Tellus SERIES B CHEMICAL AND PHYSICAL METEOROLOGY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

High-latitude volcanic eruptions in the Norwegian Earth System Model: the effect of different initial conditions and of the ensemble size

By FRANCESCO S. R. PAUSATA[1]*, ALF GRINI[2], RODRIGO CABALLERO[1], ABDEL HANNACHI[1] and ØYVIND SELAND[2], [1]*Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, SE-106 91 Stockholm, Sweden;* [2]*Norwegian Meteorological Institute, NO-0313 Oslo, Norway*

(Manuscript received 25 November 2014; in final form 21 May 2015)

ABSTRACT

Large volcanic eruptions have strong impacts on both atmospheric and ocean dynamics that can last for decades. Numerical models have attempted to reproduce the effects of major volcanic eruptions on climate; however, there are remarkable inter-model disagreements related to both short-term dynamical response to volcanic forcing and long-term oceanic evolution. The lack of robust simulated behaviour is related to various aspects from model formulation to simulated background internal variability to the eruption details. Here, we use the Norwegian Earth System Model version 1 to calculate interactively the volcanic aerosol loading resulting from SO_2 emissions of the second largest high-latitude volcanic eruption in historical time (the Laki eruption of 1783). We use two different approaches commonly used interchangeably in the literature to generate ensembles. The ensembles start from different background initial states, and we show that the two approaches are not identical on short-time scales (<1 yr) in discerning the volcanic effects on climate, depending on the background initial state in which the simulated eruption occurred. Our results also show that volcanic eruptions alter surface climate variability (in general increasing it) when aerosols are allowed to realistically interact with circulation: Simulations with fixed volcanic aerosol show no significant change in surface climate variability. Our simulations also highlight that the change in climate variability is not a linear function of the amount of the volcanic aerosol injected. We then provide a tentative estimation of the ensemble size needed to discern a given volcanic signal on surface temperature from the natural internal variability on regional scale: At least 20–25 members are necessary to significantly detect seasonally averaged anomalies of 0.5°C; however, when focusing on North America and in winter, a higher number of ensemble members (35–40) is necessary.

Keywords: volcano-climate interactions, climate variability, atmospheric processes

1. Introduction

The potential climate effects of large volcanic eruptions have received wide attention in the climate science community since they can have major – albeit short-lived – impacts on the energy balance (e.g. Robock, 2000; Cole-Dai, 2010). In fact, large volcanic eruptions have been a major natural driver of climate variability during the last millennium (Bindoff et al., 2013). However, important aspects of both short- and long-term dynamical climate responses to large volcanic eruptions remain unclear (e.g. Robock, 2002; Timmreck, 2012). The frequency of such volcanic events is low, and so direct observations of their impacts are limited, while coupled climate models show little robustness in their response to volcanic forcing. In addition, climate models have shown poor performance in capturing the post-tropical-volcanic strengthening (underestimated) of the Northern Hemisphere (NH) polar vortex, as well as the cooling (overestimated) in the tropical troposphere compared to observations (Driscoll et al., 2012). However, the differences between models and observations may be related to the effects of internal noise: Removing the El Niño/Southern Oscillation (ENSO) signal, for example, considerably improves the agreement between the observed and simulated surface temperature (Santer et al., 2014).

*Corresponding author.
email: francesco.pausata@misu.su.se
Responsible Editor: Annika Ekman, Stockholm University, Sweden.

Citation: Tellus B 2015, **67**, 26728, http://dx.doi.org/10.3402/tellusb.v67.26728

The lack of robust simulated behaviour among climate models may depend on various aspects of the model formulation (e.g. Shindell, 2004; Cagnazzo and Manzini, 2009), on the simulated background internal climate variability (e.g. Thomas et al., 2009; Zanchettin et al., 2012, 2013; Berdahl and Robock, 2013), and also on how the models implement volcanic eruptions [e.g. basic top of the atmosphere (TOA) shortwave forcing, to the more complex direct SO_2 injection]. In addition, the details of the eruption, including magnitude (Timmreck, 2012), latitude (Schneider et al., 2009) and season (Kravitz and Robock, 2011; Toohey et al., 2011), play an important role in the consequent climate impacts. Finally, there is no commonly accepted protocol for simulating the climate impacts of volcanic eruptions, including prescriptions for ensemble size, for how the ensemble is produced (from a single year or from different years), and the type of reference period chosen to compare the volcano simulations. The ensemble size should be large enough to account for the range of variability depicted by the dominant processes influencing interannual and decadal climate variability. However, no general agreement has yet been reached on this point. Some studies have used only three ensemble members (e.g. Oman et al., 2006a; Schneider et al., 2009), others 5 or 6 (Meronen et al., 2012; Man et al., 2014), others 10 (e.g. Highwood and Stevenson, 2003; Oman et al., 2006b), and a few studies 20 members (e.g. Oman et al., 2005; Kravitz and Robock, 2011) or more (e.g. Wegmann et al., 2014). In any case, the necessary ensemble size strongly depends on the type of impacts investigated and the magnitude of the forcing. When looking at global changes, such as impacts on radiative forcing (Oman et al., 2006b) and global surface temperature (Shepherd, 2014), or very large forcing/climate response (Schneider et al., 2009) a small number of ensemble members may be sufficient, whereas a large number of ensemble members is needed to establish statistical significance on regional scales, especially at mid- and high-latitudes due to high internal natural variability (Oman et al., 2005; Kravitz and Robock, 2011). Another important aspect is how to measure the volcanic signal. When comparing model results of volcanic impact on climate to reconstructions or observations, it is necessary to express the influence of the eruption as the difference between the period following the eruption and a reference period around the eruption. When instead the purpose is not to compare against reconstructions, but rather to investigate the general influence of the volcanic eruptions on climate, two alternatives are possible and the influence of the eruption can be expressed as either:

(1) The difference between a set of simulations of the volcanic eruption (volcanic ensemble) generated from a single year, i.e. slightly different atmospheric states but same ocean conditions, and another ensemble starting from the same initial conditions but with no-volcanic eruption (e.g. Schneider et al., 2009); or
(2) The difference between a volcanic ensemble and equivalent no-volcanic ensemble where each member is generated from a different year, i.e. different atmospheric and oceanic states (e.g. Highwood and Stevenson, 2003).

However, to our knowledge no study has compared these two methods. The aims here are four-fold: (1) test the performance of the Norwegian Earth System Model version 1 (NorESM1-M) in simulating the volcanic aerosol production from SO_2 released by the Lakagígar or Laki eruption in 1783–1784, the second largest high-latitude volcanic eruption in historical time (after the Eldgjá eruption in 934) and the consequent changes in the radiative forcing and climate; (2) investigate how the choice of generating the ensemble members from a single year or from different years influences the simulated climate impacts; (3) investigate how adopting interactive aerosols may impact post-volcanic surface climate variability and (4) suggest a lower bound estimate of ensemble size necessary to significantly separate the volcanic signal from internal natural variability for an eruption of this type.

The paper is structured as follows. In Section 2 we provide a description of the coupled model used and the experimental set-up adopted to simulate the Laki-type eruption. In Section 3 we evaluate the model results against other studies, and in Section 4 we discuss the differences between the two ensemble approaches. In Section 5 we investigate the changes in surface climate variability associated to the volcanic eruption; while in Section 6 we quantify the number of ensemble members necessary to isolate the volcanic signal from the background variability in terms of surface temperature on regional scale; discussion and conclusions are provided in Section 7.

2. Methods

2.1. Model description

The simulations were conducted using the coupled atmospheric–ocean–aerosol model NorESM1-M (Bentsen et al., 2013; Iversen et al., 2013). NorESM1-M is an Earth System Model that uses a modified version of the Community Atmospheric Model version 4 (CAM4, Neale et al., 2013) – CAM4–Oslo – for the atmospheric part of the model, with an updated aerosol model, with online calculation of aerosols and their direct effect, and the first and second indirect effects on warm clouds. The aerosol model includes the life cycle of sea salt, mineral dust, particulate sulphate, black and organic carbon (Kirkevåg et al., 2013). CAM4–Oslo has a separate representation of aerosols,

aerosol–radiation and aerosol–cloud interactions. The model uses the finite volume dynamical core for transport calculations, with horizontal resolution 1.9° (latitude) × 2.5° (longitude) and 26 vertical levels, as in the original CAM4. CAM4–Oslo is coupled with the updated version of the Miami Isopycnic Coordinate Ocean Model – MICOM (Assmann et al., 2010). The sea ice and land models are basically the same as in the Community Climate System Model version 4 – CCSM4 (Gent et al., 2011).

The model can adequately represent the modern climate variability (Iversen et al., 2013): The northern and southern annular modes, the Madden-Julian Oscillation and the El Niño Southern Oscillation are well-captured in the model. The Atlantic Meridional Overturning Circulation strength is in the upper range found in models contributing to CMIP3 and above the range estimated from synthesised observational records (Medhaug and Furevik, 2011). A detailed analysis of NorESM performances is provided in Iversen et al. (2013).

The model calculates the radiative properties and the size of hygroscopic sulphate aerosols, which depend explicitly on the local relative humidity. The aerosol scheme is based on a sectional model creating look-up tables for aerosol properties (see Table 1 in Kirkevåg et al., 2013). Look-up tables allow using background aerosol modes and tabulated properties for condensed and coagulated aerosol mass. The sulphuric acid produced condenses on pre-existing particles or nucleates to form new particles. All particles at all levels are subject to gravitational settling. This is an important removal process for particles at high altitudes. The model's wet deposition is fully coupled to the model's hydrologic cycle, including the convection scheme. The washout takes into account the aerosol mixing state, so hydrophilic aerosols (e.g. sulphate) are removed more easily than hydrophobic aerosols (e.g. dust). An in-depth description and validation of the performance of the model version used in this study for current climate conditions can be found in Bentsen et al. (2013), Iversen et al. (2013) and Kirkevåg et al. (2013).

2.2. *Experimental set-up*

We simulate a hypothetical multistage Laki-type eruption, basing our SO_2 emission and frequency of the eruptions on the study of Thordarson and Self (2003). The 1783–1784 Laki flood lava eruption started on June 8th, 1783 and lasted for 8 months, injecting about 122 Tg of SO_2 into the atmosphere. About 95 Tg of SO_2 were injected into the upper troposphere/lower stratosphere between 9 and 13 km, and the total mass of the Laki tephra was roughly 110 Tg. About 95% of the total SO_2 emission took place in the first 4 months of activity (see Fig. 2 in Thordarson and Self, 2003), with 9 out of 10 eruption episodes featuring a short-lived (0.5–4 d) explosive phase. The last 3.5 months were characterised by quiet emission of lava and gas. While this is the best available estimate of the SO_2 amount emitted by Laki, there are uncertainties estimated to be at least ±20% (Oman et al., 2006b).

For simplicity, we simulate the multistage Laki eruption by injecting 100 Tg of both SO_2 and dust – as an analogue for the ash injection – over eight eruptions with a 2-week frequency, starting on June 1st. Each sub-eruption lasts 4 d, for an overall duration of 4 months (Table 1). The SO_2 and dust particles are evenly distributed within each vertical domain shown in Table 1. The SO_2 injection heights follow the percentage distribution reported in the work of Thordarson and Self (2003).

To examine the role of the background initial state in affecting the impact of volcanic eruptions on climate, we apply the two approaches commonly used in the literature (denoted with the letters A and B) to generate a total of four ensembles (see Fig. 1):

(1) Ensemble ENS_{volc}–A: We simulate the volcanic eruption starting from a specific model year (1934) and generate the ensemble members by perturbing the initial condition of that specific year. We generate 20 ensemble members and run them for 5 yr each. The year is selected from a transient 105-yr climatology (1860–1964).
(2) Ensemble $ENS_{no\text{-}volc}$–A: We generate an equivalent ensemble with volcanic aerosol set to background conditions[1] by perturbing the initial condition of the year selected for the ENS_{volc}–A, i.e. 20 reference ensemble members and run them for 5 yr.
(3) Ensemble ENS_{volc}–B: We simulate the volcanic eruption and the ensemble members starting each member from a different year at 5-yr intervals (1860, 1865, …, 1960; we have skipped 1935 because too close to the selected year in ENS–A). We generate 20 ensemble members each 5-yr long.
(4) Ensemble $ENS_{no\text{-}volc}$–B: We use for each member the five unperturbed years from the climatology corresponding to the 5 yr of the simulated volcanic eruption (i.e. eruption members starting in 1860, 1865, …, 1960; selected unperturbed years from the climatology 1860 to 1964). This ensemble is not – in the strictest sense – an ensemble and is usually referred in the literature as a climatological or reference simulation.

The difference in the two approaches is shown schematically in Fig. 1. All ensemble members are initialised on the

[1]Historical emissions for the aerosol concentrations are taken from IPCC AR5 datasets, see Kirkevåg et al. (2013).

Table 1. Total amount of SO_2 (Tg) emitted for each of the 4-d long eruption, and its vertical distribution

Eruption date	Jun 1st	Jun 15th	Jul 1st	Jul 15th	Aug 1st	Aug 15th	Sep 1st	Sep 15th
Total SO_2 (Tg)	42	11	11	15	9	5	4	3
$100 < p < 150$ hPa	5	1	1	2	1	0	0	0
$150 < p < 300$ hPa	29	8	8	10	6	4	3	2
$p > 300$ hPa	8	2	2	3	2	1	1	1

Maximum emission pressure height $p = 100$ hPa (~ 15 km).

day of the eruption (June 1st). We use the initial conditions of the days preceding or following the date of the eruption to perturb the model state at the date of the volcanic eruption (June 1st) and generate the ensemble members.

Studies in the existing literature either compare ENS_{volc}–A with $ENS_{no\text{-}volc}$–A, or ENS_{volc}–B with the climatology/reference simulation (i.e. $ENS_{no\text{-}volc}$–B). Here we do both, allowing us to investigate whether or not the different approaches used in the literature are equivalent. In other words, if the climate perturbation (Δ_{volc}) induced by the volcanic eruption is measured as:

$$\Delta_{volc} = STATE_{volc} - STATE_{no\text{-}volc} \quad (1)$$

where $STATE_{no\text{-}volc}$ is the unperturbed climate state and $STATE_{volc}$ is the climate state induced by the eruption, we would like to investigate whether the change Δ_{volc} is the same using either approach.

The year 1934 of the reference simulation was selected as the eruption year for ENS_{volc}–A (eruption year: number 01), which is roughly in the middle of the climatology period and features a strong El Niño event, as was the case in 1783/84 during the Laki eruption (Cook and Krusic, 2004; D'Arrigo et al., 2011). Each ensemble member in ENS_{volc}–B starts from different initial conditions. Prior to analysis, the temperature trend in the $ENS_{no\text{-}volc}$–B climatology (1860–1964) is linearly removed.[2] The same trend is removed from the ENS_{volc}–B.

The use of a transient historical simulation rather than a reference equilibrium simulation is unlikely to significantly affect our results because the trend in temperature is small compared to the amplitude of interannual variability. Furthermore, our analyses are performed comparing volcanic and non-volcanic ensembles of the same type (i.e. ENS_{volc}–A vs. $ENS_{no\text{-}volc}$–A or ENS_{volc}–B vs. $ENS_{no\text{-}volc}$–B), further limiting the potential effects of such trend. We only compare the ENS_{volc}–A to $ENS_{no\text{-}volc}$–B (i.e. climatology) when we evaluate the model against historical and simulated data. In this case, the uncertainties in temperature reconstruction for the Laki event and the climate anomalies (temperature and precipitation) induced by the volcanic eruptions (see discussion below) in the months following the eruptions are quite large and any anthropogenic forcing is negligible. Therefore, the exact estimate of the temperature trend does not impact our model evaluation.

The two ensemble sets also allow testing whether or not the choice of one or the other set-up (ENS–A or ENS–B) may require a different number of ensemble members to discern the volcanic effect from the interannual natural climate variability.

We analyse monthly means of model outputs, and all differences discussed in this study are assessed using the Student's t test. Significance is reported at the 95% confidence level unless otherwise noted.

3. Model validation

Before going into the details of the two different approaches adopted here to model the Laki eruption and the number of necessary ensemble members required to distinguish the volcano signal from natural variability, we first validate the large-scale changes simulated by NorESM with the climate impacts of Laki known from the literature. This is the first study where NorESM has been used to compute volcanic aerosols interactively from SO_2 emissions. In previous studies (e.g. CMIP5 experiments), the simulations were performed using prescribed aerosol data (Ammann, 2003). We compare our results with the information available about the Laki eruption that mainly comes from the work of Thordarson and Self (2003) and a few other model studies (Highwood and Stevenson, 2003; Oman et al., 2006b; Schmidt et al., 2012).

We first analyse how NorESM performs in simulating the SO_4 burden, the aerosol optical depth (AOD) at 550 nm and radiative forcing, and then the NH surface temperature (TS) and precipitation.

As mentioned in the introduction, when evaluating the model results against reconstructions, we need to compare the ENS_{volc}–A (El-Niño background state as in 1783) to the climatology ($ENS_{no\text{-}volc}$–B). In any case, when analysing global or hemispheric scale changes associated to the volcanic eruption, the changes are so large that comparing

[2]The linear estimate for the period 1860–1964 is $\sim +0.2$°C on both global and NH scale; no trend is detected in precipitation.

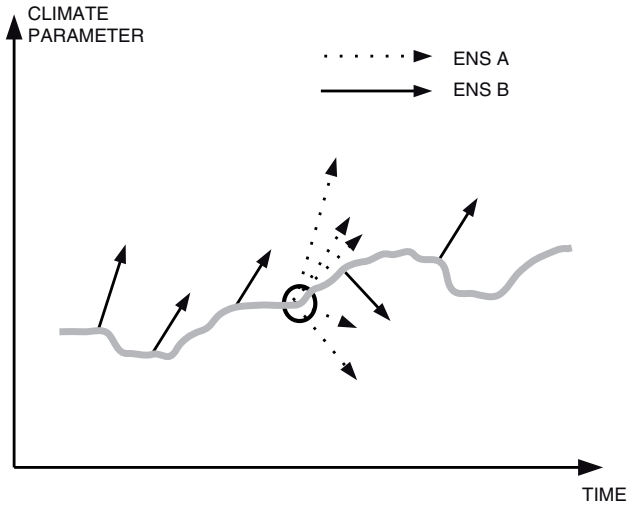

Fig. 1. Schematic representation of the two different ensembles used in this study. The dotted arrows show ENS–A that is generated starting each member from the same background state; the solid arrows show ENS–B that is generated starting each member from a different background state.

the volcano ensemble (either ENS_{volc}–A or ENS_{volc}–B) with the climatology ($ENS_{no\text{-}volc}$–B) or ENS_{volc}–A with the $ENS_{no\text{-}volc}$–A gives the same qualitative results (not shown). When instead investigating regional changes the differences can be remarkable (see Section 4).

3.1. *Optical depth and radiative forcing*

The simulated monthly ensemble average peak of SO_2 gas loading in our experiment was just over 22 Tg in June. The NH zonal monthly ensemble mean of sulphate (SO_4) concentrations peaks above the polar circle at over 42 ppbv in August with a total sulphate loading of 51 Tg (Fig. 2). Our study compares well with the results from Oman et al. (2006b) that show a sulphate loading peaking in late August 1783 at 60 Tg. The e-folding time of the sulphate aerosol is approximately 3.5 months in good agreement with the 4 months shown by Oman et al. (2006b). Using a global aerosol microphysics model, Schmidt et al. (2010) have shown a similar SO_2 gas loading maximum of 29 Tg, resulting in a SO_4 burden peak of only 14 Tg. The difference in SO_4 burden must be related to a quicker particle removal in their model.

The SO_4 AOD perturbation averaged over the entire NH peaks in August with an ensemble mean anomaly of 1.53 (Fig. 3). The AOD anomalies have almost vanished 7–8

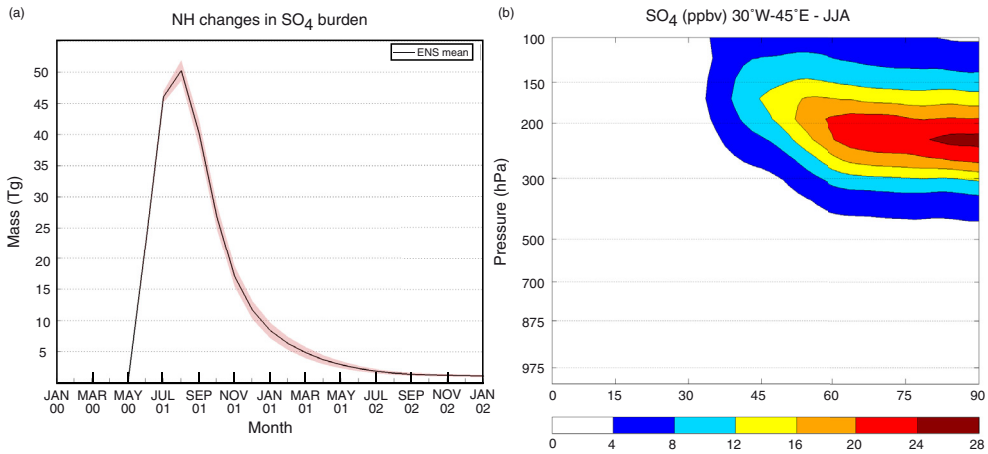

Fig. 2. Ensemble mean change (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–B) in SO_4 (Tg) mass loading (left) and 30°W to 45°E mean SO_4 concentrations (ppbv) for the summer of the eruption (JJA year 01) over the Northern Hemisphere (NH) from the surface to 100 mbar (right). The 30°W to 45°E interval has been chosen for direct comparison to Fig. 8 in Oman et al. (2006b). The shading in the left panel represents the standard deviation of the ensemble difference.

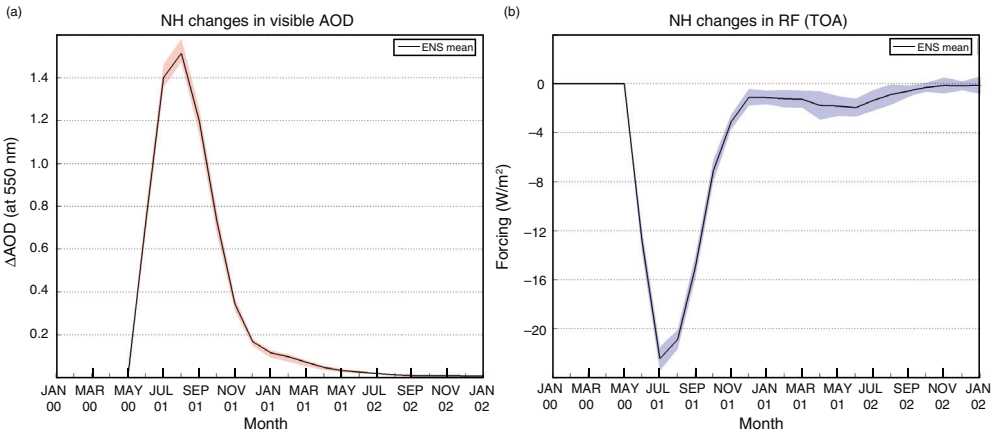

Fig. 3. Ensemble mean change (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–B) in NH visible (550 nm) aerosol optical depth – AOD (left) and net radiative forcing (shortwave + longwave) at the top of the atmosphere – TOA (right) due to a Laki-type volcanic eruption. The shadings represent the standard deviation of the ensemble difference.

months after the end of the eruption. A similar behaviour is shown by the radiative forcing that peaks in July with a NH TOA mean of $-22\,W/m^2$ (Fig. 3). The global forcing is exactly half this value since the entire radiative anomaly is confined to the NH (not shown). Our anomalies are almost three times larger compared to other studies. For example, Oman et al. (2006a) simulate a total NH forcing of $-8\,W/m^2$; Highwood and Stevenson (2003) simulate different injection height scenarios (low: SO_2 is injected up to 9 km; high: SO_2 is injected up to 13 km) and the magnitude of the radiative forcing varies depending on the scenario with maximum global mean forcing of $-3\,W/m^2$ for the high altitude scenario ($-6\,W/m^2$ for the NH). However, they argued that their value could be an underestimate and that the real direct effect could be 1.6 times larger ($\sim -5\,W/m^2$ global/$\sim -10\,W/m^2$ NH) mainly due to the assumptions about relative humidity effects on the aerosol scattering properties.

The differences in the AOD are most likely due to the differences in effective radii: Smaller particles scatter light more efficiently. The specific surface area, i.e. surface area seen by radiation is inversely proportional to the effective radius ($\sim 1/2r_{eff}$). In Oman et al. (2006b) simulations, the average UT/LS effective radius is between 0.54 and 0.61 μm. In our simulations, the effective radius is approximately 0.2 μm, and this would lead to greater optical depth by a factor of 2–3.

The reason for such small particles is that CAM4–Oslo does not simulate growth by self-coagulation (coagulation of Aitken-mode particles combining to form larger particles). CAM4–Oslo simulates nucleation and growth to Aitken-mode size, and also growth of background aerosols (which were there prior to the eruption). Self-coagulation is an important mechanism after an eruption when a massive amount of sulphate is injected as described in Pinto et al. (1989) and English et al. (2013). In our simulations, new aerosol particles are created via nucleation, and they grow to the Aitken-mode size, then slowly into the accumulation mode since the only growth mechanism is condensation, and intra-modal coagulation is not taken into account. However, given the same amount of SO_2 emitted by a volcanic eruption, there can be large differences in the effective radius (Pinto et al., 1989). Since very little is known about the effective radius of particles for the Laki eruption, based on Pinto et al. (1989), Oman et al. (2006b) assumed a 20% larger effective radius for the Laki eruption compared to the Mount Pinatubo. The Pinatubo eruption, which emitted around 15–20 Tg of SO_2 in 1 d, had an effective radius of around 0.2 to 0.4 μm in the first 4 months after the eruption as suggested by a compilation of observations done mainly over California and Wyoming (Russell et al., 1996). Pinto et al. (1989) have shown using a microphysical model that in their set-up 10 Tg SO_2 leads to an effective radius peak of 0.3–0.4 μm after 13 months and 100 Tg SO_2 emission provides an effective radius peak of 0.6 μm after 8 months from the eruption. However, no microphysical modelling studies investigating the Laki particle size are yet available. Furthermore, the Laki eruption was a multistage eruption where the 122 Tg of SO_2 were injected over 8 months rather than a few days, as was the case of the Pinatubo eruption. Therefore, an effective radius of 0.6 μm for the Laki eruption, which is greater than the Pinatubo one, may well be an overestimate.

To verify that the reasons for such differences in AOD and radiative forcing between our model and the other studies were indeed related to the particle size, we added different percentages of the sulphate emissions readily as coarse mode aerosols. The sensitivity tests show that indeed the AOD and radiative forcing decrease (not shown). This suggests that the decrease is due to both the coarse aerosols

being less scattering in the visible range and the coarse aerosols falling out more quickly.

To conclude, our model is able to reproduce the SO_4 concentrations as simulated by Oman et al. (2006b) whereas it produces a smaller particle size and consequently a stronger radiative forcing induced by the volcanic SO_2 emission. Given that no information is available on the effective radius for the Laki eruption, it is difficult to assess the impact of neglecting intra-modal coagulation in our model simulation of the eruption. We are currently considering including intra-modal coagulation in the aerosol module to investigate its influence on the effective radius for multistage Laki-type eruptions. However, this improvement would require a significant rewrite of the aerosol model, which is beyond the scope of the current work.

3.2. *NH temperature and precipitation*

The large radiative forcing in our study produces a NH cooling almost three times more intense than in other studies (Highwood and Stevenson, 2003; Oman et al., 2006b). Our simulations show a NH cooling peak of about -2.8°C in September accompanied by a significant drop in precipitation (Fig. 4). The cooling lasts around 5 yr whereas the precipitation anomalies seem to recover slightly faster (Fig. 4). This long-lasting cooling may arise from changes in the ocean heat content as also suggested for example by Kravitz and Robock (2011). The cooling relative to the climatology (ENS_{volc}–A – $ENS_{no\text{-}volc}$–B) is concentrated over the NH continents with peaks around -8°C in northeastern Asia in the summer of the eruption. In Europe the simulated summer cooling ranges between -2°C in the western side and -6°C over the eastern side (Fig. 5), about double the magnitude shown by other model studies (Highwood and Stevenson, 2003; Oman et al., 2006a).

Observations from 29 stations spread over Europe and northeastern United States, however, do not show any cooling for the summer of the Laki eruption (Thordarson and Self, 2003). The data show a summer that was actually warmer than average, especially the month of July over Europe (Grattan and Brayshay, 1995; Grattan and Sadler, 1999; Thordarson and Self, 2003; Luterbacher et al., 2004). To our knowledge, no model has been able to reproduce the observed warming, and given the scarcity of the records and their limited spatial distribution, it is challenging to attribute the warm European summer of 1783 to a specific forcing related to the Laki eruption. Warmer-than-average conditions may have simply been caused by prevailing southerly flow over western Europe due to natural variability. The observed European warming contrasts with other areas of the NH, which experienced significant negative temperature anomalies, especially over large portions of Asia and North America. Reconstructed summer temperatures using tree ring maximum latewood density data found that the summer of 1783 was the coldest of the last 400 yr in northwestern Alaska (Jacoby et al., 1999). Radial growth of tree rings in the Polar Urals and Yamal Peninsula in northwest Siberia was the smallest in about 500–600 yr (Hantemirov et al., 2004). Our study shows a cooling of around 3–4°C and 5–6°C for Alaska and Polar Urals/Yamal Peninsula respectively, which is consistent with the tree ring data.

Several studies (Soden et al., 2002; Oman et al., 2005, 2006a; Iles et al., 2013) have shown that global mean precipitation is sensitive to volcanic forcing. However, the precipitation is highly variable; therefore, it is more difficult to detect significant anomalies at regional scale. For example, Oman et al. (2005, 2006a) have shown that both

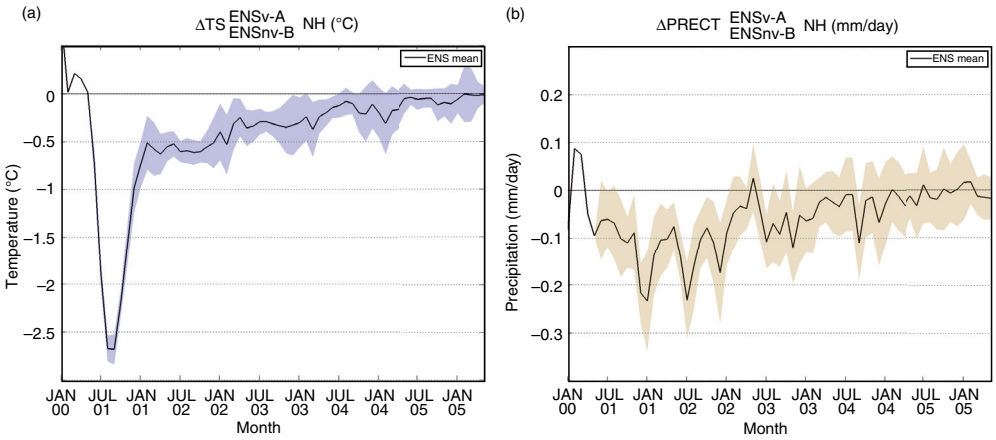

Fig. 4. Ensemble mean change (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–B) in NH surface temperature – TS (left) and precipitation – PRECT (right). The shadings represent the standard deviation of the ensemble difference.

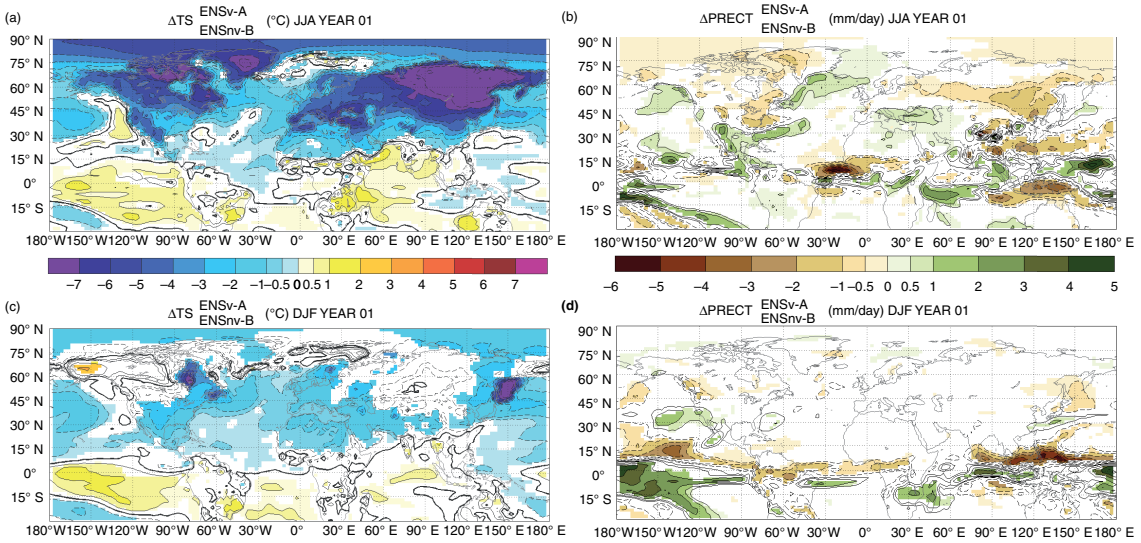

Fig. 5. Ensemble mean change (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–B) in surface temperature – TS (left) and in precipitation – PRECT (right) for the summer (JJA, upper panels) and winter (DJF, lower panels) of the year of the eruption. The colour shading displays significant anomalies at 95% confidence level. The contour intervals (dashed = negative anomalies; solid = positive anomalies) follow the colour bar scale. The bold line indicates the 0°C anomaly.

the Katmai (Alaska) eruption of 1912 and the Laki eruption of 1783 led to a reduction of monsoon precipitation; however, the effects of Katmai were barely distinguishable from weather noise (Oman et al., 2005). On the other hand, the changes in African Monsoon following the Laki eruption seem to be more robust in the model experiments: Oman et al. (2006a) have shown a decrease of 1–3 mm/d over the Sahel in the summer after the eruption. Our results show a significant decrease of NH precipitation that lasts around 3 yr (Fig. 4). The reduction is concentrated over the tropics and subtropics with 1–2 mm/d decrease in precipitation over Sahel and negative peaks of $-5/-6$ mm/d over the Indian basin for the summer of the eruption (Fig. 5). The model clearly shows a southward shift of the Inter-Tropical Convergence Zone in the summer over the Western Pacific and Atlantic Oceans and in all basins in the winter of the eruption. This shift is consistent with the expected response to a NH cooling (Kang et al., 2008; Hwang et al., 2013). These model results are supported by historical observations – severe droughts were indeed reported across both the Nile River watershed (Hassan et al., 1998) and India (Mooley and Pant, 1981) – and a recent study (Haywood et al., 2013) suggesting that volcanic eruptions in the NH can cause severe droughts over Sahel.

4. Climate impacts: comparison between the two ensemble approaches

As mentioned above, on global or hemispheric scales the modelled changes in temperature and precipitation are so large that comparing the volcano ensemble (either ENS_{volc}–A or ENS_{volc}–B) with the climatology ($ENS_{no\text{-}volc}$–B) or ENS_{volc}–A with the $ENS_{no\text{-}volc}$–A gives the same qualitative results (not shown). However, this is not true when looking at the regional scales due to the strong influence of the initial conditions (cf. Figs. 5 and 6). For example, when comparing the volcano ensemble mean (ENS_{volc}–A) for the summer of the eruption with the climatology ($ENS_{no\text{-}volc}$–B) a pronounced warm anomaly is seen over the central-eastern Tropical Pacific and Indian Oceans, but these anomalies are not present when comparing the volcano (ENS_{volc}–A) with the no-volcano ensemble mean of that specific year ($ENS_{no\text{-}volc}$–A), indicating that those changes are specific to the chosen year (which was an El Niño year).

In principle, as mentioned previously, the comparison between ENS_{volc}–A and $ENS_{no\text{-}volc}$–B is only used when evaluating the model against reconstructions and not the general volcanic influence on climate, given the specific background conditions of the selected year in ENS_{volc}–A. To study the volcanic impacts on climate, it is necessary to compare either ENS_{volc}–A or ENS_{volc}–B with the equivalent unperturbed ensemble, $ENS_{no\text{-}volc}$–A or $ENS_{no\text{-}volc}$–B, respectively. In the literature, these two approaches have been both widely used without a rigorous comparison between them. Given the non-linear nature of the climate system, it is reasonable to expect different impacts depending on the background conditions before the eruption. However, it is not obvious that these differences are statistically significant. Therefore, in order to understand whether the two approaches are equivalent, we compare

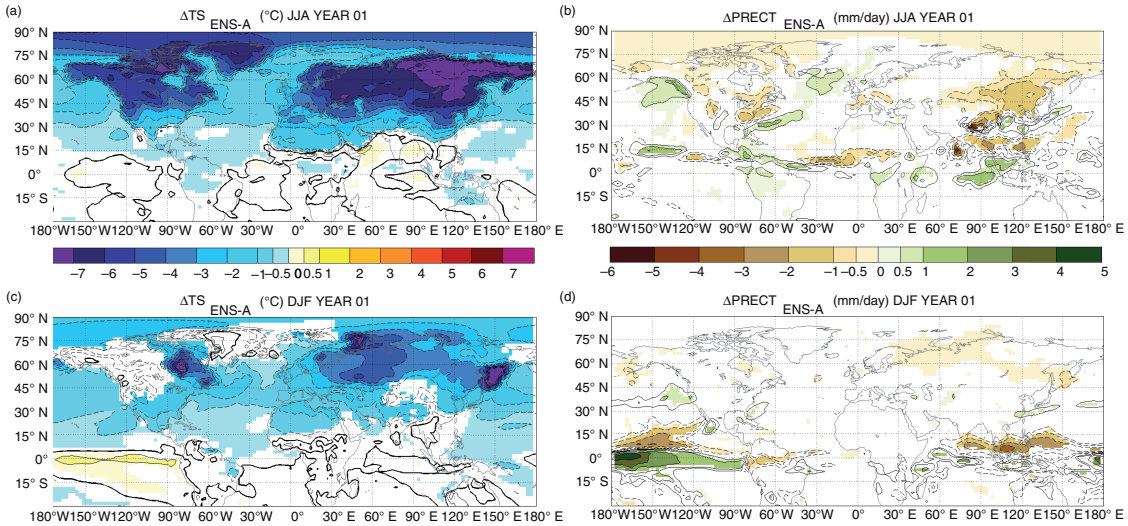

Fig. 6. Ensemble mean change (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–A) in surface temperature – TS (left) and in precipitation – PRECT (right) for the summer (JJA, upper panels) and winter (DJF, lower panels) following the eruption. The colour shading displays significant anomalies at 95% confidence level. The contour intervals (dashed = negative anomalies; solid = positive anomalies) follow the colour bar scale. The bold line indicates the 0°C anomaly.

the temperature and precipitation anomalies in the ENS–A (ENS_{volc}–A minus $ENS_{no\text{-}volc}$–A) with the ENS–B (ENS_{volc}–B minus $ENS_{no\text{-}volc}$–B), focusing only on the summer and winter of the year of the eruption.

In the summer of the eruption, broad areas with statistically significant differences in temperature anomaly are evident over the NH continents, with differences of up to 2–2.5°C in Eastern Europe (Fig. 7a). Differences in precipitation appear mainly over the tropical regions, especially the South Asian Monsoon basin. In winter, the differences in temperature change are most remarkable over North America, with anomalies warmer by up to 4–5°C in ENS–A compared to ENS–B. Warmer anomalies are also present over eastern Greenland and the Barents Sea. Widespread negative anomalies are instead shown over great part of Eurasia; however, only in a relative narrow region in central Asia is the difference significant. The regions that display significant changes in precipitation difference in winter are eastern Tropical Pacific, Indian Ocean, Pacific Northwest and eastern Greenland.

The changes highlighted above are unlikely to occur by chance or be induced by a different ocean response in the two ensemble approaches, and are more likely related to the different initial background state of the ensemble set-ups. In fact, the ocean response to the volcanic forcing is the same in both ensemble set-ups with a temporary weakening of the Meridional Overturning Circulation and a similar decrease in the ocean heat content (not shown). Furthermore, most of the significant differences concern regions known to be influenced by the ENSO. For example, ENSO-induced variability in the Aleutian low often leads to increased precipitation over the Alaskan/Pacific coastal range and above-average temperature over northwestern North America. Therefore, the El Niño background initial state in ENS–A influences the anomalies induced by volcanic eruptions.

In conclusion, our results suggest that the two approaches, ENS–A and ENS–B are not equivalent, particularly in regions strongly tied to ENSO, and caution must be used when interpreting volcano impacts on climate using the ENS–A approach.

5. Post-eruption surface climate variability

Here we investigate whether using interactive aerosols in the simulations has impacts on the variability of the ensembles, and specifically on surface temperature. We analyse the standard deviations of the area-averaged surface temperature (TS) for the ENS–A and ENS–B for three large regions: Europe – EU (35°–72.5°N; 10°W–45°E), North America – NA (35°–72.5°N; 60°–150°W) and Northern Asia – AS (35°–72.5°N; 45°–180°E). We have considered only the land-surface temperature. The domains are chosen to represent the three macro-regions showing the most pronounced cooling and covering all together the entire NH land mass above 35°N. We focus on surface temperature because it is of most direct relevance to volcanic impacts, in addition to having a long historical record that can be compared to model simulations. We have analysed the summer (JJA) and winter (DJF) of the year of the eruption (year 01). To test whether the volcanic eruption significantly impacts temperature variability,

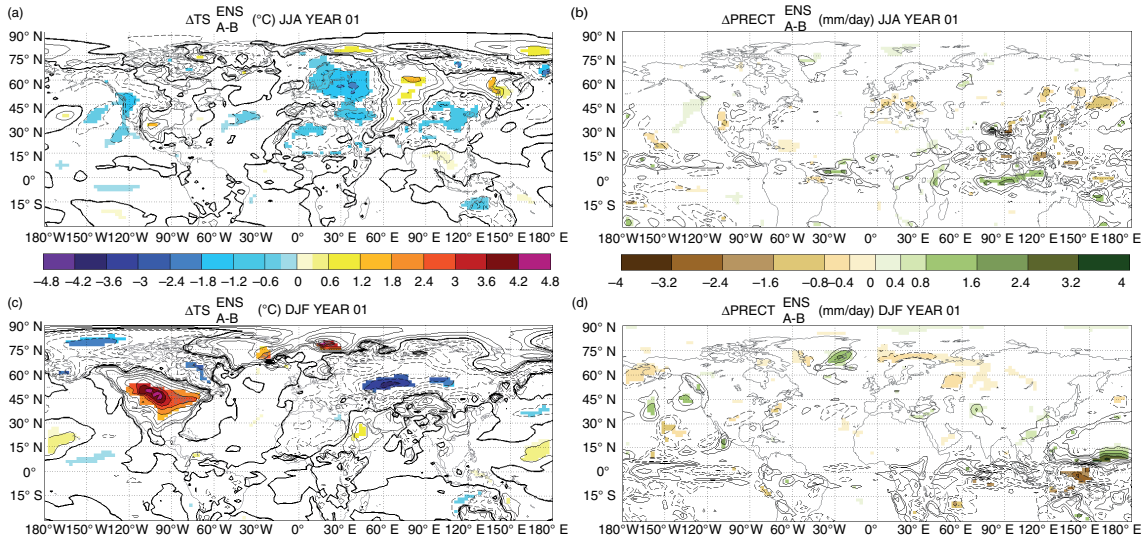

Fig. 7. Differences in surface temperature – TS (left) and precipitation – PRECT (right) anomalies for the summer (JJA, upper panels) and winter (DJF, lower panels) following the eruption between the two different approaches [e.g. ΔTS(ENS–A) – ΔTS(ENS–B)]. The contour intervals (dashed = negative anomalies; solid = positive anomalies) follow the colour bar scale. The bold line indicates the 0°C anomaly. The colour shading displays significant differences at 95% confidence level.

we generated a standard deviation distribution for each of the two $ENS_{no\text{-}volc}$ experiments: Through a bootstrap technique, we generate 1000 sub-samples of 10 members for the ENS–A and ENS–B, randomly selecting from the original pool of 20 members (e.g. TS^{NA} is a 10 × 1000 matrix). We then calculate the standard deviation for each sub-sample: The standard deviation of bootstrapped quantities is by definition an empirical estimate of the standard error of the mean (SEM) as also shown by Toohey and von Clarmann (2013) (e.g. SEM^{TS}_{NA} is 1 × 1000 matrix). We generate the $ENS_{no\text{-}volc}$ SEM distributions that allow quantifying whether the ENS_{volc} variability is within the range of no-volcanic natural variability. It also allows a comparison between the two $ENS_{no\text{-}volc}$ approaches. We consider the changes outside the 5th and 95th percentiles relative to the $ENS_{no\text{-}volc}$ to be statistically significant.

In the no-volcano ensembles, the land-surface temperature SEMs are in general comparable over Europe and Asia, whereas they are significantly larger (twice) in the $ENS_{no\text{-}volc}$–B compared to the $ENS_{no\text{-}volc}$–A over North America in winter (Table 2). $ENS_{no\text{-}volc}$–A members are generated from a specific year, i.e. a particular ENSO background state. On the other hand, the $ENS_{no\text{-}volc}$–B members include a wider range of initial ENSO states and given that North American climate is highly sensitive to ENSO, this set-up consequently leads to a greater variability compared to a single fixed state. Europe and Northern Asia are less sensitive to ENSO variability and hence, the changes between $ENS_{no\text{-}volc}$–A and $ENS_{no\text{-}volc}$–B are smaller.

Comparing the volcano to no-volcano experiments, both ENS_{volc} ensembles generally exhibit larger variability than the corresponding $ENS_{no\text{-}volc}$ ensembles, especially over North America and Europe (Table 2). Interestingly, though, the variability over North America in winter (DJF) in the ENS–B set-up actually decreases from the no-volcano to the volcano ensemble, contrary to what happens in the ENS–A. We attribute this anomalous behaviour to the fact that the eruption triggers a warming response in the equatorial Pacific in the winter of the eruption year (see Fig. 6c). A detailed examination of the ENS–B simulations shows that this warm anomaly is in

Table 2. Land-surface temperature SEMs (°C) for $ENS_{no\text{-}volc}$ and its relative changes (in %) for ENS_{volc} for the summer (JJA) and the winter (DJF) of the year of the eruption in both ensemble set-ups

		SEM_{TS} (°C) ENS–A		SEM_{TS} (°C) ENS–B	
		JJA	DJF	JJA	DJF
North America	$ENS_{no\text{-}volc}$	0.31	0.98	0.38	2.09
	ENS_{volc}	**+50%**	**+18%**	*+22%*	*−27%*
Europe	$ENS_{no\text{-}volc}$	0.40	1.03	0.44	1.04
	ENS_{volc}	+11%	**+45%**	**+23%**	*+19%*
Asia	$ENS_{no\text{-}volc}$	0.46	1.49	0.36	1.13
	ENS_{volc}	−11%	+4%	**+35%**	**+30%**

Changes in bold indicate anomalies above the 95th or below the 5th percentile, in italics anomalies outside the interquartile range (<25th or >75th).

fact considerably stronger in La Niña years than in El Niño years (not shown). Therefore, ENSO variability is effectively reduced in the ENS_{volc}–B compared to $ENS_{no\text{-}volc}$–B, and so are the teleconnected temperature anomalies over North America. This is not the case, in ENS–A given the fixed ENSO state. However, the number of El-Niño/La-Niña events is small (6–8 events), and further investigations are required to confirm our results.

The general increase in the ensemble variability over the NH likely arises from the variability in the radiative forcing induced by the volcanic sulphate aerosols, which in our model are interactively computed from the emitted SO_2. To test this hypothesis, we perform two additional sensitivity experiments, one with SO_2 injections reduced by a factor of 3 ($ENSred_{volc}$) and another with *fixed* SO_2 distribution ($ENSfix_{volc}$); both ensembles use the ENS–A set-up and are composed of 10 members each. The $ENSfix_{volc}$ members are generated simply by reading back the monthly mean concentrations from one of the ensemble members into the model's default volcanic aerosol concentration and scaling it to get approximately the same forcing as the original experiments (ENS_{volc}–A). These two sensitivity experiments should in principle enable us to get some insights on whether or not the change in variability is a function of the strength of the eruption ($ENSred_{volc}$–A) and whether it depends on the interactive aerosol ($ENSfix_{volc}$–A).

The $ENSred_{volc}$ members show a cooling that is almost 1/3 of that in ENS_{volc}–A, especially in the summer of the eruption. The magnitude of the cooling in $ENSfix_{volc}$ is almost as large as ENS_{volc}–A in the summer of the eruption, but recovers faster (Table 3). $ENSred_{volc}$–A shows increased variability compared to $ENS_{no\text{-}volc}$–A over North America (both in summer and winter) and Europe (only in winter) with a magnitude similar to ENS_{volc}–A (Table 4). This sensitivity experiment suggests that the increased variability is not a linear function of the volcanic eruption strength.

The second sensitivity experiment ($ENSfix_{volc}$–A) does not show any significant change in variability compared to the $ENS_{no\text{-}volc}$–A, and the sign of the changes is often not consistent. This experiment points towards how the interactive aerosols and the changing SO_2 spatial distribution are likely to be the primary drivers of the changes in surface temperature variability in the volcano experiments.

Table 3. Changes in land-surface temperature (°C) relative to $ENS_{no\text{-}volc}$–A for the summer (JJA) and the winter (DJF) of the year of the eruption in the ENS_{volc}–A and the sensitivity experiments ($ENSred_v$ and $ENSfix_v$)

	ΔTS (°C) ENS_{volc}–A		ΔTS (°C) $ENSred_{volc}$–A		ΔTS (°C) $ENSfix_{volc}$ – A	
	JJA	DJF	JJA	DJF	JJA	DJF
North America	−4.61	−1.43	−1.41	−0.16	−3.81	+0.16
Europe	−4.66	−2.16	−1.51	−0.96	−3.78	−1.10
Asia	−5.73	−2.45	−2.00	−1.28	−4.49	−1.03

Table 4. Changes (in %) in land-surface temperature SEMs (°C) relative to $ENS_{no\text{-}volc}$–A for the summer (JJA) and the winter (DJF) of the year of the eruption in the sensitivity experiments ($ENSred_v$ and $ENSfix_v$) for the ENS–A set-up

	ΔSEM_{TS} (°C) $ENSred_{volc}$–A		ΔSEM_{TS} (°C) $ENSfix_{volc}$–A	
	JJA	DJF	JJA	DJF
North America	**+32%**	**+29%**	*+22%*	−15%
Europe	−4%	**+33%**	+10%	*+18%*
Asia	**+65%**	*+18%*	*−19%*	+2%

Changes in bold indicate anomalies above the 95th or below the 5th percentile, in italics anomalies outside the interquartile (<25th or >75th).

In summary, our results suggest that volcanic eruptions alter and in general increase surface temperature variability. Climate models that use prescribed aerosol forcing may not be able to capture such potential changes in variability.

6. Required number of ensemble members

Here we investigate a further important methodological issue that has not been systematically addressed in previous studies, namely that of determining the number of ensemble members necessary to discern the volcano signal from the underlying noise of the natural internal variability on regional scale. We analyse the surface temperature anomalies for the ENS–A and ENS–B as a function of the number of ensemble members considered for the same three large regions defined in the previous section.

We calculate the area-averaged temperature anomalies relative to the all-member ENS mean in the three domains for an increasing number of ensemble members. The area-averaged temperature anomalies are defined as follows [eq. (2)]:

$$TS^{ENS*}_{anomaly} = \overline{TS}^{n-members}_{ENS*} - \overline{TS}^{all-members}_{ENS*} \quad (2)$$

where $\overline{TS}^{n-members}_{ENS*}$ is the surface temperature mean of a given ENS set-up for a size n of ensemble members with n going from one to the total ensemble size and $\overline{TS}^{all-members}_{ENS*}$ is the surface temperature mean of a given ENS set-up for all members (i.e. 20). This $TS^{ENS*}_{anomaly}$ is a measure of how an n-member TS mean can be offset from the all-member ('true') ensemble mean (i.e. how the mean temperature obtained – for example – using only three members differs from the 'true' mean calculated using 20 members). Hence, it gives

information on how biased a regional scale temperature estimate can be as a function of the number of members used. However, the $TS^{ENS*}_{anomaly}$ decreases as n increases and is zero when the number of members equals all members by construction. To overcome this limitation, we use the bootstrap technique with replacement, where we sequentially draw n-members from the same pool. For example, for $n=3$ we randomly draw the first ensemble member from the entire pool of 20 members, then we do the same thing with the second one, and then with the third one, each drawn from the same entire pool. We perform this calculation 1000 times in order to obtain a distribution of $\overline{TS}^{\text{n-members}}_{\text{ENS*}}$ and hence, $TS^{ENS*}_{anomaly}$ for each n. The bootstrap technique with replacement also enables extending the analysis by choosing n larger than the size of the pool. Hence, we have considered a total population size of 50 members. We have then calculated the 5th and 95th percentile of the $TS^{ENS*}_{anomaly}$ distributions for each n (n from 1 to 50) and for each ensemble. This analysis is independent from the strength of the volcanic forcing since each ensemble has been considered separately and provides a lower bound estimation of the number of ensemble members necessary to capture a given temperature anomaly/volcanic signature (e.g. 0.2 or 0.5°C) for the considered regions and for the summer and winter seasons following the eruption. For example, to be able to capture a 1°C temperature anomaly over North America for the winter following the eruption, about 15 members are necessary using the ENS–B and only five using the ENS–A approach (Fig. 8). To capture a 0.5°C anomaly, the number of

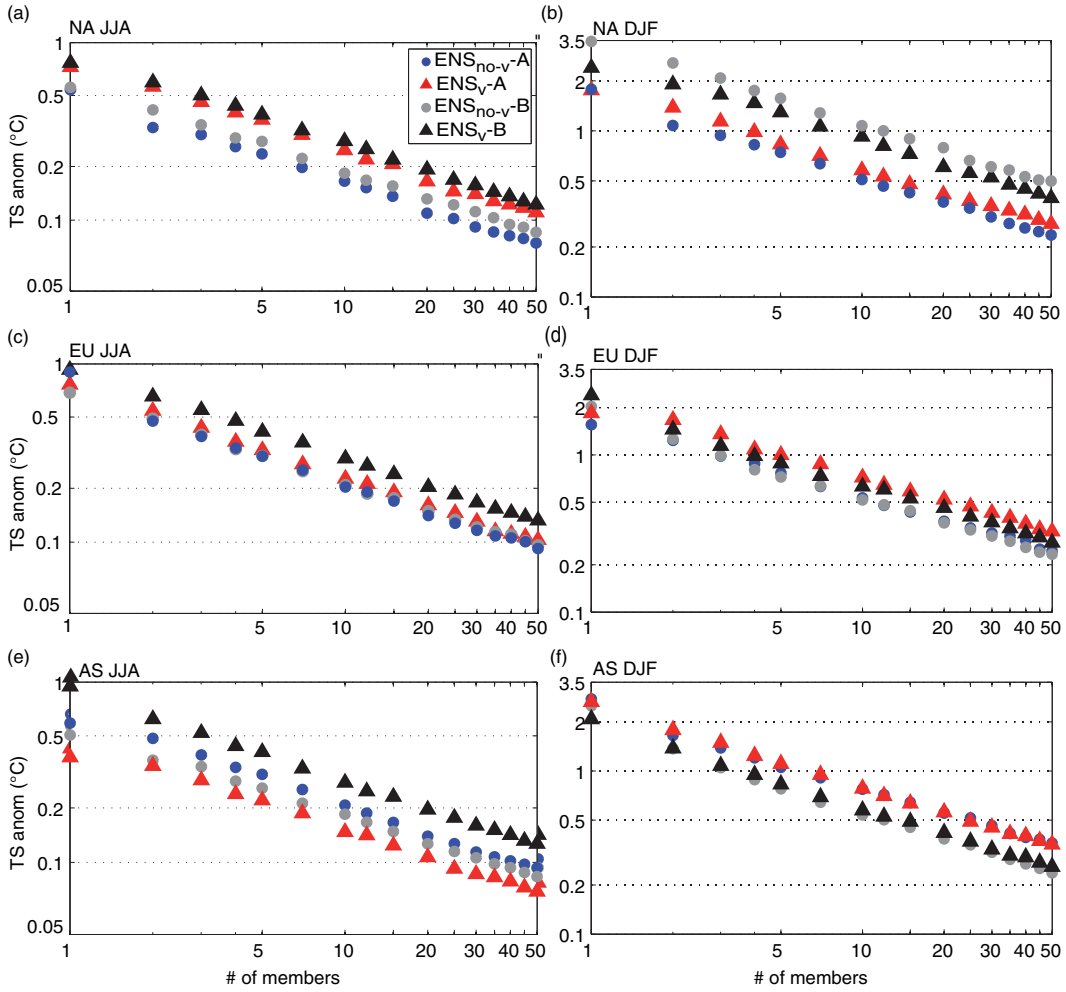

Fig. 8. Average between the absolute values of 5th and 95th percentile of $TS^{ENS*}_{anomaly}$ distributions as a function of the number of ensemble members (n) for North America (NA), Europe (EU) and Asia (AS) for the summer (JJA) and winter (DJF) of the eruption, calculated for each ensemble. The 5th and 95th anomalies are fairly symmetric around the mean for a number of members greater than 5: The difference in absolute value between the 5th and 95th percentiles is below 0.1°C. The curves show the temperature anomalies that can be detected at the *95% confidence level* as a function of n. The anomalies are computed relative to the all-member mean of each ensemble.

members triples. For the other regions, the difference in the number of necessary members between ENS–A and ENS–B is smaller and to capture a 0.5°C anomaly in winter 15–25 members are needed. As mentioned in the previous section, the reason why the ENS–B requires a higher number of ensemble members for North America is related to the high sensitivity of this region to ENSO.

To summarise, our analyses provide a lower bound estimate of the number of ensemble members necessary to capture the volcano signal on a regional scale in terms of surface temperature: In general, they show the need of selecting more members when using the ENS–B compared to the ENS–A approach to account for regions in which the initial background state (e.g. ENSO) is particularly important – such as North America. We must note that further in-depth studies are warranted to provide a more conclusive estimate of the necessary ensemble size, as it is likely that the results may be model-dependent. This dependency, however, is only related to the ability of the model to reproduce natural internal variability. For example, models that have a lower internal variability require fewer ensemble members compared to others with larger internal variability.

7. Discussions and conclusion

The present study evaluates NorESM performance in simulating the climate impacts of the second largest high-latitude eruption in historical time, using interactive SO_2 emissions and aerosol formation. While the model is able to reproduce the correct SO_4 concentrations, it seems to underestimate – with respect to prior model studies – the particle size since self-coagulation growth is not simulated in CAM4–Oslo. Self-coagulation is an important process following an eruption when a massive amount of sulphate is emitted, as described in Pinto et al. (1989) and English et al. (2013). However, for a given amount of emitted SO_2 there can be large differences in the effective radius obtained (Pinto et al., 1989). The lack of information on the effective particle size in the Laki eruption hampers a quantification of the potential biases introduced by not taking into account self-coagulation growth. A future step to better understand the role played by self-coagulation in a Laki-type eruption and to improve NorESM performance in simulating interactively the formation of sulphate from volcanic eruptions will be to include intra-modal coagulation to allow SO_4 particle growth into the coarse mode.

Some studies (e.g. Oman et al., 2006b; Schneider et al., 2009) have used three ensemble members to isolate the volcanic signal from the background internal variability, given the large volcanic forcing. While this is reasonable when looking at changes in large-scale patterns (Figs. 3 and 4) such as global temperature, AOD (e.g. Oman et al., 2006b), and radiative forcing it is less true when looking for example at regional temperature changes (Fig. 8). Our study highlights the importance of having a large ensemble size in order to be able to detect significant changes on regional scale (Fig. 8). For example, up to 40 ensemble members are necessary to detect a significant change in surface temperature of 0.5°C over North America in the winter following the volcanic eruption (Fig. 8). Our analyses provide a lower bound estimate of the number of ensemble members necessary in order to discern the volcano signal from the natural internal variability on regional scale: At least 20–25 ensemble members, but possibly more (35–40) especially when looking at changes over North America, are needed to significantly detect seasonally averaged anomalies of 0.5°C. Furthermore, our results show that volcanic eruptions are likely to generally increase the internal variability of the climate system. This implies that models using prescribed aerosol forcing may need a slightly lower number of ensemble members; however, it is likely that these models will not be able to capture potential changes in variability caused by the eruption.

Finally, our study compares the two most common approaches used in the literature to construct ensembles to simulate volcano eruptions. Twin ensembles of volcano and reference runs are conducted starting either from one specific year (ENS–A: 1934) slightly modifying the initial conditions between one member and the next, or every 5 yr (ENS–B: 1860, 1865, . . ., 1960). The response of the climate system to volcanic forcing is calculated as the ensemble mean over the volcano runs minus the ensemble mean over the reference runs. In principle, the two approaches might be considered equivalent because when selecting one specific year and then running an ensemble for the volcano and no-volcano set-up, the specific initial background state is removed when subtracting the ensemble mean of the no-volcano from the volcano ensemble mean (Eq. 1). For the ENS-B approach, generating an eruption every 5 yr randomises the initial state in which the eruption takes place and hence makes the changes independent from the initial state. Nevertheless, our study shows that the two approaches, ENS–A and ENS–B, are not equivalent, especially in regions strongly tied to ENSO. Yang and Schlesinger (2002) have shown, using composite and singular value decomposition analyses to separate the volcano signal from the ENSO, that for the 1991 Pinatubo eruption 50% of the temperature response over North America was due to the warm ENSO phase, whereas in Europe the ENSO contribution was weak. Furthermore, Zanchettin et al. (2013) have shown that under different background conditions, simulated atmospheric and especially oceanic dynamics may evolve significantly differently after large tropical eruptions. Our study extends these conclusions to high-latitude

volcanic eruptions and relative to atmospheric dynamics: Background conditions are not merely a source of additive noise, but actively influence the mechanisms involved in the post-eruption interannual evolution. Therefore, caution must be used when interpreting volcano impacts on climate using the ENS–A approach, i.e. picking one specific year to simulate the volcano eruption. The ENS–A approach should be adopted to specifically study the effects of volcanic forcing superimposed on a particular initial state (e.g. warm/cold ENSO state). ENS–B should be used, instead, when trying to isolate the volcano signal from the background internal variability.

In light of the relevance of background climate conditions, we should point out that in simulating past volcanic eruptions and comparing them with reconstructions, a proper initialisation of the model with similar large-scale modes of variability such as El Niño – if known – is recommended. For example, the positive anomalies in the summer of the eruption over the Tropical Pacific Ocean shown in Fig. 5 (ENS_{volc}–A vs. $ENS_{no\text{-}volc}$–B/climatology) are not due to the volcanic eruption but rather to the simulated El Niño phase. Therefore, when comparing results with historical and/or proxy reconstructions some of the reported/recorded climate anomalies may not be due to the volcanic eruption itself but to the pre-existing atmosphere/ocean conditions.

Given the growing interest in an initiative concerning the climate response to very strong volcanic eruptions (such as the Model Intercomparison Project on the climatic response to Volcanic forcing – VolMIP) and the differences among models in reproducing volcanic cooling, we envision the need for common guidelines to be adopted in experimental design related to simulating volcanic eruptions. Our study provides an overview of crucial aspects that need to be taken into account when developing common initiatives, such as the type of ensemble approach, the choice of the background conditions and the necessary ensemble size. It also highlights the importance of using climate models with an interactive aerosol module rather than prescribed aerosol forcing to better capture potential changes in climate variability induced by the eruption. Finally, it provides a tentative guideline related to the number of ensemble members that should be adopted for seasonal regional anomalies and for given volcanically induced temperature anomalies.

8. Acknowledgements

The authors would like to thank Matthew Toohey for insightful discussions, and Marc Chiacchio and the two anonymous reviewers for helpful comments on the manuscript. A.G. and Ø.S. have been supported through the European Commission FP7 projects PEGASOS (FP7-ENV-2010–265148) and ACCESS FP7-ENV-2010-265863) and by the Research Council of Norway through project EVA "Earth system modelling of climate Variations in the Anthropocene" (grant no. 229771) as well as NOTUR project "CPU time for EVA" (nn2345k) and NorStore project "Storage for EVA" (ns2345k). Part of the simulations was also performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC. Data for this paper can be requested to F.S.R.P. (francesco.pausata@misu.su.se)

Authors' contributions

F.S.R.P. conceived the study, designed and carried out the experiments, processed and analysed the model results, and wrote the manuscript. A.G. modified the code, carried out the experiments and edited the manuscript. R.C. and A.H. provided support in the analysis of the results and edited the manuscript. Ø.S. provided support in the model set-up.

References

Ammann, C. M. 2003. A monthly and latitudinally varying volcanic forcing dataset in simulations of 20th century climate. *Geophys. Res. Lett.* **30**(12), 1657. DOI: 10.1029/2003GL016875.

Assmann, K. M., Bentsen, M., Segschneider, J. and Heinze, C. 2010. Model development an isopycnic ocean carbon cycle model. *Geosci. Model Dev.* **3**, 143–167.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A. and co-authors. 2013. The Norwegian Earth System Model, NorESM1-M – part 1: description and basic evaluation of the physical climate. *Geosci. Model Dev.* **6**(3), 687–720. DOI: 10.5194/gmd-6-687-2013.

Berdahl, M. and Robock, A. 2013. Northern Hemispheric cryosphere response to volcanic eruptions in the Paleoclimate Modeling Intercomparison Project 3 last millennium simulations. *J. Geophys. Res. Atmos.* **118**(22), 12359–12370. DOI: 10.1002/2013JD019914.

Bindoff, N. L., Stott, P. A., AchutaRao, K. M., Allen, M. R., Gillett, N. and co-authors. 2013. Detection and attribution of climate change: from global to regional. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, and co-editors). Cambridge University Press, Cambridge, UK; New York, NY. pp. 867–952.

Cagnazzo, C. and Manzini, E. 2009. Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European Region. *J. Clim.* **22**(5), 1223–1238. DOI: 10.1175/2008JCLI2549.1.

Cole-Dai, J. 2010. Volcanoes and climate, Wiley Interdiscip. *Rev. Clim. Chang.* **1**(6), 824–839. DOI: 10.1002/wcc.76.

Cook, E. and Krusic, P. 2004. *North American Drought Atlas.* NOAA Paleoclimatol., Boulder, CO.

D'Arrigo, R., Seager, R., Smerdon, J. E., LeGrande, A. N. and Cook, E. R. 2011. The anomalous winter of 1783–1784: was the Laki eruption or an analog of the 2009–2010 winter to blame? *Geophys. Res. Lett.* **38**(5), L05706. DOI: 10.1029/2011GL046696.

Driscoll, S., Bozzo, A., Gray, L. J., Robock, A. and Stenchikov, G. 2012. Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. *J. Geophys. Res.* **117**(D17), 1–45.

English, J. M., Toon, O. B. and Mills, M. J. 2013. Microphysical simulations of large volcanic eruptions: Pinatubo and Toba. *J. Geophys. Res. Atmos.* **118**(4), 1880–1895. DOI: 10.1002/jgrd.50196.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C. and co-authors. 2011. The community climate system model version 4. *J. Clim.* **24**(19), 4973–4991. DOI: 10.1175/2011JCLI4083.1.

Grattan, J. and Brayshay, M. 1995. An amazing and portentous summer: environmental and social responses in Britain to the 1783 eruption of an Iceland Volcano. *Geograph. J.* **161**, 125–134. DOI: 10.2307/3059970.

Grattan, J. and Sadler, J. 1999. Regional warming of the lower atmosphere in the wake of volcanic eruptions: the role of the Laki fissure eruption in the hot summer of 1783. *Geol. Soc. Lond. Spec. Publ.* **161**(1), 161–171. DOI: 10.1144/GSL.SP.1999.161.01.11.

Hantemirov, R. M., Gorlanova, L. A. and Shiyatov, S. G. 2004. Extreme temperature events in summer in northwest Siberia since AD 742 inferred from tree rings, Palaeogeogr. *Palaeoclimatol. Palaeoecol.* **209**(1–4), 155–164. DOI: 10.1016/j.palaeo.2003.12.023.

Hassan, F. A., Romain, E., Hermon, E. and Bar, S. 1998. Climate change, Nile floods and civilization. *Nat. Resour.* **34**, 34–40.

Haywood, J. M., Jones, A., Bellouin, N. and Stephenson, D. 2013. Asymmetric forcing from stratospheric aerosols impacts Sahelian rainfall. *Nat. Clim. Chang.* **3**(7), 660–665. DOI: 10.1038/nclimate1857.

Highwood, E. J. and Stevenson, D. S. 2003. Atmospheric impact of the 1783 – 1784 Laki eruption: part II climatic effect of sulphate aerosol. *Atmos. Chem. Phys.* **3**, 1177–1189.

Hwang, Y.-T., Frierson, D. M. W. and Kang, S. M. 2013. Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century. *Geophys. Res. Lett.* **40**(11), 2845–2850. DOI: 10.1002/grl.50502.

Iles, C. E., Hegerl, G. C., Schurer, A. P. and Zhang, X. 2013. The effect of volcanic eruptions on global precipitation. *J. Geophys. Res. Atmos.* **118**(16), 8770–8786. DOI: 10.1002/jgrd.50678.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A. and co-authors. 2013. The Norwegian Earth System Model, NorESM1-M – part 2: climate response and scenario projections. *Geosci. Model Dev.* **6**(2), 389–415. DOI: 10.5194/gmd-6-389-2013.

Jacoby, G. C., Workman, K. W. and D'Arrigo, R. D. 1999. Laki eruption of 1783, tree rings, and disaster for northwest Alaska Inuit. *Quat. Sci. Rev.* **18**(12), 1365–1371. DOI: 10.1016/S0277-3791(98)00112-7.

Kang, S. M., Held, I. M., Frierson, D. M. W. and Zhao, M. 2008. The response of the ITCZ to extratropical thermal forcing: idealized slab-ocean experiments with a GCM. *J. Clim.* **21**(14), 3521–3532. DOI: 10.1175/2007JCLI2146.1.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E. and co-authors. 2013. Aerosol-climate interactions in the Norwegian Earth System Model – NorESM1-M. *Geosci. Model Dev.* **6**, 207–244. DOI: 10.5194/gmd-6-207-2013.

Kravitz, B. and Robock, A. 2011. Climate effects of high-latitude volcanic eruptions: role of the time of year. *J. Geophys. Res.* **116**(D1), D01105. DOI: 10.1029/2010JD014448.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. and Wanner, H. 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science.* **303**(5663), 1499–1503. DOI: 10.1126/science.1093877.

Man, W., Zhou, T. and Jungclaus, J. H. 2014. Effects of large volcanic eruptions on global summer climate and East Asian monsoon changes during the last millennium: analysis of MPI-ESM simulations. *J. Clim.* **27**(19), 7394–7409. DOI: 10.1175/JCLI-D-13-00739.1.

Medhaug, I. and Furevik, T. 2011. North Atlantic 20th century multidecadal variability in coupled climate models: sea surface temperature and ocean overturning circulation. *Ocean Sci.* **7**(3), 389–404. DOI: 10.5194/os-7-389-2011.

Meronen, H., Henriksson, S. V., Räisänen, P. and Laaksonen, A. 2012. Climate effects of northern hemisphere volcanic eruptions in an Earth System Model. *Atmos. Res.* **114–115**, 107–118. DOI: 10.1016/j.atmosres.2012.05.011.

Mooley, D. A. and Pant, G. B. 1981. Droughts in India over the last 200 years, their socio-economic impacts and remedial measures for them. In: *Climate and History: Studies Past Climates and Their Impact Man* (eds. T. M. L. Wigley, M. J. Ingram, and G. Farmer). Cambridge University Press, Cambridge, United Kingdom. Online at: http://www.cambridge.org/se/academic/subjects/earth-and-environmental-science/climatology-and-climate-change/climate-and-history-studies-past-climates-and-their-impact-man

Neale, R. B., Richter, J., Park, S., Lauritzen, P. H., Vavrus, S. J. and co-authors. 2013. The Mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully coupled experiments. *J. Clim.* **26**(14), 5150–5168. DOI: 10.1175/JCLI-D-12-00236.1.

Oman, L., Robock, A., Stenchikov, G. L., Schmidt, G. A. and Ruedy, R. 2005. Climatic response to high-latitude volcanic eruptions. *J. Geophys. Res.* **110**(D13), D13103. DOI: 10.1029/2004JD005487.

Oman, L., Robock, A., Stenchikov, G. L. and Thordarson, T. 2006a. High-latitude eruptions cast shadow over the African monsoon and the flow of the Nile. *Geophys. Res. Lett.* **33**(18), L18711. DOI: 10.1029/2006GL027665.

Oman, L., Robock, A., Stenchikov, G. L., Thordarson, T., Koch, D. and co-authors. 2006b. Modeling the distribution of the volcanic aerosol cloud from the 1783–1784 Laki eruption. *J. Geophys. Res.* **111**(D12), D12209. DOI: 10.1029/2005JD006899.

Pinto, J. P., Turco, R. P. and Toon, O. B. 1989. Self-limiting physical and chemical effects in volcanic eruption clouds. *J. Geophys. Res.* **94**(D8), 11165. DOI: 10.1029/JD094iD08p11165.

Robock, A. 2000. Volcanic eruptions and climate. *Rev. Geophys.* **38**(2), 191–219.

Robock, A. 2002. Blowin' in the wind: research priorities for climate effects of volcanic eruptions. *EOS.* **83**(42), 472. DOI: 10.1029/2002EO000333.

Russell, P. B., Livingston, J. M., Pueschel, R. F., Bauman, J. J., Pollack, J. B. and co-authors. 1996. Global to microscale

evolution of the Pinatubo volcanic aerosol derived from diverse measurements and analyses. *J. Geophys. Res.* **101**(D13), 18745. DOI: 10.1029/96JD01162.

Santer, B. D., Bonfils, C., Painter, J. F., Zelinka, M. D., Mears, C. and co-authors. 2014. Volcanic contribution to decadal changes in tropospheric temperature. *Nat. Geosci.* **7**(3), 185–189. DOI: 10.1038/ngeo2098.

Schmidt, A., Carslaw, K. S., Mann, G. W., Wilson, M., Breider, T. J. and co-authors. 2010. The impact of the 1783–1784 AD Laki eruption on global aerosol formation processes and cloud condensation nuclei. *Atmos. Chem. Phys.* **10**(13), 6025–6041. DOI: 10.5194/acp-10-6025-2010.

Schmidt, A., Thordarson, T., Oman, L. D., Robock, A. and Self, S. 2012. Climatic impact of the long-lasting 1783 Laki eruption: inapplicability of mass-independent sulfur isotopic composition measurements. *J. Geophys. Res.* **117**(D23), D23116. DOI: 10.1029/2012JD018414.

Schneider, D. P., Ammann, C. M., Otto-Bliesner, B. L. and Kaufman, D. S. 2009. Climate response to large, high-latitude and low-latitude volcanic eruptions in the community climate system model. *J. Geophys. Res.* **114**(D15), D15101. DOI: 10.1029/2008JD011222.

Shepherd, T. G. 2014. Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.* **7**(10), 703–708. DOI: 10.1038/ngeo2253.

Shindell, D. T. 2004. Dynamic winter climate response to large tropical volcanic eruptions since 1600. *J. Geophys. Res.* **109**(D5), D05104. DOI: 10.1029/2003JD004151.

Soden, B. J., Wetherald, R. T., Stenchikov, G. L. and Robock, A. 2002. Global cooling after the eruption of Mount Pinatubo: a test of climate feedback by water vapor. *Science*. **296**(5568), 727–730. DOI: 10.1126/science.296.5568.727.

Thomas, M. A., Timmreck, C., Giorgetta, M. A., Graf, H.-F. and Stenchikov, G. 2009. Simulation of the climate impact of Mt. Pinatubo eruption using ECHAM5 – part 1: sensitivity to the modes of atmospheric circulation and boundary conditions. *Atmos. Chem. Phys.* **9**(2), 757–769. DOI: 10.5194/acp-9-757-2009.

Thordarson, T. and Self, S. 2003. Atmospheric and environmental effects of the 1783–1784 Laki eruption: a review and reassessment. *J. Geophys. Res.* **108**(D1), 4011. DOI: 10.1029/2001JD002042.

Timmreck, C. 2012. Modeling the climatic effects of large explosive volcanic eruptions, Wiley Interdiscip. *Rev. Clim. Chang.* **3**(6), 545–564. DOI: 10.1002/wcc.192.

Toohey, M., Krüger, K., Niemeier, U. and Timmreck, C. 2011. The influence of eruption season on the global aerosol evolution and radiative impact of tropical volcanic eruptions. *Atmos. Chem. Phys.* **11**(23), 12351–12367. DOI: 10.5194/acp-11-12351-2011.

Toohey, M. and von Clarmann, T. 2013. Climatologies from satellite measurements: the impact of orbital sampling on the standard error of the mean. *Atmos. Meas. Tech.* **6**(4), 937–948. DOI: 10.5194/amt-6-937-2013.

Wegmann, M., Brönnimann, S., Bhend, J., Franke, J., Folini, D. and co-authors. 2014. Volcanic influence on European summer precipitation through monsoons: possible cause for "years without summer." *J. Clim.* **27**(10), 3683–3691. DOI: 10.1175/JCLI-D-13-00524.1.

Yang, F. and Schlesinger, M. 2002. On the surface and atmospheric temperature changes following the 1991 Pinatubo volcanic eruption: a GCM study. *J. Geophys. Res.* **107**(D8), 4073. DOI: 10.1029/2001JD000373.

Zanchettin, D., Bothe, O., Graf, H. F., Lorenz, S. J., Luterbacher, J. and co-authors. 2013. Background conditions influence the decadal climate response to strong volcanic eruptions. *J. Geophys. Res. Atmos.* **118**(10), 4090–4106. DOI: 10.1002/jgrd.50229.

Zanchettin, D., Timmreck, C., Graf, H.-F., Rubino, A., Lorenz, S. and co-authors. 2012. Bi-decadal variability excited in the coupled ocean–atmosphere system by strong tropical volcanic eruptions. *Clim. Dyn.* **39**(1–2), 419–444. DOI: 10.1007/s00382-0.