



Interactive Stratospheric Aerosol models response to different amount and altitude of SO₂ injections during the 1991 Pinatubo eruption

Ilaria Quaglia¹, Claudia Timmreck², Ulrike Niemeier², Daniele Visioni³, Giovanni Pitari¹, Christoph Brühl⁴, Sandip Dhomse⁵, Henning Franke^{2,6}, Anton Laakso⁷, Graham Mann^{5,8}, Eugene Rozanov^{9,10}, and Timofei Sukhodolov^{9,10}

¹Department of Physical and Chemical Sciences, Università dell'Aquila, 67100 L'Aquila, Italy

²Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany

³Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, USA

⁴Max Planck Institute for Chemistry, Mainz, Germany

⁵School of Earth and Environment, University of Leeds, Leeds, U.K

⁶International Max Planck Research School on Earth System Modelling, Bundesstr. 53, 20146 Hamburg, Germany

⁷Finnish Meteorological Institute, Atmospheric Research Centre of Eastern Finland, 70200 Kuopio, Finland

⁸UK National Centre for Atmospheric Science, University of Leeds, Leeds, UK

⁹Physikalisch-Meteorologisches Observatorium Davos and World Radiation Center, Davos, Switzerland

¹⁰Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

Correspondence: Ilaria Quaglia (ilaria.quaglia@aquila.infn.it)

Abstract. Recent model inter-comparison studies highlighted model discrepancies in reproducing the climatic impacts of large explosive volcanic eruptions, calling into question the reliability of global aerosol model simulations for future scenarios. Here, we analyse the simulated evolution of the stratospheric aerosol plume following the well observed June 1991 Mt. Pinatubo eruption by six interactive stratospheric aerosol microphysics models in comparison to a range of observational data sets.

5 Our primary focus is on the uncertainties regarding initial SO₂ emission following the Pinatubo eruption in 1991, as prescribed in the Historical Eruptions SO₂ Emission Assessment experiments (HErSEA), in the framework of the Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP). Six global models with interactive aerosol microphysics took part in this study: ECHAM6-SALSA, EMAC, ECHAM5-HAM, SOCOL-AERv2, ULAQ-CCM and UM-UKCA. Model simulations are performed by varying SO₂ injection amount (ranging between 5 and 10 Tg-S), and the altitude of injection
10 (between 18-25 km).

We find that the common and main weakness among all the models is that they can not reproduce the persistence of the sulfate aerosols in the stratosphere. Most models show a stronger transport towards the extratropics in the northern hemisphere, at the expense of the observed tropical confinement, suggesting a much weaker subtropical barrier in all the models, that results in a shorter e-folding time compared to the observations. Moreover, the simulations in which more than 5 Tg-S of SO₂ are injected
15 show a large surface area density a few months after the eruption compared to the values measured in the tropics and the in-situ measurements over Laramie. This results in an overestimation of the number of particles globally during the build-up phase,



and an underestimation in the Southern Hemisphere, which draws attention to the importance of including processes as the ash injection and the eruption of Cerro Hudson.

1 Introduction

20 Large magnitude volcanic eruptions can emit sulfur dioxide (SO₂) and other gases directly into the stratosphere. An abrupt increase in stratospheric SO₂ creates a long-lived volcanic aerosol cloud that scatters incoming solar radiation, absorbs longwave radiation, and can affect the composition of the stratosphere.

Such volcanically induced enhancement of the stratospheric aerosol layer exerts strong direct effects on climate because they influence Earth radiation budget, cooling the surface via the reduced insolation (McCormick et al., 1995; Soden et al., 25 2002); they also show a range of indirect effects, due to the volcanic aerosol effects on stratospheric circulation, dynamics and chemistry (e.g., Robock et al., 2009; Timmreck et al., 2012; Kremser et al., 2016).

Here we investigate the evolution of the volcanic aerosol cloud after Mt. Pinatubo (15.1°N, 120.4° E) eruption in June 1991 by analysing co-ordinated simulations within the HErSEA (Historical Eruptions SO₂ Emission Assessment) experiments, in the frame of the Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP, Timmreck et al., 2018).

30 Initial injection rate estimates based on satellite observations indicated that the eruption of Mt. Pinatubo injected between 14-23 Tg of sulfur dioxide (SO₂) or 7 to 11.5 Tg sulfur (Guo et al., 2004) into the stratosphere between 20-30 km (Bluth et al., 1992) and it produced a sulfuric acid cloud that peaked at 7 Tg-S in September (Lambert et al., 1993; Baran and Foot, 1994).

Several model studies tried to reproduce the observation based sulfate loading. While some managed to reproduce it by injecting mean amount of observational SO₂ estimates (Niemeier and Timmreck, 2009; Toohey et al., 2011; Brühl et al., 35 2015), others had to inject an amount of SO₂ below the lower boundary of the observed one (Dhomse et al., 2014; Sheng et al., 2015; Mills et al., 2016). Other uncertainties pertain to the vertical extension of the SO₂ plume: lidar measurements made several weeks after the eruption showed the largest aerosol concentration between 20-25 km altitude with much lower amounts between 15-20 km (Winker and Osborn, 1992a, b; DeFoor et al., 1992).

Previous model intercomparison projects related to stratospheric aerosols and their impact on climate revealed large differences in the simulation of the aerosol radiative forcing. In the Model Intercomparison Project on the climatic response to 40 Volcanic forcing (VolMIP, Zanchettin et al., 2016), the aerosol optical properties (stratospheric aerosol extinction, single scattering albedo, and asymmetry factor) after large magnitude explosive volcanic eruptions are constrained across participating models and coupled ocean-atmosphere models are used to explore the causes of inter-model differences minimising the differences in the applied “volcanic forcing”. This approach for the Pinatubo-like eruption (Zanchettin et al., 2022) has been shown 45 to reduce discrepancies in reproducing surface air temperature and precipitation anomalies, while they may still be present in the radiative fluxes.

The pre-study experiment of VolMIP (VolMIP-Tambora ISA ensemble) simulated the volcanic aerosol cloud from the eruption of Mt. Tambora in 1815 by prescribing a set of injection parameters and was initially designed to create a common forcing data set for the VolMIP volc-long-eq experiment and then used to evaluate inter-model differences (Marshall et al., 2018;



50 Clyne et al., 2021). The Tambora eruption 1815 was strong, but measurements were not available. Marshall et al. (2018) tried to overcome this problem by comparing simulated sulfate deposition at polar regions to ice core data. The differences to ice core data and between the models are partially related to different meridional transport and stratosphere-troposphere exchange. Clyne et al. (2021) focused on the stratospheric global mean aerosol optical depth, showing that the main reason for discrepancies between models is due to differences in the effective radius, which affects the scattering efficiency and the amount of the
55 stratospheric aerosols.

The Geoengineering Model Intercomparison Project Phase 6 (GeoMIP6, Kravitz et al., 2015) also includes experiments where SO₂ in specific amounts is injected into the tropical stratosphere in multiple models to determine the surface impact. The amount of SO₂ required to achieve the proposed cooling varies by a factor of 2 between models and results in a different temporal and latitudinal distribution of aerosols that affects surface temperature and local precipitation differently (Visioni
60 et al., 2021).

Hence, ISA-MIP was developed to determine which set of initial conditions allows models to reproduce the available measurements, and understand how their different chemical and microphysical schemes, stratospheric dynamics, and radiative transfer treatment influence these choices. Specifically the ISA-MIP HErSEA focuses on the uncertainty in the initial volcanic emission in terms of amount and injection altitude of SO₂ for the recent large-magnitude volcanic eruptions in the last 100
65 years (Mt. Agung 1963, Mt. El Chichón 1982, Mt. Pinatubo 1991); multiple interactive stratospheric aerosol simulations of each of the volcanic aerosol clouds with common upper-, mid- and lower-estimate amounts and injection altitudes of sulfur dioxide were performed.

Here we investigate the evolution of the volcanic aerosol cloud after Mt. Pinatubo eruption by analysing co-ordinated simulations within the HErSEA framework. In contrast to the aforementioned model intercomparison studies, the HErSEA experiments offer a test of the reliability of these models by allowing a direct comparison of the simulated volcanic enhancement
70 of the stratospheric aerosol layer with observation data sets, especially during the Mt. Pinatubo eruption, for which several satellite and in-situ measurements are available. In particular, we ask whether previous results in inter-model differences are confirmed in this new MIP; the presence of multiple injection settings common between all models will also allow an exploration of the reason for these differences, based on the models abilities to reproduce observations with different sets of initial
75 conditions.

The experimental design, the main features of the participating models and the observational data sets are described in Section 2. Section 3 shows model results of the optical and microphysical properties of the volcanic aerosol cloud, which are summarised and discussed in Section 4.



2 Methods and Data

80 2.1 Methods

2.1.1 Experimental Protocol

The HErSEA Pinatubo experiment design includes five different emission scenarios considering different amounts and altitudes of injection of SO₂, as summarised in Figure 1. The first three emission scenarios describe injections at medium altitude (between 21–23 km) of an amount of SO₂ that varies from the lowest values of 5 Tg-S (Low-22km), to medium of 7 Tg-S
85 (Med-22km), to the highest of 10 Tg-S (High-22km). The medium injection scenario (7 Tg-S of SO₂) has three different injection altitude settings: Med-22km, as discussed, another shallow one at lower altitudes (18–20 km, Med-19km) and one over deep altitude-range (18–25 km, Med-18–25km).

The Mt. Pinatubo-like eruption is timed on June 15, 1991. SO₂ is injected in models in a single grid-cell close to the Pinatubo location (15°N, 120° E) and at the prescribed altitudes, with the precision given by the vertical and horizontal model resolution
90 (table S1). Differently from the protocol, SO₂ is injected in UM-UKCA at Mt. Pinatubo longitude and latitudes between 0 and 15°N. In EMAC, SO₂ is entered as a 3D-mixing ratio perturbation derived from satellite observations (Schallock et al., 2021).

All models are radiatively coupled to the volcanically enhanced stratospheric aerosol in order to resolve the composition–radiation–dynamics interactions. Previous model studies (e.g., Young et al., 1994; Timmreck et al., 1999; Aquila et al., 2012; Sukhodolov et al., 2018) showed that inclusion of the interaction between volcanic sulfate aerosol and radiation is
95 essential for a reliable simulation of the transport of the volcanic cloud.

We performed transient AMIP-type (Gates et al., 1999) runs of the Mt. Pinatubo eruption in which sea surface temperatures and sea ice extent are prescribed as monthly climatologies from the MetOffice Hadley Center Observational data set (Rayner et al., 2003). Boundary conditions are prescribed also for greenhouse gases and ozone depleting substances as recommended for the SPARC CCM1 hindcast scenario REF-C1SD (Eyring et al., 2013), in order to match those for the time period. All
100 groups except ULAQ-CCM submitted a 3-member ensemble for each different injection setting. Unless otherwise specified, all results shown refer to the ensemble mean.

2.1.2 Participating Models

The ISA-MIP multi-model ensemble includes simulations from six global aerosol models: ECHAM6-SALSA (Sect. 2.1.2.1), ECHAM5-HAM (Sect. 2.1.2.2), SOCOL-AERv2 (Sect. 2.1.2.3), ULAQ-CCM (Sect. 2.1.2.4), UM-UKCA (Sect. 2.1.2.5). In
105 addition closely related simulations from the EMAC model are considered (Sect. 2.1.2.6). The main characteristics of the participating models are reported in Table 1.

2.1.2.1 ECHAM6-SALSA

ECHAM6-SALSA (ECHAM6.3-HAM2.3-MOZ1.0) is an aerosol-chemistry-climate model. The host model is the ECHAM6.3 general circulation model (Stevens et al., 2013) which is interactively coupled with the HAMMOZ aerosol-chemistry model

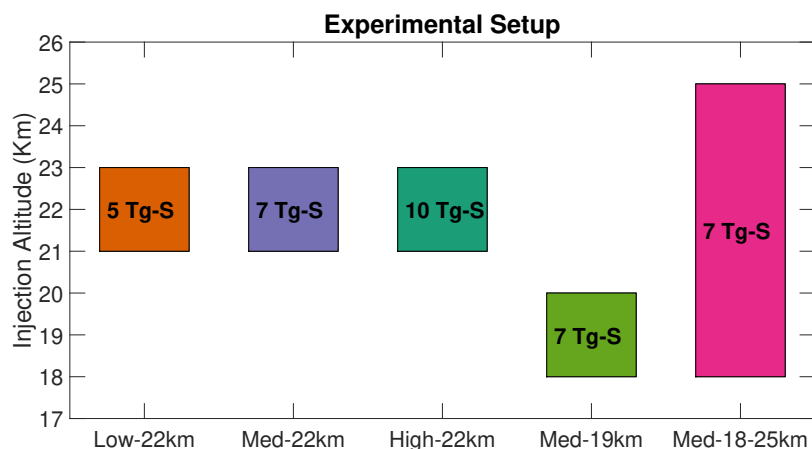


Figure 1. Graphical representation of injection setting parameters. The orange box represents the injection of 5 Tg-S of SO₂ centred at 22 km; the purple, green and violet boxes represent the injection of 7 Tg-S of SO₂ for two shallow injection altitudes centred at 22 km and 19 km, respectively, and one deep injection between 18 and 25 km; the teal box represents the injection of 5 Tg-S of SO₂ centred at 22 km.

110 (Schultz et al., 2018). HAMMOZ is a combination of the Hamburg Aerosol Model (HAM) and the Model for OZone And
Related chemical Tracers (MOZART) chemistry model. However MOZART was not used in the simulations of this study
and OH and ozone concentrations were prescribed by a monthly mean climatology and simplified sulfate chemistry scheme
of HAM was used. Aerosol model HAM calculates the emissions, removal, and radiative properties of aerosol. It simulates
five major global aerosol compounds: sulfate, organic carbon, black carbon, sea salt and mineral dust. The aerosol emissions
115 from anthropogenic sources were based on the Community Emission Data System (CEDS) for CMIP6 anthropogenic emission
inventory. Sea salt and dust emissions were calculated online. Aerosol microphysics were calculated by the sectional aerosol
module SALSA. A detailed description of the Model is given in Kokkola et al. (2018). SALSA describes aerosols using 10
size bins in size space and the seven largest bins are separated into externally mixed soluble and insoluble populations. T63L95
resolution was used in ECHAM6-SALSA simulations, which corresponds to an approximately 1.9°×1.9° horizontal grid and
120 95 vertical layers reaching up to 80 km. Quasi-biennial oscillation (QBO) was internally resolved by the model (Laakso et al.,
2022). Ensemble members were produced by using insignificantly different values for one of the tuning parameters (the rate of
snow formation by aggregation).

2.1.2.2 ECHAM5-HAM

The driving general circulation model (GCM) is ECHAM5 (Giorgetta et al., 2006) which was used as a high-top model in the
125 middle atmosphere (MA) version, with maximum altitude at 0.01 hPa (about 80 km), 90 levels and an interactive simulation of
the QBO. The horizontal resolution was about 2.8°, in a spectral truncation at wave number 42 (T42), with 91 vertical layers up
to 0.01 hPa. Interactively coupled to ECHAM is the aerosol microphysical model HAM (Stier et al., 2005), which calculates the
oxidation of sulfur and sulfate aerosol formation, including nucleation, accumulation, condensation, and coagulation processes.



The width of the HAM modes has been adapted to the conditions under a high sulfur load. The aerosols are prescribed in three
130 modes with a fixed width (Niemeier and Timmreck, 2009). HAM was further adopted to stratospheric conditions by applying
a simple stratospheric sulfur chemistry above the tropopause (Timmreck, 2001; Hommel et al., 2011). ECHAM prescribes
oxidant fields of OH, NO₂, and O₃ on a monthly basis, as well as photolysis rates of OCS, H₂SO₄, SO₂, SO₃, and O₃. The
sulfate was radiatively active for both SW and LW radiation and coupled to the radiation scheme of ECHAM. Further details
are described in Niemeier et al. (2021).

135 2.1.2.3 SOCOL-AERv2

SOCOL-AERv2 is an interactive aerosol-chemistry-climate model that is also based on the ECHAM5 GCM but coupled to the
MEZON chemistry (Egorova et al., 2003) and AER sulfate aerosol microphysics (Weisenstein et al., 1997) modules. The model
version used here has a horizontal resolution of about 2.8° in longitude and latitude (T42) and 39 vertical layers up to 0.01
hPa. The chemistry module calculates interactions of 89 chemical species of the oxygen, hydrogen, nitrogen, carbon, chlorine,
140 bromine, and sulfur groups in gas-phase, photolysis, and heterogeneous reactions, including reactions in/on aqueous sulfuric
acid aerosols. The sulfate aerosol module resolves the aerosol particles in 40 size bins, ranging in dry radius from 0.39 nm to
3.2 μm, and calculates nucleation, condensational growth, evaporation, coagulation, and sedimentation of sulfate aerosol bins.
H₂SO₄ weight percent is calculated online based on actual temperature and relative humidity. Dry and wet deposition of species
are interactively calculated based on actual meteorological conditions in the model. Modelled aerosols and chemical species
145 are coupled with the short-wave and long-wave radiation schemes. Aerosol radiative properties are treated following a look-
up-table approach with precalculated values using Mie theory for actual H₂SO₄ weight percent and temperature. All boundary
conditions follow the recommendations of ISA-MIP (Timmreck et al., 2018). Compared to other participating models, SOCOL-
AERv2 has the highest aerosol size resolution of 40 size bins, however a rather low vertical resolution (1.5 km in the lower
stratosphere). Because of the coarse vertical resolution, the QBO is also nudged to the observed equatorial wind profiles.

150 2.1.2.4 ULAQ-CCM

ULAQ-CCM (University of L'Aquila Chemistry Climate Model) is a global scale climate-chemistry coupled model extending
from the surface to the mesosphere (0.04 hPa) with 126 log pressure levels (approximate pressure altitude increment of 568 m)
and with a T21 horizontal resolution (5°x6°). The chemistry module includes medium and short-lived species (O_x, NO_y, NO_x,
CHO_x, Cl_y, Br_y, SO_x) and the major component of stratospheric and tropospheric aerosols (sulfate, nitrate, organic and black
155 carbon, soil dust, sea salt, PSCs). The microphysical code for aerosol formation and growth includes gas-particle conversion
scheme, homogeneous and heterogeneous nucleation, coagulation, condensation and evaporation (Pitari et al., 2002, 2016).
It also includes heterogeneous chemical reactions on sulfuric acid aerosols and polar stratospheric cloud particles; both het-
erogeneous and homogeneous upper tropospheric formation processes are also included (Visioni et al., 2018a). The aerosol
module calculates the aerosol extinction, asymmetry factor, and single scattering albedo, given the calculated size distribution
160 of the particles for different wavelength and they are passed daily to the radiative transfer module that is a two-stream delta-



Eddington approximation model (Toon et al., 1989). The QBO is nudged to climatological values (Morgenstern et al., 2017) and its future values are repeated from the historical time series.

2.1.2.5 UM-UKCA

UM-UKCA model simulations are performed using Global Atmosphere 4.0 configuration (Walters et al., 2014, GA4) of the
165 UK Met Office Unified Model (UM v8.4) general circulation model with the UK Chemistry and Aerosol chemistry–aerosol
sub-model (UKCA). The GA4 atmosphere model has a horizontal resolution of $1.875^{\circ} \times 1.25^{\circ}$ and 85 vertical levels (N96L85)
ranging from the surface to about 85 km. UM-UKCA configuration adapts GA4 with aerosol radiative effects from the in-
teractive GLOMAP aerosol microphysics scheme and ozone radiative effects from the whole-atmosphere chemistry that is a
combination of the detailed stratospheric chemistry and simplified tropospheric chemistry schemes (Archibald et al., 2020).
170 GLOMAP stratospheric aerosol microphysics scheme is described in Dhomse et al. (2014), and model setup is described
in Dhomse et al. (2020). Briefly, the model uses the GLOMAP aerosol microphysics module coupled with troposphere-
stratosphere chemistry scheme and modelled aerosols are coupled with the radiation scheme. Model also uses Greenhouse
gas (GHG) and ozone-depleting substance (ODS) concentrations from Ref-C1 scenario used in the CCMI-1 (Morgenstern
et al., 2017) activity. Simulations are performed in atmosphere-only mode, and we use CMIP6 recommended sea-surface tem-
175 peratures and sea-ice concentration that are obtained from <https://esgf-node.llnl.gov/projects/cmip6/> (last access: 25 March
2021).

2.1.2.6 EMAC

EMAC is the ECHAM5 general circulation model coupled with Modular Earth Submodel System Atmospheric Chemistry
(Brühl et al., 2015, 2018; Schallock et al., 2021). It contains comprehensive gas-phase and heterogeneous chemistry. The
180 resolution is T63/L90, i.e. about 1.9 degrees latitude and longitude and 90 layers up to about 80 km with a vertical res-
olution of about 500 m near the tropopause. Below 100 hPa dynamics and temperature are nudged to ERA-Interim. The
applied aerosol module GMXE (Pringle et al., 2010) accounts for seven modes using lognormal size distributions (nucle-
ation mode, soluble and insoluble Aitken, accumulation and coarse modes). The boundary between accumulation mode and
coarse mode, a model parameter, is set at a dry particle radius of $1.6 \mu\text{m}$ to avoid too fast sedimentation of a too large coarse
185 mode fraction in case of major volcanic eruptions. Optical properties for the types sulfate, dust, organic carbon and black
carbon (OC and BC), sea salt, and aerosol water are calculated using Mie-theory-based lookup tables consistent with the se-
lected size distribution widths of the modes. The resulting optical depths, single scattering albedos and asymmetry factors are
used in radiative transfer calculations which feedback to atmospheric dynamics. Volcanic SO_2 injections are entered as 3D-
mixing ratio perturbations derived from satellite data using an inventory for the period 1990 to 2019 (Schallock et al., 2021,
190 https://doi.org/10.26050/WDCC/SSIRC_3). For the Pinatubo period also the eruptions of Cerro Hudson, Spurr and Lascar are
included.



Table 1. Main chemical, microphysical and dynamic characteristics of the participating models.

Model	Injection region	Interactive OH	Stratospheric aerosol components	Aerosol dynamics scheme	Simulated aerosol in het. chem.	Nucleation scheme	QBO
ECHAM6-SALSA	Point	N	Sulfate, Dust, OC, BC and SS	2-moment sectional, 10 bins	N	Vehkamäki et al. (2002)	Internally generated
ECHAM5-HAM	Point	N	Sulfate	2-moment modal, 7 modes	N	Vehkamäki et al. (2002)	Internally generated
EMAC	3D-plume	Y	Sulfate, Dust, OC, BC, aerosol water	Modal, 7 modes	Y	Vehkamäki et al. (2002)	Internally generated but slightly nudged
SOCOL-AERv2	Point	Y	Sulfate	Sectional, 40 size bins	Y	Vehkamäki et al. (2002)	Nudged
ULAQ-CCM	Point	Y	Sulfate (also other components in troposphere)	Sectional, 22 bins	Y	Pitari et al. (1993)	Nudged
UM-UKCA	Band	Y	Sulfate and Meteoric Smoke particles	2-moment modal, 7 modes	N	Mann et al. (2010)	Internally generated

2.2 Observation data sets

2.2.1 AVHRR

The Advanced Very High Resolution Radiometer (AVHRR/2) is a space-borne sensor that measures the reflectance of the Earth in five spectral bands covering visible and infrared wavelengths (0.63, 0.86, 3.7, 11, 12 μm). AVHRR/2 instrument was on board of the polar-orbiting satellites (POES) NOAA-11 that provided global coverage data with a resolution of 1.1 km and a frequency of earth scans twice per day (https://www.avl.class.noaa.gov/release/data_available/avhrr/index.htm). The data used here are on a $1^\circ \times 1^\circ$ grid as monthly averages (as archived at NOAA's National ClimateData Center). As in Long and Stowe (1994) and Aquila et al. (2012), the stratospheric optical depth at 0.5 μm is calculated by removing monthly mean background values (June 1989 to May 1991) from the AVHRR observations. The optical depth at 0.5 μm is retrieved through a radiative transfer surface/atmosphere model (RAO et al., 1989) therefore, combined with the previous assumption, AVHRR can not detect the changes of stratospheric AOD smaller than 0.01 but can detect values up to 2.0 (Russell et al., 1996).

2.2.2 SAGE II

The Stratospheric Aerosol and Gas Experiment II (SAGE II) is a satellite-based sun photometer that was launched in October 1984 aboard the Earth Radiation Budget Satellite (ERBS) and retired in August 2005. The instrument measures the extinction of the solar radiation through the limb of the Earth's atmosphere in 7 channels ranging from 0.385 to 1.02 μm , with a global coverage from 80°S to 80°N latitude and a vertical resolution of 1 km for the retrieved data (Mauldin et al., 1985). We used the SAGE II version 7.0 (Damadeo et al., 2013; NASA/LARC/SD/ASDC, 2012b) that included, beside the primary data products, also the effective radius and the surface area density of aerosol particles, computed from the aerosol extinction at 525 and 1020 nm (Thomason and Burton, 2008). The stratospheric sulfate burden is taken from the SAGE-3 λ data set (ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/) that was compiled for phase 6 of the Coupled Model Intercomparison Project (CMIP6). SAD, mean radius, volume density and H₂SO₄ density are obtained from the extinction coefficients at 3 wavelength (452, 525 and 1024 nm) from SAM, SAGE I, SAGE II, CALIPSO, OSIRIS instrument and, for gap-filling following Pinatubo eruption, CLAES measurements (Revell et al., 2017).



215 2.2.3 HIRS

The High Resolution Infrared Radiation Sounder (HIRS) is an infrared scanning radiometer that has been onboard of several NOAA platforms starting with the first satellite of the Television Infrared Observation Satellite series (TIROS-N), followed by NOAA-6 up to NOAA-19 (Borbas and Menzel, 2021). It measures the reflectance of the earth in 19 infrared channels (3.7 to 15 μm) and one solar channel (0.69 μm) with a spatial resolution at nadir of 20.4 km on HIRS/2. Baran and Foot (1994) used HIRS/2 cloud-cleared radiances at 8.3 μm (NOAA-10/12) and 12.5 μm (NOAA-11) to retrieve the sulfuric acid aerosols from May 1991 to November 1993, assuming a composition for the stratospheric aerosols of 75% H_2SO_4 and 25% H_2O . The data cover the latitudes from 80°N to 80°S and all longitudes with 5° of resolution and are affected by a systematic error of 10% due to the sensitivity of the retrieved method and uncertainties in the background.

2.2.4 OPC

225 The University of Wyoming balloon-borne Optical Particle Counter (OPC) is a spectrometer that measures the light-extinction cross section of the particles using a broadband incandescent light source, developed by Rosen (1964). From the measurements, the number concentration, surface area and volume density are retrieved. The stratospheric aerosol measurements from 1991 to 2012 are made over Laramie (Wyoming) with the so-called OPC40, that can detect particles throughout the size range 0.1-10.0 μm , distinguished in 8 or 12 channels, depending on the instruments (Deshler, 2003). Here we used the revised data set (UWv2.0, http://www.atmos.uwyo.edu/deshler/Data/Aer_Meas_Wy_read_me.htm) of the OPC measurements (Kovilakam and Deshler, 2015; Deshler et al., 2019).

2.2.5 GloSSAC

The Global Space-based Stratospheric Aerosol Climatology (GloSSAC) is a global and gap-free data set of stratospheric aerosol optical properties (focused on aerosol extinction coefficient at 525 and 1020 nm) from 1976-2018. It is mainly based on the Aerosol and Gas Experiment (SAGE), and on the Optical Spectrograph and InfraRed Imager System (OSIRIS) and the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO). Ground, airborne and balloon-based instruments were used to fill major gaps in the data set (Thomason et al., 2018). Here, we used the updated version v2 (NASA/LARC/SD/ASDC, 2012a) from Kovilakam et al. (2020).

3 Results

240 The various sets of initial conditions of SO_2 injections result in an aerosol cloud with different optical properties depending on the dispersion of the cloud over time and the size of the aerosols produced.

In the following section, we start by analysing the aerosol optical depth (AOD) and how the models reproduce the measured AOD with different initial conditions. Since the amount of attenuation depends on the particle number concentrations and size, we then investigated both the magnitude and distribution of the sulfate burden and the size of the sulfate aerosols.



245 3.1 Aerosol optical depth

The stratospheric AOD simulated by the different interactive aerosol microphysical models is evaluated by comparing it with satellite observations from AVHRR and GloSSAC (Fig. 2). The AOD is calculated at a wavelength of 550 nm in EMAC, ECHAM5-HAM and ULAQ-CCM, 533 nm in ECHAM6-SALSA, 525 nm in SOCOL-AERv2 and GloSSAC, and 600 nm in AVHRR.

250 GloSSAC provides a coarse (horizontal resolution of about 550 km) but uniform spatio-temporal coverage up to the year 1994. As it is mostly based on SAGE II measurements, the instrument saturates for optical depth of about 0.15, therefore it is less accurate in the tropical core in the first months after the eruption (Russell et al., 1996). Conversely, AVHRR can only measure stratospheric AOD larger than 0.01. Because of the paucity of measures, "global values" when comparing against AVHRR are calculated between 60°S-60°N.

255 Figure 2 shows the time evolution of the zonal stratospheric AOD for each model and ensemble mean. It is clear that medium and high injection of SO₂ (Med-22km and High-22km, respectively) overestimate the stratospheric AOD in the tropics and in the Northern Hemisphere (NH) extratropics compared to both observations. The ability to reproduce the observed values in the Southern Hemisphere (SH) extratropics depends both on the model and the injection parameters. UM-UKCA and EMAC, contrary to other models, show more southward transport; this may be due to the different injection settings, band and 3D-
260 plume injection, respectively (see section 2.1.1) compared to the prescribed protocol. Moreover, EMAC also included smaller eruptions such as that of Cerro Hudson in the southern hemisphere in August 1991 (45.9°S, 72.9°W).

A quantitative comparison with the observations is shown with the use of Taylor diagrams (see Appendix A) in Figure 3. Model results are compared for the first year after the eruption with both AVHRR and GloSSAC (first row and second row, respectively) and for the second year only with GloSSAC.

265 Three-member ensembles, when provided, are represented with smaller circles of the same colour with respect to the ensemble mean of a specific simulation. In ECHAM6-SALSA, the differences between members of the same scenario are greater than those between scenarios because of differences in local winds at the time of the eruption in each ensemble-member. The impact of local winds is weaker when SO₂ is injected over the deep altitude-range between 19 and 25 km (pink circles in panels 3 (a) and (g)).

270 There are various sets of initial conditions that, depending on the model, are close to the observations. In the first year after the eruption, in ECHAM6-SALSA, SOCOL-AERv2 and ULAQ-CCM, the injection of 7 Tg-S of SO₂ is a good compromise between the too high and too low stratospheric AOD produced in the tropics by an injection of 5 and 10 Tg-S of SO₂, respectively. In ECHAM6-SALSA the deep injection between 18-25 km allows both southward and northward transport which otherwise would be too fast in the north hemisphere when injecting all the SO₂ in a few km near the tropopause. SOCOL-
275 AERv2 and ULAQ-CCM need to inject close to the tropopause in order to have a better poleward transport and a lower tropical AOD. Unlike other models, in ECHAM5-HAM 10 Tg-S of SO₂ results in a comparable stratospheric AOD both in the tropics and extratropics. In UM-UKCA, the different injection strategy produces stratospheric AOD values comparable with the observations in both hemispheres with only 5 Tg-S of SO₂



Table 2. Correlation (CORR) and root-mean-square-difference (RMSD) of the experiments that better reproduce the AVHRR and SAGE GloSSAC stratospheric AOD.

Model	AVHRR (June 91 - May 92)			GloSSAC (June 91 - May 92)			GloSSAC (June 92 - May 93)		
	Experiment	CORR	RMSD	Experiment	CORR	RMSD	Experiment	CORR	RMSD
ECHAM6-SALSA	Med-18-25km	0.74	0.07	Med-22km	0.70	0.09	Med-19km	0.79	0.02
ECHAM5-HAM	High-22km	0.74	0.08	High-22km	0.75	0.10	Med-19km	0.82	0.02
EMAC		0.79	0.06		0.55	0.10		0.67	0.03
SOCOL-AERv2	Med-19km	0.73	0.07	Med-19km	0.61	0.10	Med-19km	0.86	0.02
ULAQ-CCM	Med-19km	0.84	0.06	Med-19km	0.74	0.09	Med-19km	0.69	0.03
UM-UKCA	Low-22km	0.87	0.05	Med-22km	0.80	0.11	Low-22km	0.86	0.02

The best set of initial parameters also depends on the observation considered for comparison. In the comparison with GloS-
280 SAC during the first year after the eruption (second row of Fig. 3), ECHAM6-SALSA and UM-UKCA a different scenario
corresponds better to the observations. Moreover, for all models, correlations are smaller, ranging from 0.55 to 0.80 compared
to 0.73 to 0.87 of the comparison with AVHRR, and RMSD are larger (see Table 2).

After May 1992 (third row of Fig. 3), all models agree that Med-19km better reproduces the GloSSAC observations, with
the exception of Low-22km for UM-UKCA (for which no experiments with different injection altitudes are provided). For
285 these experiments, the global normalised stratospheric optical depth in Figure 4 shows that the decline of the volcanic cloud is
too fast in all models except UM-UKCA. The e-folding time (calculated as the time between the maximum and the 1/e value)
is 10 months in EMAC, 11 in ECHAM5-HAM and ECHAM6-SALSA, 12 in SOCOL-AERv2, and 14 in ULAQ-CCM and 20
in UM-UKCA, compared to 13 months in AVHRRa and 15 months in GloSSAC.

3.2 Sulfate Burden

290 Figure 5 shows the time evolution of the global and tropical stratospheric sulfate burden of different injection set-ups for each
model. The results of each model are compared with satellite measurements from HIRS and the SAGE-3 λ data set. Large
differences are evident in the temporal evolution of the sulfate burden between the aerosol model simulation on one hand and
the satellite data set on the other. In contrast to the AOD, the two satellite observations show similar values and a similar
temporal evolution for the sulfate burden. In general, the models show a larger peak and a faster decay than the satellite data.

295 In the six months following the eruption (July-December, termed the build-up phase), all models except EMAC best match
the global stratospheric sulfate burden of HIRS and SAGE-3 λ (first column of Fig. 5) with the injection 5 Tg-S of SO₂
(Low-22km). For SOCOL-AERv2, Med-19km also shows values within the uncertainties of the HIRS measurements. All
the experiments with larger amounts of injected SO₂, including the EMAC experiment with 8.5 Tg-S of SO₂, overestimate
the measured global sulfate burden. However, Low-22km underestimates the tropical burden (second column of Fig. 5) in



Stratospheric AOD

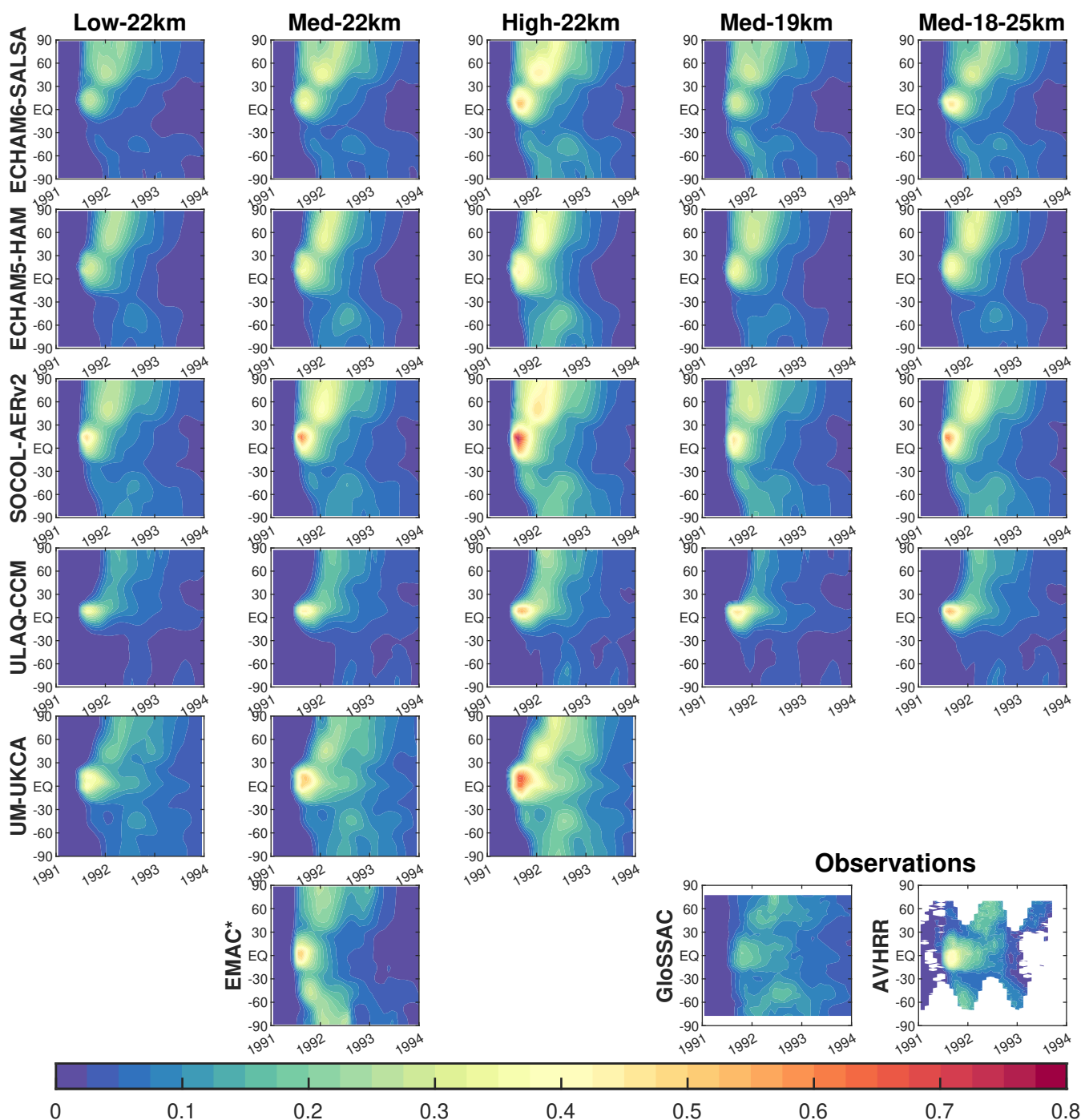


Figure 2. Time evolution of zonal stratospheric AOD for all models, in Low-22km (first column), Med-22km (second column), High-22km (third column), Med-19km (fourth column), Med-18-25km (fifth column). The last row includes the different scenario simulated by EMAC and the two observations used for comparison: GloSSAC and AVHRR. AOD is calculated at a wavelength of 550 nm in ECHAM5-HAM, EMAC and ULAQ-CCM, 533 nm in ECHAM6-SALSA, 525 nm in SOCOL-AERv2, 525 nm in GloSSAC, 600 nm in AVHRR.

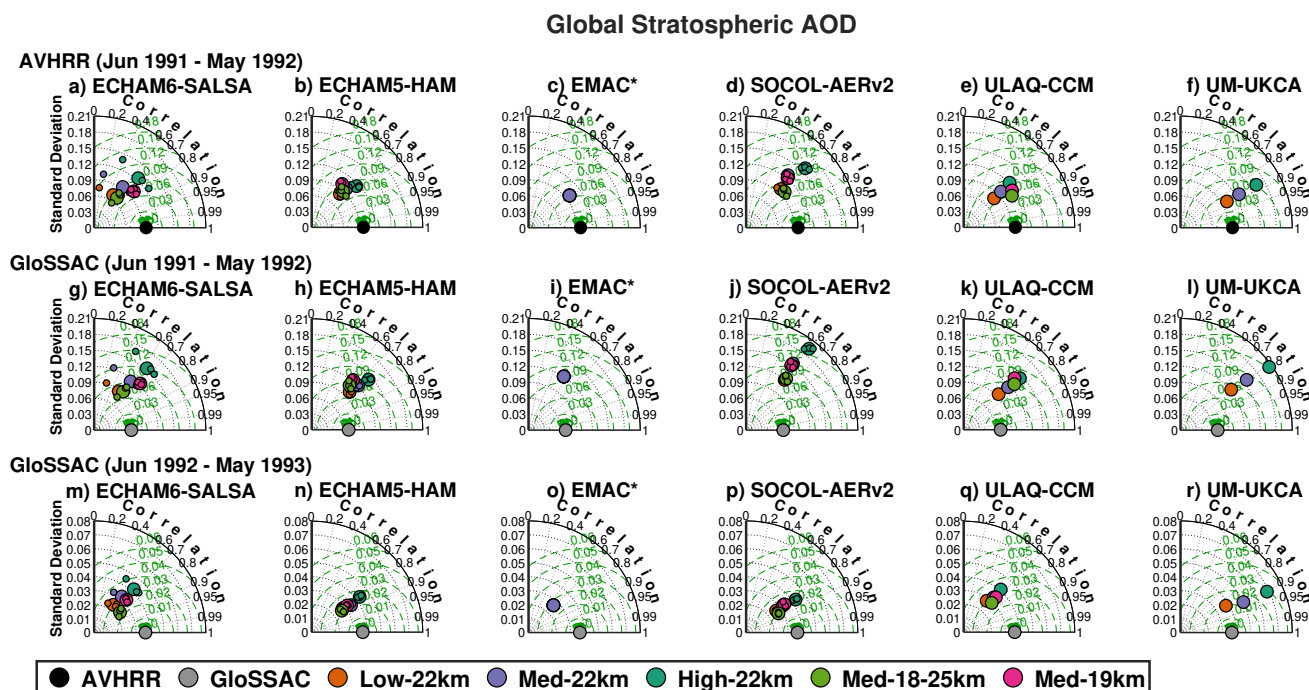


Figure 3. Taylor diagrams for the global stratospheric AOD. Zonal monthly mean values for different time periods have been used to calculate the standard deviation, correlation, centred root mean square difference between model experiments and measurements. Model results are compared in the 1st row with respect to AVHRR over the period June 1991 to May 1992; in the second row with respect to GloSSAC over the period June 1991 to May 1992; and in the third row with respect to GloSSAC over the period June 1992 to May 1993.

300 ECHAM6-SALSA, ECHAM5-HAM and SOCOL-AERv2 while in ULAQ-CCM the values are even higher than observed. Same happens in Med-19km. In general, all scenarios in ULAQ-CCM and the single scenario in EMAC overestimate the tropical burden, while ECHAM6-SALSA, ECHAM5-HAM and SOCOL-AERv2 overestimate the burden in the NH extratropics, for which Med-22km is able to reproduce tropical values.

In the build-up phase, SAGE-3 λ assumes the lowest values and slowly reaches a peak of 5.0 Tg-S in December, compared to 5.4 Tg-S of HIRS in September. Lower values in SAGE-3 λ are related to the saturation effects of the limb-occultation instrument, therefore HIRS measurements are to be considered more reliable for this initial period (Sukhodolov et al., 2018). For EMAC, the injection of 8.5 Tg-S of SO₂ produces a sulfate aerosol cloud that peaks in September at 7.0 Tg-S, a value comparable to the results of the Med-22km experiment (performed by the other models), in which 7 Tg-S of SO₂ is injected. For SOCOL-AERv2, Med-19km shows the best agreement with HIRS in terms of peak and timing of the peak (September),
 310 whereas in Low-22km and the other experiments it is reached one month later. This is followed by ECHAM6-SALSA in October (November only in High-22km) and ULAQ-CCM in November. ECHAM5-HAM is more sensitive to the altitude of injection: it peaks between October in Med-19km, November in Med-18-25km and December in the experiments with the



Global Normalised Stratospheric AOD

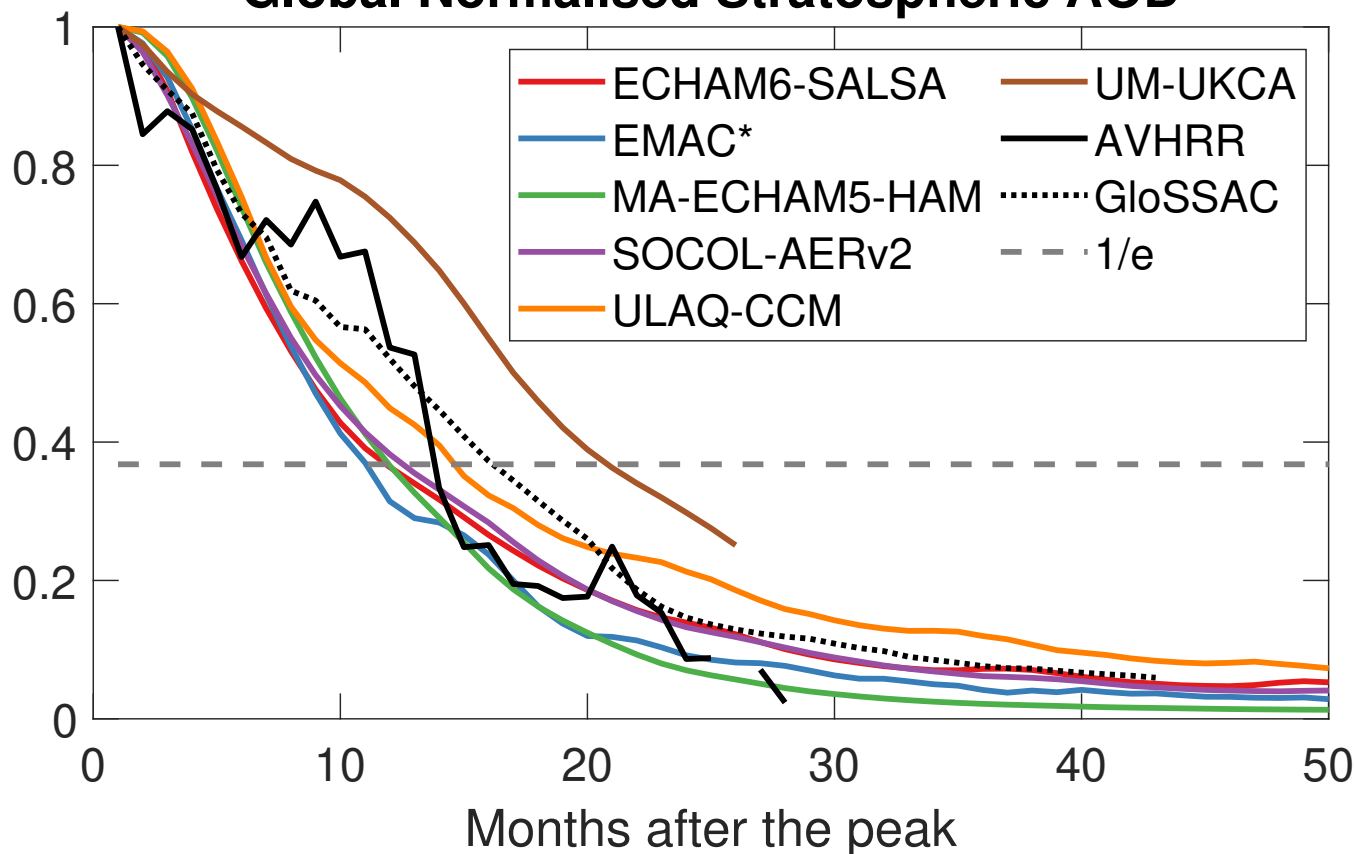


Figure 4. Time evolution of monthly values of the normalised global stratospheric AOD for models (colored lines) and AVHRR and GloSSAC observations (black lines). The dashed gray line represents the $1/e$ value. The experiment shown is Med-19km for ECHAM6-SALSA, ECHAM5-HAM, SOCOL-AERv2 and ULAQ-CCM, and Low-22km for UM-UKCA. For EMAC, it refers to the only experiment provided.

same altitude of injection (Low-21km, Med-21km and High-21km); the values of the peak is 14.3% lower in Med-19 km and 7.1% lower in Med-18-25km compared to Med-22km.

315 The sensitivity to injection altitude depends on the model: the Med-18-25km and Med-22km curves coincide in ECHAM6-SALSA and SOCOL-AERv2, and, compared to these experiments, the values in Med-19km are up to 9% and 20% smaller for each model, respectively. In ULAQ-CCM, as in ECHAM5-HAM, the more SO_2 is injected at lower altitudes the smaller is the value of the peak but for ULAQ-CCM the peak is 1% and 6% lower in Med-18-25km and Med-19km compared to Med-22km. Value and time of the peak for all models and experiments are summarised in table 2. In general, when the amount of SO_2
320 injected is exclusively or even in the lowest levels (Med-19km and Med-18-25km respectively), the sulfate burden is lower, and therefore this effect is less pronounced at Med-18-25km, as the aerosol distribution is more dependent on the balance between gravitational sedimentation in the lower stratosphere and the force of vertical transport by the BDC.



Differences among models and experiments in terms of amount and timing during the build-up phase are influenced by the oxidation of SO₂ by the hydroxyl radical (OH) that determines the timescale for aerosol formation (Clyne et al., 2021).
325 For this reason, we distinguish between models with prescribed OH (ECHAM6-SALSA and ECHAM5-HAM) and those with interactive OH (SOCOL-AERv2, ULAQ-CCM, UM-UKCA) when looking at the SO₂ evolution. The global normalised SO₂ burden curves (fig. S2 (a)) coincide for all models with prescribed OH. An exception of Med-19km in ECHAM6-SALSA, which has lower values, and this might depend on the different vertical concentrations of OH in the model. This indicates that, as expected, the stratospheric rate of SO₂ conversion is similar amongst these models, and differences in the subsequent
330 evolution of the stratospheric burden of SO₂ has to be due to differences in the removal processes. Due to the availability of only monthly values, some observations of the SO₂ behaviour at a finer resolved temporal scale are not possible here. In ULAQ-CCM, however, when comparing High-22km with Low-22km we find that an higher injection rate produce a longer initial e-folding time for SO₂. The same applies when comparing injections concentrated in a few kilometres (Med-22km and Med-19km), i.e. where SO₂ oxidation depletes OH more quickly (Mills et al., 2017), with those where the same amount of SO₂ is
335 injected over a wider altitude band. Consequently, initial values of the stratospheric sulfate burden in Med-18-25km are slightly higher compared to Med-22km and Med-19km. However, in the models, SO₂ lifetime depends on both OH concentration and transport and mixing into adjacent grid boxes. Therefore, when comparing different models, the timing of the peak cannot be simply related to the treatment of OH.

In order to better understand the models sensitivity to the different emission scenarios and eventual non-linearities, in Fig. 6
340 we normalise the resulting global sulfate burden by the amount of SO₂ injected: thus, in the build-up phase we would expect all the curves for all experiments to reach a value of 1, since no sulfate has yet been removed from the atmosphere. When the curves overlap, it means there are no differences in the global rate of production and removal of the aerosols. Looking at the subsequent phases, this will highlight differences in removal depending on the injection altitude and differences in microphysical growth.

345 Not all models and experiments however reach the value of 1: ECHAM5-HAM in Med-19km and Med-18-25km, ULAQ-CCM in Med-19km, and ECHAM6-SALSA and SOCOL-AERv2 in all experiments. This is due to the use of monthly means for our analyses, and because of gravitational settling that can quickly remove the formed particles, especially in Med-19km and Med-18-25km experiments, where the SO₂ is injected closer to the tropopause. To confirm this, we observe that this is particularly evident in Med-19km with the lowest injection height. The curves of the experiments with injection between 21-23
350 km coincide in the build-up phase and the differences emerge later, after 1992: the aerosol lifetime decreases by increasing the injection rate (table S2), which corresponds to the increase of the aerosol size in all models. In ECHAM6-SALSA the lifetime increases when increasing the injection rate. However, Figure 3 shows, the differences in results between ensemble members of the same scenarios are larger in ECHAM6-SALSA than in other models. This indicates that differences in aerosol lifetimes between Low-19km, Med-19km and High-19km scenarios are probably not statistically significant in ECHAM6-SALSA.

355 Among all models and experiments, the e-folding time of the global stratospheric sulfate burden ranges between a minimum value of 8 months in EMAC and maximum of 14 months in SOCOL-AERv2, while HIRS and SAGE-3λ have a longer e-folding time of 21 and 20 months, respectively. The e-folding time of the tropical stratospheric sulfate burden is 12 and 13



months in HIRS and SAGE-3 λ and half for the models with the exception of ECHAM5-HAM with a longer duration of 9 months. No model can reproduce the observed slow descent phase during 1992 of the stratospheric sulfate burden, and only the High-22km scenario approaches for all models the measured values at the end of 1992, while strongly overshooting them in the preceding months. Overall, we find that Low-22km and High-22km are the experiments that, in all models, better reproduce the observations in the build-up and descent phase, respectively (fig. 5, S3).

The spatial-temporal development of the sulfate burden (fig. S3) reflects in general the one of the AOD (Figs. 2, 3). In the SH the stratospheric burden shown in SAGE-3 λ is not reproduced by the models in Low-22km, therefore more SO₂ (High-22km) must be injected for the aerosol cloud to persist for as long as in SAGE-3 λ and reach the same values. In this way, however, the burden in the NH is overestimated. The absence of the simulations of Cerro Hudson eruption in August 1991 that injected about 0.75 Tg-S of SO₂ between 12 and 16 km (Barton et al., 1992; Saxena et al., 1995; Bluth et al., 1997), may partially explain the lack of sulfate aerosol in the SH that we observe for all model scenarios except EMAC, which included it.

There are clear differences in the peak location that is located between 5-20°N in the models, but around 5°S-10°N in the observations pointing to differences in the meridional transport in the early phase after the eruption. To discuss the meridional transport, Figure 7 shows the ratio of the simulated sulfate burden in the tropics (20°N-20°S), in the northern mid-latitudes (35°-60°N) and in the southern mid-latitudes (35°-60°S) with respect to the global value, for SAGE-3 λ (black line), for all models and scenarios (first row for the different injections amount, second row for the different injection altitudes).

Tropical confinement (panels 7a and d), as shown in the observations, is not captured by any model: ECHAM6-SALSA, ECHAM5-HAM, SOCOL-AERv2 and EMAC underestimate the tropical ratio, resulting in a stronger transport to the NH for the first three models and to the SH for EMAC, while ULAQ-CCM overestimates the ratio during the first six months after the eruption and becomes comparable thereafter. The different tropical confinement can be affected by a different vertical advection between ULAQ-CCM and the other models, based on the same dynamical core ECHAM5 or ECHAM6. Here, the tropical confinement is independent from the different horizontal resolution (Niemeier et al., 2020) while the particular definition of the tropical pipe (see Waugh et al., 2018) may also strongly affect this conclusion. The vertical resolution of a model can also affect the transport from the tropics to high northern latitudes: Brodowsky et al. (2021) showed for the SOCOL-AER model that a longer tropical confinement was found with increased vertical resolution.

The fraction of burden for the NH midlatitudes (panels 7c and f) is overestimated with differences of up to 20% compared to SAGE-3 λ (panel 7h), while for the SH (panels 7c and f) it is underestimated but to a smaller extent with differences of 10% compared to SAGE-3 λ (panel 7i). Overall, NH transport is favoured in all models at the expense of tropical confinement.

In most models, varying the injection rate does not affect the fraction of aerosols transported out of the tropics towards both hemispheres (panels 7a, b and c). This is consistent with the findings of Young et al. (1994) and Aquila et al. (2012) where the aerosol heating by absorption of the infrared radiation induces a lofting and a divergent motion that affects only the initial transport (within one month) of the aerosols towards both north and south tropics. However, the radiative impact of volcanic aerosol doesn't affect the transport from the tropics to the midlatitudes. The only exception is ECHAM6-SALSA, where an increased injection rate increases the tropical confinement, especially in the first six months after the eruption, especially in



the first six months after the eruption, probably due to a stronger radiative interaction from the absorption of more long-wave radiation by larger particles.

Differences emerge when SO₂ is injected at different altitudes. Labitzke and McCormick (1992), based on SAGE II measurements, show an upper transport regime (above 20 km) in which aerosols remain confined to the tropical reservoir spreading between 30°N and 10°S, and lower transport regime (below 20 km) in which aerosols mainly spread to northern high latitudes. Between August and September, aerosols above 20 km spread across most of the SH, reaching latitudes of 50°S, followed in November and December by an enhancement in the NH due to the transition from summer to winter circulation in the middle and upper stratosphere. Therefore, the meridional transport in the models depends on the vertical wind structure and on the vertical distribution of the simulated volcanic cloud in the first months after the eruption, and a faster transport in the NH is favoured when aerosol are mainly distributed in the lower transport regime (Timmreck et al., 1999).

All models, except ULAQ-CCM, show that the tropical confinement is reduced in favour of transport towards both hemispheres when SO₂ is injected below 20 km (Med-19km). Compared to high altitude injection settings (>20 km), Med-19km has the greatest transport in SH. The increase of altitude of injection (Med-22km and Med-18-25km) produces a higher confinement in the tropics with a consequent reduced transport toward both hemispheres in ECHAM6-SALSA and SOCOL-AERv2. In ECHAM5-HAM, the strongest confinement is achieved in Med-22km, while Med-18-25km shows a similar behaviour to Med-19km as most of the sulfate aerosols found below 20 km. In ULAQ-CCM differences among the injection settings emerge six months after the eruption and the injection at lower altitudes (Med-19km) shows a more efficient polewards transport, especially towards the NH.

3.3 Effective radius and Surface area density

Figure 8 shows the time evolution of the observed and simulated stratospheric effective radius in the tropics (20°S-20°N) and over Laramie (41°N-105° W). In the tropics (first row) the stratospheric effective radius is calculated as the average between 21-27 km because of a paucity of tropical measures below 21 km in SAGE II. Over Laramie (second row), the stratospheric effective radius is defined as the average between 14-30 km in order to compare it with in-situ OPC measurements (Deshler et al., 2019). Models results are calculated as the value of the nearest grid-cell to Laramie; therefore, the ability to reproduce the OPC measurements is more influenced by atmospheric circulation patterns and depends also on the horizontal resolution (see Table S1).

Before the eruption, the simulated evolution of the tropical mean effective radius is almost flat compared to SAGE II. Only ULAQ-CCM reproduces the observed seasonal variation and matches the pre-eruption measurements, resulting in particles with a radius of 0.27 μm, as for SAGE II (calculated over the 5 months before the eruption). The other models have smaller background particles with a constant value of 0.14 in ECHAM6-SALSA, 0.17 in ECHAM5-HAM, 0.17 in EMAC, 0.15 in SOCOL-AERv2 and 0.09 in UM-UKCA. Over Laramie, ECHAM6-SALSA, ECHAM5-HAM, EMAC and SOCOL-AERv2 have comparable radii to the OPC ones, while ULAQ-CCM and UM-UKCA lay outside the uncertainty range with larger and smaller radii, respectively. The causes of these differences are unclear; however, an in-depth exploration of the background



Stratospheric Sulfate Burden

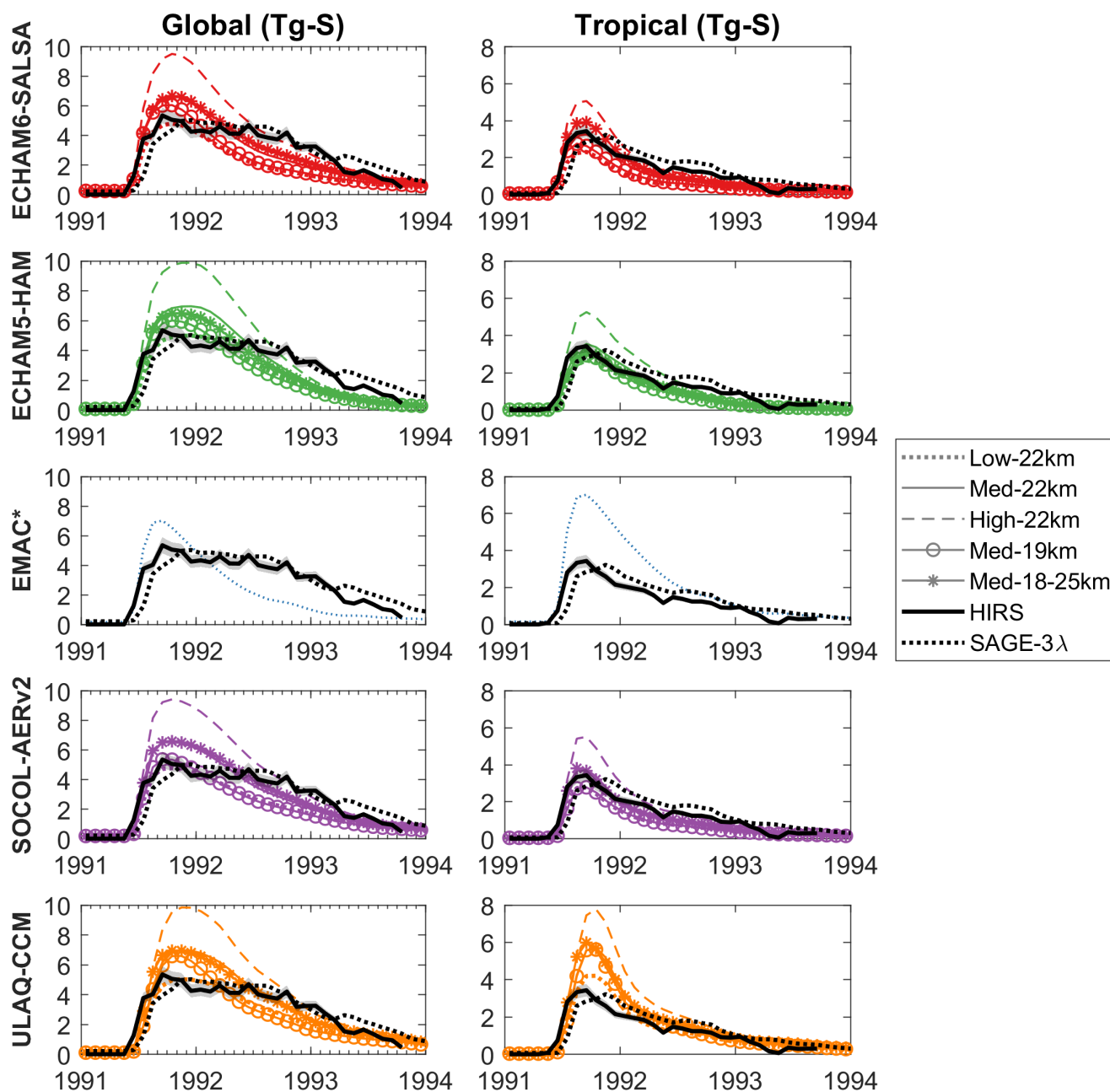


Figure 5. Time evolution of monthly values of global and tropical stratospheric sulfate burden in Tg-S (first and second column, respectively). Each panel refers to the respective model in which the different results of the experiments (coloured lines; different line styles for different experiments, see legend on the left) are compared with the HIRS and SAGE-3 λ data sets (black lines, see legend on the right).



Stratospheric Sulfate Burden normalised to the injected

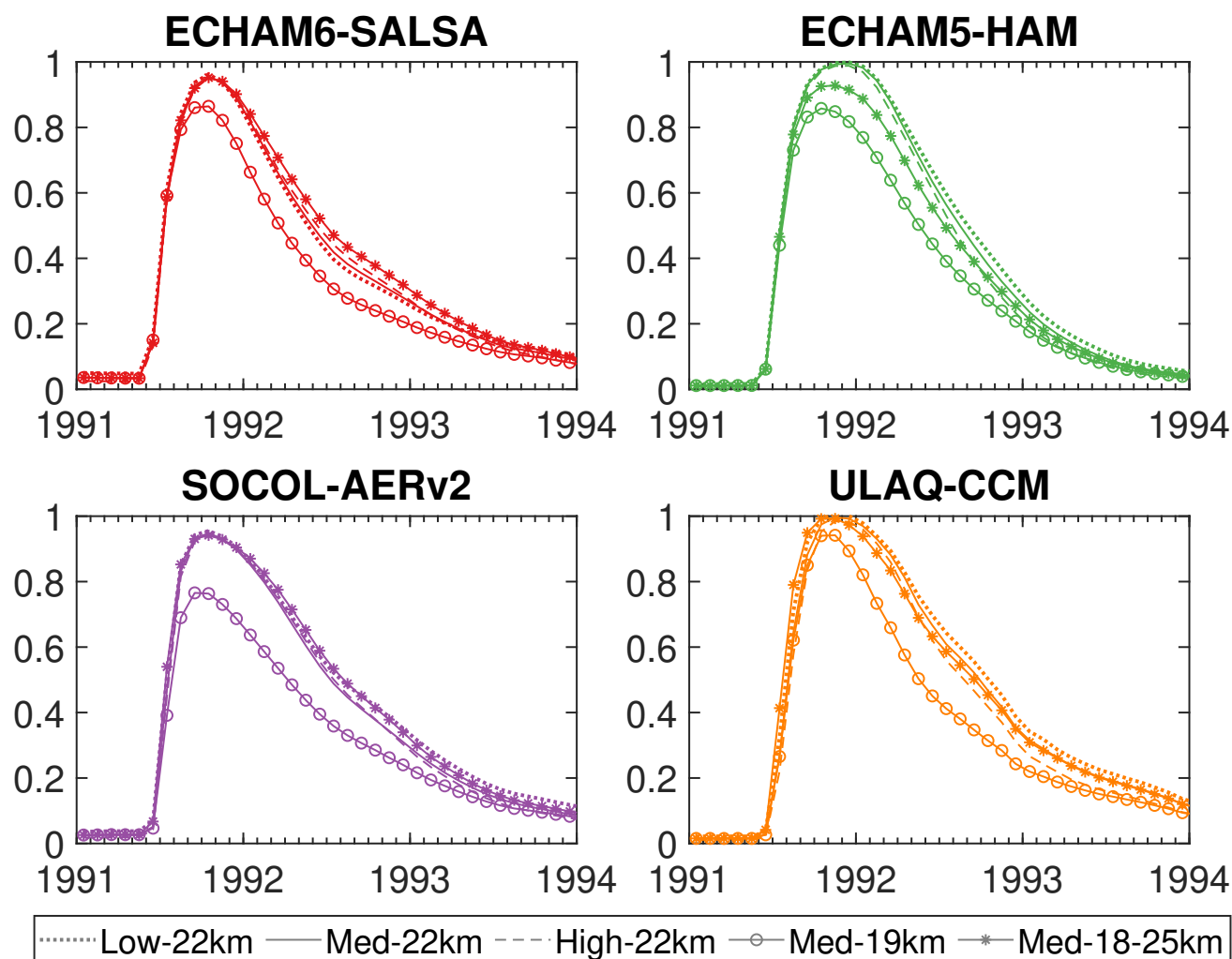


Figure 6. Time evolution of global stratospheric sulfate burden normalised to the amount of injected SO_2 . Each panel refers to the respective model in which the different experiments are compared.

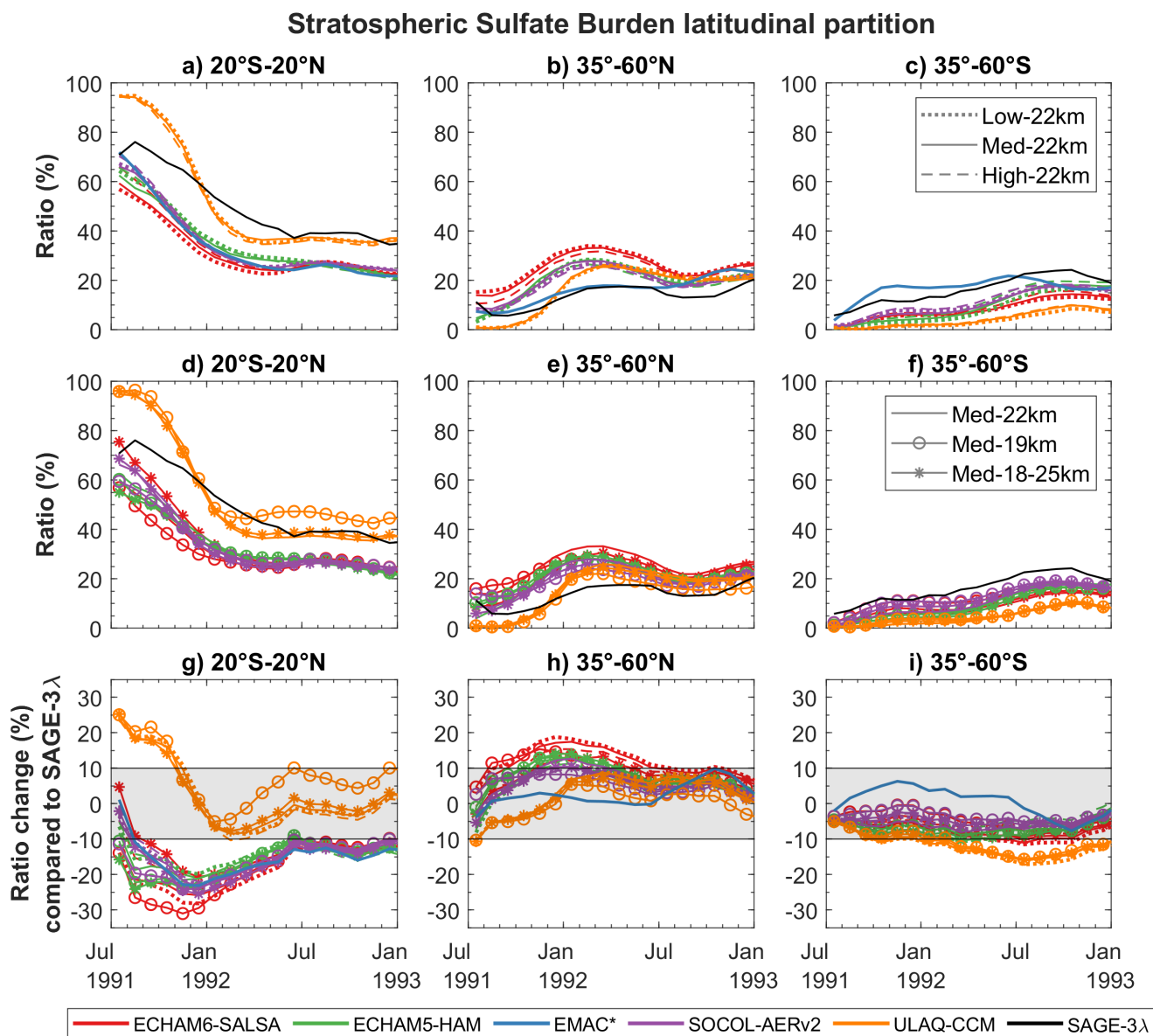


Figure 7. Time evolution of the latitudinal partition of the stratospheric sulfate burden ratio. The ratio is calculated with respect to the global average burden, for the tropical average (20°N-20°S, first column, a, d, g), the average over the northern mid-latitudes (35°-60°N, second column, b, e, h) and the average over the southern mid-latitudes (35°-60°S, third column, c, f, i). The first row includes the experiments with different injection rate, the second row experiments with different injection altitudes. The third shows the percentage change of the latitudinal partition for all model experiments compared to SAGE-3λ.



425 behaviour is out of this experiment and could be better understood by studies aimed at the analysing the Background (BG) experiment, which focuses on microphysics and transport under volcanically quiescent conditions (Timmreck et al., 2018).

After the eruption, all models, except EMAC, are able to capture the same decay rate as the SAGE II measurements and produce a comparable tropical effective radius for approximately a couple of years (with different injection settings), since the simulated decay onset time is anticipated. The timing of the peak is also similar and is reached between November and Decem-
430 ber, months from which the effective radius remains flat around the peak value. The models agree that particle size increases with increasing the injection rate, with differences from the medium injection scenario within 15% in ECHAM6-SALSA and 10% in ECHAM5-HAM, SOCOL-AERv2, ULAQ-CCM and UM-UKCA. The differences are larger when comparing different injection altitude scenarios and corresponding increase of the particle size is model-dependent. In ECHAM6-SALSA and SOCOL-AERv2, High-22km shows a tropical stratospheric effective radius within 10% of SAGE II until the end of 1993,
435 peaking, respectively at 0.47 and 0.49 μm compared to 0.51 in SAGE II. In ECHAM5-HAM, all experiments except High-22km, which fits best the observed AOD (see Section 3.1), produce similar effective radii, ranging between 0.46 and 0.51 μm , and are comparable with SAGE II until the end of 1992. High-22km differs by larger radii reaching a maximum of 0.56 μm . One year after the eruption, the differences among the different ECHAM5-HAM experiments disappear and the effective radius decreases more rapidly than in SAGE II. In EMAC, the effective radius grows more rapidly compared to other models
440 and the observations, peaking at 0.48 in August, and shows the fastest decay. In ULAQ-CCM, the effective radius of Med-19km reproduces the SAGE II measurements with a similar time decrease, as differences stay within 10% until the end of 1995, while other experiments produce larger particles, with peaks ranging between 0.53 and 0.71 μm . UM-UKCA shows the slowest decay that is comparable to SAGE II, but with smaller radii ranging between 0.35 and 0.45 after reaching the peak.

Over Laramie, all experiments of ECHAM6-SALSA and SOCOL-AERv2 produce radii within the estimated uncertainties
445 of the OPC measurements for all five years considered. However, only in SOCOL-AERv2 the month of the peak agrees with the observations, between April and May 1992, whereas in the other models it is 2 to 5 months earlier. ECHAM5-HAM and EMAC show comparable values during the pre-eruption phase but during the build-up phase radii rise faster compared to the observation and other models. In the descent phase, in ECHAM5-HAM the effective radius returns to values comparable with the measurements, while in EMAC, after reaching a peak that is about 30% smaller than the OPC one, the radii assume
450 smaller values, below the uncertainty. In ULAQ-CCM, all the experiments overestimate the OPC measurements and, here, Med-19km shows larger particles in the first months after the eruption, peaking at 0.78 μm in November 1991. In UM-UKCA, the pre-eruption values are underestimated and become comparable for all scenarios three months after the eruption.

Figure 9 summarises the information regarding the vertical distribution of the effective radius, SAD and extinction at 0.5
455 μm for the Med-22km experiment, in the tropical area (20°S-20°N) and over Laramie, six months after the eruption. A corresponding figure including all available experiments is shown in Figure S6. Looking to the vertical profiles of various quantities, biases that are hidden in integrated variables emerge. Figure 9 reveals that the vertical profiles differ not only between models and observations but also strongly between the observations themselves.

In the tropics, the effective radius peaks between 100-50 hPa in ECHAM6-SALSA, EMAC and ULAQ-CCM and between 50-20 hPa in ECHAM5-HAM and UM-UKCA as in SAGE II, with values within 30% of that measured, except for ULAQ-



460 CCM where the radii are up to 4 times larger. SOCOL-AERv2 shows good agreement with SAGE II between 100-20 hPa with values that remain constant around $0.44 \mu\text{m}$ above 70 hPa. The tropical SAD simulated by the models follows the same vertical distribution as that of SAGE II, and all models have a peak between 50-20 hPa, with the exception of EMAC whose peak is around 50 hPa. However, in that range of altitudes the values of the SAD is comparable with the observations for SOCOL-AERv2 and ULAQ-CCM for most of the altitudes, and is up to 2 times larger in the other models.

465 The tropical extinction follows the same distribution of the SAD. In this case, the extinction is compared with SAGE II and GloSSAC and large differences exist between them: below 20 hPa the extinction in GloSSAC is larger than in SAGE II and the differences increase with decreasing height up to 100% compared to SAGE II because of its gap-filling with ground-based measurements (Thomason et al., 2018; Kovilakam et al., 2020). Above 70 hPa, around the lower bound of the injection altitude, models extinction is even larger than GloSSAC: ECHAM6-SALSA, SOCOL-AERv2 and ULAQ-CCM approaches
470 the measurements at limit of maximum of uncertainty around 70-25 hPa, EMAC between 40-20 hPa, while ECHAM5-HAM and UM-UKCA overestimate measurements up to double their value. Below 70 hPa, all models underestimate the GloSSAC data, but the models extinction is still larger than that of SAGE II, with the exception of EMAC, which shows the greatest extinction below 50 hPa, where it peaks. Considering that the SAD depends on the size and the number of particles, we can assume, for the models that show a comparable radius and a larger SAD compared to SAGE II in the tropics, that they
475 overestimate the number of optically active particles and therefore show a larger extinction (ECHAM5-HAM, UM-UKCA).

Over Laramie, the vertical distribution of the effective radius is within the errorbar of the OPC measurements in ECHAM6-SALSA, ECHAM5-HAM, EMAC and SOCOL-AERv2 up to 20 hPa, while ULAQ-CCM produces larger particles especially below 50 hPa. Only EMAC is able to reproduce the vertical profile of the SAD from OPC measurements in terms of the position of the maximum and values: the models that showed faster transport in the northern mid-latitudes overestimate the
480 observations for most of the altitudes.

The ability to reproduce the observations also depends on the period considered (fig. S4 and S5): in the first months after the eruption models and observations show large differences, especially for SAD and extinction, which are overestimated in both the latitudes considered. This may be related both to the sensitivity to the actual meteorological conditions that climate models are unable to accurately replicate, and to the absence in HERSEA simulations of volcanic ash injection that could remove some
485 of the initial SO_2 gas or affect the local winds and the SO_2 dispersion (Ayris et al., 2013; Zhu et al., 2020; Dhomse et al., 2020; Kloss et al., 2021; Niemeier et al., 2021). This sensitivity to the initial conditions decreases the more time passes after the eruption. One year after the eruption, the models still show a vertical profile of the effective radius comparable to observations, while the simulated SAD starts to decrease everywhere after six months from the eruption, underestimating tropical values but still overestimating OPC measurements.

490 4 Discussion

With the use of Taylor diagrams, we find the experiments that better match the observations in terms of stratospheric AOD, in two different time periods, based on the reliability of the measurements. Each model requires different injection scenarios

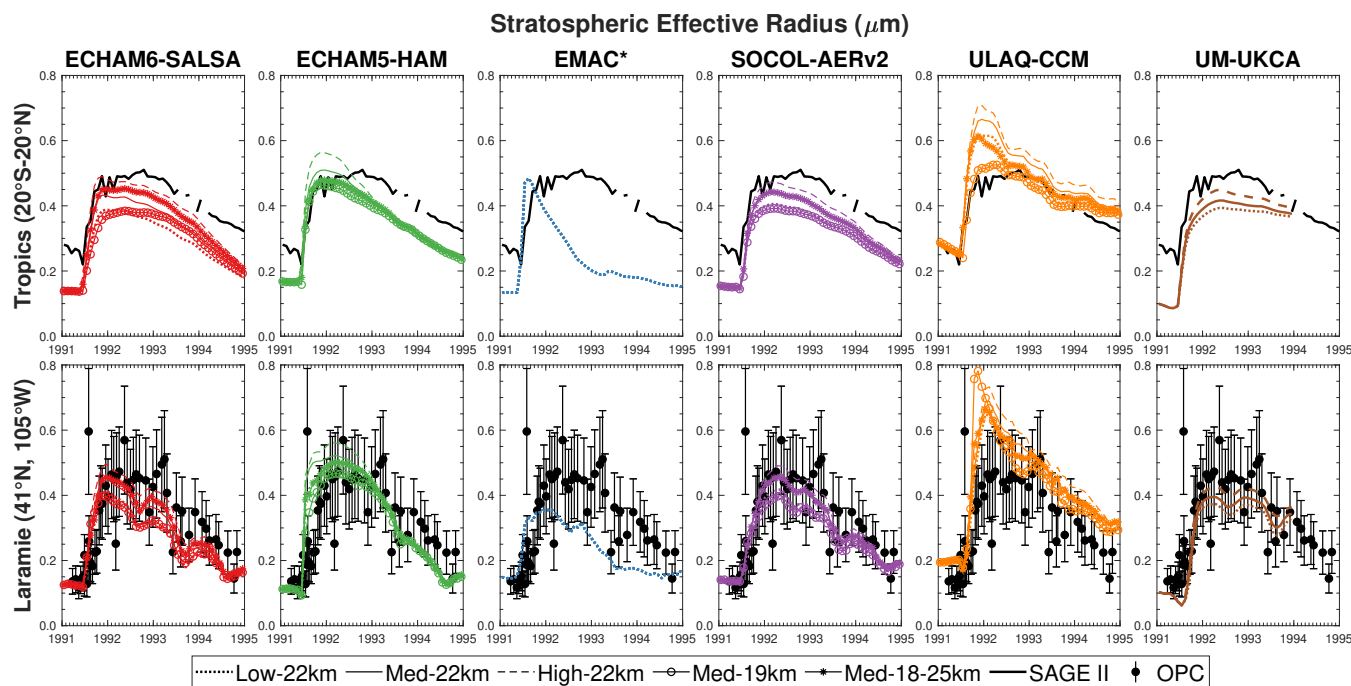


Figure 8. Time evolution of stratospheric effective radius (μm) in the tropics (first row) and over Laramie (41°N , 105°W , second row). In the panels of the first row, the effective stratospheric radius of the models is calculated between 21–27 km (50–20 hPa) to be compared with the available SAGE II observations. In the panels of the second row, it is calculated between 14–30 km (130–30 hPa) to be compared with the OPC observations.

to reproduce the observations, due to differences in the transport and microphysical mechanism and their mutual interaction. Even considering the best set of initial parameters based on AOD, differences with observations more or less persist in the models, and we can not unequivocally define a “best” model as that varies depending on the variable considered and the timing of the observation.

Comparing the results of the models between the experiments with the same injection setup (not shown), we observe a large difference between models in reproducing the stratospheric optical depth compared to the similar evolution of the global stratospheric sulfate burden. It is hard to disentangle the transport and the microphysics contribution on the differences in the considered variables, i.e. what fraction of it depends on microphysical schemes or different dispersion of the aerosol cloud. However, we find a common problem in transport, either too fast from the tropics to high northern latitudes (ECHAM6-SALSA, ECHAM5-HAM, SOCOL-AERv2) or too confined to the tropics (ULAQ-CCM), and strong removal processes such as sedimentation and transport through the tropopause, as shown in the comparison of the time evolution of the stratospheric burden between models and observations in the descent phase. Moreover, models show a different response and sensitivity to the altitude of injection that identify the main source of inter-model variability in the transport of sulfur gases and aerosol that consequently affects their formation and growth.

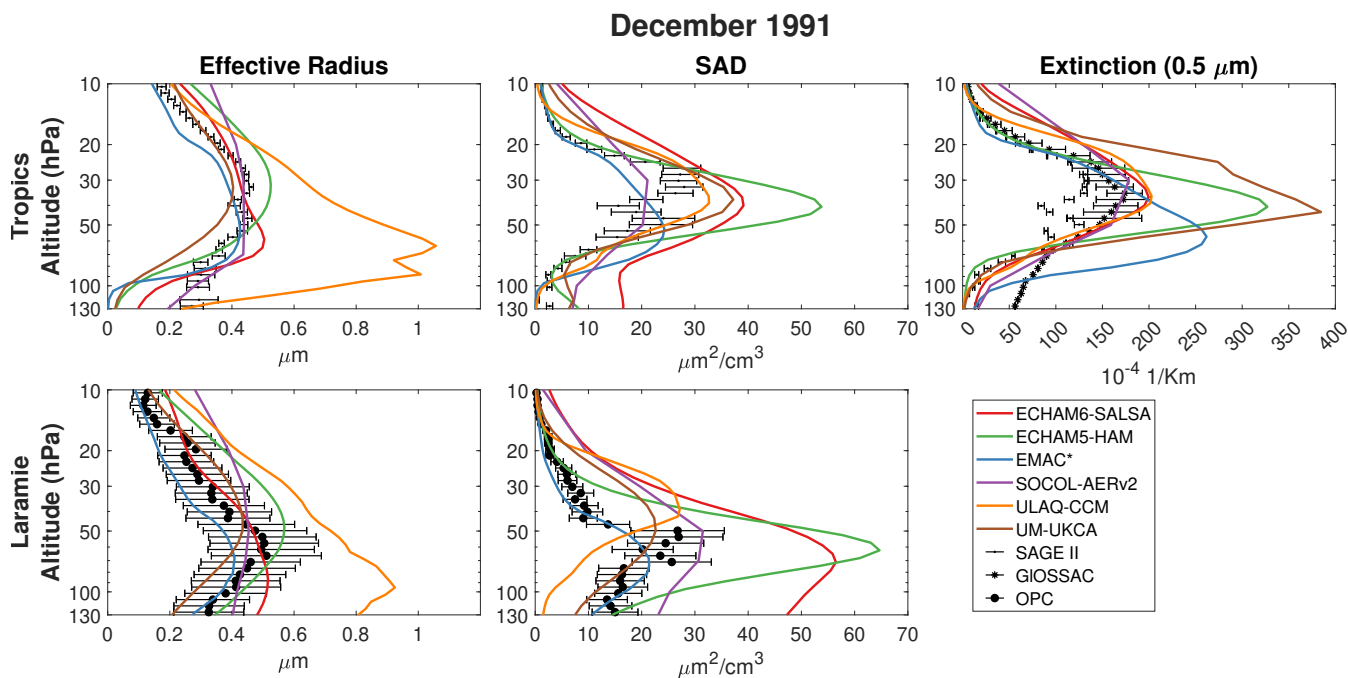


Figure 9. Vertical profile of the effective radius in μm (left panels), surface area density (SAD) in $\mu\text{m}^2/\text{cm}^3$ (middle panels), and extinction at $0.5 \mu\text{m}$ in $1/\text{km}$ (right panel) in the tropics (first row) and over Laramie (second row) for Med-22km in December 1991. Model results are compared with SAGE II and GloSSAC in the tropics and with OPC over Laramie.

Even when models and measurements look comparable for the integrated variables (optical depth, integrated effective radius and burden), these similarities hide the models inability to reproduce the observed vertical structure depending on the latitude and time period after the eruption under consideration. Most models take up to six months before they can reproduce the vertical structure of effective radius, SAD and extinction in the tropics, and up to a year at mid-latitudes.

The vertical distribution of SAD and effective radius in three moments identifying the build-up, maximum and descent phase of the evolution of the sulfate burden (September and December 1991 and June 1992, respectively) show an initial overestimation of the number concentration and the underestimation one year after the eruption. The lack of ash co-emission, a process not included in HErSEA simulations, would be crucial in the first days/month to better reproduce the initial plume evolution (Mills et al., 2017). On one hand, the ash may have removed part of the initial sulfur cloud through the SO_2 or H_2SO_4 uptake on these coarse particles, which have a significant fall velocity (Zhu et al., 2020); on the other hand, the presence of smaller ash particles causes greater heating and vertical lofting of the volcanic cloud, which results in a longer lifetime of stratospheric volcanic aerosols, affecting their transport (Kloss et al., 2021; Niemeier et al., 2021). The faster transport to the northern mid-latitudes than observed may have removed most of the stratospheric particles, so that the aerosol lifetime in the models is about half that observed. In addition to different transport and microphysical mechanisms, there could be



other missing processes that cause these divergences: the absence of the Cerro Hudson eruption two months after the Pinatubo eruption (Barton et al., 1992; Long and Stowe, 1994) could explain the lack of sulfate burden in the southern extratropics.

The inter-models differences may depend on numerous factors that interact with one another; this makes it hard to group models by perceived similarities, for instance a similar modal scheme, similarities in the large scale transport or an absence of interactive stratospheric chemistry. Laakso et al. (2022), for instance, used the same climate model (ECHAM-HAMMOZ) with two different aerosol microphysics schemes, one sectional and one modal: even with just that difference, the sectional scheme produced up to 52% larger effective radius than in the modal scheme simulation with the same amount of SO₂ injected for stratospheric aerosol intervention. Further, Niemeier et al. (2020) showed that, in two models with a similar modal scheme but different vertical advection (CESM-WACCM-110L and MAECHAM-HAM), the resulting vertical distribution of the aerosol plume can be substantially different. Even in the same model (CESM-WACCM), Richter et al. (2017) showed that the presence or not of interactive chemistry could strongly affect the local stratospheric warming, and thus the w* changes, due to feedback from the changing ozone. In our case, all of these differences are compounded, therefore it is hard to identify which exactly is the cause of the disagreement. Furthermore, in all the works cited above, SO₂ was injected continuously for a number of years rather than in an impulsive way, whereas in the case of a volcanic eruption, the synoptical conditions at the time of the eruption play an important role (Thomas et al., 2009; Toohey et al., 2014; Niemeier et al., 2021).

5 Conclusions

The ISA-MIP/HErSEA protocol allows us to investigate the weaknesses and the abilities of a group of climate models, all with interactive stratospheric aerosol microphysical treatment, by comparing them with measurements after the Mt. Pinatubo eruption in 1991. This is done through a well-defined experimental protocol with different sets of initial parameters for the stratospheric SO₂, both in terms of magnitude (5, 7 or 10 Tg-S injected) and altitude of the SO₂ plume (18-20, 21-23, 18-23 km, uniformly distributed). We found that different sets of initial conditions of the volcanic emissions are required in different models to reproduce available observations from that time period. The experiment protocol that best fits the observation also changes in some models depending on the variables to be considered (aerosol optical depth, effective radius, sulfate burden, surface area density).

The main reason for the disagreement with observations is stratospheric transport, which is too fast towards the northern mid-latitudes for some models or results in stronger tropical confinement in others. The transport consequently influences the growth of sulfate aerosols and their global distribution, which in turn affects the persistence of aerosols in the stratosphere, with a feedback on the transport itself (Brühl et al., 2015; Niemeier and Schmidt, 2017; Vioni et al., 2018b). Other reasons could be related to the absence of processes such as the absence of the Cerro Hudson eruption in the southern extratropics two months after the Pinatubo eruption, which may partly explain the lack of sulfate aerosols in the southern hemisphere, and the omission of ash injection which would be crucial in the early days/months to better reproduce the initial evolution of the plume. Our results highlight the need for some specific experiments that might be needed to disentangle the different components that contribute to the overall uncertainty. For instance, simulations that nudge stratospheric transport to reanalyses



(as done in Schmidt et al., 2018, in CESM(WACCM)) in multiple models could clarify the role of different microphysical
555 schemes. Similarly, consistently turning interactive stratospheric chemistry on and off in multiple models could highlight the
importance of ozone feedback (as done in Richter et al., 2017).

Even considering the best set of initial parameters, differences between models and observations remain and the inter-model
differences are still large, as found before in other multi model-experiments of explosive volcanic clouds (i.e. Tambora in
Marshall et al., 2018; Clyne et al., 2021).

560 We also note that the observations themselves show disagreement, sometimes as high as inter-model differences, because
of various issues with the saturation or sensitivity of the particular instrument. Our observations around the reliability of the
measurements during the Pinatubo event highlights the future need for more observation in order to be better prepared for future
explosive volcanic eruptions (Newhall et al., 2018), both for understanding short and long term impacts and as a benchmark
test for current Earth System models.

565 *Data availability.* Data used in this work will be made available through the Cornell eCommons platform before publication.

A Analysis of model output

A1 Taylor Diagrams

In section 3.1 we used Taylor diagrams (Taylor, 2001) in order to summarise all the information regarding the reproducibility
of the stratospheric optical depth simulated compared to satellite observations. Taylor diagrams provide a concise statistical
570 summary of how well patterns from simulations and observations match each other in terms of their correlation (COR, az-
imuthal angle), their root-mean-square difference (RMSD, proportional to the distance between the observations - grey and
black circles on the x axis- and the reference field - colored circle), and the ratio of their variances (STD, x and y axis). STDs,
RMSs and CORs are calculated for zonal values of the stratospheric AOD for two different time periods (first year and second
year after the eruption). Therefore, Higher CORR and lower RMSD mean similar amplitudes of variation in terms of latitudinal
575 distribution and time evolution.

A2 Effective radius

The effective radius is calculated as the ratio of the third and second moments of the number size distribution of the aerosol
particles. This results in Eq. A1 for models with a sectional scheme; in this case, the sum is over the bins, n_i is number of
particles and r_i is radius of particles in each bin. In models with a modal scheme, the effective radius is calculated as the sum
580 over the modes as in Eq. A2, where SAD_i is the surface area density and vol_i is the volume density.

$$r_{eff} = \frac{\sum_i n_i \cdot r_i^3}{\sum_i n_i \cdot r_i^2} \quad (A1)$$



$$reff = \frac{3 \cdot \sum_i voli}{\sum_i SAD_i} \quad (A2)$$

The stratospheric effective radius ($reff_{strat}$) for the models and SAGE II is calculated in Eq. A3 by integrating the provided effective radius ($reff$) from the tropopause to the top of the atmosphere weighted with the SAD. The thickness of the vertical layer (h) is calculated from the hypsometric equation (Eq A4)

$$reff_{strat} = \frac{\sum_z (SAD \cdot h \cdot reff)_z}{\sum_z (SAD \cdot h)_z} \quad (A3)$$

$$h = \frac{R \cdot T}{g} \cdot \ln \frac{P_{z+1}}{P_z} \quad (A4)$$

For the OPC measurements, we calculate the stratospheric effective radius (Eq. A5) as in Kleinschmitt et al. (2017) for the updated UWv2.0 data set. The measurement error bars consider a 40% uncertainty in SAD and vol and assume a correlation coefficient of 0.5 between SAD at different altitudes, vol at different altitudes and SAD and vol at the same altitude.

$$reff_{strat} = \frac{3 \cdot \sum_z vol_z}{\sum_z SAD_z} \quad (A5)$$

Author contributions. IQ led the analysis and wrote the paper with contributions by CT, UN and DV. CT, GM, CB, and SD designed the study. HF ran the ECHAM5-HAM simulations and provided the output data. All authors contributed to discussion and finalisation of the article.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgements. Claudia Timmreck and Ulrike Niemeier were supported by the Deutsche Forschungsgemeinschaft Research Unit Vol-Impact (FOR2820 (grant no. 398006378)) and use resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bm855 “ISA-MIP”.

The paper is part of the “Interactive Model Intercomparison Project” from the WCRP/SPARC activity “Stratospheric Sulfur and its Role in Climate (SSiRC)”.



References

- Aquila, V., Oman, L. D., Stolarski, R. S., Colarco, P. R., and Newman, P. A.: Dispersion of the volcanic sulfate cloud from a Mount Pinatubo-like eruption: DISPERSION OF A MOUNT PINATUBO CLOUD, *Journal of Geophysical Research: Atmospheres*, 117, n/a–n/a, <https://doi.org/10.1029/2011JD016968>, 2012.
- 605 Ayris, P., Lee, A., Wilson, K., Kueppers, U., Dingwell, D., and Delmelle, P.: SO₂ sequestration in large volcanic eruptions: High-temperature scavenging by tephra, *Geochimica et Cosmochimica Acta*, 110, 58–69, <https://doi.org/10.1016/j.gca.2013.02.018>, 2013.
- Baran, A. J. and Foot, J. S.: New application of the operational sounder HIRS in determining a climatology of sulphuric acid aerosol from the Pinatubo eruption, *Journal of Geophysical Research*, 99, 25 673, <https://doi.org/10.1029/94JD02044>, 1994.
- Barton, I. J., Prata, A. J., Watterson, I. G., and Young, S. A.: Identification of the Mount Hudson volcanic cloud over SE Australia, *Geophysical Research Letters*, 19, 1211–1214, <https://doi.org/10.1029/92GL01122>, 1992.
- 610 Bluth, G., Doiron, S., Schnetzler, C., Krueger, A., and Walter, L.: Global tracking of the SO₂ clouds from the June, 1991 Mount Pinatubo eruptions, *Geophysical Research Letters*, 19, <https://doi.org/10.1029/91GL02792>, 1992.
- Bluth, G., Rose, W., Sprod, I., and Krueger, A.: Stratospheric Loading of Sulfur From Explosive Volcanic Eruptions, *The Journal of Geology*, 105, 671–684, <https://doi.org/10.1086/515972>, 1997.
- 615 Borbas, E. E. and Menzel, P. W.: Observed HIRS and Aqua MODIS Thermal Infrared Moisture Determinations in the 2000s, *Remote Sensing*, 13, 502, <https://doi.org/10.3390/rs13030502>, 2021.
- Brodowsky, C., Sukhodolov, T., Feinberg, A., Höpfner, M., Peter, T., Stenke, A., and Rozanov, E.: Modeling the Sulfate Aerosol Evolution After Recent Moderate Volcanic Activity, 2008–2012, *Journal of Geophysical Research: Atmospheres*, 126, e2021JD035 472, <https://doi.org/https://doi.org/10.1029/2021JD035472>, e2021JD035472 2021JD035472, 2021.
- 620 Brühl, C., Lelieveld, J., Tost, H., Höpfner, M., and Glatthor, N.: Stratospheric sulfur and its implications for radiative forcing simulated by the chemistry climate model EMAC, *Journal of Geophysical Research: Atmospheres*, 120, 2103–2118, <https://doi.org/10.1002/2014JD022430>, 2015.
- Brühl, C., Schalllock, J., Klingmüller, K., Robert, C., Bingen, C., Clarisse, L., Heckel, A., North, P., and Rieger, L.: Stratospheric aerosol radiative forcing simulated by the chemistry climate model EMAC using Aerosol CCI satellite data, *Atmospheric Chemistry and Physics*, 18, 12 845–12 857, <https://doi.org/10.5194/acp-18-12845-2018>, 2018.
- 625 Clyne, M., Lamarque, J.-F., Mills, M. J., Khodri, M., Ball, W., Bekki, S., Dhomse, S. S., Lebas, N., Mann, G., Marshall, L., Niemeier, U., Poulain, V., Robock, A., Rozanov, E., Schmidt, A., Stenke, A., Sukhodolov, T., Timmreck, C., Toohey, M., Tummon, F., Zanchettin, D., Zhu, Y., and Toon, O. B.: Model physics and chemistry causing intermodel disagreement within the VolMIP-Tambora Interactive Stratospheric Aerosol ensemble, *Atmos. Chem. Phys.*, p. 27, 2021.
- 630 Damadeo, R. P., Zawodny, J. M., Thomason, L. W., and Iyer, N.: SAGE version 7.0 algorithm: application to SAGE II, *Atmospheric Measurement Techniques*, 6, 3539–3561, <https://doi.org/10.5194/amt-6-3539-2013>, 2013.
- DeFoor, T. E., Robinson, E., and Ryan, S.: Early lidar observations of the June 1991 Pinatubo eruption plume at Mauna Loa Observatory, Hawaii, *Geophysical Research Letters*, 19, 187–190, <https://doi.org/https://doi.org/10.1029/91GL02791>, 1992.
- Deshler, T.: Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments, *Journal of Geophysical Research*, 108, 4167, <https://doi.org/10.1029/2002JD002514>, 2003.
- 635



- Deshler, T., Luo, B., Kovilakam, M., Peter, T., and Kalnajs, L. E.: Retrieval of Aerosol Size Distributions From In Situ Particle Counter Measurements: Instrument Counting Efficiency and Comparisons With Satellite Measurements, *Journal of Geophysical Research: Atmospheres*, 124, 5058–5087, <https://doi.org/10.1029/2018JD029558>, 2019.
- 640 Dhomse, S. S., Emmerson, K. M., Mann, G. W., Bellouin, N., Carslaw, K. S., Chipperfield, M. P., Hommel, R., Abraham, N. L., Telford, P., Braesicke, P., Dalvi, M., Johnson, C. E., O'Connor, F., Morgenstern, O., Pyle, J. A., Deshler, T., Zawodny, J. M., and Thomason, L. W.: Aerosol microphysics simulations of the Mt.~Pinatubo eruption with the UM-UKCA composition-climate model, *Atmospheric Chemistry and Physics*, 14, 11 221–11 246, <https://doi.org/10.5194/acp-14-11221-2014>, 2014.
- Dhomse, S. S., Mann, G. W., Antuña Marrero, J. C., Shallcross, S. E., Chipperfield, M. P., Carslaw, K. S., Marshall, L., Abraham, N. L., and Johnson, C. E.: Evaluating the simulated radiative forcings, aerosol properties, and stratospheric warmings from the 1963
645 Mt Agung, 1982 El Chichón, and 1991 Mt Pinatubo volcanic aerosol clouds, *Atmospheric Chemistry and Physics*, 20, 13 627–13 654, <https://doi.org/10.5194/acp-20-13627-2020>, 2020.
- Egorova, T., Rozanov, E., Zubov, V., and Karol, I.: Model for investigating ozone trends (MEZON), *Izvestiya - Atmospheric and Ocean Physics*, 39, 277–292, 2003.
- Eyring, V., Lamarque, J.-F., Hess, P., Arfeuille, F., Bowman, K., Duncan, B., Fiore, A., Gettelman, A., Giorgetta, M. A., Granier, C., Hegglin,
650 M., Kinnison, D., Kunze, M., Langematz, U., Luo, B., Martin, R., Matthes, K., Newman, P. A., Peter, T., Robock, A., Ryerson, T., Saiz-Lopez, A., Salawitch, R., Schultz, M., Shepherd, T. G., Shindell, D., Staehelin, J., Thomason, L., Tilmes, S., Vernier, J.-P., Waugh, D. W., and Young, P. J.: Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) Community Simulations in Support of Upcoming Ozone and Climate Assessments, p. 19, 2013.
- Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler, P. J., Hnilo, J. J., Marlais,
655 S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E., and Williams, D. N.: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I), *Bulletin of the American Meteorological Society*, 80, 29–55, <http://www.jstor.org/stable/26214897>, 1999.
- Giorgetta, M. A., Manzini, E., Roeckner, E., Esch, M., and Bengtsson, L.: Climatology and Forcing of the Quasi-Biennial Oscillation in the MAECHAM5 Model, *Journal of Climate*, 19, 3882–3901, <https://doi.org/10.1175/JCLI3830.1>, 2006.
- 660 Guo, S., Rose, W. I., Bluth, G. J. S., and Watson, I. M.: Particles in the great Pinatubo volcanic cloud of June 1991: The role of ice: JUNE 1991 PINATUBO VOLCANIC CLOUDS, *Geochemistry, Geophysics, Geosystems*, 5, n/a–n/a, <https://doi.org/10.1029/2003GC000655>, 2004.
- Hommel, R., Timmreck, C., and Graf, H. F.: The global middle-atmosphere aerosol model MAECHAM5-SAM2: comparison with satellite and in-situ observations, *Geoscientific Model Development*, 4, 809–834, <https://doi.org/10.5194/gmd-4-809-2011>, 2011.
- 665 Kleinschmitt, C., Boucher, O., Bekki, S., Lott, F., and Platt, U.: The Sectional Stratospheric Sulfate Aerosol module (S3A-v1) within the LMDZ general circulation model: description and evaluation against stratospheric aerosol observations, *Geoscientific Model Development*, 10, 3359–3378, <https://doi.org/10.5194/gmd-10-3359-2017>, 2017.
- Kloss, C., Berthet, G., Sellitto, P., Ploeger, F., Taha, G., Tidiga, M., Eremenko, M., Bossolasco, A., Jégou, F., Renard, J.-B., and Legras, B.: Stratospheric aerosol layer perturbation caused by the 2019 Raikoke and Ulawun eruptions and their radiative forcing, *Atmospheric
670 Chemistry and Physics*, 21, 535–560, <https://doi.org/10.5194/acp-21-535-2021>, 2021.
- Kokkola, H., Kühn, T., Laakso, A., Bergman, T., Lehtinen, K. E. J., Mielonen, T., Arola, A., Stadtler, S., Korhonen, H., Ferrachat, S., Lohmann, U., Neubauer, D., Tegen, I., Siegenthaler-Le Drian, C., Schultz, M. G., Bey, I., Stier, P., Daskalakis, N., Heald, C. L., and



- Romakkaniemi, S.: SALSA2.0: The sectional aerosol module of the aerosol–chemistry–climate model ECHAM6.3.0-HAM2.3-MOZ1.0, *Geoscientific Model Development*, 11, 3833–3863, <https://doi.org/10.5194/gmd-11-3833-2018>, 2018.
- 675 Kovilakam, M. and Deshler, T.: On the accuracy of stratospheric aerosol extinction derived from in situ size distribution measurements and surface area density derived from remote SAGE II and HALOE extinction measurements, *Journal of Geophysical Research: Atmospheres*, 120, 8426–8447, <https://doi.org/10.1002/2015JD023303>, 2015.
- Kovilakam, M., Thomason, L. W., Ernest, N., Rieger, L., Bourassa, A., and Millán, L.: The Global Space-based Stratospheric Aerosol Climatology (version 2.0): 1979–2018, *Earth System Science Data*, 12, 2607–2634, <https://doi.org/10.5194/essd-12-2607-2020>, 2020.
- 680 Kravitz, B., Robock, A., Tilmes, S., Boucher, O., English, J. M., Irvine, P. J., Jones, A., Lawrence, M. G., MacCracken, M., Muri, H., Moore, J. C., Niemeier, U., Phipps, S. J., Sillmann, J., Storelvmo, T., Wang, H., and Watanabe, S.: The Geoengineering Model Inter-comparison Project Phase 6 (GeoMIP6): simulation design and preliminary results, *Geoscientific Model Development*, 8, 3379–3392, <https://doi.org/10.5194/gmd-8-3379-2015>, 2015.
- Kremser, S., Thomason, L. W., von Hobe, M., Hermann, M., Deshler, T., Timmreck, C., Toohey, M., Stenke, A., Schwarz, J. P., Weigel, R.,
685 Fueglistaler, S., Prata, F. J., Vernier, J.-P., Schlager, H., Barnes, J. E., Antuña-Marrero, J.-C., Fairlie, D., Palm, M., Mahieu, E., Notholt, J., Rex, M., Bingen, C., Vanhellemont, F., Bourassa, A., Plane, J. M. C., Klocke, D., Carn, S. A., Clarisse, L., Trickl, T., Neely, R., James, A. D., Rieger, L., Wilson, J. C., and Meland, B.: Stratospheric aerosol-Observations, processes, and impact on climate: Stratospheric Aerosol, *Reviews of Geophysics*, 54, 278–335, <https://doi.org/10.1002/2015RG000511>, 2016.
- Laakso, A., Niemeier, U., Vioni, D., Tilmes, S., and Kokkola, H.: Dependency of the impacts of geoengineering on the stratospheric sulfur
690 injection strategy – Part 1: Intercomparison of modal and sectional aerosol modules, *Atmospheric Chemistry and Physics*, 22, 93–118, <https://doi.org/10.5194/acp-22-93-2022>, 2022.
- Labitzke, K. and McCormick, M. P.: Stratospheric temperature increases due to Pinatubo aerosols, *Geophysical Research Letters*, 19, 207–210, <https://doi.org/10.1029/91GL02940>, 1992.
- Lambert, A., Grainger, R. G., Remedios, J. J., Rodgers, C. D., Corney, M., and Taylor, F. W.: Measurements of the evolution of the Mt.
695 Pinatubo aerosol cloud by ISAMS, *Geophysical Research Letters*, 20, 1287–1290, <https://doi.org/10.1029/93GL00827>, 1993.
- Long, C. S. and Stowe, L. L.: using the NOAA/AVHRR to study stratospheric aerosol optical thicknesses following the Mt. Pinatubo
Eruption, *Geophysical Research Letters*, 21, 2215–2218, <https://doi.org/10.1029/94GL01322>, 1994.
- Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.:
700 Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, *Geoscientific Model Development*, 3, 519–551, <https://doi.org/10.5194/gmd-3-519-2010>, 2010.
- Marshall, L., Schmidt, A., Toohey, M., Carslaw, K. S., Mann, G. W., Sigl, M., Khodri, M., Timmreck, C., Zanchettin, D., Ball, W. T.,
Bekki, S., Brooke, J. S. A., Dhomse, S., Johnson, C., Lamarque, J.-F., LeGrande, A. N., Mills, M. J., Niemeier, U., Pope, J. O.,
Poulain, V., Robock, A., Rozanov, E., Stenke, A., Sukhodolov, T., Tilmes, S., Tsigaridis, K., and Tummon, F.: Multi-model compar-
705 ison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora, *Atmospheric Chemistry and Physics*, 18, 2307–2328, <https://doi.org/10.5194/acp-18-2307-2018>, 2018.
- Mauldin, L. E., I., Zaun, N. H., McCormick, M. P., J., Guy, J. H., and Vaughn, W. R.: Stratospheric Aerosol And Gas Experiment II
Instrument: A Functional Description, *Optical Engineering*, 24, 307, <https://doi.org/10.1117/12.7973473>, 1985.
- McCormick, M. P., Thomason, L. W., and Trepte, C. R.: Atmospheric effects of the Mt Pinatubo eruption, *Nature*, 373, 399–404, <https://doi.org/10.1038/373399a0>, 1995.



- 710 Mills, M. J., Schmidt, A., Easter, R., Solomon, S., Kinnison, D. E., Ghan, S. J., Neely, R. R., Marsh, D. R., Conley, A., Bardeen, C. G., and Gettelman, A.: Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1(WACCM), *Journal of Geophysical Research: Atmospheres*, 121, 2332–2348, <https://doi.org/10.1002/2015JD024290>, 2016.
- Mills, M. J., Richter, J. H., Tilmes, S., Kravitz, B., MacMartin, D. G., Glanville, A. A., Tribbia, J. J., Lamarque, J., Vitt, F., Schmidt, A., Gettelman, A., Hannay, C., Bacmeister, J. T., and Kinnison, D. E.: Radiative and Chemical Response to Interactive Stratospheric Sulfate Aerosols in Fully Coupled CESM1(WACCM), *Journal of Geophysical Research: Atmospheres*, 122, <https://doi.org/10.1002/2017JD027006>, 2017.
- 715 Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marécal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.:
720 Review of the global models used within phase I of the Chemistry–Climate Model Initiative (CCMI), *Geoscientific Model Development*, 10, 639–671, <https://doi.org/10.5194/gmd-10-639-2017>, 2017.
- NASA/LARC/SD/ASDC: Stratospheric Aerosol and Gas Experiment (SAGE) II Version 7.0 Aerosol, O₃, NO₂ and H₂O Profiles in binary format, https://doi.org/10.5067/ERBS/SAGEII/SOLAR_BINARY_L2-V7.0, 2012a.
- NASA/LARC/SD/ASDC: Surface Radiation Budget (SRB) Release 3.0 QC Shortwave monthly data in netcdf format, https://doi.org/10.5067/SRB/REL3.0_LPSA_MONTHLY_NC_L3, 2012b.
- 725 Newhall, C., Self, S., and Robock, A.: Anticipating future Volcanic Explosivity Index (VEI) 7 eruptions and their chilling impacts, *Geosphere*, 14, 572–603, <https://doi.org/10.1130/GES01513.1>, 2018.
- Niemeier, U. and Schmidt, H.: Changing transport processes in the stratosphere by radiative heating of sulfate aerosols, *Atmospheric Chemistry and Physics*, 17, 14 871–14 886, <https://doi.org/10.5194/acp-17-14871-2017>, 2017.
- 730 Niemeier, U. and Timmreck, C.: Initial fate of fine ash and sulfur from large volcanic eruptions, *Atmos. Chem. Phys.*, p. 15, 2009.
- Niemeier, U., Richter, J. H., and Tilmes, S.: Differing responses of the quasi-biennial oscillation to artificial SO₂ injections in two global models, *Atmospheric Chemistry and Physics*, 20, 8975–8987, <https://doi.org/10.5194/acp-20-8975-2020>, 2020.
- Niemeier, U., Riede, F., and Timmreck, C.: Simulation of ash clouds after a Laacher See-type eruption, *Climate of the Past*, 17, 633–652, <https://doi.org/10.5194/cp-17-633-2021>, 2021.
- 735 Pitari, G., Rizi, V., Ricciardulli, L., and Visconti, G.: High-speed civil transport impact: Role of sulfate, nitric acid trihydrate, and ice aerosols studied with a two-dimensional model including aerosol physics, *Journal of Geophysical Research: Atmospheres*, 98, 23 141–23 164, <https://doi.org/https://doi.org/10.1029/93JD02600>, 1993.
- Pitari, G., Mancini, E., Rizi, V., and Shindell, D. T.: Impact of Future Climate and Emission Changes on Stratospheric Aerosols and Ozone, *JOURNAL OF THE ATMOSPHERIC SCIENCES*, 59, 27, 2002.
- 740 Pitari, G., Cionni, I., Di Genova, G., Visioni, D., Gandolfi, I., and Mancini, E.: Impact of Stratospheric Volcanic Aerosols on Age-of-Air and Transport of Long-Lived Species, *Atmosphere*, 7, 149, <https://doi.org/10.3390/atmos7110149>, 2016.
- Pringle, K. J., Tost, H., Message, S., Steil, B., Giannadaki, D., Nenes, A., Fountoukis, C., Stier, P., Vignati, E., and Lelieveld, J.: Description and evaluation of GMX_e: a new aerosol submodel for global simulations (v1), *Geoscientific Model Development*, 3, 391–412, <https://doi.org/10.5194/gmd-3-391-2010>, 2010.
- 745 RAO, C. R. N., STOWE, L. L., and McCLAIN, E. P.: Remote sensing of aerosols over the oceans using AVHRR data Theory, practice and applications, *International Journal of Remote Sensing*, 10, 743–749, <https://doi.org/10.1080/01431168908903915>, publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/01431168908903915>, 1989.



- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research: Atmospheres*, 108, <https://doi.org/10.1029/2002JD002670>, 2003.
- 750 Revell, L., Stenke, A., Luo, B., Kremser, S., Rozanov, E., Sukhodolov, T., and Peter, T.: Chemistry-climate model simulations of the Mt. Pinatubo eruption using CCM1 and CMIP6 stratospheric aerosol data, preprint, *Aerosols/Atmospheric Modelling/Stratosphere/Chemistry (chemical composition and reactions)*, <https://doi.org/10.5194/acp-2017-633>, 2017.
- Richter, J. H., Tilmes, S., Mills, M. J., Tribbia, J. J., Kravitz, B., MacMartin, D. G., Vitt, F., and Lamarque, J.: Stratospheric Dynamical Response and Ozone Feedbacks in the Presence of SO₂ Injections, *Journal of Geophysical Research: Atmospheres*, 122, <https://doi.org/10.1002/2017JD026912>, 2017.
- 755 Robock, A., Ammann, C. M., Oman, L., Shindell, D., Levis, S., and Stenchikov, G.: Did the Toba volcanic eruption of ~74 ka B.P. produce widespread glaciation?, *Journal of Geophysical Research*, 114, D10 107, <https://doi.org/10.1029/2008JD011652>, 2009.
- Rosen, J. M.: The vertical distribution of dust to 30 kilometers, *Journal of Geophysical Research*, 69, 4673–4676, <https://doi.org/10.1029/JZ069i021p04673>, 1964.
- 760 Russell, P. B., Livingston, J. M., Pueschel, R. F., Bauman, J. J., Pollack, J. B., Brooks, S. L., Hamill, P., Thomason, L. W., Stowe, L. L., Deshler, T., Dutton, E. G., and Bergstrom, R. W.: Global to microscale evolution of the Pinatubo volcanic aerosol derived from diverse measurements and analyses, *Journal of Geophysical Research: Atmospheres*, 101, 18 745–18 763, <https://doi.org/10.1029/96JD01162>, 1996.
- 765 Saxena, V. K., Anderson, J., and Lin, N.-H.: Changes in Antarctic stratospheric aerosol characteristics due to volcanic eruptions as monitored by the Stratospheric Aerosol and Gas Experiment II satellite, *Journal of Geophysical Research*, 100, 16 735, <https://doi.org/10.1029/95JD01537>, 1995.
- Schallock, J., Brühl, C., Bingen, C., Höpfner, M., Rieger, L., and Lelieveld, J.: Radiative forcing by volcanic eruptions since 1990, calculated with a chemistry-climate model and a new emission inventory based on vertically resolved satellite measurements, preprint, *Aerosols/Atmospheric Modelling/Stratosphere/Chemistry (chemical composition and reactions)*, <https://doi.org/10.5194/acp-2021-654>, 2021.
- 770 Schmidt, A., Mills, M. J., Ghan, S., Gregory, J. M., Allan, R. P., Andrews, T., Bardeen, C. G., Conley, A., Forster, P. M., Gettelman, A., Portmann, R. W., Solomon, S., and Toon, O. B.: Volcanic Radiative Forcing From 1979 to 2015, *Journal of Geophysical Research: Atmospheres*, 123, 12 491–12 508, <https://doi.org/10.1029/2018JD028776>, 2018.
- 775 Schultz, M. G., Stadtler, S., Schröder, S., Taraborrelli, D., Franco, B., Krefting, J., Henrot, A., Ferrachat, S., Lohmann, U., Neubauer, D., Siegenthaler-Le Drian, C., Wahl, S., Kokkola, H., Kühn, T., Rast, S., Schmidt, H., Stier, P., Kinnison, D., Tyndall, G. S., Orlando, J. J., and Wespes, C.: The chemistry–climate model ECHAM6.3-HAM2.3-MOZ1.0, *Geoscientific Model Development*, 11, 1695–1723, <https://doi.org/10.5194/gmd-11-1695-2018>, 2018.
- Sheng, J.-X., Weisenstein, D. K., Luo, B.-P., Rozanov, E., Stenke, A., Anet, J., Bingemer, H., and Peter, T.: Global atmospheric sulfur budget under volcanically quiescent conditions: Aerosol-chemistry-climate model predictions and validation, *Journal of Geophysical Research: Atmospheres*, 120, 256–276, <https://doi.org/10.1002/2014JD021985>, 2015.
- 780 Soden, B. J., Wetherald, R. T., Stenchikov, G. L., and Robock, A.: Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor, *Science*, 296, 727–730, <https://doi.org/10.1126/science.296.5568.727>, 2002.



- 785 Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornbluh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, *Journal of Advances in Modeling Earth Systems*, 5, 146–172, <https://doi.org/https://doi.org/10.1002/jame.20015>, 2013.
- Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, *Atmospheric Chemistry and Physics*, 5, 1125–1156, <https://doi.org/10.5194/acp-5-1125-2005>, 2005.
- 790 Sukhodolov, T., Sheng, J.-X., Feinberg, A., Luo, B.-P., Peter, T., Revell, L., Stenke, A., Weisenstein, D. K., and Rozanov, E.: Stratospheric aerosol evolution after Pinatubo simulated with a coupled size-resolved aerosol–chemistry–climate model, SOCOL-AERv1.0, *Geoscientific Model Development*, 11, 2633–2647, <https://doi.org/10.5194/gmd-11-2633-2018>, 2018.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, 106, 7183–7192, <https://doi.org/10.1029/2000JD900719>, 2001.
- 795 Thomas, M. A., Giorgetta, M. A., Timmreck, C., Graf, H.-F., and Stenchikov, G.: Simulation of the climate impact of Mt. Pinatubo eruption using ECHAM5 – Part 2: Sensitivity to the phase of the QBO and ENSO, *Atmospheric Chemistry and Physics*, 9, 3001–3009, <https://doi.org/10.5194/acp-9-3001-2009>, 2009.
- Thomason, L. W. and Burton, S. P.: SAGE II measurements of stratospheric aerosol properties at non-volcanic levels, *Atmos. Chem. Phys.*, p. 13, 2008.
- 800 Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, *Earth System Science Data*, 10, 469–492, <https://doi.org/10.5194/essd-10-469-2018>, 2018.
- Timmreck, C.: Three-dimensional simulation of stratospheric background aerosol: First results of a multiannual general circulation model simulation, , 106, 28,313–28,332, <https://doi.org/10.1029/2001JD000765>, 2001.
- 805 Timmreck, C., Graf, H.-F., and Kirchner, I.: A one and half year interactive MA/ECHAM4 simulation of Mount Pinatubo Aerosol, *Journal of Geophysical Research: Atmospheres*, 104, 9337–9359, <https://doi.org/https://doi.org/10.1029/1999JD900088>, 1999.
- Timmreck, C., Graf, H.-F., Zanchettin, D., Hagemann, S., Kleinen, T., and Krüger, K.: Climate response to the Toba super-eruption: Regional changes, *Quaternary International*, 258, 30–44, <https://doi.org/10.1016/j.quaint.2011.10.008>, 2012.
- Timmreck, C., Mann, G. W., Aquila, V., Hommel, R., Lee, L. A., Schmidt, A., Brühl, C., Carn, S., Chin, M., Dhomse, S. S., Diehl, T., English, J. M., Mills, M. J., Neely, R., Sheng, J., Toohey, M., and Weisenstein, D.: The Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP): motivation and experimental design, *Geoscientific Model Development*, 11, 2581–2608, <https://doi.org/10.5194/gmd-11-2581-2018>, 2018.
- 810 Toohey, M., Krüger, K., Niemeier, U., and Timmreck, C.: The influence of eruption season on the global aerosol evolution and radiative impact of tropical volcanic eruptions, *Atmospheric Chemistry and Physics*, 11, 12 351–12 367, <https://doi.org/10.5194/acp-11-12351-2011>, 2011.
- Toohey, M., Krüger, K., Bittner, M., Timmreck, C., and Schmidt, H.: The impact of volcanic aerosol on the Northern Hemisphere stratospheric polar vortex: mechanisms and sensitivity to forcing structure, *Atmospheric Chemistry and Physics*, 14, 13 063–13 079, <https://doi.org/10.5194/acp-14-13063-2014>, 2014.
- Toon, O. B., McKay, C. P., Ackerman, T. P., and Santhanam, K.: Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres, *Journal of Geophysical Research: Atmospheres*, 94, 16 287–16 301, <https://doi.org/https://doi.org/10.1029/JD094iD13p16287>, 1989.
- 820



- Vehkamäki, H., Kulmala, M., Napari, I., Lehtinen, K. E. J., Timmreck, C., Noppel, M., and Laaksonen, A.: An improved parameterization for sulfuric acid–water nucleation rates for tropospheric and stratospheric conditions, *Journal of Geophysical Research: Atmospheres*, 107, AAC 3–1–AAC 3–10, <https://doi.org/https://doi.org/10.1029/2002JD002184>, 2002.
- 825 Visionsi, D., Pitari, G., di Genova, G., Tilmes, S., and Cionni, I.: Upper tropospheric ice sensitivity to sulfate geoengineering, *Atmospheric Chemistry and Physics*, 18, 14 867–14 887, <https://doi.org/10.5194/acp-18-14867-2018>, 2018a.
- Visioni, D., Pitari, G., Tuccella, P., and Curci, G.: Sulfur deposition changes under sulfate geoengineering conditions: quasi-biennial oscillation effects on the transport and lifetime of stratospheric aerosols, *Atmospheric Chemistry and Physics*, 18, 2787–2808, <https://doi.org/10.5194/acp-18-2787-2018>, 2018b.
- 830 Visionsi, D., MacMartin, D. G., Kravitz, B., Boucher, O., Jones, A., Lurton, T., Martine, M., Mills, M. J., Nabat, P., Niemeier, U., Séférian, R., and Tilmes, S.: Identifying the sources of uncertainty in climate model simulations of solar radiation modification with the G6sulfur and G6solar Geoengineering Model Intercomparison Project (GeoMIP) simulations, *Atmospheric Chemistry and Physics*, 21, 10 039–10 063, <https://doi.org/10.5194/acp-21-10039-2021>, 2021.
- Walters, D. N., Williams, K. D., Boutle, I. A., Bushell, A. C., Edwards, J. M., Field, P. R., Lock, A. P., Morcrette, C. J., Stratton, R. A., 835 Wilkinson, J. M., Willett, M. R., Bellouin, N., Bodas-Salcedo, A., Brooks, M. E., Copsey, D., Earnshaw, P. D., Hardiman, S. C., Harris, C. M., Levine, R. C., MacLachlan, C., Manners, J. C., Martin, G. M., Milton, S. F., Palmer, M. D., Roberts, M. J., Rodríguez, J. M., Tennant, W. J., and Vidale, P. L.: The Met Office Unified Model Global Atmosphere 4.0 and JULES Global Land 4.0 configurations, *Geoscientific Model Development*, 7, 361–386, <https://doi.org/10.5194/gmd-7-361-2014>, 2014.
- Waugh, D. W., Grise, K. M., Seviour, W. J. M., Davis, S. M., Davis, N., Adam, O., Son, S.-W., Simpson, I. R., Staten, P. W., Maycock, A. C., 840 Ummenhofer, C. C., Birner, T., and Ming, A.: Revisiting the Relationship among Metrics of Tropical Expansion, *Journal of Climate*, 31, 7565–7581, <https://doi.org/10.1175/JCLI-D-18-0108.1>, 2018.
- Weisenstein, D. K., Yue, G. K., Ko, M. K. W., Sze, N.-D., Rodriguez, J. M., and Scott, C. J.: A two-dimensional model of sulfur species and aerosols, *Journal of Geophysical Research: Atmospheres*, 102, 13 019–13 035, <https://doi.org/https://doi.org/10.1029/97JD00901>, 1997.
- Winker, D. M. and Osborn, M. T.: Airborne lidar observations of the Pinatubo volcanic plume, *Geophysical Research Letters*, 19, 167–170, 845 <https://doi.org/https://doi.org/10.1029/91GL02867>, 1992a.
- Winker, D. M. and Osborn, M. T.: Preliminary analysis of observations of the Pinatubo volcanic plume with a polarization-sensitive lidar, *Geophysical Research Letters*, 19, 171–174, <https://doi.org/https://doi.org/10.1029/91GL02866>, 1992b.
- Young, R. E., Houben, H., and Toon, O. B.: Radiatively forced dispersion of the Mt. Pinatubo volcanic cloud and induced temperature perturbations in the stratosphere during the first few months following the eruption, *Geophysical Research Letters*, 21, 369–372, 850 <https://doi.org/10.1029/93GL03302>, 1994.
- Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., LeGrande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6, *Geoscientific Model Development*, 9, 2701–2719, 855 <https://doi.org/10.5194/gmd-9-2701-2016>, 2016.
- Zanchettin, D., Timmreck, C., Khodri, M., Schmidt, A., Toohey, M., Abe, M., Bekki, S., Cole, J., Fang, S.-W., Feng, W., Hegerl, G., Johnson, B., Lebas, N., LeGrande, A. N., Mann, G. W., Marshall, L., Rieger, L., Robock, A., Rubineti, S., Tsigaridis, K., and Weierbach, H.: Effects of forcing differences and initial conditions on inter-model agreement in the VolMIP volc-pinatubo-full experiment, *Geoscientific Model Development*, 15, 2265–2292, <https://doi.org/10.5194/gmd-15-2265-2022>, 2022.



- 860 Zhu, Y., Toon, O. B., Jensen, E. J., Bardeen, C. G., Mills, M. J., Tolbert, M. A., Yu, P., and Woods, S.: Persisting volcanic ash particles impact stratospheric SO₂ lifetime and aerosol optical properties, *Nature Communications*, 11, 4526, <https://doi.org/10.1038/s41467-020-18352-5>, 2020.