Replies are in blue colour and italics

Review of “Evaluating the simulated radiative forcings, aerosol properties and stratospheric warmings from the 1963 Agung,1982 El Chichón and 1991 Mt Pinatubo volcanic aerosol clouds” by Sandip S. Dhomse et al.

Review #1 by Daniel Visioni

This article gives an overview of the results from the UM-UKCA model simulations of the three biggest volcanic eruptions of the 20th century, and compares against available datasets. All simulations are run following the design of ISA-MIP. In light of both CMIP6 and the release of the new generation of models, and also of ISA-MIP, of which this study is most likely the first showing results of the simulations described in Timmreck et al. (2018), I believe this study to be of great importance and a very good fit for ACP. I have some suggestions to improve the presentation of the results and the discussion in this paper before it can be published. After these minor comments are addressed, the study can certainly be published in ACP.

--> We thank Dr Visioni for these positive comments.

Some broad comments:

“Evaluation dataset” section: this section is a bit confused and hard to follow. I suggest a table for the supplementary (similar to Table 1), at least, that sums up all of this information, including columns for timespan, type of observation and link to the dataset.

--> Thank you for a very useful suggestion. We decided to add the suggested table into the main article (new Table 2) rather than the supplementary, so as to provide a summary of the important details for each observation dataset (wavelengths, data source) and to reference the papers and/or web-links to the individual datasets.

We have pasted below the Table 2 added to the dataset, the process also alerting us to correct some aspects of the text of the evaluation datasets section (see track-changes manuscript). For example, we improved the text re: the Pinatubo period in GloSSAC to read:

```
“For the Pinatubo period, GloSSAC is an updated version of the gap-
filled dataset described in \citet[][chapter 4]{SPARC2006},
combining SAGE II aerosol extinction (in the solar part of the
spectrum), with HALOE and CLAES aerosol extinction in the infra-red
\citep[see][]{Thomason2018}. For the period where the SAGE-II signal
was saturated (e.g. Thomason, 1992), GloSSAC applies an improved
gap-fill method in mid-latitudes, but in the tropics is still based
on the composite dataset from \citet[][pages 140-147]{SPARC2006},
combining with ground-based lidar measurements from Mauna Loa,
Hawaii (19.5°N, \citep{Barnes1997}), and after January 1992 also
with lidar measurements from Camaguey, Cuba \citep[23°N,
see][]{Antuna1996}).”
```

Table 2: Some important aspects of the evaluation dataset.

Aerosol property	Key Aspects
Global stratospheric sulphur burden	
High-resolution Infrared Radiation Sounder (HIRS)	1.1 Derived from HIRS measurements onboard NOAA-10, -11, -12 satellites 1.2 Aqueous sulphuric acid aerosol retrieval using 8.2μm and 12.5 μm HIRS water vapour channels (Baran et al., 1993). 1.3 Derived sulphur burden based on assumed aerosol composition of 75% weight aqueous sulphuric acid solution droplets. 1.4 Global sulphur burden dataset is digitized from Figure 3 of Baran and Foot (1994)
Stratospheric AOD ($sAOD_{550}$) and extinction (ext_{550}) at 550 nm	
The CMIP6-GloSSAC forcing dataset https://eosweb.larc.nasa.gov/project/glossac/glossac)	2.1 Pinatubo aerosol cloud primarily from SAGE-II, HALOE and CLAES observations (Thomason et al., 2018). 2.2 Improved Pinatubo gap-fill in mid-latitudes from combining SAGE-II with CLAES 2.3 Tropical Pinatubo gap-fill from combining SAGE-II with Mauna Loa lidar (SPARC, 2006). 2.4 El Chichon aerosol cloud mostly derived from high-latitude SAM-II data (64°N-82°N and 64°S-84°S) 2.5 Tropical and mid-latitude El Chichon cloud from combining SAM-II with lidar data from 5 aircraft surveys (13°N to 80°N) and from Hampton, Virginia (37°N).
The CMIP6-AER2D forcing dataset (ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/)	3.1 From 2D interactive stratospheric aerosol simulations (Arfeuille et al., 2014). 3.2 Primarily from the 8 major eruption clouds in 1850-1979 (29 in 1600-present dataset). 3.3 Additional minor eruption clouds from Stothers (1996) are also included.
The CMIP5-Sato forcing dataset (https://data.giss.nasa.gov/modelforce/strataer/)	4.1 NASA GISS observation-based forcing data for 1850-2012 ($sAOD_{550}$ only). 4.2 Satellite era, uses SAGE-I, SAM-II, SAGE-II and OSIRIS measurements. 4.3 Pre-satellite era uses syntheses of different measurements 4.4 Surface radiation measurement dataset for Agung (Dyer and Hicks, 1968) highly uncertain in the tropics (Stothers, 2001).
The CMIP5-Ammann dataset (ftp://ftp.ncdc.noaa.gov/pub/data/paleo/climate_forcing/volcanic_aerosols/ammann2003b_volcanics.txt)	5.1 Simple model-based dataset for 13 eruption clouds 1880-2000 ($sAOD_{550}$ only). 5.2 Based on parameterisation for meridional dispersion from tropical reservoir to mid-latitudes, determined by Brewer-Dobson circulation seasonal cycle. 5.3 12-month e-folding timescale for decay of tropical volcanic aerosol reservoir 5.4 Peak $sAOD_{550}$ for each eruption scaled to match aerosol loading from Stothers (1996); Hofmann and Rosen (1983b); Stenchikov et al. (1998), assuming Reff = 0.42 μm.
The post-Agung Lexington lidar dataset from Grams (1966) (see Supplementary Information)	6.1 694nm backscatter ratio profiles from Lexington, Massachusetts (42°N, 71°W Fiocco and Grams, 1964; Grams and Fiocco, 1967) (ext_{550} only) 6.2 1-km dataset for 66 lidar soundings (Jan 1964 to July 1965) in Table A1 of Grams (1966). 6.3 Backscatter ratio timeseries at 15km, 20km and 24km tabulated into ASCII file. 6.4 Conversion to ext_{550} using extinction-to-backscatter ratio from Jäger and Deshler (2003).
Vertical profile evolution of Effective Radius (Reff) and Surface Area Density (SAD)	
CMIP6-GloSSAC (Pinatubo and El Chichon) and CMIP6-AER2D (Agung)	7.1 SAD for Pinatubo and El Chichon aerosol clouds from GloSSAC, using SAGE-II 3-λ method 7.2 SAD for Agung aerosol cloud from 2D interactive stratospheric aerosol model simulations 7.3 Volume concentration for each cloud derived from same method 7.4 Effective radius from 3 times ratio of volume concentration to SAD
Vertical profile of tropical stratospheric temperature anomaly	
From ECMWF reanalysis data	8.1 Temperature anomaly based on difference from 5-year mean starting in year of eruption 8.2 T-anomalies for Pinatubo and El Chichón from the ERA-interim re-analysis (Dee et al., 2011) 8.3 For the Agung period, anomaly derived from the ERA-40 year dataset (Uppala et al., 2005)

Minor comments: Careful rewording

a) Supplementary: the reference is missing at line 4. In general, I suggest a more careful check of the grammar of the manuscript: some phrases seem to be written in haste, and it could make for a much more enjoyable read if the style was a bit easier to understand. I offer some examples below:

b) Lines 277-280: this phrase needs a bit of rewording, it's confusing.

c) This again confirms that the more SO2 injection leads to the faster particle growth, hence quicker removal within first few months after the eruption.

d) Line 288: "the" lower end.

e) Line 341: I think here you might be referring to the other Pitari et al. (2016) paper (Stratospheric Aerosols from Major Volcanic Eruptions: A Composition-Climate Model Study of the Aerosol Cloud Dispersal and e-folding Time) that discusses the effects of the QBO phase on the cloud dispersal.

--> We agree with the reviewer. Some of the sentences were confusing and had some grammatical errors. We apologise for this. We have had a careful read and worked on the flow of the manuscript. The reference has been corrected.

Lines 343-345: While true that both cited papers mention the low altitude of the aerosols formed after the Hudson eruption, both remark that indeed the effect of that eruption was clearly distinguishable from the one from Pinatubo. From the conclusions of Pitts and Thomason (1993): "Below 15 km, Cerro Hudson aerosols were transported poleward during September and remained a persistent feature beneath the vortex throughout the spring" I understand that the experiments shown in this paper are part of ISA-MIP and thus part of a strict protocol, but I would just not be so quick in dismissing the Hudson eruption, especially in explaining the differences shown in Fig. 2 against the CMIP6 database, that are much larger in the southern hemisphere (where the Hudson eruption had more effect). I would like to see this discussed a little bit more in the manuscript (and, as a curiosity, see how the results change if this eruption is included, but I'm not suggesting to the authors do that for this work).

--> We agree with the reviewer that we cannot dismiss the influence of Mt. Hudson eruption, and our wording was somewhat dismissive, hence we have reworded the sentence. As GloSSAC V2 data became available, we have updated Figure 2 and, as reviewer pointed out, differences are indeed significant. Hence, we expanded discussion about Mt. Hudson eruption.

Line 371: "the more SO2 is injected"? and then, "within the first few months"

--> We reworded it as:

"This again confirms that the more SO2 injection leads to the faster particle growth, hence quicker removal within the first few months after the eruption."

Line 375: the first three months

--> Done.

Additional comments added after the original upload of the above reply to reviewers

Shortly after uploading our replies to the reviewers (including AC1 above), and when finalising the revised manuscript, we discovered two subtle but important mistakes in the Python code used to generate the figures in the ACP-Discussions manuscript:

1) *Figure 4: Typo in the code used to calculate effective radius: assigned the accumulation-soluble mode H_2SO_4 mmr to accumulation-insoluble H_2SO_4 mmr.*

   A subtle typo in the code used to calculate effective radius caused an error in the initial assignment of modal H_2SO_4 component mass mixing ratios. The typo caused the calculation of total particle volume (PVOL) to double-count the accumulation-soluble mode H_2SO_4 mass mixing ratios (mmr), the accumulation-insoluble H_2SO_4 mode mmr not used in the calculations as a consequence. Specifically, the excerpt of code:

```
H2SO4_nucsol_mmr=nc_fid.variables['NUCLEATION_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_Aitsol_mmr=nc_fid.variables['AITKEN_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_accsol_mmr=nc_fid.variables['ACCUMULATION_MODE__SOL__H2SO4_MMR'][:]
H2SO4_corsol_mmr=nc_fid.variables['COARSE_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_accins_mmr=nc_fid.variables['ACCUMULATION_MODE__SOL__H2SO4_MMR'][:]
```

   should have been:

```
H2SO4_nucsol_mmr=nc_fid.variables['NUCLEATION_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_Aitsol_mmr=nc_fid.variables['AITKEN_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_accsol_mmr=nc_fid.variables['ACCUMULATION_MODE__SOL__H2SO4_MMR'][:]
H2SO4_corsol_mmr=nc_fid.variables['COARSE_MODE__SOLUBLE__H2SO4_MMR'][:]
H2SO4_accins_mmr=nc_fid.variables['ACCUMULATION_MODE__INS__H2SO4_MMR'][:]
```

   For the major volcanic aerosol cloud simulations analysed here, the majority of sulphuric acid mass is in that double-counted accumulation-soluble mode.

   *Hence the typo caused the particle volume PVOL, used in the calculation of the model's effective radius (=3*PVOL/SAREA) to be much higher.*

   This affected the effective radius values shown in Figures 4 a), b), d), e) in the ACPD article to be much higher than their true values.

2) *Figures 2d), 8d),11d): Error in sAOD calculated for CMIP6 dataset (depth error).*

   The stratospheric AOD values shown for the CMIP6 representations of the Pinatubo, El Chichon and Agung aerosol in Figures 2d, 8d, 11d of the ACPD article are factor 2 too high, due to an error in the depth used in the calculations.

   The depth error arises within our code to integrate to sAOD the altitude-resolved aerosol extinction dataset provided with the CMIP6 volcanic aerosol dataset. When calculating the sum over vertical levels, the depth assigned when integrating the aerosol extinction to stratospheric Aerosol Optical Depth (sAOD) was set at 1.0 km rather than 0.5 km, calculated sAOD was then a factor of 2 too high.

We apologise for both bugs. Whereas the error 1) was more subtle, and the Reff from the model being too high was not obvious, we should have realised that the CMIP6 sAOD values shown in those figures were a factor-2 too high. Even though there is no documentation paper for the pre-satellite part of the CMIP6 dataset (CMIP6-AER2D) sAOD, we should still have realised this error when preparing the manuscript. We are relieved to have found this error during the review process and in the revised manuscript, the Figures 2d, 8d and 11d show the correct $sAOD_{525}$ values.

The typo explained in 1) is now remedied, and the simulated Reff values shown in Figures 4c) to f) of the revised manuscript represent the model predictions correctly.

We also added the two extra panels requested by Reviewer 2 to additionally show the 20Tg simulation Reff field at 25km (Figure 4a) and 20km (Figure 4b).

Note that the Reff figure in the Supplementary Material (Figure S6) has also been updated since the ACPD article to show the correct values. There only the 10Tg and 20Tg runs are shown to match the runs used in Dhomse et al. (2014), enabling comparison with the corresponding figure in that paper.

With these changes to the simulated Reff values in Figure 4 and Figure S6, the Section 4.1 text analysing the Reff variations (lines 389-419 in the ACPD article) has been re-written to:

"Next, we evaluate the meridional, vertical and temporal variations in effective radius (Reff) in the Pinatubo UM-UKCA datasets. The particle size variations in these interactive simulations of the Pinatubo cloud reflect the chemical and microphysical processes resolved by the chemistry-aerosol module, in association with the stratospheric circulation and dynamics occurring in the general circulation model. We analyse these model-predicted size variations also alongside those in the benchmark observation-based Reff dataset from CMIP6-GloSSAC, which applies the 3-lamda size retrieval from the 453nm, 525nm and 1020nm aerosol extinction measurements from SAGE-II (Thomason et al., 1997a, 2018).

Figure 4 shows zonal mean Reff at 25km, within the altitude range of the volcanic SO_2 injection, and at 20km, underneath the main volcanic cloud, results shown from 3-member means from the 10, 14 and 20Tg SO_2 emission runs (Pin10, Pin14 and Pin20). For comparability with the equivalent Figure from Dhomse et al. (2014), we also show in the Supplementary Material (Figure S6) the updated comparison to the Bauman et al. (2003) Reff dataset, for the corresponding Pin10 and Pin20 runs. Overall, the model captures the general spatio-temporal progression in the Reff variations seen in the GloSSAC dataset. However, whereas the 10Tg and 14Tg simulations agree best with the HIRS-2 sulphur-burden (Figure 1) and the GloSSAC sAOD and extinction (Figures 2 and 3), the magnitude of the Reff enhancement is best captured in the 20Tg run (Pin20). The comparisons suggest the low bias in simulated Reff seen in the previous UM-UKCA Pinatubo study (Dhomse et al., 2014) continues to be the case here. However, this low-bias in particle size/growth may simply be reflecting the required downward-adjustment of the Pinatubo SO_2 emission, a larger Reff enhancement in the 20Tg simulation clearly apparent. It is possible that the two-moment modal aerosol dynamics in GLOMAP-mode may affect its predicted Reff enhancement. However, the model requirement for reduced SO_2 emission is attributed to likely be due to a missing, or poorly resolved, model loss pathway, such as accommodation onto co-emitted volcanic ash. The sustained presence of ash within the Pinatubo cloud (e.g. Winker and Osborne, 1992) will likely have altered particle size and growth rates in the initial months after the eruption.

In the tropics, where Reff increases are largest, the timeseries of Reff is noticeably different in the core of the tropical reservoir (10°S to 10°N) to that in the edge regions (10°N-20°N and 10°S-20°S), at both 20km and 25km. The Reff increases in these edge regions occur when tropics to mid-latitude transport is strongest, in phase with the seasonal cycle of the Brewer-Dobson circulation, which tends to transport air towards the winter pole (Butchart, 2014). The Reff increases are due primarily to particle growth from coagulation and condensation, and the simulations also illustrate how the simulated Pinatubo cloud comprises much smaller particles at 25km than at 20km. The 25km level is in the central part of the Pinatubo cloud, particles there being younger (and smaller), because the oxidation of emitted volcanic SO_2 that occurs at that level, triggers extensive new particle formation in the initial months after the eruption (e.g. Dhomse et al., 2014). By contrast, at the 20km level, particles there will almost exclusively have sedimented from the main cloud, and therefore be larger. There is a slow but sustained increase in average particle size in the equatorial core of the tropical Pinatubo cloud, with the 20km level reaching peak Reff values only during mid-1992, in contrast to the peak S-burden and $sAOD_{550}$, which have already peaked at this time, being in decay phase since the start of 1992 (see Figures 1 and 2).

Whereas the simulated peak Reff enhancement occurs by mid-1992 in the tropics, the peak Reff in NH mid-latitudes occurs at the time of peak meridional transport, the Reff variation there reflecting the seasonal cycle of the Brewer-Dobson circulation, as also seen in the tropical reservoir edge region. The different timing of the volcanic Reff enhancement in the tropics and mid-latitudes is important when interpreting or interpolating the in-situ measurement record from the post-Pinatubo OPC soundings from Laramie (Deshler et al., 2003). Russell et al. (1996) show the Reff values derived from Mauna Loa ground-based remote sensing are substantially larger than those from the dust-sonde measurements at Laramie. The interactive Pinatubo simulation here confirm this expected meridional gradient in effective radius, with the chemical, dynamical and microphysical processes also causing a vertical gradient in the tropical to mid-latitude Reff ratio. The current ISA-MIP activity (Timmreck et al., 2018) brings a potential opportunity to identify a consensus among interactive stratospheric aerosol models for the expected broad-scale spatio-temporal variations in uncertain volcanic aerosol metrics such as effective radius."

With the 1km-depth error in the integration of the CMIP6 aerosol extinction, and the subsequent correction to the sAOD shown for CMIP6-GloSSAC/CMIP6-AER2D in Figures 2d, 8d and 11d, there have also been some minor changes to interpret the evaluation of the UM-UKCA volcanic simulations. The revisions here are only minor changes in emphasis re: the comparisons to GloSSAC sAOD, and since the text is mainly analysing the sAOD variation, these changes are only minor.

The text changes for this CMIP6 sAOD correction are the lines 371-415, 562-583 and 623-631 of the revised (lines 446-491, 672-694 and 735-755 in a tracked change version) manuscript, corresponding to lines 316-357, 484-504 and 542-563 of the ACP Discussions article