Reply on RC2
Helen Mackay et al.

Author comment on "The 852/3 CE Mount Churchill eruption: examining the potential climatic and societal impacts and the timing of the Medieval Climae Anomaly in the North Atlantic region" by Helen Mackay et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2021-170-AC3, 2022

Author Response to RC2

- We thank the reviewer for their helpful and detailed comments, which have helped to strengthen the manuscript. We have provided a response to each comment below and stated the corresponding changes that have been made to the manuscript.

Specific questions

A main result is that the tree-ring reconstructed temperature suggests a significant cooling but the reconstructed sulfur emission and hence 'climate forcing potential' is suggested to be moderate, and the simulated cooling is smaller.

However, the authors only run climate model simulations based on SAOD reconstructions using the best-estimate of 5 Tg SO₂. Although it is briefly discussed that the climate forcing could have been underestimated due to missed interactions with halogens or ash, I think it would be useful to also include more discussion on the sulfur emission uncertainty. The SO₂ emission uncertainty is included for the EVA_H reconstruction in Figure 3, but not EVA - what is the peak SAOD if EVA is run using 7.5 Tg SO₂? This might be a useful addition to Figure 3. Is a higher EVA SAOD discounted because the EVA_H SAOD estimate is much lower? Given the limited ice core records could the emission estimate have been even higher? Some discussion on the spatial SAOD pattern would also be useful - would these results be different if an aerosol-microphysical model was used such as in Toohey et al. (2019) who show strong confinement of the aerosol to the NH for an eruption at 56°N. Would this lead to a stronger NH forcing and temperature response?

- RESPONSE: We have now included the 95% confidence interval on the EVA global mean SAOD prediction in Figure 3.b and quoted the corresponding number in the text (Line 285-287):

“The EVA(eVolv2k) reconstructed stratospheric aerosol optical depth (SAOD) at 550 nm for the 852/3 CE eruption is relatively moderate, with a peak aerosol optical depth perturbation of 0.049 (95% confidence interval 0.021–0.085) in terms of global monthly mean, and 0.078 in terms of NH monthly mean (Fig. 3a-b).”
We have also corrected a mistake: the uncertainty on the injected SO$_2$ mass is 3.3 Tg of SO$_2$ (Line 175), not 2.5 as wrongly reported in the initial submission. To further discuss uncertainty, we provide lower and upper-end estimate of the peak global mean radiative forcing (-1.7- -0.33 W m$^{-2}$) (Lines 303-306) based on the upper and lower end SAOD estimates as well as the range of estimates for the forcing efficiency depending on the eruption season (Marshall et al., 2019):

“The upper-end SAOD estimate from EVA(eVolv2k), obtained from a winter eruption (which would maximize the forcing efficiency, Marshall et al., 2019), has a global monthly mean radiative forcing peak of -1.7 W m$^{-2}$. Conversely, the lower-end SAOD estimate from EVA_H, obtained from a summer eruption, has a mean peak forcing of -0.33 W m$^{-2}$.”

Furthermore, we now present in greater detail these uncertainties, and those related to the dispersion of the aerosol cloud, on Lines 667-672 (please see the quoted text below), and we discuss how they would affect the consistency between the simulated NH summer cooling and that reconstructed from tree rings. Although we cannot exclude that the mass of SO$_2$ emitted could be even higher than the upper-end used, we note that it is not necessary to invoke such arguments as our simulations show that the combination of the forced response and natural variability could lead to the reconstructed NH cooling.

“Another possibility, which would relax the requirement for a rather strong contribution of natural variability, would be that the volcanic aerosol forcing was in reality stronger than that used here. With the upper-end reconstructed EVA(eVolv2k) SAOD estimate being 66% higher than the best estimate used in our model simulations, the forced model response could be even higher (Figure 3.b). Furthermore, a stronger restriction of aerosols to the NH, not simulated in the simple SAOD reconstruction methods but compatible with interactive stratospheric aerosol model simulations (e.g., Toohey et al., 2019) may also contribute to stronger aerosol forcing over the NH than used here.”

L69 – I suggest changing 'precisely' to 'even when the eruption has a small age uncertainty' or similar

- RESPONSE: Change incorporated: wording changed as suggested above (now Line 71)

L124 – 1991 Mt. Pinatubo estimates are now slightly lower; perhaps say 'around a third’?

- RESPONSE: Change incorporated: wording changed as suggested above (now Line 128)

L173 and Appendix B – this description is a bit confusing - how are the ensemble members generated? Do they consider different ENSO states or is a small perturbation introduced for each? Could background conditions play a role in reconciling the model results and tree-ring reconstructions?

- RESPONSE: The ensemble members are generated by introducing a small perturbation for each simulation starting from the year 845 CE. We added a sentence clarifying the description in lines 187-188:

"Each ensemble member is branched off at 845 CE with a small perturbation in the atmosphere introduced at the first time step. Then, the simulations are seamlessly run until 859 CE."

All simulations initially start from the same background conditions, but by the year of the eruption (853 CE), the range of ENSO states is diverse among the simulations as the SST is not fixed. Hence, ensembles members are considering different ENSO and initial conditions. Background conditions, not only from the ENSO but associated with
internal variability, may play a role in reconciling the model and proxies. We have briefly mentioned the observed model-proxy discrepancy due to the internal variability in Lines 686-688.

" These deviations (changes in temperature amplitudes and in spatial patterns) are expected as the ensemble means of the simulations focus on the signal of the volcanic eruption by reducing internal climate system variability."

Why does the model prescribe sulfate aerosol mass rather than the optical properties from EVA? This seems inconsistent with the focus on reconstructing the forcing up to this point. How does the model treat the aerosol-radiation interactions and what does the model then simulate for SAOD and radiative forcing and how does this compare to Figure 3?

- RESPONSE: CESM1.2 uses a prescribed monthly mean sulfate aerosol mass on a predefined latitudinal and vertical grid as an input volcanic forcing. The aerosol is then used in the radiation code of the model, i.e., optical properties are estimated within the model assuming that the aerosol mass is comprised of 75% sulfuric acid and 25% water and has a constant log-normal size distribution with a constant effective radius and following Neely et al (2016). We included these details in Appendix B, Lines 919-923.

L183 – I’m a bit confused regarding the anomalies - are the volcanic anomalies with respect to the 845-859 period or pre-eruption (as specified for the statistical significance)? Please could you clarify. How sensitive are the anomalies to this reference period vs. for example the 5 years pre-eruption? Also for the tree-ring reconstructions – are the volcanic anomalies robust if a different reference period is taken such as the preceding 5 years?

- RESPONSE: The anomalies are calculated with respect to the 845-859 CE period, which is the entire simulated period (Lines 193-194 and 931-933). The statistical tests are, however, performed with respect to the pre-volcanic period 845-852 CE (Lines 198-202 and 940-946). The reason is that we would like to identify the strength of the changes in temperature and precipitation caused by the Churchill eruption compared to the eruption-free period, which is the period 845-852 CE.
- Using a different period, for instance the pre-eruption period (845-852 CE), increases slightly the mean of the reference period, but the difference between global mean is rather minimal (Please see attached Fig. R1). Note that the analysis between the pre- and post-eruption temperature and precipitation is not affected by the choice of the period calculating the anomalies.

L278 – It would be useful to have the radiative forcing simulated by CESM for comparison.

- RESPONSE: To diagnose the effective radiative forcing (ERF) from CESM, we would need to run simulations with fixed SSTs, which is outside the scope of this work. For this reason we still only present the scaling-based ERF estimates in our revised manuscript.

Fig. 4 – Please introduce NVOLC in the main text. The ylabels are also inconsistent (change vs anomaly). A legend on panels (c) and (d) would also be useful.

- RESPONSE: Changes are incorporated: we have introduced NVOLC to the main text (now Line 207). The legend for panels (c) and (d) is now included and the ylabels have been standardised.

Fig. 5 – Please specify that (b) (c) (e) and (f) are simulated anomalies. Why is (d) in a different projection?
changes incorporated: “simulated” anomalies is highlighted in the titles and captions for Fig. 5 (b), (c), (e) and (f). The projection of (d) has been altered to match the other projections in this figure.

L521 – consider rephrasing to ‘climate model simulations run with/using estimates of the stratospheric aerosol’

- RESPONSE: change incorporated: rephrased to "climate model simulations using estimates of the stratospheric sulfate aerosol..” as suggested above (now Line 625).

L558 – 561: I think it would be useful here to also briefly discuss dynamically driven precipitation changes

- RESPONSE: Change incorporated: we have included a brief discussion in Lines 706-719 (it is brief to reflect the reviewers comment and because we do not see a consistent response in precipitation).

"In principle, two possible processes might lead to precipitation changes after an eruption: thermodynamic or dynamic affects. The direct thermodynamics effect is related to the Clausius–Clapeyron relationship, which predicts that the water-holding capacity of the atmosphere decreases by approximately 7% for every 1°C cooling (Held and Soden, 2006). Therefore, moisture changes associated with the 852/3 CE Churchill eruption would be expected to be in the order of ca. <5%. Some observational and modelling studies have, however, reported a reduction in global precipitation following explosive volcanic eruptions (e.g. Robock and Liu, 1994; Iles and Hegerl, 2014, 2015). Beyond the thermodynamic affects, volcanic eruptions may also generate hydroclimate anomalies through changes in large-scale ocean-atmosphere circulation, including shifts in the latitudinal position of the Intertropical Convergence Zone (ITCZ; Haywood et al., 2013; Colose et al., 2016), an anomalously positive NAO or Northern Annular Mode (e.g. Christiansen, 2008; Stenchikov et al., 2006; Raible et al., 2016), a poleward jet shift (Barnes et al., 2016), and/or a narrowing of the Hadley Circulation (Ménégoz et al., 2018). The potential hydroclimate response to a high latitude eruption is therefore complex, reflecting the multiple, combined, and interacting effects of direct radiative forcing, feedbacks in those through changes in ocean-atmosphere circulation, and internal stochastic variability.”

L725 – consider also adding the relationship for winter extratropical eruptions, which gives a slightly higher forcing of -1 Wm\(^{-2}\). More importantly, what does CESM simulate? Is the RF different to that based on this scaling?

- RESPONSE: We now provide the range of forcing efficiency in Appendix A (Lines 911-912):

"The scaling pre-factor may vary between -20.9 and -17.4 W m\(^{-2}\) depending on the eruption season."

- Furthermore, we include the full range of possible forcing (Lines 303-306, please refer to our response to the first comment) accounting for the uncertainty on both SAOD and on forcing efficiency due to season of eruption. Please see our reply to your comment on L278 above for the CESM ERF response.

L738 - Where does the scaling of 1.49 come from? Why not run the model with the prescribed optical properties?

- RESPONSE: We followed a similar procedure used by Zhong et al. (2018) to create CESM-mountable EVA forcing, which uses the scaling of 1.67 based on a set of
sensitivity experiments for the 1815 Tambora eruption. The value used here is derived based on sensitivity tests using the 1991 Pinatubo eruption by comparing the EVA-generated forcing and available CESM forcing (Amann et al. 2003), in order to attain similar atmospheric responses (vertical and surface mean temperatures, surface radiative balances) from both forcings. We have included more detail to explain this in Lines 928-931. Also please consider our response to the comment above about L173.

L741 – are there eruptions that have been excluded or are there none during this time?

- RESPONSE: An eruption occurring in the Southern Hemisphere in 853 CE is excluded in the simulations (Line 943).

L764 – 766 – it would be useful to include this in the main text so that it is clear the volcanic anomalies are not just with reference to 1961-1990

- RESPONSE: change incorporated as suggested: the main text has been updated to include the full detail (now Lines 210-212)

**Technical corrections**

L167 –insert ‘the’ in front of EVA_H. parameter --> parameters

- REPONSE: change incorporated: included “the model parameters” now in Line 174.

L267 - add ‘at 550 nm’

- REPONSE change incorporated as suggested, now Line 285.

L278 – 2019 --> 2020

- REPONSE: change incorporated as suggested, now Line 301.

L286 – add also injection height?

- REPONSE: change incorporated (“injection altitude” is added in to Line 313).

L504 – Northern Hemisphere --> NH

- REPONSE: change incorporated as suggested and all other occurrences have been checked and updated.

L719 – add ‘for Phase 4 of PMIP’

- REPONSE: change incorporated as suggested, now Line 904.

L728 – remove ‘simulation’

- REPONSE: change incorporated as suggested, now Line 914.

L742 – insert ‘A’

- REPONSE: change incorporated (although changed to plural rather than singular): it now states “Mann-Whitney U-tests were used..” in Line 940.

L789 – location --> locations
• REPONSE: change incorporated as suggested, now Line 989.

Appendix F – what do panel a and b refer to? Blue/red labels are inconsistent in caption vs. legend.

• REPONSE: changes made to incorporate comments: have removed mention of panel a and b (as these are not required here) and have altered the caption to ensure that all is consistent within this figure (now Lines 1000-1003).

Please check Mount vs. Mt throughout.

• RESPONSE: change incorporated: have adopted the common approach of removing ‘Mount’ or ‘Mt’ when referring to the eruption after the first use in paper. All other ‘Mt’ are standardised to ‘Mount’.

Yours sincerely,

Helen Mackay (on behalf of all authors)

Please also note the supplement to this comment: