robust quantitative analysis for specific eruptions thus requires an ad hoc design. Our results are based on the strongest volcanic event of the past 500 years; hence, they provide an overall strong signal-to-noise ratio in the volcanic surface imprint. Yet simulated posteruption winter anomalies of North American and European temperature only occasionally exceed the range of internal variability, encompassing positive and negative values. For eruptions smaller than Tambora, hence more frequent but with lower signal-to-noise ratio (e.g., Pinatubo-like), the continental winter responses generated by different forcing magnitudes would be even more mixed by the initial conditions. In a future prediction perspective, beyond the near-term prediction horizon where the prediction system is initialized through the assimilation of observed data, our results thus demonstrate that accurate sampling of initial conditions can predominate over specific choices made regarding the volcanic forcing in terms of surface temperature spread in some regions and particularly for the Northern Hemisphere winter season.

Finally, the summer-winter asymmetry in the ensemble spread of posteruption anomalies reflects the most prominent contribution of internal dynamics to the winter response. Its realism thus depends on the quality of the simulated intrinsic climate variability. This stresses the need to improve understanding of internal climate variability and how its ongoing phase can modulate the strength of the feedbacks initiated by the volcanic forcing, as a crucial aspect to identify and understand mechanisms determining the climatic response to volcanic eruptions.

### Acknowledgments

Computations were performed at the German Climate Computer Center (DKRZ). D. Z. acknowledges support by the DAIS-IRIDE project (VolClim). C. T. acknowledges support from the research programme “MiKlip” (FKZ:01LP1517B) of the German federal Ministry of Education (BMBF) and the European Union project StratoClim (FP7-ENV.2013.6.1-2). J. J. acknowledges support from the JPI-Belmont Forum’s project PaCMEDy-Paleo Constraints on Monsoon Evolution and Dynamics (BMBF FKZ: 01LP1607B). M. T. acknowledges support by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority programme “Antarctic Research with comparative investigations in Arctic ice areas” (grant TO 967/1-1). The spatially integrated data used in this study are provided as supporting information.

### References

- Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., et al. (2016). The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geoscientific Model Development*, *9*(10), 3751–3777. <https://doi.org/10.5194/gmd-9-3751-2016>
- Cole-Dai, J., Ferris, D., Lanciki, A., Savarino, J., Baroni, M., & Thiemens, M. H. (2009). Cold decade (AD 1810–1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption. *Geophysical Research Letters*, *36*, L22703. <https://doi.org/10.1029/2009GL040882>
- Gagné, M.-E., Kirchmeier-Young, M. C., Gillett, N. P., & Fyfe, J. C. (2017). Arctic sea ice response to the eruptions of Agung, El Chichón and Pinatubo: Arctic sea ice response to volcanoes. *Journal of Geophysical Research: Atmospheres*, *122*, 8071–8078. <https://doi.org/10.1002/2017JD027038>
- Guevara-Murua, A., Williams, C. A., Hendy, E. J., Rust, A. C., & Cashman, K. V. (2014). Observations of a stratospheric aerosol veil from a tropical volcanic eruption in December 1808: Is this the unknown ~1809 eruption? *Climate of the Past*, *10*(5), 1707–1722. <https://doi.org/10.5194/cp-10-1707-2014>
- Illing, S., Kadow, C., Pohlmann, H., & Timmreck, C. (2018). Assessing the impact of a future volcanic eruption on decadal predictions. *Earth System Dynamics*, *9*(2), 701–715. <https://doi.org/10.5194/esd-9-701-2018>
- Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., et al. (2017). The PMIP4 contribution to CMIP6—Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 past 1000 simulations. *Geoscientific Model Development*, *10*(11), 4005–4033. <https://doi.org/10.5194/gmd-10-4005-2017>
- Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*, 422–446. <https://doi.org/10.1002/jame.20023>
- Jungclaus, J. H., Lohmann, K., & Zanchettin, D. (2014). Enhanced 20th-century heat transfer to the Arctic simulated in the context of climate variations over the last millennium. *Climate of the Past*, *10*(6), 2201–2213. <https://doi.org/10.5194/cp-10-2201-2014>
- 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>
- Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Luterbacher, J., Werner, J. P., Smerdon, J. E., Fernández-Donado, L., González-Rouco, F. J., Barriopedro, D., et al. (2016). European summer temperatures since Roman times. *Environmental Research Letters*, *11*(2), 024001. <https://doi.org/10.1088/1748-9326/11/1/024001>
- Marotzke, J., & Forster, P. M. (2015). Forcing, feedback and internal variability in global temperature trends. *Nature*, *517*(7536), 565–570. <https://doi.org/10.1038/nature14117>
- Marshall, L., Schmidt, A., Toohey, M., Carslaw, K. S., Mann, G. W., Sigl, M., et al. (2018). Multi-model comparison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora. *Atmospheric Chemistry and Physics*, *18*(3), 2307–2328. <https://doi.org/10.5194/acp-18-2307-2018>
- Moreno-Chamarro, E., Zanchettin, D., Lohmann, K., & Jungclaus, J. H. (2015). Internally generated decadal cold events in the northern North Atlantic and their possible implications for the demise of the Norse settlements in Greenland. *Geophysical Research Letters*, *42*, 908–915. <https://doi.org/10.1002/2014GL062741>
- Moreno-Chamarro, E., Zanchettin, D., Lohmann, K., & Jungclaus, J. H. (2017). An abrupt weakening of the subpolar gyre as trigger of Little Ice Age-type episodes. *Climate Dynamics*, *48*(3–4), 727–744. <https://doi.org/10.1007/s00382-016-3106-7>
- Moreno-Chamarro, E., Zanchettin, D., Lohmann, K., Luterbacher, J., & Jungclaus, J. H. (2017). Winter amplification of the European Little Ice Age cooling by the subpolar gyre. *Scientific Reports*, *7*(1), 9981. <https://doi.org/10.1038/s41598-017-07969-0>
- PAGES Hydro2k Consortium (2017). Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past*, *13*, 1851–1900. <https://doi.org/10.5194/cp-13-1851-2017>
- PAGES 2k-PMIP3 group (2015). Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Climate of the Past*, *11*, 1673–1699. <https://doi.org/10.5194/cp-11-1673-2015>
- Pausata, F. S. R., Karamperidou, C., Caballero, R., & Battisti, D. S. (2016). ENSO response to high-latitude volcanic eruptions in the Northern Hemisphere: The role of initial conditions. *Geophysical Research Letters*, *43*, 8694–8702. <https://doi.org/10.1002/2016GL069575>
- Raible, C. C., Brönnimann, S., Auchmann, R., Brohan, P., Frölicher, T. L., Graf, H. F., et al. (2016). Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. *WIREs Climate Change*, *7*(4), 569–589. <https://doi.org/10.1002/wcc.407>
- Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics*, *38*(2), 191–219. <https://doi.org/10.1029/1998RG000054>
- 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*(7562), 543–549. <https://doi.org/10.1038/nature14565>
- Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., et al. (2013). Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*, 146–172. <https://doi.org/10.1002/jame.20015>
- Stevenson, S., Fasullo, J., Otto-Bliessner, B. L., Tomas, R., & Gao, C. 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>
- 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>
- 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 Planetary Change*, *150*, 24–45. <https://doi.org/10.1016/j.gloplacha.2017.01.006>

- Swingedouw, D., Ortega, P., Mignot, J., Guilyardi, E., Masson-Delmotte, V., Butler, P. G., & Khodri, M. (2015). Bidecadal North Atlantic Ocean circulation variability controlled by timing of volcanic eruptions. *Nature Communications*, 6(1), 6545. <https://doi.org/10.1038/ncomms7545>
- Timmreck, C. (2012). Modeling the climatic effects of large volcanic eruptions. *WIREs Climate Change*, 3(6), 545–564. <https://doi.org/10.1002/wcc.192>
- 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>
- Toohey, M., Stevens, B., Schmidt, H., & Timmreck, C. (2016). Easy Volcanic Aerosol (EVA v1.0): An idealized forcing generator for climate simulations. *Geoscientific Model Development*, 9(11), 4049–4070. <https://doi.org/10.5194/gmd-9-4049-2016>
- Zanchettin, D. (2017). Aerosol and solar irradiance effects on decadal climate variability and predictability. *Current Climate Change Reports*, 3(2), 150–162. <https://doi.org/10.1007/s40641-017-0065-y>
- Zanchettin, D., Bothe, O., Graf, H. F., Lorenz, S. J., Luterbacher, J., Timmreck, C., & Jungclaus, J. H. (2013). Background conditions influence the decadal climate response to strong volcanic eruptions. *Journal of Geophysical Research: Atmospheres*, 118, 4090–4106. <https://doi.org/10.1002/jgrd.50229>
- Zanchettin, D., Bothe, O., Lehner, F., Ortega, P., Raible, C. C., & Swingedouw, D. (2015). Reconciling reconstructed and simulated features of the winter Pacific/North American pattern in the early 19th century. *Climate of the Past*, 11(6), 939–958. <https://doi.org/10.5194/cp-11-939-2015>
- 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>
- Zhong, Y., Miller, G. H., Otto-Bliesner, B. L., Holland, M. M., Bailey, D. A., Schneider, D. P., & Geirsdottir, A. (2011). Centennial-scale climate change from decadal-paced explosive volcanism: A coupled sea ice–ocean mechanism. *Climate Dynamics*, 37(11–12), 2373–2387. <https://doi.org/10.1007/s00382-010-0967-z>