Warning: If you are getting these scripts and source datasets from GitHub, the GloSSAC_V2.22 was too big to upload. You can get a copy from:

Aubry, T. (2025). Scripts and source datasets for CMIP7 stratospheric aerosol forcing datasets (v2.2.1) (version 2.2.1). Zenodo. https://doi.org/10.5281/zenodo.17295697

Please cite the CMIP7 stratospheric aerosol dataset preprints or papers (submission expected Oct 2025, contact thomas.aubry@earth.ox.ac.uk should you be unable to find them) should you use any part of the provided scripts or data generated from it. If you use the EVA_H version 2 model, please also cite the EVA and EVA_H paper:

Aubry, T. J., M. Toohey, L. R. Marshall, A. Schmidt, and A. M. Jellinek (2020) A new volcanic stratospheric sulfate aerosol forcing emulator (EVA_H): Comparison with interactive stratospheric aerosol models, Journal of Geophysical Research: Atmospheres, 125, https://doi.org/10.1029/2019JD031303

Toohey, M., Stevens, B., Schmidt, H., and Timmreck, C.: Easy Volcanic Aerosol (EVA v1.0): an idealized forcing generator for climate simulations, Geosci. Model Dev., 9, 4049–4070, https://doi.org/10.5194/gmd-9-4049-2016, 2016.

1-Overview

The scripts and datasets provided were used to produce version 2.2.1 of the CMIP7 stratospheric aerosol dataset, i.e. the stratospheric aerosol optical properties (historical, 1750-2023, and pre-industrial climatology) and stratospheric volcanic SO2 emission dataset (historical, 1750-2023). All scripts are written in Matlab and run with version R2024a. The following are provided:

- Script ***build_CMIP7_emissions_v221_GMD.m***: Produces a preliminary version of the emission dataset (*prelim_CMIP7_volcano_S_emissions_v2_2_1.xlsx*). Further details on this script are provided below.
- Script ***build_CMIP7_SAOP_v221_GMD.m***: Produces the final version of the emission dataset (*utsvolcsulfur_input4MIPs_emissions_CMIP7_2.2.1_grid_label_1750-2023.nc* and *utsvolcsulfur_input4MIPs_emissions_CMIP7_2.2.1_grid_label_1750-2023.xlsx*) and of the aerosol optical property dataset (*strat_aer_opt_prop_input4MIPs_type_CMIP7_2.2.1_grid_label_1750-2023.nc*). This scripts needs to be run after build_CMIP7_emissions_v221_GMD.m. Further details on this script are provided below.
- Functions *emis_sanchecks_num.m* (to do sanity checks on the emission dataset), *emistonetcdf.m* (to save the emission dataset spreadsheet to netcdf format), *makeclimatology.m* (to calculate pre-industrial control climatologies from the historical aerosol optical property dataset), and *varfromname.m* (to load a variable from a netcdf file). These functions are used to run *build_CMIP7_SAOP_v221_GMD.m* and/or *build_CMIP7_emissions_v221_GMD.m*.
- Sub-folder *CMIP6*: Contains the CMIP6 historical extinction to compare with CMIP7 in *build_CMIP7_SAOP_v221_GMD.m*.
- Sub-folder ***EVA_H_v2***: Contains all functions and datasets required to run version 2 of EVA_H, which is itself required to run *build_CMIP7_SAOP_v221_GMD.m*. Further details on this sub-folder are provided below.

- Sub-folder ***source_datasets***: Contains all source datasets required to run scripts *build_CMIP7_emissions_v221_GMD.m* and *build_CMIP7_SAOP_v221_GMD.m*. Further details on this sub-folder are provided below.

2- build_CMIP7_emissions_v221_GMD.m

Note that this script has a couple prompts requiring a y/n answer from the user. The script contains the following key steps:

1. Import the MSVOLSO2L4 satellite-based volcanic emission inventory (Carn, 2022; all references in the CMIP7 papers)
2. Import eVolv2k emission inventory (Toohey and Sigl, 2017), covering up to year 1900 using a bipolar ice-core array
3. Import the Sigl et al (2015) emission inventory, used to cover years 1900-1978 and based on a bipolar ice-core array
4. Import the D4i emission inventory (Fang et al., 2023), covering up to year 1900 using a single high-resolution core from Greenland
5. Combine eVolv2k (for 1750-1900), Sigl et al. 2015 for (1901-1978) and D4i (for 1730-1900, excluding any eruption in eVolv2k)
6. Load Global Volcanism Program (GVP) Volcanoes of the World (VOTW) dataset for confirmed eruptions
7. Assess rates of GVP under-recording for VEI 3-6
8. Load the eruption match spreadsheet
9. Load the eruption source parameter (ESP) spreadsheet
10. Pre-allocate CMIP7 parameters and filter GVP data to only keep relevant eruptions in terms of time period and explosivity.
11. Match pre-satellite ice-core sulfate deposition events to eruptions in the GVP database, and assign ESPs when constrained.
12. Attribute ESPs to remaining GVP events and merge with CMIP7 emission dataset
13. Combine pre-satellite (step 2-12) and satellite (step 1) datasets, and pre-allocate arrays for all CMIP7 emission variables.
14. Create a dataset of combined MSVOLSO2L4 events, i.e. combine events belonging to the same eruption
15. Load dataset of pre-1750 ice-core sulfate deposition events geochemically matched to known eruptions, and their ESPs. These are considered high-confidence matches
16. Analyze the relationship between ice-core sulfate deposition asymettry vs latitude, using only CMIP7 events that are satellite-era or high-confidence matches, or pre-1750 high-confidence matches. Derive an empirical relationship to attribute a latitude to ice-core sulfate deposition events not matched to a historical eruption.
17. Repeat 16 but to analyze relationship between SO_2 mass and SO_2 injection height. Derive an empirical relationship to attribute a height to events (mostly pre-satellite) for which height is not-constrained.
18. Create additional variables and fill all gaps, either using empirical relationships derived at step 16 and 17, or using ad-hoc assumptions.
19. Sanity checks, sort chronologically and save this preliminary version of the CMIP7 emission as xlsx. Note GVP eruptions not matched to an ice-core event have no constrain on injected SO_2 mass and have default masses. Initial guesstimate for appropriate masses are refined in the script *build_CMIP7_emissions_v221_GMD.m*, where the final version of the emission dataset is created.

Commented [IS1]:

3- build_CMIP7_SAOP_v221_GMD.m

The script contains the following key steps:

1. Load and pre-process the preliminary CMIP7 emission dataset (created with previous script)
2. Load and pre-process GloSSAC v2.22
3. Define default SO_2 mass for VEI 4 and 5 eruptions originating from GVP eruptions not matched to an ice-core event, i.e. with no constrain on injected SO_2 mass. The VEI 5 default mass is chosen as the average of VEI 5 eruptions with know SO_2 masses in the dataset. The VEI 4 default mass is chosen so that the global mean mid-visible SAOD anomaly at 525nm is the same for: i) 1998-2023 (observed SAOD), characterized by eruptions of VEI up to 5 and masses up to 2 Tg SO2; and ii) 1850-1978 (SAOD derived from emission dataset using EVA_H v2 model), when considering only eruptions injecting less than 3 Tg SO_2.
4. Update GVP-derived events SO_2 masses in the emission dataset.
5. Save final emission dataset as xlsx (full dataset, including e.g. confidence match flags) and netcdf (only variables required to run interactive stratospheric aerosol model).
6. Run the volcanic aerosol model EVA_H v2 using the final CMIP7 inventory, where we only modified the Agung 1963 latitude to ensure forcing mostly over the Southern Hemisphere.
7. Harmonize the GloSSAC (1979-2023) and emission-derived EVA_H v2 product (1750-2023) over 1979-1981 and get full aerosol optical properties for GloSSAC
8. Combine the EVA_H product (1750-1981, including GloSSAC harmonization period) and GloSSAC (1982-2023) products
9. Make some basic plots comparing CMIP6 and CMIP7 stratospheric aerosol optical property datasets
10. Write CMIP7 historical dataset in a netcdf file
11. Calculate pre-industrial aerosol optical property climatologies from 1850-2021 average and write in same netcdf file as historical

4- EVA_H_v2

The EVA_H_v2 folder contains all scripts and data required to run version 2 of EVA_H. We only provide a brief description of each file here, with further comments given in each script, and full model details provided in the CMIP7 paper and Aubry et al. (2020). To actually run EVA_H, users can see code examples in sections 3 or 5 of the script build_CMIP7_emissions_v221_GMD.m. The key steps are to: i) format a volcanic SO_2 emission dataset to EVA_H required format using function *so2injection_8boxes_cmip7.m*. ii) run script *parameterfile.m* to load all EVA_H parameter values. iii) solve the model equations contained in function *eightboxequations_CMIP7.m* using the previously defined model parameter values and volcanic SO_2 injections, as well as a background, non-volcanic stratospheric aerosol climatology defined by the user. iv) use *postproc_v5.m* to post-process the obtained sulfate aerosol mass time series in the eight boxes of the model to obtain full field of stratospheric aerosol properties, using in particular the Mie look-up table *eva_lut_for_SAD_bimodal_Bier75_215_extended.nc* and the EVA_H shape functions *shapefunctions.mat*.

varfromname.m : Matlab function to read a variable from a netcdf file

ncep_tropo.mat : Matlab data file containing a tropopause height climatology derived from the NCEP-NCAR reanalysis (Aubry et al., 2020).

so2injection_8boxes_cmip7.m : Matlab function taking as input a set of volcanic SO_2 emission defined by mass, height, longitude, latitude and time, and returning a matrix containing corresponding SO_2 emission as a function of time in each of the 8 stratospheric boxes of the EVA_H model.

eva_lut_for_SAD_bimodal_Bier75_215_extended.nc : Netcdf file containing the Mie look-up table used to generate full stratospheric aerosol property fields.

shapefunctions.mat : Matlab data file containing shape functions used to convert the time series of aerosol mass in eight stratospheric boxes into gridded latitude-altitude zonal mean stratospheric aerosol optical properties.

eightboxequations_CMIP7.m : Matlab function defining the EVA_H model equations.

parameterfile.m : Matlab script defining all EVA_H model parameter values.

postproc_v5.m : Matlab function to post-process the obtained sulfate aerosol mass time series in the eight boxes of the model to obtain full field of stratospheric aerosol properties, using in particular the Mie look-up table eva_lut_for_SAD_bimodal_Bier75_215_extended.nc and the EVA_H shape functions shapefunctions.mat.

postproc_glwl_v3.m : Matlab function to calculate full stratospheric aerosol optical properties from extinction at two wavelengths (the script was written with the intent to use the GloSSAC core wavelength, 525 and 1020 nm), using the same Mie look-up table as the EVA_H model.

5- source_datasets

eVolv2k_v3_ds_1.nc : Netcdf file containing the eVolv2k dataset (Toohey and Sigl, 2017), i.e. i.e. volcanic SO_2 emissions derived from a bipolar ice-core array.

Sigl2015.xlsx : Excel spreadsheet containing the pre-processed Sigl et al. (2015) dataset i.e. volcanic SO_2 emissions derived from a bipolar ice-core array. The only pre-processing done to modify the actual data is that for 20^{th}C eruptions, the lower bound of Greenland sulfate deposition flux was used instead of the best estimate. Please refer to the CMIP7 emission dataset paper for justification.

D4i_cleaned_noevolv2k.xlsx : Excel spreadsheet containing the pre-processed D4i dataset (Fang et al., 2023), i.e. volcanic SO_2 emissions derived from a single, high-resolution Greenland ice-core. Pre-processing consisted in removing any eruption contained in eVolv2k, and reinstating original ice-core parameters where parameters originating from an eruption match in Fang et al. (2023). The latter was done because all matches were re-assessed for CMIP7.

MSVOLSO2L4_v04-00-2024m0129_noheader.txt : Excel spreadsheet containing the pre-processed MSVOLSO2L4 dataset (Carn, 2022), i.e. volcanic SO_2 emissions derived from satellite observations. The only pre-processing applied was to remove headers to facilitate reading of the dataset.

GVP_ConfirmedEruption_Search_Result_202501121044.xlsx : Excel spreadsheet containing the GVP dataset, obtained from a search for all confirmed eruptions on the GVP Volcanoes of the World database on Jan 12 2025.

CMIP7_presat_list_match.xlsx: Excel spreadsheet containing a list of ice-core derived volcanic SO_2 emission characterized by parameters such as SO_2 mass and date of sulfate deposition, and the proposed eruption match (if any) characterized by parameters such as the GVP volcano and eruption number. It also contains explanations and references to justify the proposed match, as well as a flag indicating the match confidence. This spreadsheet is used in *build_CMIP7_emissions_v221_GMD.m* to match ice-core events to specific eruptions. Note that the same script saves a file *CMIP7_presat_list.xlsx* in the same folder, containing the list of all ice-core-derived, pre-satellite events. This spreadsheet, which is not used directly in the script, was the basis to produce the match spreadsheet *CMIP7_presat_list_match.xlsx*.

CMIP7_ESP.xlsx : Excel spreadsheet containing the list of eruption source parameters, such as volcano latitude and longitude, eruption date, and SO_2 injection height, for as many events as possible among the pre-satellite events used in the CMIP7 database. This includes source references and explanations on ESP choices where required.

geochemical_match_pre1750.xlsx : Excel spreadsheet containing parameters (e.g. SO_2 mass, deposition asymmetry, height) for pre-1750 ice-core sulfate deposition events that were geochemically matched to a known volcanic eruption (considered high-confidence matches). This data is used in *build_CMIP7_emissions_v221_GMD.m* to constrain empirical relationships between parameters constrained from ice-core (e.g. sulfate deposition asymmetry) and parameters for which a value must be attributed if unconstrained (e.g., SO_2 injection latitude).

marshalletal2021.xlsx : Excel spreadsheet containing simulated deposition for a range of eruptions in Marshall et al. (2021). This dataset is not used to build the CMIP7 volcanic emission inventory, but in script *build_CMIP7_emissions_v221_GMD.m,* we compare the ice-core derived relationship between deposition asymmetry and eruption latitude to that simulated by the interactive stratospheric aerosol model used in Marshall et al. (2021).

GloSSAC_V2.22.nc : Netcdf file containing version 2.22 of the GloSSAC dataset (Thomason et al., 2018; Kovilakam et al., 2020).