Environmental Research Letters

LETTER • OPEN ACCESS

A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures

View the article online for updates and enhancements.

Environmental Research Letters

CORRIGENDUM

Corrigendum: A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures (2017 *Environ. Res. Lett.* **12** 094005)

L Schneider[1,6] , J E Smerdon[2], F Pretis[3,4], C Hartl-Meier[5] and J Esper[5]

1   Department of Geography, Justus Liebig University, Gießen, Germany
2   Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, United States of America
3   Programme for Economic Modelling, INET at the Oxford Martin School, University of Oxford, Oxford, United Kingdom
4   Department of Economics, University of Oxford, Oxford, United Kingdom
5   Department of Geography, Johannes Gutenberg University, Mainz, Germany
6   Author to whom any correspondence should be addressed.

E-mail: lea.schneider@geogr.uni-giessen.de

In the original version of this paper, figure 2 reports the time-integrated radiative forcing units as Jm^{-2}. While this unit is dimensionally correct, to be numerically correct the units should be $Wm^{-2}yr$. For forcing estimates in the main text, extracted from the volcanic forcing record, the units similarly require correction. The error arose from an attempt to correct the units of the time-integrated forcing estimates from the original source (Sigl *et al* 2015), where values were presented in Wm^{-2} instead of $Wm^{-2}yr$. A version of figure 2 and the section 'The relationship between forcing magnitude and temperature response', both with the corrected units, are presented below. The radiative forcing values themselves were not affected, nor are the results, discussion or conclusions of this paper.

The relationship between forcing magnitude and temperature response

Volcanic cooling is sensitive to the altitude, latitude, and character of the volcanic eruption (Hansen *et al* 1997). The relationship between cooling patterns and forcing estimates is thus expecetd to be variable. The ice core sulfate records agree, for example, on an enormous peak in 1258, but the reconstructed cooling is less extreme (figure 2(*c*)). Measurement and calibration uncertainties associated with single events in ice core and tree-ring derived reconstructions further complicate such comparisons and impede the verification of the forcing magnitude based on the climatic impact. This is assessed by fitting linear regressions between the detected breaks and sulfate peaks from G08, C13 and S15 (figure 2(*b*)). G08 and C13 cohere relatively well with the break coefficients, but their regression models are based on large intercepts (22.1 and 10.4 $Wm^{-2}yr$) which were physically expected to be zero. The S15 volcanic forcing is not significantly correlated with the break coefficients, mainly due to the differences in 1258 and 1453. The more minor breaks are associated with very small forcing events in G08 and C13 (smallest forcing = -0.5 $Wm^{-2}yr$ and -0.1 $Wm^{-2}yr$, respectively) and more substantial events in S15 (smallest forcing = -3.3 $Wm^{-2}yr$). The latter better explains a temperature response outside the range of internal variability and results in a linear regression with an intercept close to the origin for S15. The relatively weak forcing associated with the maximum cooling in 1453 points to another major inconsistency between forcing and temperature reconstructions in the 1450s. Although now off by 5 years, an integrated forcing of -20 $Wm^{-2}yr$ (in 1458) would better explain the strong 1453 cooling observed in the tree-ring reconstruction and would be well in line with the linear regression model. The pronounced radiative forcing in 1258 has been discussed previously (Timmreck *et al* 2009) and is likely too large due to nonlinear aerosol microphysics in the volcanic plume of that eruption.

ORCID iDS

L Schneider https://orcid.org/0000-0002-8208-7300

References

Sigl M *et al* 2015 Timing and climate forcing of volcanic eruptions for the past 2500 years *Nature* **523** 543–9
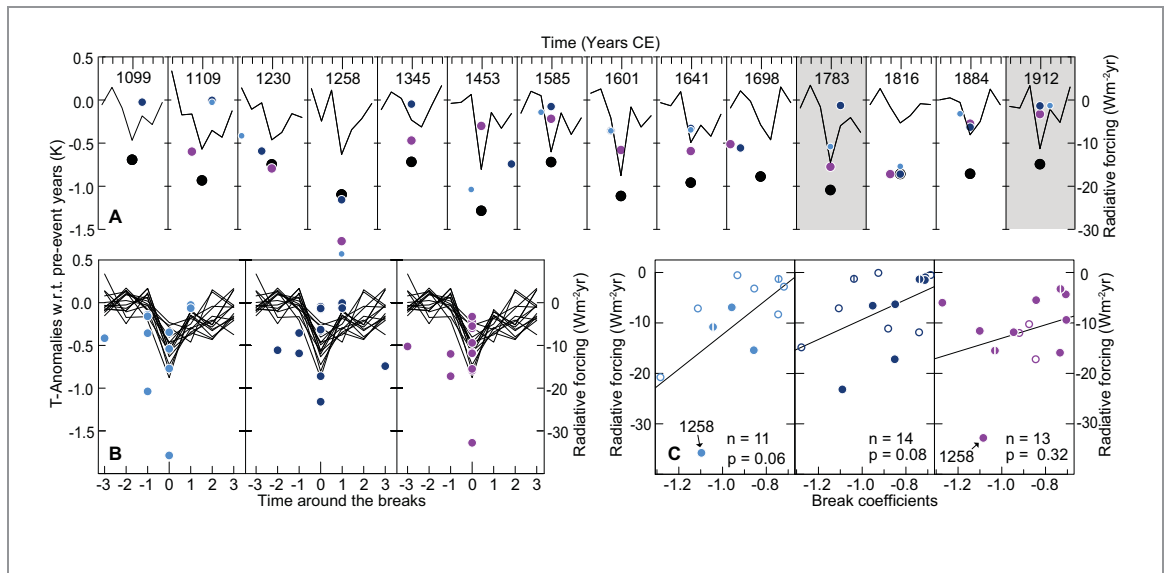
Environmental Research Letters

LETTER

A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures

L Schneider[1,6], J E Smerdon[2], F Pretis[3,4], C Hartl-Meier[5] and J Esper[5]

1   Department of Geography, Justus Liebig University, Gießen, Germany
2   Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, United States of America
3   Programme for Economic Modelling, INET at the Oxford Martin School, University of Oxford, Oxford, United Kingdom
4   Department of Economics, University of Oxford, Oxford, United Kingdom
5   Department of Geography, Johannes Gutenberg University, Mainz, Germany
6   Author to whom any correspondence should be addressed.

E-mail: lea.schneider@geogr.uni-giessen.de

Abstract

Information about past volcanic impact on climate is mostly derived from historic documentary data and sulfate depositions in polar ice sheets. Although these archives have provided important insights into the Earth's volcanic eruption history, the climate forcing and exact dating of many events is still vague. Here we apply a new method of break detection to the first millennium-length maximum latewood density reconstruction of Northern Hemisphere summer temperatures to develop an alternative record of large volcanic eruptions. The analysis returns fourteen outstanding cooling events, all of which agree well with recently developed volcanic forcing records from high-resolution bipolar ice cores. In some cases, however, the climatic impact detected with our new method peaks in neighboring years, likely due to either dating errors in the polar ice cores or uncertainty in the interpretation of atmospheric aerosol transport to polar ice core locations. The most apparent mismatches between forcing and cooling estimates occur in the 1450s and 1690s. Application of the algorithm to two additional and recently developed reconstructions that blend maximum latewood density and ring width data reproduces twelve of the detected events among which eight are retrieved in all three of the dendroclimatic reconstructions. Collectively, the new estimates of volcanic activity with precise age control provide independent evidence for forcing records during the last millennium. Evaluating the cooling magnitude in response to detected events yields an upper benchmark for the volcanic impact on climate. The average response to the ten major events in the density derived reconstruction is $-0.60\,°C \pm 0.13\,°C$. Other last millennium temperature records from proxies and model simulations reveal higher cooling estimates, which is, to some degree, related to the very different high frequency variance in these timeseries.

Introduction

Knowledge about past volcanism improves our understanding of the sensitivity of the climate system to exogenous radiative forcings. Associating large volcanic eruptions with climate anomalies and estimating their magnitude of impact requires a comprehensive and well-dated record of past volcanism (Esper *et al* 2013a). The majority of known large eruptions in the past millennium have been dated by

historical observations[7] (Global Volcanism Program 2013). In a climatological context, however, the amount and character of volcanic aerosols dispersed into the atmosphere are of greatest consequence and not well characterized by historical observations. In

---

7 The Global Volcanism Program lists 259 total events between 1000 CE and the present with VEI greater than or equal to 4 for the Northern Hemisphere and the tropics. Among these, 204 are dated using historical observations. Other dates are mainly derived from radiocarbon and tephrochronology.

particular, radiation absorbing sulfate aerosols injected into the stratosphere undergo multiyear transport and distribution resulting in significant alterations of the earth's energy budget (Robock 2000). In contrast to historical observations, volcanic particles deposited in polar ice sheets have been successfully used to estimate the amount and composition of sulfate aerosols from volcanic eruptions before the onset of modern observations. Combined with a model for atmospheric dispersion, ice core deposition records are used to derive radiative forcing reconstructions to evaluate the impacts of volcanic eruptions in climate model simulations (Schmidt *et al* 2011).

A concern for ice core derived estimates of volcanic events is their potential for dating errors due to uncertainties in the age-depth models or false assignment of reference horizons to certain eruptions (Baillie and McAneney 2015, Sigl *et al* 2015). Even for a correctly dated ash layer, the time lag between ash injection, atmospheric perturbations and polar deposition can be up to 2.5 years and therefore can result in misinterpretations of environmental effects (Plummer *et al* 2012). Irregular snow accumulation causes high spatial variability in the amount of deposited volcanic material and, hence, additional large uncertainties in the magnitude of sulfate records (Hegerl *et al* 2006, Sigl *et al* 2014). Finally, translating the sulfurous ash deposits into radiative forcing estimates requires several physical assumptions about the character of aerosols, their visual properties and atmospheric transport, and thus again adds further uncertainty to the quantification of volcanic history (Toohey *et al* 2016). These shortcomings motivate the need for alternative and independent approaches to reconstructing the timing and climatic impact of volcanic eruptions.

Previous studies of annually resolved climate proxies (Briffa *et al* 1998a), model simulations (Atwood *et al* 2016) and even early observational data (Jones *et al* 2003) have revealed volcanic eruptions to cause severe surface temperature cooling at hemispheric scales. The pulse-like forcing of these events triggers structural breaks (shifts in the mean) in time series of hemispheric or global mean temperatures, the magnitude of which exceed the natural range of climate variability (Naveau *et al* 2003). Our study is therefore motivated by the assumption that detecting and separating such breaks without prior knowledge of their occurrence will generate independent evidence of past volcanism constraining the timing and climatic effects of large eruptions.

We use a new reconstruction of mean Northern Hemispheric extra tropical summer temperatures derived from maximum latewood density (MXD) records (Schneider *et al* 2015) to test this hypothesis. Some recent work has suggested that tree-ring records smooth and underestimate the cooling associated with volcanic eruptions (Mann *et al* 2012). But these findings were rebutted by numerous studies that demonstrated a distinct volcanic signal in MXD data (Anchukaitis *et al* 2012, D'Arrigo *et al* 2013, Esper *et al* 2013a). In contrast, temperature reconstructions based on tree-ring width, the far more abundant dendrochronological parameter, were shown to underestimate abrupt cooling events due to biological memory effects (Esper *et al* 2015). The Schneider *et al* (2015) reconstruction is the first purely MXD-based record covering temperatures over more than a thousand summers and thus represents an ideal record for large eruptions. In this paper, we detect volcanic-induced breaks in this reconstruction independent of ice core estimates using an indicator saturation method. This technique has been described and validated in a pseudoproxy context (Pretis *et al* 2016) and is used to construct an independent chronology of volcanic eruptions that is then compared with existing ice core derived forcing records. We conclude by discussing the implications of our alternative record of past volcanism and evaluate our findings in the context of additional proxy reconstructions and climate model simulations.

Data and methods

Tree-ring reconstructions

The employed Northern Hemisphere (NH) summer temperature reconstruction is based solely on MXD data (Schneider *et al* 2015 (SCH15)). Although the MXD network provides far fewer records than one that includes tree-ring width chronologies, four North American, seven European and four Asian MXD chronologies are available, each longer than 600 years (figure 1(*a*)). Together these chronologies enable a skillful calibration and transfer into summer temperature estimates of the extra tropical NH. In contrast to the original SCH15 reconstruction, data processing herein was performed with a focus on high frequency variability: we removed age related noise from the density chronologies by calculating residuals from Hugershoff functions fit to the individual power-transformed MXD measurement series (Cook and Kairiukstis 1990, Cook and Peters 1997). This classical method (Briffa *et al* 1998b) better approximates the age trend and produces index chronologies that most accurately represent high-frequency variance, but these benefits come at the cost of losing multidecadal to multicentennial trends (Esper *et al* 2012). These low-frequency losses are nevertheless not relevant to our interest in short-term climate responses to volcanism.

We additionally use two recently developed summer temperature reconstructions derived from MXD plus tree-ring width networks (Stoffel *et al* 2015 (STO15), Wilson *et al* 2016 (WIL16)). STO15 chose site-specific detrending methods, combined local chronologies using principal components, and applied a high-pass filter to the final reconstruction for the
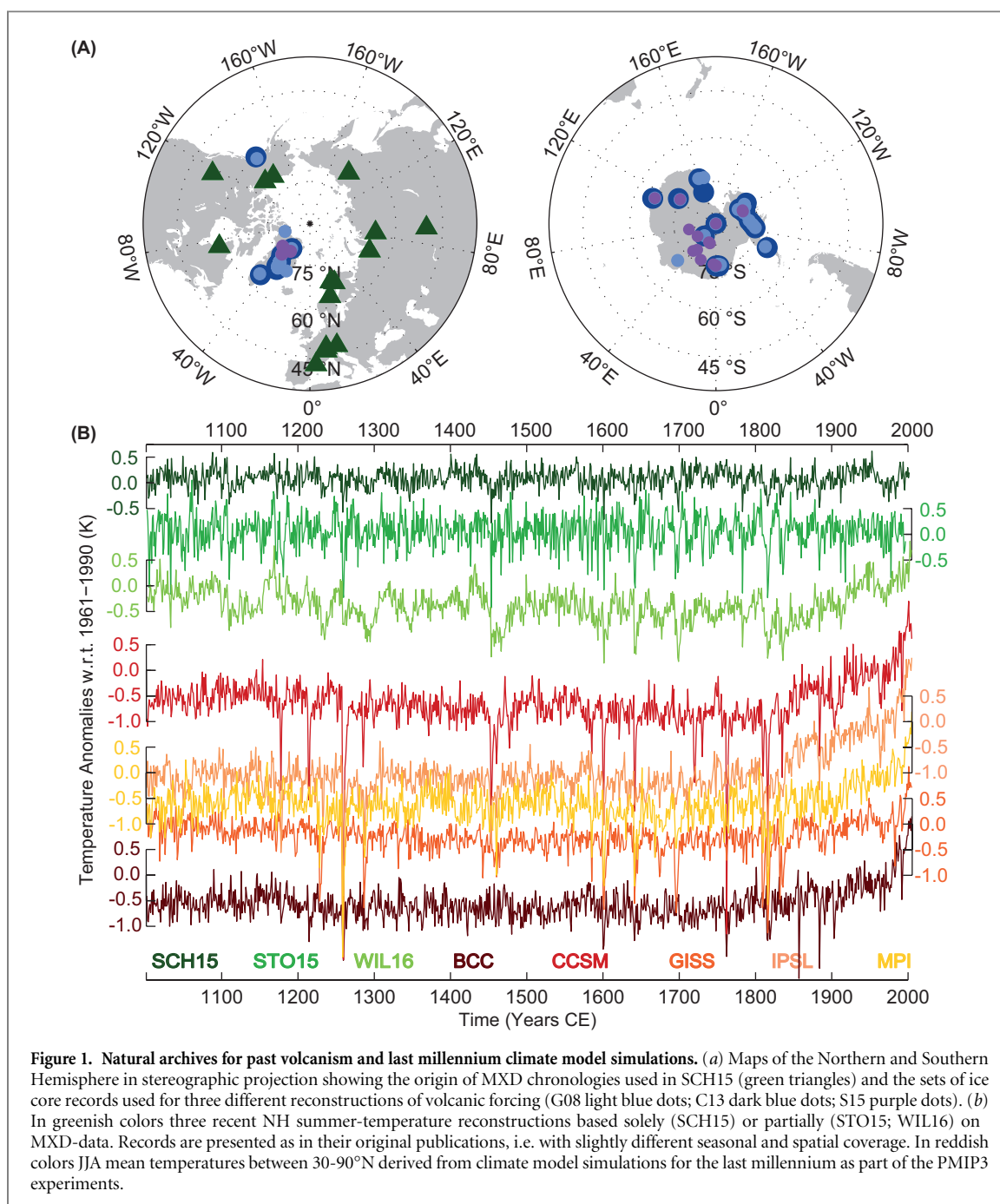
Figure 1. **Natural archives for past volcanism and last millennium climate model simulations.** (*a*) Maps of the Northern and Southern Hemisphere in stereographic projection showing the origin of MXD chronologies used in SCH15 (green triangles) and the sets of ice core records used for three different reconstructions of volcanic forcing (G08 light blue dots; C13 dark blue dots; S15 purple dots). (*b*) In greenish colors three recent NH summer-temperature reconstructions based solely (SCH15) or partially (STO15; WIL16) on MXD-data. Records are presented as in their original publications, i.e. with slightly different seasonal and spatial coverage. In reddish colors JJA mean temperatures between 30-90°N derived from climate model simulations for the last millennium as part of the PMIP3 experiments.

analysis of volcanic signals. Wilson *et al* (2016) aggregated published reconstructions, which were mostly detrended with a focus on preserving low-frequency variability.

Volcanic forcing records and model simulations
The third phase of the Paleoclimate Modelling Intercomparison Project (PMIP3) uses either the forcing timeseries of Gao *et al* (2008, G08) or Crowley and Unterman (2013, C13) as the volcanic component of the radiative boundary conditions in the last-millennium model simulations (Schmidt *et al* 2011). Both volcanic reconstructions combine sulfate concentration records from partly overlapping sets of polar ice cores (figure 1(*a*)) to model atmospheric sulfate aerosol loadings and optical depth changes.

The recently published record by Sigl *et al* (2015, S15) uses new but fewer ice core records in their radiative forcing estimates, which are not accompanied by an atmospheric transport model, but will be implemented in phase four of PMIP (Jungclaus *et al* 2016, Kageyama *et al* 2016). Despite their common purpose, the three volcanic reconstructions differ regarding the targeted volcanic eruption estimates, their spatial and temporal resolution and the estimated timing of events. For comparison with the dendrochronological records, we converted G08 and C13 into annual, global averages of radiative forcing (see SI1) using linear scalings (Stothers 1984, Wigley *et al* 2005). To further homogenize these estimates with the format of S15, we calculated time-integrated sums for each volcanic event from both the G08 and C13 records. The

integrated value was assigned to the year of peak forcing, because this is most relevant for the anticipated temperature response. In S15, atmospheric transport is not yet modeled and therefore peak forcing years cannot be determined. In this record, the timing of the integrated forcing events corresponds to the year in which the sulfate concentrations initially exceeded a pre-defined threshold (Sigl *et al* 2013, M. Sigl 2016, personal communication). Peak forcing can be expected in the same or subsequent year (Gao *et al* 2008, Crowley and Unterman 2013).

For assessing the effect of volcanic forcing in climate models, we also investigated monthly surface temperature patterns in five PMIP3 last-millennium simulations after interpolating all model fields to $5 \times 5°$ latitude-longitude grids. For comparison to the proxy reconstructions, the simulated summer (June–August) temperatures were extracted between 30 and 90° North and an area-weighted mean was calculated. Two of the simulations (BCC_CSM1.1 and CCSM4) used the G08 volcanic forcing estimate and two used C13 (GISS-E2-R and MPI-ESM-P) (Masson-Delmotte *et al* 2013). IPSL-CM5A-LR was forced with an unpublished record from Ammann *et al* (2007). Note that problems with the implementation of the forcing timeseries were reported for IPSL-CM5A-LR (Atwood *et al* 2016, see SI2). Hereinafter the last-millennium simulations will be referred to by the following abbreviations: BCC, CCSM, GISS, MPI and IPSL.

Detection algorithm
For the detection of volcanic impacts in temperature reconstructions, we applied a method adopted from econometric timeseries modeling for the detection of outliers and breaks that has been successfully demonstrated in the context of volcanic cooling (Pretis *et al* 2016). The detection of breaks (and outliers) is formulated as a problem of model selection, in which a regression model for summer temperature is saturated with a full set of break indicators representing volcanic eruptions at each step in time, of which all but significant indicators are removed (Doornik 2009, implemented in Doornik and Hendry 2013, or also implemented in the R-package 'gets', Pretis *et al* 2017. See also Hendry *et al* 2008, Castle *et al* 2015, and Johansen and Nielsen 2016 for a general overview of indicator saturation). The novelty of this approach applied to volcanic eruptions is in the design of the break function, which adapts the shape of the temperature response with an abrupt cooling followed by a smooth recovery to the mean. A geometric decay by ~50% for two more years after the initial temperature drop (corresponding to Wigley *et al* (2005) and Robock (2000)) is the simplified approximation to the modeled fading of stratospheric aerosol concentration (e.g. in G08 and C13).

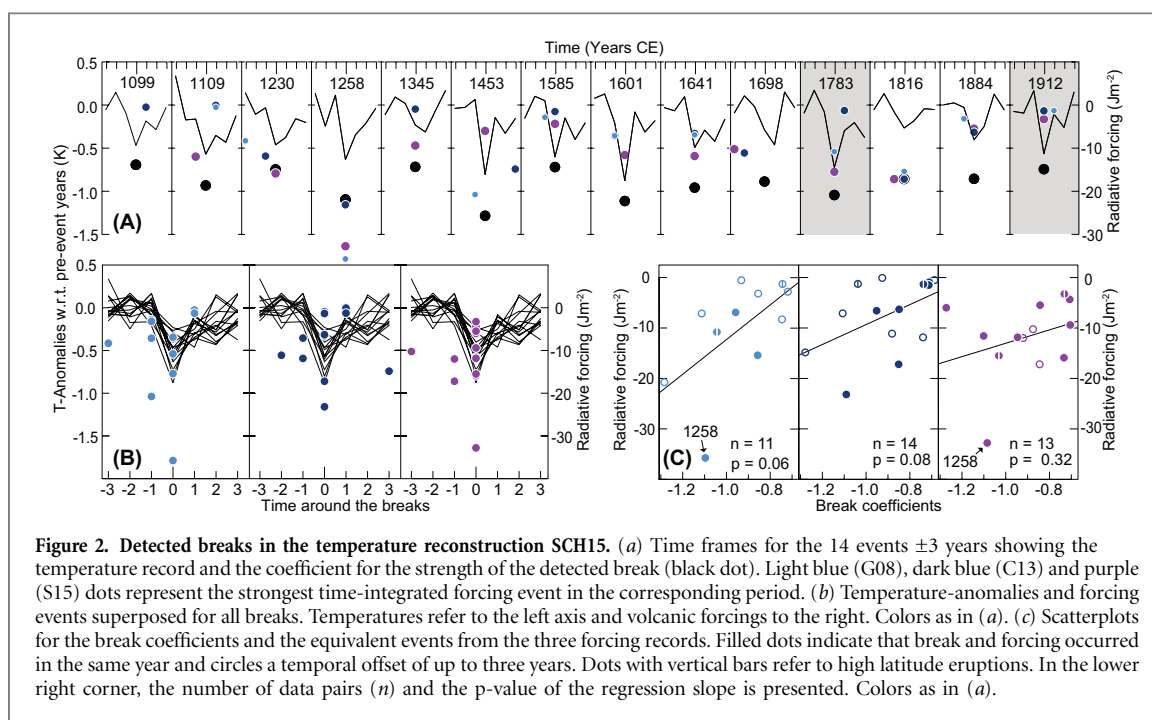The magnitude of the response is estimated jointly with selection over break functions. Its coefficient is the time-integrated temperature response over the three-year period of the volcanic function and it is allocated to the year of peak cooling. Eruptions that lead to temperature responses different from the hypothesized functional form can be approximated by linear combinations of break functions (for example single-period cooling is matched by two opposite-signed break functions). The constrained seasonal window (June–August) of the temperature data impedes inferences about the time lag between eruption and temperature response. But peak forcing seems to precede the main cooling by a maximum of one year (Wigley *et al* 2005). According to the categorization of the G08 and C13 records, forcing events are expected in the break year or in the year before.

Break detection is used to estimate the climatic impact of paleo-volcanism. Multiple break events from the temperature record are averaged in a superposed epoch analyses (SEA, Panofsky and Brier 1958). To contextualize the estimated cooling for the different temperature records, the range of high-frequency temperature variability is calculated as the average standard deviation in moving 11-year windows over the past millennium.

Results and discussion

Detection effectiveness
The algorithm automatically searches for both positive and negative breaks, where the detection of positive events likely captures outlying observations not well-approximated by the statistical model of temperatures (here modelled as autoregressive of order three). Abrupt warming should not exceed natural climate variability because there is no apparent external driver for such an event. Thus, negative spikes of the same magnitude are likely within this range of natural variability and not necessarily caused by a volcanic eruption. In order to avoid the detection of positively-signed (and spurious negative) breaks, the significance level is set conservatively ($\alpha = 0.0065$) so that only large events are picked up by the algorithm. This reduces the likelihood of a falsely identified break to less than 1% (Pretis *et al* 2016) at the expense of missing breaks caused by smaller eruptions. Using SCH15 and the prescribed significance level, we detect 14 negative breaks in the temperature record during the past millennium and zero positive breaks (figure 2(*a*), table S1(*a*) and (*c*)). Common to all detected breaks in SCH15 is a sudden temperature drop, but the subsequent recovery-pattern varies strongly among the events (figure 2(*a*)). Immediate reversion to average temperatures is recorded after 1601, for instance, while temperature further decreases after the initial drop in 1698. Despite these differing response patterns, specifying a multi-year cooling in the detection model yields more rigorous results than an assumed single-year deviation: running the algorithm with a cooling model using a single-year

Figure 2. Detected breaks in the temperature reconstruction SCH15. (*a*) Time frames for the 14 events ±3 years showing the temperature record and the coefficient for the strength of the detected break (black dot). Light blue (G08), dark blue (C13) and purple (S15) dots represent the strongest time-integrated forcing event in the corresponding period. (*b*) Temperature-anomalies and forcing events superposed for all breaks. Temperatures refer to the left axis and volcanic forcings to the right. Colors as in (*a*). (*c*) Scatterplots for the break coefficients and the equivalent events from the three forcing records. Filled dots indicate that break and forcing occurred in the same year and circles a temporal offset of up to three years. Dots with vertical bars refer to high latitude eruptions. In the lower right corner, the number of data pairs (*n*) and the p-value of the regression slope is presented. Colors as in (*a*).

impulse more often returns positive breaks, evidence for a less selective volcanic detection.

The number of detected events represents a conservative minimum of volcanic eruptions with hemispheric impact on summer temperatures during the last millennium. Because the detection model identifies events outside the range of internal climate variability we assume many smaller events to be masked. The known very large eruptions of the past millennium, including Samalas (1257), Tambora (1815) and Huaynaputina (1600), are all detected. Each of the 14 events can also be associated with a radiative forcing peak from C13, if we allow for a dating mismatch of up to ±3 years (figure 2(*a*)). Only 11 events are linked to G08 forcing events. Likewise, 13 of the 14 events are in line with peaks in S15 including nine cooling breaks ranking among the 13 strongest forcing events. The remaining four breaks correspond to global or NH forcing peaks ranking among the top 33 events in S15. These numbers are slightly weaker than suggested by previous pseudo-proxy experiments in which it was possible to detect all of the 15 largest events using the G08-forced CCSM simulation (Pretis *et al* 2016)[8]. The high success rate is, however, fostered by a very pronounced volcanic forcing in CCSM. Additionally, an inaccurate or incomplete forcing reconstruction could compromise the successful

verification of events detected in proxy derived temperatures. Other reasons for a break detection in a year that is not associated with a peak in the forcing record could be a mis-specified model in the detection algorithm or a deficient assumption about forced and internal climate variability.

Regarding the total number of events since 1001 CE documented in S15, we detect ~31% of the global eruptions (*n* = 36) and ~3% of the NH eruptions (*n* = 58) with the latter being usually much weaker and thus unable to cause cooling significantly outside the range of natural climate variability. These numbers are, however, dependent on the threshold chosen to separate background sulfate variability in the ice core record from an actual volcanic source. With 118 events during the past millennium, Sigl *et al* (2015) chose an intermediate threshold compared to the earlier forcing reconstructions G08 (67 events) and C13 (193 events).

Performance on other reconstructions and model simulations

If the STO15 and WIL16 reconstructions are used as the basis of the detection experiment, 20 and 15 volcanic events are detected, respectively. There are eight events detected in all tree-ring based reconstructions: these include 1258, 1453, 1601, 1641, 1698, 1783, 1816 and 1912 (table S1(*a*) and figure S1). Five additional events are common among two records. For SCH15, only the smallest breaks, 1345 and 1099, are neither supported by STO15 nor WIL16. A volcanic background of these temperature anomalies is, however, plausible because both can be aligned with spikes in ice core forcing reconstructions (figure 3(*a*)). Using the reconstructions that additionally include ring width data (STO15 and WIL16) yields three additional strong cooling events that are likely

---

[8] In Pretis *et al* (2016) the reported proportion of correct detection for large eruptions is only 74%. Four events of significant sulfur injection in the G08-dataset (939, 1167, 1227 and 1783) did not reliably translate in breaks in the temperature record. The reason is that the model simulation behind the pseudoproxy-reconstruction was forced with an older version of G08 that was revised and corrected later. Considering the original version of the forcing record reveals that in all of the four cases breaks were detected in the years that were erroneously equipped with large sulfur peaks, moving the correctly detected proportion of large eruptions from 74% to 100%.
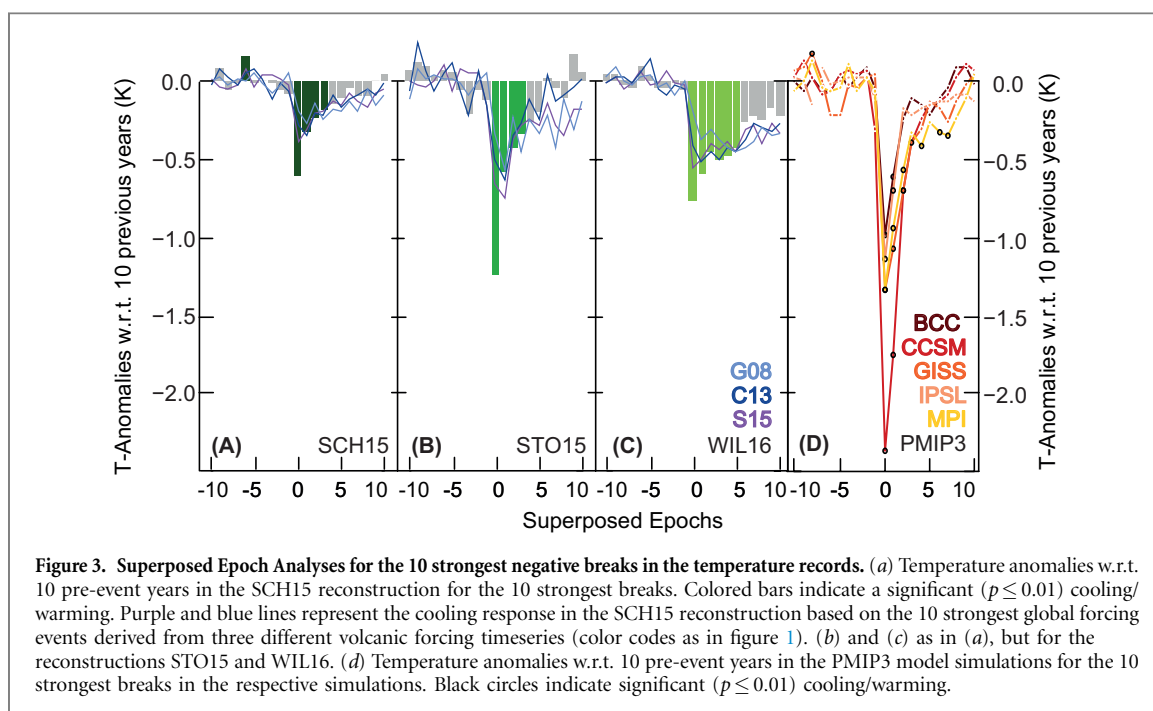
Figure 3. **Superposed Epoch Analyses for the 10 strongest negative breaks in the temperature records.** (*a*) Temperature anomalies w.r.t. 10 pre-event years in the SCH15 reconstruction for the 10 strongest breaks. Colored bars indicate a significant ($p \le 0.01$) cooling/warming. Purple and blue lines represent the cooling response in the SCH15 reconstruction based on the 10 strongest global forcing events derived from three different volcanic forcing timeseries (color codes as in figure 1). (*b*) and (*c*) as in (*a*), but for the reconstructions STO15 and WIL16. (*d*) Temperature anomalies w.r.t. 10 pre-event years in the PMIP3 model simulations for the 10 strongest breaks in the respective simulations. Black circles indicate significant ($p \le 0.01$) cooling/warming.

volcanically driven: 1169/70, 1832 and 1835/6 (top 20 in S15). At the same time, the multiple positive breaks in STO15 and WIL16 indicate that these temperature records reconstruct abrupt warming events that are similar in absolute magnitude to volcanic-forced cooling. The setup for the detection algorithm was designed to impede the detection of positive breaks in SCH15. In STO15 and WIL16, however, four and five additional positive events are detected, respectively. Positive breaks occur largely independent of their negative counterparts, but appear to be more frequent in periods with reduced spatial coverage in the tree-ring networks (table S1(*c*)). This could be related to model misspecification or increased noise during these periods. The magnitudes of positive breaks rank among the lower half of all events indicating that these are expected false-positives or small outliers with *t*-values close to the selection-significance threshold and thus likely retained by chance. With a lower significance level it is likewise possible to prevent the detection of positive breaks in STO15 and WIL16, but the number of remaining negative breaks would then shrink as well and be smaller than for SCH15. A different autoregressive structure in ring width chronologies, potentially caused by biologic memory effects (Esper *et al* 2015), make these records less applicable for the detection of volcanic events.

For comparison to the results from the temperature reconstructions, the detection algorithm is applied to forced-transient model simulations of the last millennium and yields between eight (BCC) and 37 (GISS) negative breaks in these timeseries. The wide range in the frequency of breaks is symptomatic of the internal variability associated with each model simulation and, most importantly, the very different model sensitivities to the ice core derived estimates of volcanic forcing with some of them likely over-

estimating the climatic impact (Marotzke and Forster 2015, Stoffel *et al* 2015). Major breaks in years of strong forcing indicate that CCSM, GISS and MPI translate their volcanic forcing effectively into significant temperature anomalies, which the detection algorithm, in turn, successfully detects. The top ten negative breaks in these simulations comprise five to six events that are among the ten strongest forcing events of the respective ice core forcing record. This result is similar to the agreement between the proxy-reconstructions and S15 (table S1(*a*)), although forcing and climatic response are estimated from independent sources in the real-world comparison. Among the top 20 events in CCSM, GISS and MPI, at most one event is detected in periods of zero volcanic forcing indicating a high selectivity of the algorithm, but some of the weaker breaks occur in the year neighboring peak forcing. In BCC, CCSM, and GISS very strong and persistent forcing events like Samalas in 1258 and Tambora in 1815 result in a second detected break in the previous or subsequent year. BCC breaks only in eight instances and agrees least with the G08 volcanic forcing record used in the simulation. Of particular concern is the strongest break in BCC (1857), which is not associated with any forcing in G08, and the absence of a break in BCC for 1452/3 (2nd strongest forcing in G08). Together with the generally small number of negative breaks in this simulation, there is some evidence for an underestimation of the volcanic effect in this model. For IPSL, these considerations cannot be evaluated because the original forcing data were not published.

The ranking of the simulated break magnitudes is often inconsistent with the size of the respective forcing peaks. Potential reasons are (1) forcing events in the Southern Hemisphere, which do not result in NH cooling (e.g. 1275 in G08) or extreme NH-cooling

with a less outstanding global sulfate flux (e.g. 1903 in BCC). The latter scenario applies if the volcanically driven cooling coincides with a cold phase of internal climate variability and can result in an overestimation of volcanic cooling.

With 13 (GISS) to 21 (CCSM) cases, positive breaks are more common in the model simulations than in the proxy reconstructions. In contrast to STO15 and WIL16, almost all positive breaks in the simulations overlap with negative breaks or occur during the 20th century warming. In 1000 observations and with a 1% likelihood of falsely identified breaks, some of the positive events are probably random outliers that might not be detected with a reduced significance threshold. Most of the positive breaks, however, are associated with considerably large coefficients and would not drop out before many of the negative breaks. The more regular occurrence of positive breaks (next to negative ones and during the 20th century) indicates differences in the timeseries dynamics of the simulated records compared to the reconstructions: positive breaks following negative breaks could be indicative of a more abrupt termination of the cooling signal relative to the slower volcanic recovery based on geometric decay. In CCSM and IPSL, the recovery pattern for more than half of the negative breaks is altered through this effect. For the events in the 20th century, greenhouse gas forcing is likely the driver of the anomalous positive deviations. Their detection is nonetheless surprising because the homogenous increase in greenhouse gas concentrations should be emulated by the autoregressive terms in the model. While they are beyond the scope of this study, these findings deserve more attention in subsequent analyses.

New insights on the timing of volcanic eruptions
Dendrochronological records are characterized by a precise annual age control. Reviewing the dating accuracy of the ice core records with respect to the breaks in the MXD reconstruction reveals improvement in the more recent S15 reconstruction (figure 2 (*b*)). In G08, the oldest volcanic record, peak forcing matches only four of the detected breaks in the SCH15 reconstruction, and three breaks are not reproduced by any event even when considering the ±3 year uncertainty range. In the C13 record, containing most volcanic events, the eruption year is exactly matched in 7 out of 14 years, and all detected breaks are matched by a sulfate spike within the ±3 year uncertainty window. In S15, 10 events are synchronous among the 14 detected peaks, indicating an improved age control in this most recent ice core chronology. S15 nevertheless records no acidity spike in 1099, which is detected in the MXD reconstruction. As for the mismatch of the peaks in 1108 and 1815 (1109 and 1816 in SCH15), these differences likely reflect the ambiguity in the allocation of the forcing peak to a specific year mentioned in the data description. While

S15 is so far not accompanied by a model for atmospheric aerosol dispersion, the application of the 'Easy Volcanic Aerosol' forcing generator (Toohey *et al* 2016) is in progress and will enable a more straightforward assignment of peak forcing years as calculated herein for G08 and C13. Overall, the good agreement between the breaks and the forcing events suggests the time lag between peak forcing and cooling to be rather small. Any more detailed inferences about the delay of the climatic response on a seasonal basis are impeded by the constrained temporal resolution of the ice core and tree-ring records.

The most apparent mismatch between breaks detected in this study and peaks in the forcing records occurs in the 1690s. All MXD reconstructions break in 1698, while none of the forcing records reveals a signal in that year or the previous. There is, however, evidence from the ice cores for one to two considerable forcing events in this decade (G08 in 1693; C13 in 1696; S15 in 1693 and 1695; details in SI3 and in figure S2).

Two eruptions among the ten largest forcing events from S15 do not appear as breaks in one of the three temperature records: 1458 and 1809. The timing and the interpretation of the sulfate signal currently assigned to 1458 has been modified during the last decade (Gao *et al* 2006, Plummer *et al* 2012, Cole-Dai *et al* 2013, Sigl *et al* 2013) with the latest version disagreeing with tree-ring evidences (Esper *et al* 2017). The supplementary information stacks.iop.org/ERL/12/094005/mmedia (SI section 3) further discusses these discrepancies. Like the 1458 eruption, the 1809 forcing peak is of unknown source (Cole-Dai *et al* 2009) and did not result in a break in the temperature reconstructions (figure S2(*c*)). In line with STO15 and WIL16, SCH15 reports cooling in 1809. In contrast to other events, temperatures do not return to the climatological mean after the initial drop, but remain low until the Tambora eruption in 1815. If the forcing records are correct in assigning a major sulfate loading to the 1809 eruption, then this prolonged cooling, potentially amplified by the Dalton Minimum (Raible *et al* 2016), obscured the detection of the volcanic eruption.

The relationship between forcing magnitude and temperature response
Volcanic cooling is sensitive to the altitude, latitude, and character of the volcanic eruption (Hansen *et al* 1997). The relationship between cooling patterns and forcing estimates is thus expected to be variable. The ice core sulfate records agree, for example, on an enormous peak in 1258, but the reconstructed cooling is less extreme (figure 2(*c*)). Measurement and calibration uncertainties associated with single events in ice core and tree-ring derived reconstructions further complicate such comparisons and impede the verification of the forcing magnitude based on the climatic impact. This is assessed by fitting linear

regressions between the detected breaks and sulfate peaks from G08, C13 and S15 (figure 2(*b*)). G08 and C13 cohere relatively well with the break coefficients, but their regression models are based on large intercepts (22.1 and 10.4 J m^{-2}) that were physically expected to be zero. The S15 volcanic forcing is not significantly correlated with the break coefficients, mainly due to the differences in 1258 and 1453. The more minor breaks are associated with very small forcing events in G08 and C13 (smallest forcing = -0.5 J m^{-2} and -0.1 J m^{-2}, respectively) and more substantial events in S15 (smallest forcing = -3.3 J m^{-2}). The latter better explains a temperature response outside the range of internal variability and results in a linear regression with an intercept close to the origin for S15. The relatively weak forcing associated with the maximum cooling in 1453 points to another major inconsistency between forcing and temperature reconstructions in the 1450s. Although now off by 5 years, an integrated forcing of -20 J m^{-2} (in 1458) would better explain the strong 1453 cooling observed in the tree-ring reconstruction and would be well in line with the linear regression model. The pronounced radiative forcing in 1258 has been discussed previously (Timmreck *et al* 2009) and is likely too large due to nonlinear aerosol microphysics in the volcanic plume of that eruption.

Volcanic cooling in reconstructed and simulated temperatures

To isolate the influence of non-volcanic climate deviations, we average the cooling in response to strong volcanism over multiple events during the last millennium before comparing the proxy and model derived cooling estimates (Fischer *et al* 2007). Previous studies with the same intension are challenged by the selection of key volcanic eruptions from either ice core forcing records or historical observations (Wilson *et al* 2016). If the event list is drawn from ice core forcing records, proxy reconstructions will necessarily underestimate the cooling compared to the simulated temperatures (Masson-Delmotte *et al* 2013) because model simulations are driven with exactly these forcing records. For temperature reconstructions, in contrast, the true forcing is unknown and if the event list includes years of misdated or overestimated ice core sulfate deposition the temperature signal will be blurred. Drawing the events from observed eruptions, on the other hand, yields a weaker cooling estimate in model simulations for similar reasons (Esper *et al* 2013b). To avoid such inconsistencies we compile event lists individually for the respective temperature record based on the break detection results. This allows a fair comparison between real-world reconstructions and simulated paleotemperatures. The cooling estimate is thereby maximised due to the intenional focus on the strongest deviations.

The averaged climatic response to detected volcanic breaks in the three reconstructions varies. STO15 shows a much stronger cooling signal, and

WIL16 a longer cold phase compared to SCH15. Application of SEA to SCH15, STO15 and WIL16 reveals maximum post-volcanic cooling that ranges from -0.60 ± 0.13 °C (SCH15)[9] to -1.23 ± 0.21 °C (STO15)[9] for the ten largest breaks (figures 3(*a*–*c*)). Using instead the top ten events from the S15 ice core forcing record in the SEA, the temperature peaks are 28% (WIL16) to 46% (STO15) smaller compared to the detection derived SEA (figures 3(*a*–*c*), table S3). Although we found the forcing peaks in S15 to be in relatively good agreement with the detected breaks, the estimated cooling based on these forcing events is not only weaker but also less accentuated compared to the cooling estimate derived from break-detection events. In STO15, for example, temperatures are estimated to be lowest in the year after peak forcing (figure 3(*b*)).

Our results do not confine the absolute magnitude of the expected temperature drop in response to large volcanic eruptions, because of the remarkable spread across the reconstructions. The discrepancy in absolute magnitude by a factor of two corresponds to the differences in variance among the NH-mean reconstructions (figure 1(*b*)): the standard deviation in the high-frequency domain is 0.16 °C in SCH15 and 0.30 °C in STO15 (table S3). It reflects that the reconstructions vary with respect to their instrumental targets, spatial domains, methods used for detrending and network compilation. In all reconstructions, however, the ten major volcanic events cool surface temperatures by four times this standard deviation. For comparison, the modeled temperature response to volcanic forcing exceeds the standard deviation by five (BCC, IPSL, MPI; six for GISS) to nine (CCSM) times, and is thus much stronger than the reconstructed values. The absolute values range from -0.98 °C (BCC) to -2.37 °C (CCSM) for the model simulations.

Biological memory effects do not reduce the average cooling magnitude in the large-scale compilations including ring width chronologies (STO15 and WIL16). The ratio between high-frequency variability and volcanic cooling is very similar to the purely MXD-based record SCH15 (table S3), although the detection algorithm was more effective in the latter. The temporal persistence of low temperature estimates after the initial drop in WIL16 is, however, characteristic for ring width data (figure 3(*c*)). The much faster recovery of the signal in STO15 suggests the reconstruction contains a smaller influence of ring width chronologies. This is potentially a product of the principal component analysis applied in STO15 that may suppress the generally much noisier ring width chronologies. Similar to SCH15, the cooling in STO15 is reduced by half in the year subsequent to the initial drop (figures 3(*a*–*b*)). This fast relaxation exceeds the

---

[9] The error range refers to the standard deviation of the temperature in the central year of the SEA.

duration observed in simulated temperatures in which the post-event year is on average only 30% less cool than the event year itself. This result is in contrast to previous studies in which reconstructed temperatures were largely based on ring width (e.g. Masson-Delmotte *et al* 2013).

Conclusion

Knowledge about the history of volcanism and climate is often derived from 'on-site' historical observations or polar sulfate depositions. Large-scale temperature reconstructions can serve as an additional archive for climatically relevant volcanism. The independent detection of characteristic temperature patterns complements existing records by adding insight into mid-latitude climatic impacts. In a successful application of a new algorithm we detected all of the most relevant eruptions during the last millennium. The exact timing and, to a lesser degree, the magnitude of these cooling events can add useful information to forcing reconstructions. Particular disagreement with the latest ice core derived record is found in the 1450s and 1690s.

Using an independently derived list of volcanic eruptions to estimate the climatic response reveals an average cooling signal more accentuated and stronger than the signal in response to ice core derived forcing events. Discrepancies in the magnitude between tree-ring reconstructions are ascribed to differences in the year-to-year variability of temperature. Considering the range of internal, high-frequency temperature variability at the hemispheric scale also reveals that most model simulations feature a more pronounced and more persistent volcanic signal than proxy reconstructions.

Acknowledgments

References

Ammann C M, Joos F, Schimel D S, Otto-Bliesner B L and Tomas R A 2007 Solar influence on climate during the past millennium: results from transient simulations with the NCAR climate system model *Proc. Natl Acad. Sci. USA* **104** 3713–18

Anchukaitis K J *et al* 2012 Tree rings and volcanic cooling *Nat. Geosci.* **5** 836–7

Atwood A R, Wu E, Frierson D M W, Battisti D S and Sachs J P 2016 Quantifying climate forcings and feedbacks over the last millennium in the CMIP5-PMIP3 models *J. Clim.* **29** 1161–78

Baillie M G L and McAneney J 2015 Tree ring effects and ice core acidities clarify the volcanic record of the first millennium *Clim. Past* **11** 105–14

Briffa K R, Jones P D, Schweingruber F H and Osborn T J 1998a Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years *Nature* **393** 450–5

Briffa K R, Schweingruber F H, Jones P D, Osborn T J, Shiyatov S G and Vaganov E A 1998b Reduced sensitivity of recent tree-growth to temperature at high northern latitudes *Nature* **391** 678–82

Castle J L, Doornik J A, Hendry D F and Pretis F 2015 Detecting location shifts by step-indicator saturation during model selection *Econometrics* **3** 240–64

Cole-Dai J, Ferris D, Lanciki A, Savarino J, Baroni M and Thiemens M H 2009 Cold decade (AD 1810–1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption *Geophys. Res. Lett.* **36** L22703

Cole-Dai J, Ferris D G, Lanciki A L, Savarino J, Thiemens M H and McConnell J R 2013 Two likely stratospheric volcanic eruptions in the 1450s CE found in a bipolar, subannually dated 800 year ice core record *J. Geophys. Res.: Atmos.* **118** 7459–66

Cook E and Kairiukstis L 1990 *Methods of Dendrochronology: Applications in the Environmental Sciences* (Dordrecht: Kluwer Academic Publishers)

Cook E R and Peters K 1997 Calculating unbiased tree-ring indices for the study of climatic and environmental change *Holocene* **7** 361–70

Crowley T J and Unterman M B 2013 Technical details concerning development of a 1200 yr proxy index for global volcanism *Earth Sys. Sci. Data* **5** 187–97

D'Arrigo R, Wilson R and Anchukaitis K J 2013 Volcanic cooling signal in tree ring temperature records for the past millennium *J. Geophys. Res.: Atmos.* **118** 9000–10

Doornik J A 2009 Autometrics ed Castle J and Shephard N *The Methodology and Practice of Econometrics: a Festschrift in Honour of David F. Hendry* (Oxford: Oxford University Press) pp 88–121

Doornik J A and Hendry D F 2013 *PcGive 14* (London: Timberlake Consultants)

Esper J, Büntgen U, Luterbacher J and Krusic P J 2013a Testing the hypothesis of post-volcanic missing rings in temperature sensitive dendrochronological data *Dendrochronologia* **31** 216–22

Esper J *et al* 2012 Orbital forcing of tree-ring data *Nat. Clim. Change* **2** 862–6

Esper J, Schneider L, Krusic P J, Luterbacher J, Büntgen U, Timonen M, Sirocko F and Zorita E 2013b European summer temperature response to annually dated volcanic eruptions over the past nine centuries *B. Volcanol.* **75**

Esper J, Schneider L, Smerdon J E, Schone B R and Büntgen U 2015 Signals and memory in tree-ring width and density data *Dendrochronologia* **35** 62–70

Esper J, Büntgen U, Hartl-Meier C, Oppenheimer C and Schneider L 2017 Northern Hemisphere temperature anomalies during the 1450s period of ambiguous volcanic forcing *Bull. Volcanol.* **79** 41

Fischer E M, Luterbacher J, Zorita E, Tett S F B, Casty C and Wanner H 2007 European climate response to tropical volcanic eruptions over the last half millennium *Geophys. Res. Lett.* **34** L05707

Gao C *et al* 2006 The 1452 or 1453 A D Kuwae eruption signal derived from multiple ice core records: greatest volcanic sulfate event of the past 700 years *J. Geophys. Res.: Atmos.* **111** D12107

Gao C, Robock A and Ammann C 2008 Volcanic forcing
of climate over the past 1500 years: an improved ice
core-based index for climate models *J. Geophys.
Res.* **113**

Global Volcanism Program 2013 Volcanoes of the World, v. 4.5.2
ed Venzke E (Smithsonian Institution) downloaded 27 Jul
2016 (http://volcano.si.edu/list_volcano_holocene.cfm)

Hansen J, Sato M and Ruedy R 1997 Radiative forcing and
climate response *J. Geophys. Res. Atmos.* **102** 6831–64

Hegerl G C, Crowley T J, Hyde W T and Frame D J 2006
Climate sensitivity constrained by temperature
reconstructions over the past seven centuries *Nature* **440**
1029–32

Hendry D F, Johansen S and Santos C 2008 Automatic selection
of indicators in a fully saturated regression *Comput. Stat.*
23 317–35

Johansen S and Nielsen B 2016 Asymptotic theory of outlier
detection algorithms for linear time series regression
models *Scand. J. Stat.* **43** 321–48

Jones P D, Moberg A, Osborn T J and Briffa K R 2003 Surface
Climate Responses to Explosive Volcanic Eruptions Seen in
Long European Temperature Records and Mid-to-High
Latitude Tree-Ring Density Around the Northern
Hemisphere ed Robock A and Oppenheimer C *Volcanism
and the Earth's Atmosphere* (Washington, DC: American
Geophysical Union) pp 239–54

Jungclaus J H *et al* 2016 The PMIP4 contribution to CMIP6—
Part 3: the last millennium, scientific objective and
experimental design for the PMIP4 past 1000 simulations
Geosci. Model Dev. Discuss. **2016** 1–34

Kageyama M, Braconnot P, Harrison S and Zhou T 2016
PMIP4-CMIP6: the contribution of the Paleoclimate
Modelling Intercomparison Project to CMIP6 Geoscientific
Model Development Discussions

Mann M E, Fuentes J D and Rutherford S 2012 Underestimation
of volcanic cooling in tree-ring-based reconstructions of
hemispheric temperatures *Nat. Geosci.* **5** 202–5

Marotzke J and Forster P M 2015 Forcing, feedback and internal
variability in global temperature trends *Nature* **517**
565–U291

Masson-Delmotte V *et al* 2013 Information from Paleoclimate
Archives ed Stocker T F, Qin D, Plattner G-K, Tignor M,
Allen S K, Boschung J, Nauels A, Xia Y, Bex V and
Midgley P M *Climate Change 2013: The Physical Science
Basis. Contribution of Working Group I to the Fifth
Assessment Report of the Intergovernmental Panel on
Climate Change* (Cambridge: Cambridge University Press)
pp 383–464

Naveau P, Ammann C M, Oh H-S and Guo W 2003 An
Automatic Statistical Methodology to Extract Pulse-Like
Forcing Factors in Climatic Time Series: Application to
Volcanic Events ed Robock A and Oppenheimer C
Volcanism and the Earth's Atmosphere (Washington, DC:
American Geophysical Union) pp 177–86

Panofsky H A and Brier G W 1958 *Some Applications of
Statistics to Meteorology* (University Park, PA: Mineral
Industries Extension Services, College of Mineral
Industries, Pennsylvania State University)

Plummer C T, Curran M A J, van Ommen T D, Rasmussen S O,
Moy A D, Vance T R, Clausen H B, Vinther B M and
Mayewski P A 2012 An independently dated 2000 years
volcanic record from Law Dome, East Antarctica, including
a new perspective on the dating of the 1450s CE eruption
of Kuwae, Vanuatu *Clim. Past* **8** 1929–40

Pretis F, Schneider L, Smerdon J E and Hendry D F 2016
Detecting volcanic eruptions in temperature reconstructions
by designed break-indicator saturation *J. Eco. Sur.* **30** 403–29

Pretis F, Reade J and Sucarrat G 2017 Automated general-to-
specific (GETS) modelling of the mean and variance of
regressions, and indicator saturation methods for outliers
and structural breaks *J. Stat. Soft.* in press

Raible C C *et al* 2016 Tambora 1815 as a test case for high
impact volcanic eruptions: Earth system effects *Wiley
Interdisciplinary Rev.: Clim. Change* **7** 569–89

Robock A 2000 Volcanic eruptions and climate *Rev. Geophys.* **38**
191–219

Schmidt G A *et al* 2011 Climate forcing reconstructions for use
in PMIP simulations of the last millennium (v1.0)
Geoscientific Model Development **4** 33–45

Schneider L, Smerdon J E, Büntgen U, Wilson R J S, Myglan V
S, Kirdyanov A V and Esper J 2015 Revising midlatitude
summer temperatures back to AD600 based on a wood
density network *Geophys. Res. Lett.* **42** 4556–62

Sigl M *et al* 2013 A new bipolar ice core record of volcanism from
WAIS divide and NEEM and implications for climate forcing
of the last 2000 years *J. Geophys. Res.: Atmos.* **118** 1151–69

Sigl M *et al* 2014 Insights from Antarctica on volcanic forcing
during the Common Era *Nat. Clim. Change* **4** 693–7

Sigl M *et al* 2015 Timing and climate forcing of volcanic
eruptions for the past 2500 years *Nature* **523** 543–+

Stoffel M *et al* 2015 Estimates of volcanic-induced cooling in the
Northern Hemisphere over the past 1500 years *Nat. Geosci.*
8 784

Stothers R B 1984 The great tambora eruption in 1815 and its
aftermath *Science* **224** 1191–8

Timmreck C, Lorenz S J, Crowley T J, Kinne S, Raddatz T J,
Thomas M A and Jungclaus J H 2009 Limited temperature
response to the very large AD 1258 volcanic eruption
Geophys. Res Lett. **36** L21708

Toohey M, Stevens B, Schmidt H and Timmreck C 2016 Easy
Volcanic Aerosol (EVA v1.0): An idealized forcing generator
for climate simulations *Geosci. Model Dev. Discuss.* **9** 4049–70

Wigley T M L, Ammann C M, Santer B D and Raper S C B
2005 Effect of climate sensitivity on the response to
volcanic forcing *J. Geophys. Res.: Atmos.* **110** D09107

Wilson R *et al* 2016 Last millennium northern hemisphere
summer temperatures from tree rings: Part I: The long
term context *Quatern. Sci. Rev.* **134** 1–18