Unveiling the Dominant Control of the Systematic Cooling Bias in CMIP6 Models: Quantification and Corrective Strategies

Jie Zhang[1,2*], Kalli Furtado[3*], Steven T. Turnock[4,5], Yixiong Lu[1,2], Tongwen Wu[1,2], Fang Zhang[1,2], Xiaoge Xin[1,2], Yuyun Liu[6]

1. *State Key Laboratory of Disaster Weather Science and Technology, CMA EMPC*
2. *CMA Earth System Modeling and Prediction Centre, China Meteorological Administration, Beijing, China*
3. *Centre for climate research Singapore, Singapore*
4. *Met Office Hadley Centre, Exeter, UK*
5. *University of Leeds Met Office Strategic (LUMOS) Research Group, University of Leeds, Leeds, UK*
6. *Center for Monsoon System Research, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

Corresponding to: Jie ZHANG (jiezhang@cma.gov.cn) & Kalli Furtado (Kalli_FURTADO@nea.gov.sg)

Abstract

Including sophisticated aerosol schemes in the models of the sixth Coupled Model Inter-comparison Project (CMIP6) has not improved historical climate simulations. In particular, the models underestimate the surface air temperature anomaly (SATa) when anthropogenic sulfur emissions increased in 1960-1990, making the reliability of the CMIP6 projections questionable. Biases in cooling among the models are correlated with sulfate burden anomaly and the total sulfur sink (including both sulfate and SO_2 depositions) is the process responsible. Accordingly, we define a diagnostic tool, named Sulfur Assessment Metric for Earth system models (τ_{same}), for model evaluation and improvement. Reducing the SATa biases to within the observational uncertainty is consistent with a physically plausible τ_{same} of around 1.35 days, which is overestimated by most of the CMIP6 models. Based on targeting a reduction of τ_{same}, two post-CMIP6 models show greatly improved SATa reproduction. The systematically underestimated sulfate turnover time (τ_{SO_4}) in CMIP6 models suggests that modifying SO_2 deposition rather than sulfate deposition would be a more scientific approach. The strong correlations between sulfate burden anomaly and both SATa and τ_{SO_4}, as well as between SATa and τ_{same}, persist before, during and after the 1960-

1990 period. This temporal persistence confirms the dominant influence of sulfate physical processes across all examined time periods.

1. Introduction

Atmospheric aerosols have rapidly increased since the Industrial Revolution. Over this time period, the total aerosol effective radiative forcing (ERF) was dominated by the sulfate cooling effect, and offset a substantial portion of global-mean forcing from well-mixed greenhouse gases (IPCC, 2023). Without this historical aerosol ERF, the Paris Agreement's target of limiting global warming to 1.5°C above pre-industrial levels would have already been missed in 2015 (Hienola et al., 2018). Similarly, stopping all present-day anthropogenic aerosol emissions is estimated to induce a global-mean surface heating of 0.5-1.1°C (Samset et al., 2018). The year 2024 has been confirmed as the hottest year in human history, and was the first year to breach the 1.5°C warming limit (Bevacqua et al., 2025). Moreover, recent accelerated temperature trends may be attributable to reductions in atmospheric aerosols, particularly from reduced commercial shipping emissions. Hansen et al. (2025) suggest that even small emissions in relatively pristine air have substantial effects, highlighting the crucial need to improve the representation of aerosol effects in global climate models for more reliable projections.

The observed temporal evolution of historical surface air temperature (SAT) is one of the major metrics used for evaluating the performance of climate models. However, the SAT anomalies (SATa) in the CMIP6 models are systematically lower than was observed for the 1960-1990 period, whereas the CMIP5 models, on average, track the instrumental record quite well (e.g., Flynn and Mauritsen, 2020). The 1960–1990 period, when the cooling bias prevailed, is coincident with the so-called Great Acceleration period, during which human activities intensified remarkably and led to global-scale impacts on the Earth System (Steffen et al., 2007). Recent studies hypothesized that aerosol forcing in CMIP6 is stronger than in CMIP5 and is responsible for the suppressed late 20th-century warming (e.g., Dittus et al., 2020; Smith and Forster, 2021).

The cooling anomaly points towards a problem with the sulfur cycle in recent ESMs or the emissions data (Hardacre et al., 2021; Wang et al., 2021). Considering the importance of the sulfur cycle in historical aerosol ERF, we examine the sulfur-related processes in eleven CMIP6 models with aerosol schemes in this study. All the models are forced with CMIP6 historical anthropogenic aerosol emissions (Hoesly et al., 2018), and therefore differences in their sulfate burdens are mainly due to different representations of the sulfur cycle in the models.

We will identify the key processes that determine sulfate burden in these models, and introduce a simple index for measuring the level of activity in the sulfur cycle in the models on the global scale. This index is mainly an effective diagnostic tool for global cycling of atmospheric sulfate, which can be easily calculated from time series of global means only, without the need for complex diagnostics of the sulfur-cycle processes. Our results demonstrate that the index exhibits strong correlation with both sulfate burden anomalies and SATa, allowing each model's sulfur cycle to be calibrated using historical temperature biases.

2. Model, data, and method

2.1 CMIP6 models and data

Table 1. Information of the eleven CMIP6 models with aerosol schemes.

Model	Country	Interactive Chemistry	Members	Reference
BCC-ESM1	China	Yes	3	Wu et al., (2020); Zhang et al., (2021b)
CESM2	US	No	11	Danabasoglu et al. (2020)
CESM2-FV2	US	No	3	Danabasoglu et al. (2020)
EC-Earth3-AerChem	European consortium	Yes	2	Döscher et al. (2021)
GFDL-ESM4	US	Yes	3	Dunne et al.

				(2020)
MIROC6	Japan	No	50	Tatebeet al. (2019)
MIROC-ES2L	Japan	No	30	Hajima et al. (2020)
MPI-ESM-1-2-HAM	Germany	Yes	3	Mauritsen et al. (2019)
MRI-ESM2-0	Japan	Yes	10	Yukimoto et al. (2019)
NorESM2-LM	Norway	Yes	3	Seland et al. (2020)
UKESM1-0-LL	UK	Yes	19	Sellar et al. (2019)

Eleven CMIP6 climate models with interactive aerosol schemes are utilized in this study, including seven models with interactive chemistry and four without (Table 1). The outputs from two CMIP6 experiments are used: (1) the historical experiment of climate change over the period 1850-2014, forced by time-varying external forcings that are based on observations of natural processes (e.g., solar activity, volcanic eruptions) and human-induced changes (e.g., greenhouse gas, aerosol emissions, land-use changes). All the available realizations for each model were used to minimize the uncertainty from internal variability in the climate system; (2) the 1pctCO2 simulations, in which CO_2 is gradually increased at a rate of 1% per year. The 1pctCO2 experiment is designed for studying model responses to CO_2 and is somewhat more realistic than rapidly increasing CO_2, such as in the abrupt-4×CO2 experiment. The historical experiment outputs from two post-CMIP6 models (BCC-ESM1-1 and UKESM1-1-LL) with revised SO_2 deposition parameterizations are also included.

Model outputs used in this study comprise SAT and five key sulfur-cycle variables: sulfate burden (B_{so4}), sulfate wet deposition and dry deposition rate (R_{so4}), sulfur-dioxide (SO_2) wet deposition and dry deposition rate (R_{so2}). For these sulfur-cycle variables, the inter-member variability within the historical experiment is substantially smaller than that of SAT. The standard deviation of B_{so4} in 1960-1990 across the 11 CESM2 members is only 4% of its interannual variability, compared to approximately 21% for SAT. Similar results are also evident in the 19 UKESM1 members, where the standard deviation of B_{so4} is 3% of its interannual variability, versus 32% for SAT. Therefore, given the relatively small inter-member variability in sulfur-cycle variables compared to their interannual fluctuations and to SAT variability, we utilize the first realization of the historical simulations and neglect inter-member differences for these sulfur-related quantities.

The monthly mean SAT from the Met Office Hadley Centre/Climatic Research Unit global surface temperature (HadCRUT) data version 5 from 1850 to 2014 is used for model evaluations (Morice et al., 2021). Considering the lack of long-term reliable observations in polar regions, we focus on SAT changes between 60°S to 65°N and the 'global' mean is calculated as the area-weighted mean in this latitudinal belt.

2.2 The Sulfur Assessment Metric for ESMs (τ_{same}) and sulfate turnover time

Atmospheric sulfate concentrations are determined by the emission and oxidation of sulfate precursors, as well as deposition processes. Anthropogenic SO_2 emissions are the major source of sulfate aerosol over land in polluted regions. Given that all CMIP6 models use identical anthropogenic SO_2 emissions, the inter-model differences in simulating atmospheric sulfate concentrations primarily arise from differences in SO_2-to-sulfate oxidation rates and sulfate deposition velocities. Enhanced SO_2 deposition limits precursor availability for sulfate formation, while accelerated sulfate deposition directly reduces atmospheric loading. This dual mechanism demonstrates how deposition processes govern sulfate concentrations through direct and indirect pathways.

Here we define the **S**ulfur **A**ssessment **M**etric for **E**SMs (τ_{same}) as the ratio of the sulfate burden anomaly and sulfur deposition anomaly, relative to preindustrial period. Sulfur deposition comprises the deposition fluxes of sulfate aerosol and its major precursor SO_2:

$$\tau_{same} = \frac{B_{aSO_4}}{R_{aSO_4} + R_{aSO_2}} \quad (1),$$

where B_{aSO_4} is the total sulfate burden anomaly in the atmosphere; R_{aSO_4} and R_{aSO_2} denotes the total (wet plus dry) sulfate and SO_2 deposition rate anomaly, respectively. We use the anomaly of to mitigate the influence of different model climatologies.

Sulfate turnover time is a physically meaningful index. It reflects the atmospheric residence time of sulfate aerosols in the atmosphere before being scavenged by wet or dry deposition. Sulfate turnover time is defined as:

$$\tau_{SO_4} = \frac{B_{SO_4}}{R_{SO_4}} \quad (2),$$

where τ_{SO_4} denotes the sulfate turnover time; B_{SO_4} is the total sulfate burden in the atmosphere; and R_{SO_4} is the total sulfate deposition.

2.3 The transient Climate Response (TCR) index

The transient Climate Response (TCR) index is calculated as the mean SAT anomaly of a 1pctCO2 simulation in a 20-year period centered on year-number 70, by which a doubling CO_2 concentration has occurred. It is an important metric representing CO_2-related historical warming and has been widely used for model evaluations and comparisons (Bevacqua et al., 2025; O'neill et al., 2016).

3. Results

3.1 SAT and sulfate burden

The historical evolutions of near-global mean (60°S to 65°N) SATa in the eleven CMIP6 models with interactive aerosol schemes are shown in Fig. 1a. All the models tend to underestimate SATa since the 1930s. The cooling anomaly in CMIP6 model marked a notable departure from earlier model generations, which can effectively capture the instrumental SAT record with observation falling well within model spread (e.g., Flynn and Mauritsen, 2020; Hegerl, et al., 2007).

The cooling bias is most pronounced from 1960 to 1990. The SATa is about 0.34°C in the observations. However, the multi-model mean (MMM) SATa is about 0.3°C lower with a large model spread. The SATa ranges from -0.24°C in EC-Earth3-AerChem to 0.19°C in GFDL-ESM4 and MIROC6. The cooling is noticeable at the mid to high latitude in the Northern Hemisphere (as shown in the attached SATa map in Fig.1a). The sudden drop in SATa in the early 1960s and 1990s may be due to the stronger model responses to large volcanic eruptions, Mount Agung in 1963 and Mount Pinatubo in 1991, than in the observations (Chylek et al., 2020). The cooling biases diminish in later periods corresponding to the generally high model sensitivity to greenhouse gas forcing (Smith and Forster, 2021).

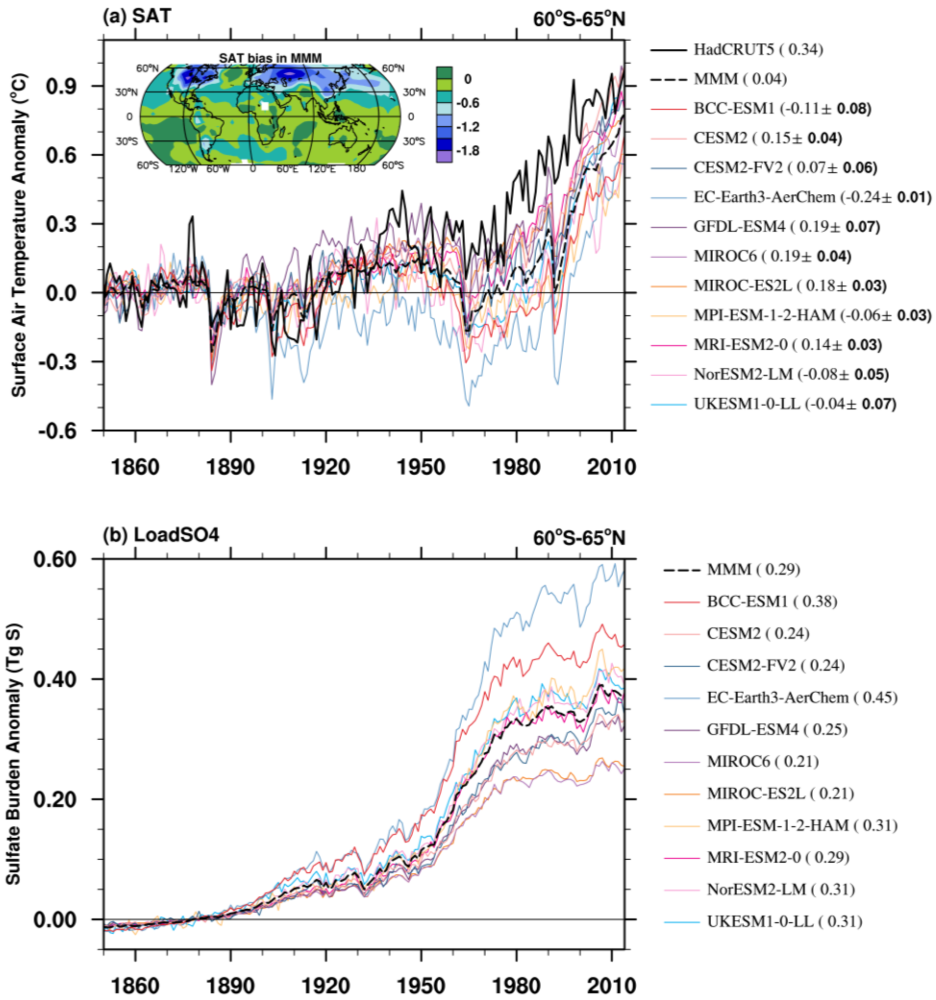

Figure 1. (a) Historical surface air temperature anomalies (SATa) relative to 1850-1900 mean for HadCRUT5 (thick black line), the ensemble mean for each CMIP6 model (solid color lines), and multi-model mean (MMM, dashed black line). The numbers in brackets are the mean results in

1960-1990, together with the inter-member spread for each model. Units: °C. (b) is the same as (a) but for sulfate burden anomalies for the first realization from each CMIP6 model (solid color lines) and MMM (dashed black line). Units: Tg S.

The cooling bias coincides with increased anthropogenic emissions, particularly of sulfate precursors such as SO_2 (Zhang et al., 2021a). Global emissions of SO_2 grew steadily after the 1950s and peaked in the 1970s at 180Tg yr^{-1}, which is about 3.6 times the 1950s' emissions (Hoesly et al., 2018). The increasing emissions of SO_2, as the primary precursor of atmospheric sulfate, have directly contributed to elevated sulfate concentrations in the troposphere. The temporal evolution of sulfate burden demonstrates significant growth trajectories with the anthropogenic emission (Fig.1b), initially driven by industrialization and further accelerated post-1950s mainly due to intensified anthropogenic SO_2 emission from industries and the energy-transformation sectors (e.g., Ohara et al., 2007; Vestreng et al., 2007). The increased sulfate burden interrupted a decades-long warming trend via the cooling effects of sulfate aerosols on climate, even though carbon dioxide emissions continued to rise (Wilcox et al., 2013). Because of the emission control policies in Europe and North America (Hand et al., 2012; Vestreng et al., 2007), such as the Gothenburg Protocol (Eb, 1999) and the 1990 Clean Air Act Amendments in the U.S. (Likens et al., 2001), global anthropogenic SO_2 emissions were suppressed after the 1980s and SAT started to increase rapidly in the observations (Aas et al., 2019). Anthropogenic SO_2 emission continued to increase across Asia due to industrial developments, but has decreased since 2006 in East Asia (Wang et al., 2021). The CMIP6 emission inventory fails to represent the early 21st century SO_2 emission reductions in East Asia. But it is outside of the 1960-1990 period, and the impact on SAT reproduction is beyond the main scope of this paper.

In the 11 CMIP6 models, sulfate concentrations increased rapidly in 1960-1990 (Fig.1b). The systematically underestimated SATa indicate excessively strong sulfate-induced cooling effect in CMIP6 models. The MIROC models exhibit the lowest sulfate burden (0.21 Tg S) and smallest cooling bias (0.15°C below HadCRUT5), while EC-Earth3-AerChem approximately doubles the sulfate loading (0.45 Tg S) and nearly four times the cooling bias (0.58°C below HadCRUT5). Generally, models with larger sulfate burden anomalies also show more pronounced SATa underestimations. As shown in Fig. 2a, the correlation coefficient between sulfate burden anomaly and SATa in 1960-1990 is -0.92, significant at the 1% level using a Student's *t*-test. Interactive chemistry may have an impact on sulfate formation and affect the sulfate aerosol burdens in the atmosphere (Mulcahy et al., 2020). As shown in Fig.2a, models with interactive chemistry (color dots) seem to have higher sulfate burden anomaly and lower SATa than models without (color circles). However, the relationship between sulfate burden anomaly and SATa is consistent among models with and without interactive chemistry. That is, there is no obvious difference in the relationship between sulfate burden anomaly and SATa for models with and without interactive chemistry.

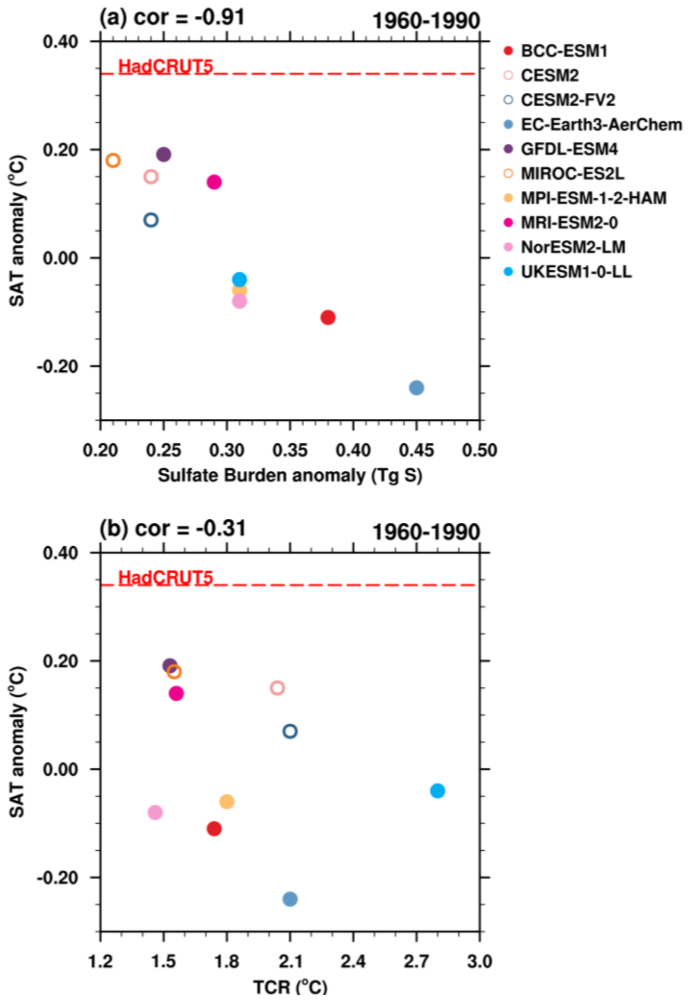

Figure 2. Scatter plots of (a) sulfate burden anomaly in 1960-1990 (Tg S, x-axis) in historical experiments, and (b) the transient climate response (TCR, ºC) versus SATa in 1960-1990 (ºC, y-axis) for each model. The corresponding correlation coefficient (cor) is shown at the top-left corner

of each panel. The anomalies are relative to 1850-1900 mean. Models with and without interactive chemistry are marked by color dots and color circles, respectively.

Greenhouse gases (GHGs) also show a rapidly increasing trend in 1960-1990. However, TCR, which can generally indicate the impact of GHGs, is insignificantly correlated with SAT anomalies in CMIP6 models, and the correlation coefficient across models is even negative (Fig.2b). Therefore, the biases of atmospheric sulfate burden and the associated sulfate aerosol cooling effects play an essential role in the cooling biases in the CMIP6 models.

It should be noticed that there are fast and slow components of global warming in response to radiative forcing changes (Held et al., 2010). A fast component with an exponential decay time scale of less than 5 years, primarily driven by rapid adjustments in the upper ocean layers. A slow component that evolves over centuries, associated with heat uptake by deeper ocean layers. The lagged oceanic and dynamical feedbacks will delay and modulate warming rates (Chen et al., 2016; Watterson and Dix, 2005). In this study, the fast response to sulfate forcing can be rapidly detected by SATa, especially when the sulfate forcing is sustained in 1960-1990. The study's global mean perspective makes the results insensitive to the impact of spatial redistribution of temperature anomalies caused by dynamical feedbacks.

3.2 Sulfur deposition and a metric for the global sulfur cycle diagnostic (τ_{same})

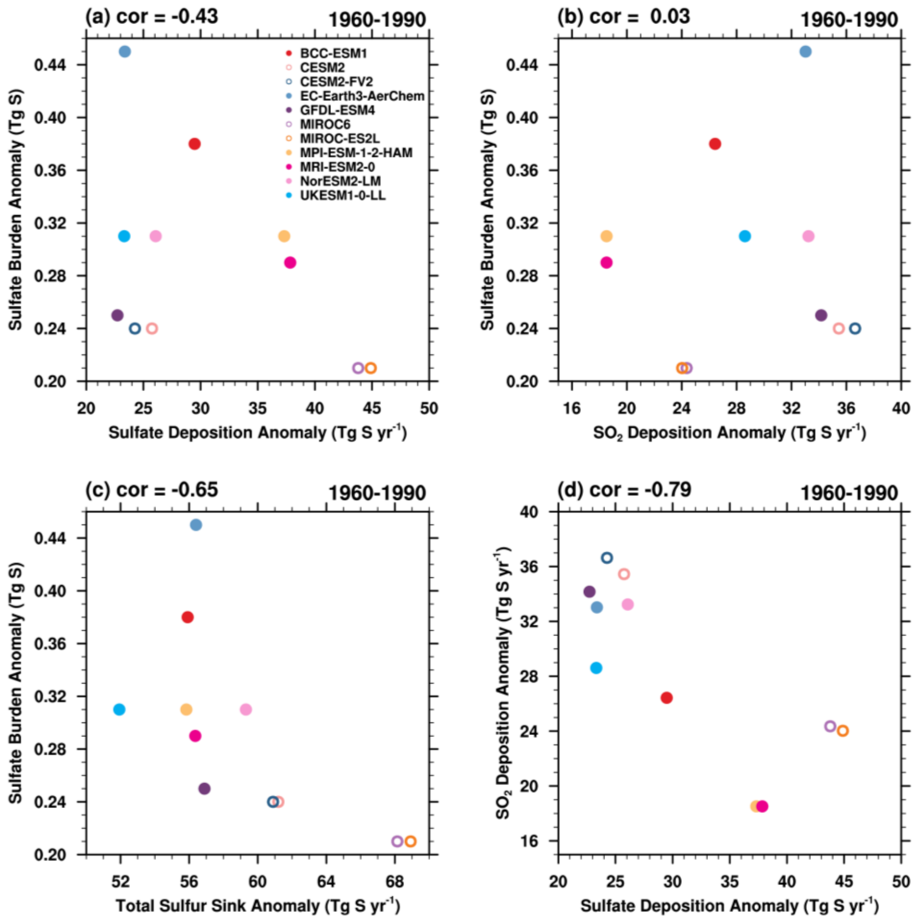

Figure 3. (a) Sulfate deposition anomaly (R_{aSO_4}), (b) SO_2 deposition anomaly (R_{aSO_2}), and (c) total sulfur sink anomaly ($R_{aSO_4} + R_{aSO_2}$) versus sulfate burden anomaly (Tg S, y-axis) in each model during 1960-1990. (d) R_{aSO_4} (x-axis) versus R_{aSO_2} (y-axis) during 1960-1990. Units for deposition anomalies are Tg S yr^{-1}.

Fig. 3 shows the inter-model relationship between global mean anomalies of total sulfate burdens and sulfur deposition during 1960-1990, relative to the 1850-1900 baseline period. As shown in Fig.3a, the sulfate burden anomaly (B_{aSO_4}) is negatively correlated with sulfate deposition anomaly (R_{aSO_4}). However, the correlation is not statistically significant, partly attributable to a subset of five models characterized by low sulfate deposition and low sulfate burden. These models degrade the robustness of the linear fit derived from the remaining models. There is no clear statistical relationship between B_{aSO_4} and SO_2 deposition anomaly (R_{aSO_2}, Fig. 3b). However, the correlation between B_{aSO_4} and total sulfur sink anomaly ($R_{aSO_4} + R_{aSO_2}$) increases to -0.65, significant at the 5% level using a Student's t-test (Fig.3c). Notably, within the subset of five models exhibiting both low B_{aSO_4} and low R_{aSO_4}, most display higher R_{aSO_2} in relative to the ensemble mean. This high R_{aSO_2} compensates for their low R_{aSO_4}, influencing the total sulfur deposition magnitude sufficiently to sustain a significant correlation with B_{aSO_4} in these models. Higher SO_2 deposition rates result in reduced atmospheric SO_2 availability for oxidation to sulfate. That is, both the sulfate deposition and the SO_2 deposition are responsible for the sulfate burden anomalies. Further examination indicates that the SO_2 deposition rate anomaly among the models is highly negatively correlated with the sulfate deposition rate anomaly with a correlation coefficient of -0.79 (Fig.3d).

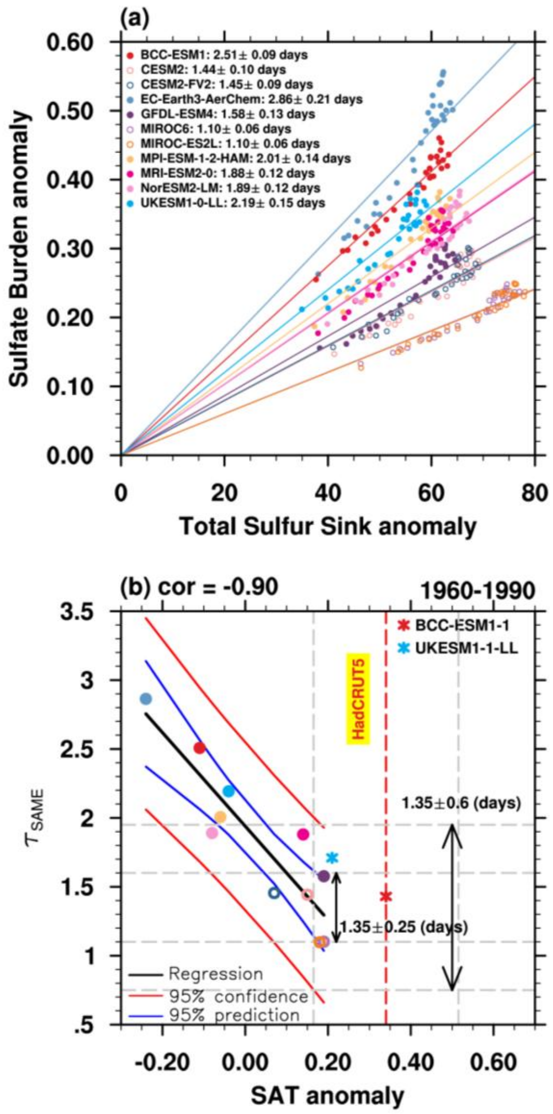

Figure 4. (a) Scatter plots of yearly total sulfur sink anomaly (Tg S yr^{-1}, x-axis) versus sulfate burden anomaly (Tg S, y-axis) in 1960-1990. Number in legend shows the mean and standard deviation of

the ratio between sulfate burden anomaly and total sulfur sink anomaly (τ_{same}, days). (b) The mean SATa (°C, x-axis) versus τ_{same} (days, y-axis) in 1960-1990 for each model. The black solid line is the linear fitting. The blue and red curves are the corresponding 95% confidence interval (CI) and 95% prediction interval (PI). SATa in HadCRUT5 and its 0.175°C boundaries are shown by the red dashed line and parallel gray dashed lines. The red and blue asterisks are the results in the two post-CMIP6 models (BCC-ESM1-1 and UKESM1-1-LL).

Considering the significant influence of total sulfur sink anomalies on sulfate burden anomalies, Fig.4a examines their interannual variability during 1960-1990 in each model. Generally, the sulfate burden anomalies and total sink anomalies are positively correlated and co-vary almost linearly in all the models. The ratio between sulfate burden anomalies and total sulfur sink anomalies is defined as τ_{same} in Section 2.2. The mean τ_{same} in 1960-1990 ranges from 1.1 days in MIROC models to 2.86 days in EC-Earth3-AerChem. τ_{same} is generally longer in models with interactive chemistry (color dot) than without (color circle).

The standard deviation of τ_{same} for each model in 1960-1990 ranges from 0.03 to 0.12 days, about 3.0% of the mean τ_{same}. That is, although the sulfate burden increased significantly in 1960-1990, τ_{same} hardly changed. This is an important sign that τ_{same} is a robust index for evaluating the sulfur cycle in model development. Fig. 4b shows τ_{same} and SATa in 1960-1990 in each model. The SATa is highly correlated with τ_{same} with a correlation coefficient of -0.90. That is, τ_{same} is capable of characterizing the cooling anomaly for each model, making it a convenient target for model tuning.

3.3 The recommended τ_{same} range and performances in the two post-CMIP6 models

Tuning based on τ_{same} requires an empirical best-estimate τ_{same} to aim for. Therefore, a further question is how to estimate the reasonable values for τ_{same}. Here we try to constrain τ_{same} using the SATa in observations. The SATa in 1960-1990 is 0.34°C. Considering the internal variability in the climate system and the uncertainty in observation, the observed uncertainty is suggested to be 0.175°C. The observed uncertainty is estimated as the standard deviation of the observed annual mean globally averaged SAT in HadCRUT5 from 1850 to 2014 after removing the least squares linear trend. SATa in seven CMIP6 models falls beyond the observational range. SATa closely approaches the lower bound of observation in the remaining four models, giving a range of τ_{same} between 1.1 to 1.58 days.

Here we use this metric to modify the sulfur cycle in BCC-ESM1, more specifically, we quadruple the SO_2 dry deposition over land and multiply the SO_2 dry deposition over the ocean by 1.5. This effect is similar to that in UKESM1-0-LL by modifying SO_2 dry deposition parameterization (Hardacre et al., 2021; Mulcahy et al., 2023). The impact of changes to the SO_2 dry deposition parameterization in UKESM1-0-LL is an increase of SO_2 dry deposition by a factor of 2 to 4. As indicated by the red and blue asterisks in Fig.4b, τ_{same} reduced from 2.51 to 1.43 days in BCC-ESM1-1 accompanied by a 0.45°C SATa increase, and reduced from 2.19 to 1.71 days in UKESM1-1-LL with a concurrent 0.25°C SATa rise. The SATa from both post-CMIP6 models falls within the observational uncertainty ranges.

We perform linear fitting between SATa and τ_{same} (black line in Fig. 4b), along with the 95% confidence interval (CI, blue curves), and the 95% prediction interval (PI, red curves) across the eleven CMIP6 models. Given that most models underestimate SATa relative to observations, extrapolating τ_{same} for SATa exceeding the lower bound of observation becomes highly uncertain. Results from BCC-ESM1-1 suggest that the rate of decrease in τ_{same} predicted by the regression line may not hold for SATa values above the observed lower bound (0.165°C). Therefore, we estimate τ_{same} by the linear fitting at the observed lower bound, yielding a central τ_{same} estimate of 1.35 days. Critically, this value carries inherent uncertainties that must be quantified:

- The 95% confidence interval (CI) of ±0.25 days (i.e., 1.10–1.60 days).

- The wider 95% prediction interval (PI) of ±0.6 days (i.e., 0.75–1.95 days).

The substantial difference between the CI and PI ranges underscores the challenge in precisely constraining τ_{same}. We advise using the PI for applications requiring robustness against individual model deviations.

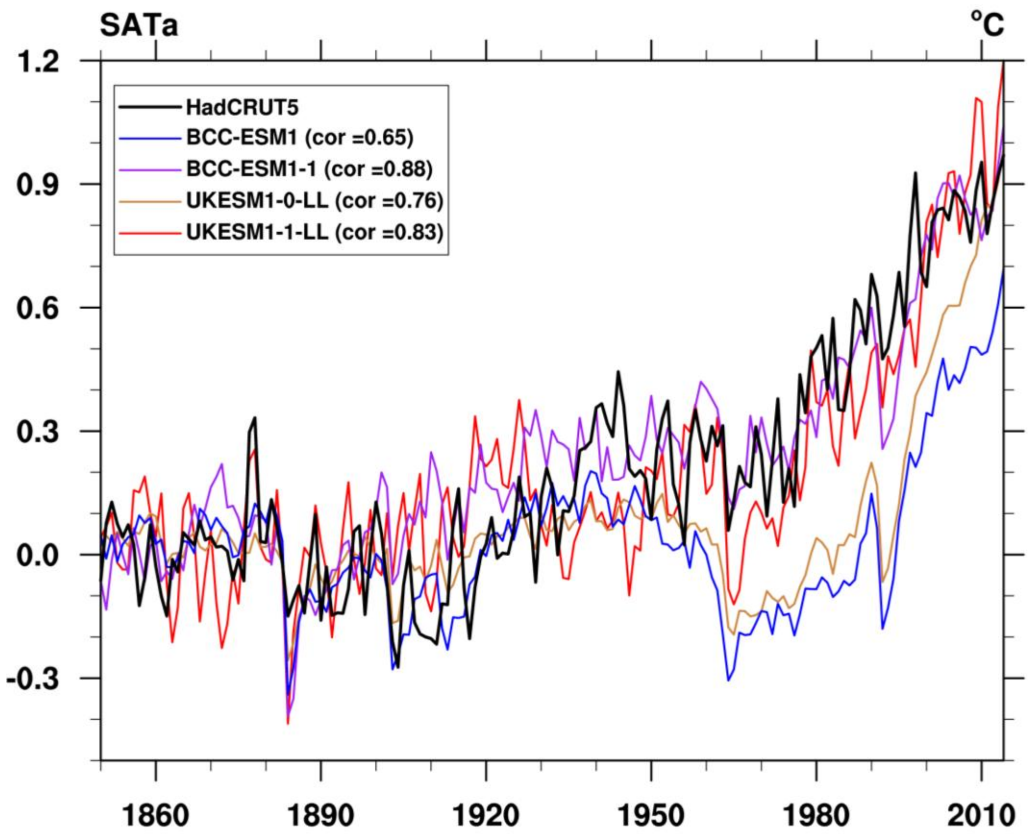

Figure 5. Evolutions of SATa relative to 1850-1900 mean for HadCRUT5, BCC-ESM models, and UKESM models. The numbers in legend are the corresponding correlation coefficients with HadCRUT5.

As demonstrated by the global mean SATa in BCC-ESM1-1 and UKESM1-1-LL models (Fig.5), both models on average tracked the instrumental record quite well with statistically higher correlation coefficients with observation (HadCRUT5). That is, improvements in sulfur deposition parameterizations, which reduced τ_{same}, improved the representation of historical surface temperature evolution.

3.4 Sulfate turnover time and dominant sulfur deposition

τ_{same} can be derived from fundamental model output variables, ensuring straightforward calculation across all models. While we recognize that τ_{same} is a composite metric incorporating both SO_2 and sulfate deposition rather than a physical timescale, here we examine two metrics with clear physical interpretations to identify the dominant physical processes.

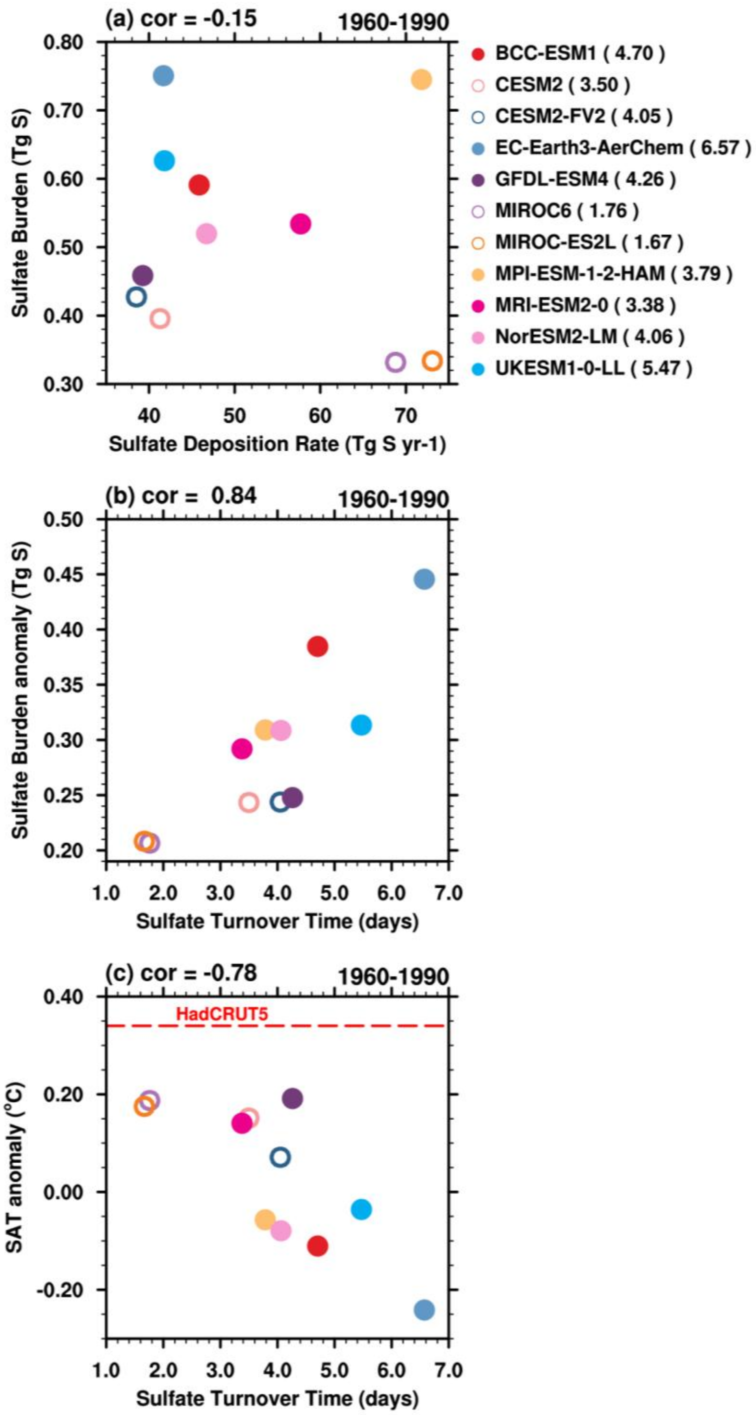

Figure 6. (a) Sulfate deposition versus sulfate burden in 1960-1990. (b) Sulfate turnover time versus sulfate burden anomaly in 1960-1990. (c) Sulfate turnover time versus SATa in 1960-1990.

Figure 6 presents the sulfate deposition and sulfate burden in 1960-1990. A weak negative correlation (-0.15) indicates that sulfate deposition alone cannot fully explain sulfate burden differences in these simulations. Sulfate turnover time is critical for validating the model capability in representing the sulfate cycle. It is quantified following Eq. (2) in Section 2.2 as the ratio of sulfate burden to sulfate deposition, representing the average atmospheric residence time of sulfate aerosols. The sulfate turnover time exhibits considerable inter-model variability, ranging from 1.67 days in MIROC-ES2L to 6.57 days in EC-Earth3-AerChem. These results generally agree with most aerosol models, which typically simulate sulfate lifetimes of around 4 days (e.g., Textor et al.,2006; Liu et al., 2012; Matsui and Mahowald, 2017; Tegen et al., 2019). However, sulfate turnover time in models is notably shorter than observational estimates, 7.3 days (0.02 yr) in Charlson et al. (1992) and 10-14 days in Kristiansen et al. (2012). This discrepancy may stem from premature removal processes, inadequate poleward transport, or incomplete chemical representations (e.g., Croft et al., 2014).

The inter-model variations in sulfate turnover time exhibit a strong correlation with sulfate burden anomalies and SAT anomaly during the 1960-1990 period, with a correlation coefficient of 0.84 and -0.78 (Fig.6b and Fig.6c). This suggests that differences in sulfate turnover time may account for both the sulfate burden anomaly variations and the consequent surface temperature differences among models. However, CMIP6 models systematically overestimate sulfate burden anomalies, implying that these models should exhibit shorter lifetimes to produce lower sulfate burden anomalies (Fig.6c). This would further exacerbate the existing underestimation of sulfate turnover time in CMIP6 models. Thus, enhancing sulfate deposition to mitigate burden anomalies is not an appropriate solution.

Sulfate turnover time in the two post-CMIP6 models, 8.53 days in BCC-ESM1-1 and 5.77 days in UKESM1-1-LL, is generally longer than that of their CMIP6 versions. The longer sulfate lifetimes in the two post-CMIP6 models may be due to lower SO_2 in these revised models, but also could be due to physical climate changes (e.g., temperatures, clouds, rainfall).

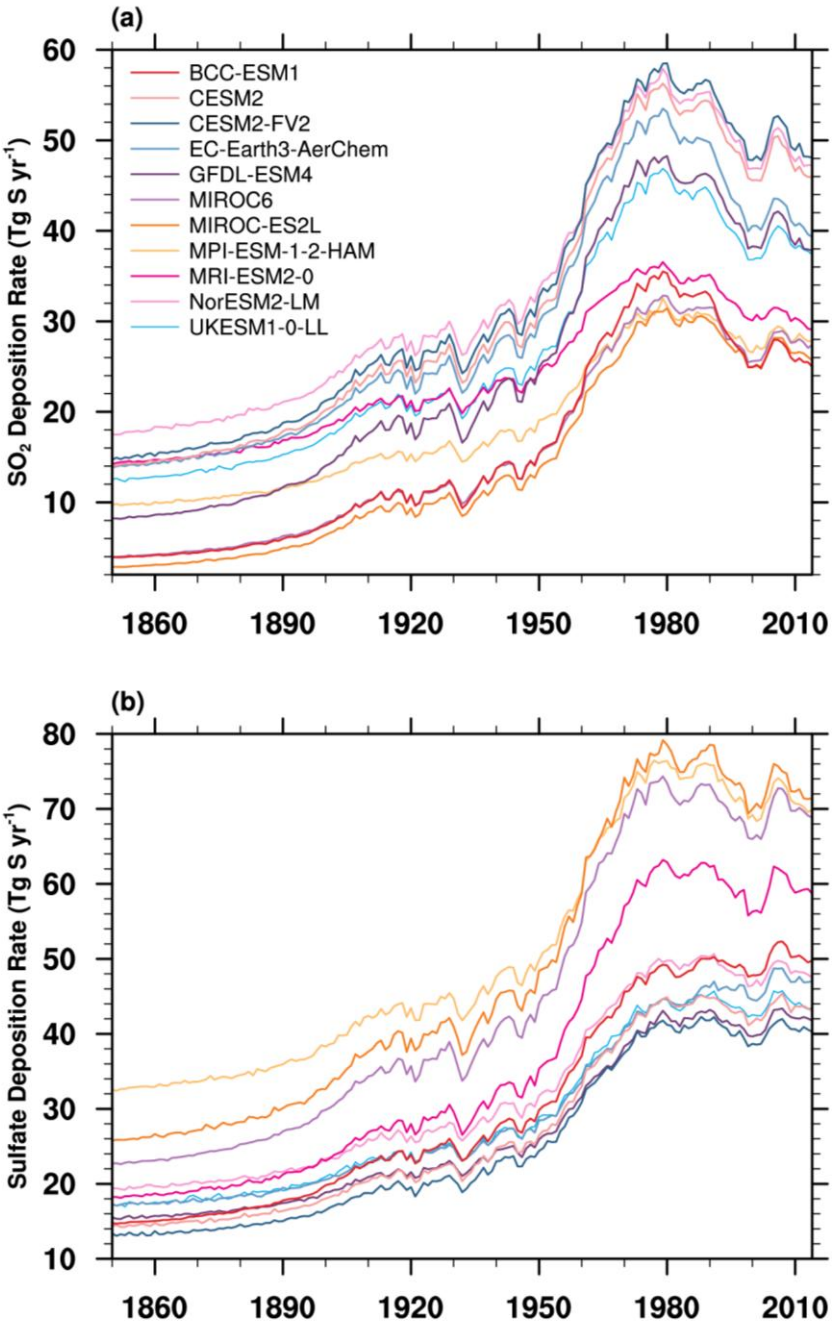

Figure 7. Evolution of (a) the SO_2 deposition rate (R_{SO_2}), and (b) the sulfate deposition rate (R_{SO_4}).

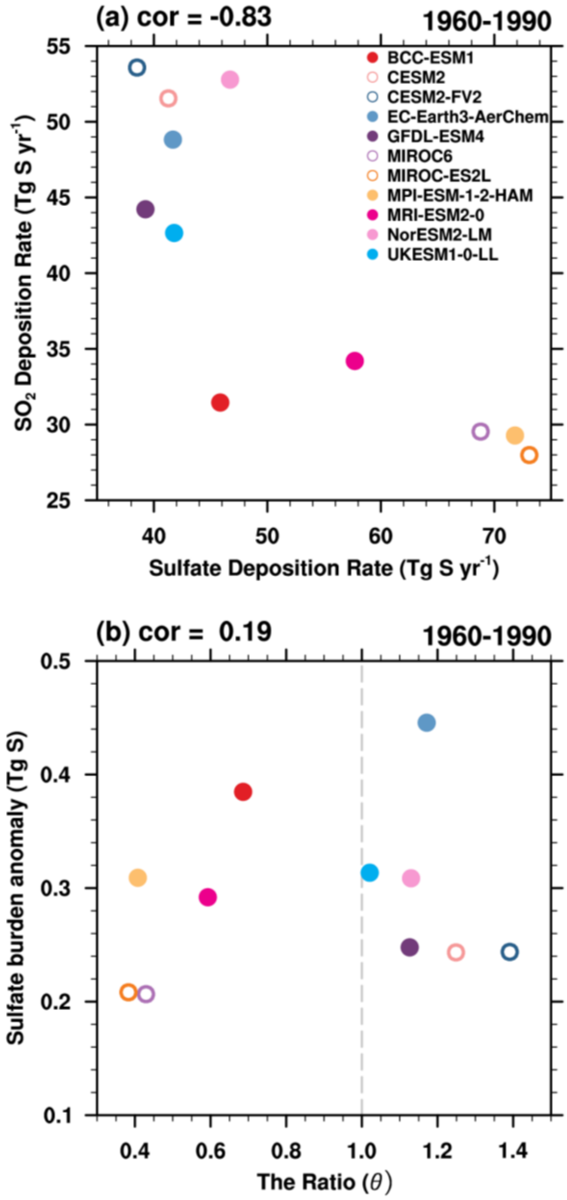

Figure 8. (a) SO_2 deposition rate (R_{SO_2}) versus sulfate deposition rate (R_{SO_4}) in 1960-1990. (b) Ratio between R_{SO_2} and R_{SO_4} (θ, *x*-axis) versus sulfate burden anomaly (y-axis) in 1960-1990.

Considering the importance of sulfate and SO_2 deposition to sulfate burden changes, we further examine their dominance in CMIP6 models. The temporal evolutions of SO_2 and sulfate depositions exhibit a clear dependence on the anthropogenic SO_2 emission across the 11 CMIP6 models (Fig.7). Notably, models with higher sulfate deposition rates generally show lower SO_2 deposition rates. Their correlation is significant with a correlation coefficient of -0.83 in 1960-1990 (Fig. 8a).

We examine the dominance of sulfate and SO_2 deposition by their ratio:

$$\theta = \frac{R_{SO_2}}{R_{SO_4}} \quad (3),$$

where R_{SO_2} and R_{SO_4} denotes the deposition rate of SO_2 and sulfate, respectively. The ratio between SO_2 deposition and sulfate deposition (θ) also varies across CMIP6 models (Fig.8b). In 5 out of 11 models, sulfate removal dominates over SO_2 removal, while the opposite is true in the remaining four. The ratio θ, which reflects the relative dominance of sulfur deposition, does not exhibit a clear relationship with sulfate burden anomalies. This suggests that θ is not a major source of inter-model discrepancy in sulfate burden anomalies.

Sulfur oxidation rate is also a physically meaningful metric, which quantifies the efficiency of atmospheric conversion of SO_2 to sulfate and indicates the degree of secondary aerosol formation. However, it cannot be evaluated precisely since this variable is not available for most CMIP6 models.

4. Relative changes preceding and following the 1960-1990 period

Our analysis reveals a robust correlation between sulfate burden anomalies and SATa in 1960-1990 (Fig. 2a). To assess temporal consistency, we examined these relationships before and after this period. Since their relationship reflects clear underlying physics, we expected similar correlations across periods. As shown in Fig.9, statistically significant correlations are evident in both periods, suggesting that sulfate burden anomaly was overestimated prior to the 1960-1990 interval, and this overestimation continued to influence SATa in subsequent years. The SATa is underestimated by 0.11°C during 1930-1959 and by 0.31°C during 1991-2014 in MMM relative to HadCRUT5. The correlations between sulfate burden anomaly and SATa (-0.79 and -0.78) are weaker than that in 1960–1990 (-0.91). This may be potentially due to the smaller sulfate burden bias during the 1930-1959 interval. The combined effects of high model sensitivity in CMIP6 models (Hausfather et al., 2022) and the atmospheric CO_2 accumulation since the Industrial Revolution may partially offset the cooling bias during the 1991-2014 period.

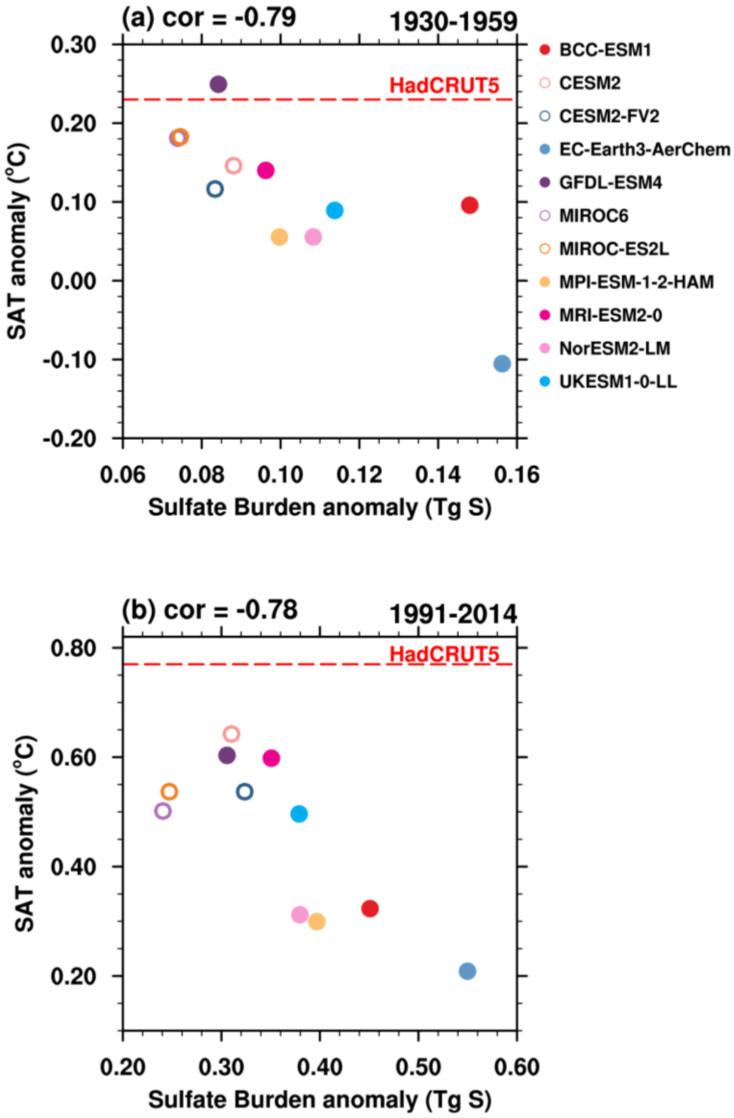

Figure 9. Correlate sulfate burden anomalies with SATa in (a) 1930-1959, and (b) 1991-2014.

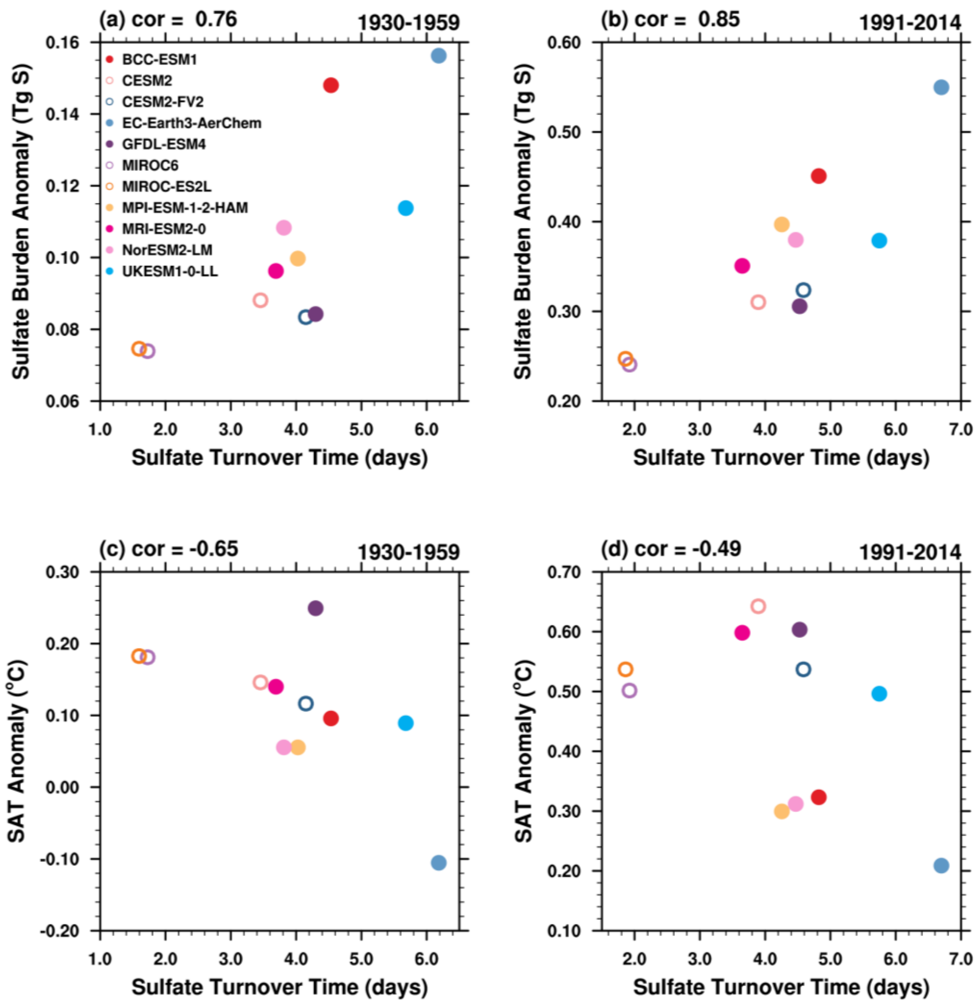

Figure 10. Correlation between sulfate turnover time (τ_{SO_4}) and: (a, b) sulfate burden anomalies, and (c, d) SATa for the periods 1930-1959 and 1991-2014.

Sulfate turnover time serves as a key parameter influencing sulfate burden variations and exhibits a strong correlation with SATa in 1960–1990 (Figs. 6b and 6c). Significant correlations between sulfate turnover time and both sulfate burden anomalies and SATa persist both before and after this period (Fig. 10), confirming the dominant role of sulfate physical processes across all examined time periods.

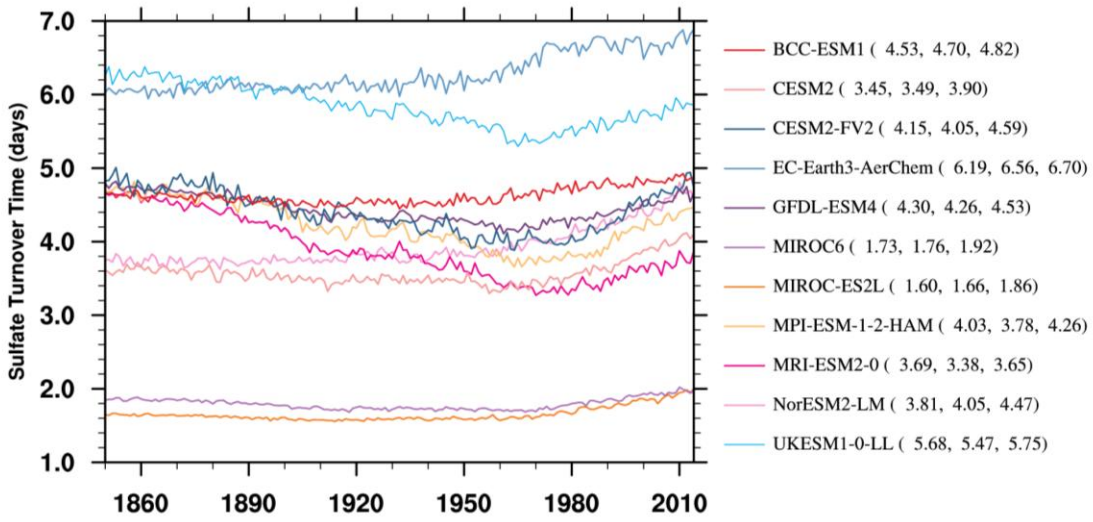

Figure 11. Temporal evolution of sulfate turnover time (τ_{SO_4}) in CMIP6 models. Numerical labels denote mean τ_{SO_4} value during 1930-1959, 1960-1990, and 1991-2014.

We also examine the temporal evolution of sulfate turnover time (Fig.11). Its temporal variability, indicated by the standard deviation ($\sigma < 0.5$ days), is substantially smaller than the inter-model spread. For 1930-1959, models exhibit divergent changes with 5 of 11 models simulating reduced turnover times in the following period. Conversely, all models show prolonged turnover times during 1991-2014 relative to earlier periods. This may be partly due to the shift in the regional distribution of sulfur emissions, with an increasing proportion of emissions from Asia and stringent emission control policies in Europe and North America.

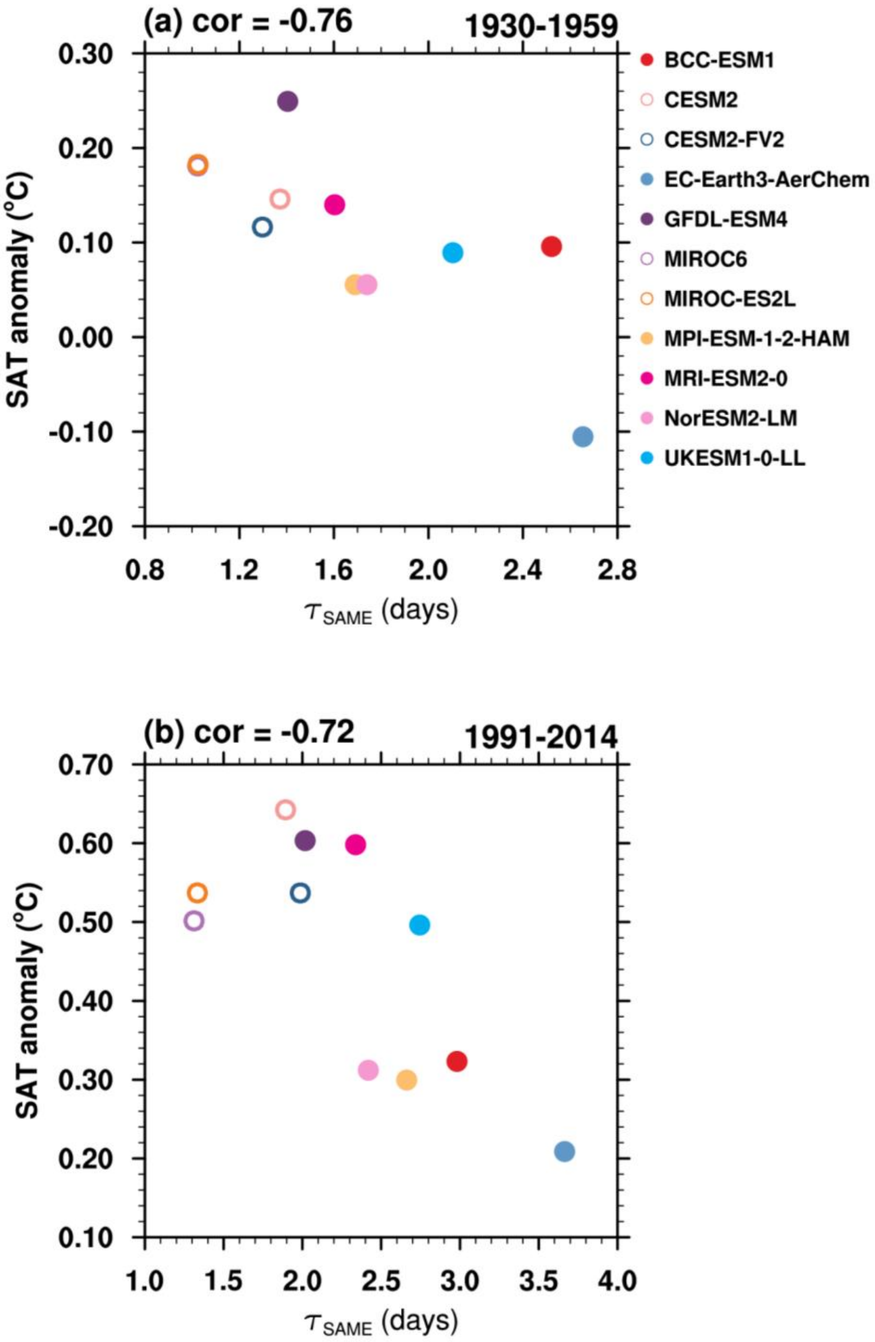

Figure 12. Correlation between τ_{same} and SATa over (a) the 1930-1959 period, and (b) the 1991-2014 period.

The τ_{same} in CMIP6 models show comparable magnitudes between 1930-1959 and 1960-1990, followed by a modest amplification during 1991-2014 (Fig. 12). This may partially reflect spatial shifts in emission sources, analogous to sulfate turnover time (Fig.11). Notably, τ_{same} maintains more significant correlations with SATa in both periods (-0.76 and -0.72) compared to sulfate turnover time (-0.65 and -0.49 in Fig.10c and Fig.10d), mirroring the relationship seen during the 1960-1990 period (-0.90 in Fig.4b versus -0.78 in Fig. 6c).

5. Conclusions

Aerosol cooling effect is considered to be the second most important anthropogenic forcing over the 20th Century. Our study, based on the 11 CMIP6 models with aerosol schemes, demonstrates that the cooling bias in 1960-1990 is closely related to the sulfate burden changes in the atmosphere. Sulfate burden anomaly in the models, and hence the strength of the cooling bias, is determined by sulfur deposition. We introduce a metric, called τ_{same}, which incorporates the effects of sulfur removal processes on sulfate concentration. The index is highly correlated with cooling and can be used to constrain sulfur removal processes in models, on a global scale. Sulfate turnover time is critical for validating the model's physical realism and is further examined to ensure model credibility. Analysis of sulfate turnover time compared with observational measurement demonstrates that increasing sulfate deposition to reduce sulfate burden anomalies is not a reasonable approach.

A constraint on τ_{same}, derived from observed SATa, is used to improve SO_2 deposition parameterizations in models. The modifications in BCC-ESM1 and UKESM1-0-LL lead to shortened τ_{same} values and improved SATa simulations. The optimal τ_{same} is 1.35 days, with a 95% confidence interval (CI) of ±0.25 days and a 95% prediction interval (PI) of ±0.6 days.

Analyses both preceding and following the 1960-1990 period indicate the persistent dominance of sulfate physical processes across all examined time periods. Therefore, the models are likely to underestimate the rate of warming in future climate projections. This has potential implications for the use of CMIP6 in scenarios that incorporate clean-air measures to inform the Paris Agreement goals of limiting warming to below 2 or 1.5°C, i.e., SSP1-2.6 and SSP1-1.9 in CMIP6 (O'neill et al., 2016). Generally, τ_{same} introduced in this study provides a tunable measurement, which can be directly calculated from basic model output. It can effectively guide modifications to sulfur processes, ensuring that models do not overestimate the sulfate cooling effect over the historical period, as was the case in CMIP6 and is a current concern for model performance in the upcoming CMIP7.

Code availability

All data processing codes are available if a request is sent to the corresponding authors.

Data availability

The HadCRUT5 dataset is accessible through Met Office Hadley Centre observations database (https://www.metoffice.gov.uk/hadobs/hadcrut5/). All the model data can be freely downloaded from the Earth System Grid Federation (ESGF) nodes (https://aims2.llnl.gov/search/cmip6/).

Author contributions

The main ideas were formulated by J.Z. and K.F. J.Z. wrote the original draft. The results were supervised by K.F. and S.T.T. All the authors discussed the results and contributed to the final manuscript.

Competing interests

The authors declare no competing financial and/or non-financial interests.

Acknowledgements

We sincerely appreciate the constructive and insightful comments from Dr. Stephen E. Schwartz and anonymous reviewers, which significantly improved the quality of this manuscript. We acknowledge all data developers, their managers, and funding agencies for the datasets used in this study, whose work and support are essential to us.

Financial support

This work was jointly supported by the National Natural Science Foundation of China (Grant no. 42341202 and 42230608) and the UK–China Research and Innovation

Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund.

References

Aas, W., Mortier, A., Bowersox, V., Cherian, R., Faluvegi, G., Fagerli, H., Hand, J., Klimont, Z., Galy-Lacaux, C., Lehmann, C. M. B., Myhre, C. L., Myhre, G., Olivié, D., Sato, K., Quaas, J., Rao, P. S. P., Schulz, M., Shindell, D., Skeie, R. B., Stein, A., Takemura, T., Tsyro, S., Vet, R., and Xu, X.: Global and regional trends of atmospheric sulfur, Scientific Reports, 9, 953, 10.1038/s41598-018-37304-0, 2019.

Bevacqua, E., Schleussner, C., and Zscheischler, J.: A year above 1.5 °C signals that Earth is most probably within the 20-year period that will reach the Paris Agreement limit, Nature Climate Change, 1-4, 10.1038/s41558-025-02246-9, 2025.

Chen, J. P., Chen, I. J. and Tsai, I. C.: Dynamic Feedback of Aerosol Effects on the East Asian Summer Monsoon, Journal of Climate, 29, 6137-6149. 2016.

Chylek, P., Folland, C., Klett, J. D., and Dubey, M. K.: CMIP5 Climate Models Overestimate Cooling by Volcanic Aerosols, Geophysical Research Letters, 47, e2020GL087047, https://doi.org/10.1029/2020GL087047, 2020.

Croft, B., Pierce, J. R. and Martin, R. V.: Interpreting aerosol lifetimes using the GEOS-Chem model and constraints from radionuclide measurements, Atmospheric Chemistry and Physics, 14, 4313-4325. 2014.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Syst., 12, 35, 10.1029/2019ms001916, 2020.

Dittus, A. J., Hawkins, E., Wilcox, L. J., Sutton, R. T., Smith, C. J., Andrews, M. B., and Forster, P. M.: Sensitivity of Historical Climate Simulations to Uncertain Aerosol Forcing, Geophysical Research Letters, 47, e2019GL085806, 10.1029/2019gl085806, 2020.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., Bernadello, R., Bousetta, S., Caron, L. P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J. P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6, Geosci. Model Dev. Discuss., 2021, 1-90, 10.5194/gmd-2020-446, 2021.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, J. Adv. Model. Earth Syst., 12, 10.1029/2019ms002015, 2020.

EB, U.: Protocol to Abate Acidification, Eutrophication and Ground-level Ozone, 1999.

Flynn, C. M. and Mauritsen, T.: On the climate sensitivity and historical warming evolution in recent coupled model ensembles, Atmos. Chem. Phys., 20, 7829-7842, 10.5194/acp-20-7829-2020, 2020.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, Geoscientific Model Development, 13, 2197-2244, 10.5194/gmd-13-2197-2020, 2020.

Hand, J. L., Schichtel, B. A., Malm, W. C., and Pitchford, M. L.: Particulate sulfate ion concentration and SO2 emission trends in the United States from the early 1990s through 2010, Atmos. Chem. Phys., 12, 10353-10365, 10.5194/acp-12-10353-2012, 2012.

Hansen, J., Kharecha, P., Sato, M., Tselioudis, G., Kelly, J., Bauer, S., Ruedy, R., Jeong, E., Jin, Q., Rignot, E., Velicogna, I., Schoeberl, M., Schuckmann, K., Amponsem, J., Cao, J., Keskinen, A., Li, J., and Pokela, A.: Global Warming Has Accelerated: Are the United Nations and the Public Well-Informed?, Environment: Science and Policy for Sustainable Development, 67, 6-44, 10.1080/00139157.2025.2434494, 2025.

Hardacre, C., Mulcahy, J. P., Pope, R. J., Jones, C. G., Rumbold, S. T., Li, C., Johnson, C., and Turnock, S. T.: Evaluation of SO2, SO42- and an updated SO2 dry deposition parameterization in the United Kingdom Earth System Model, Atmospheric Chemistry and Physics, 21, 18465-18497, 10.5194/acp-21-18465-2021, 2021.

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W. and Zelinka, M.: Climate simulations: recognize the 'hot model' problem, Nature, 605, 26-29. 2022.

Hegerl, G.C., F. W. Zwiers, P. Braconnot, N.P. Gillett, Y. Luo, J.A. Marengo Orsini, N. Nicholls, J.E. Penner and P.A. Stott, 2007: Understanding and Attributing Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge

University Press, Cambridge, United Kingdom and New York, NY, USA.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F. R. and Vallis, G. K.: Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing, Journal of Climate, 23, 2418-2427. 2010.

Hienola, A., Partanen, A.-I., Pietikäinen, J.-P., O'Donnell, D., Korhonen, H., Matthews, H. D., and Laaksonen, A.: The impact of aerosol emissions on the 1.5 °C pathways, Environmental Research Letters, 13, 044011, 10.1088/1748-9326/aab1b2, 2018.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J. I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369-408, 10.5194/gmd-11-369-2018, 2018.

IPCC. Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 10.1017/9781009157896, 2023.

Likens, G. E., Butler, T. J., and Buso, D. C.: Long- and short-term changes in sulfate deposition: Effects of the 1990 Clean Air Act Amendments, Biogeochemistry, 52, 1-11, 10.1023/a:1026563400336, 2001.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J. F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G. and Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geoscientific Model Development, 5, 709-739. 2012.

Matsui, H. and Mahowald, N.: Development of a global aerosol model using a two-dimensional sectional method: 2. Evaluation and sensitivity simulations, Journal of Advances in Modeling Earth Systems, 9, 1887-1920. 2017.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Flaeschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Moebis, B., Muellner, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, J. Adv. Model. Earth Syst., 11, 998-1038, 10.1029/2018ms001400, 2019.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, Journal of Geophysical Research-Atmospheres, 126, 10.1029/2019jd032361, 2021.

Mulcahy, J. P., Jones, C. G., Rumbold, S. T., Kuhlbrodt, T., Dittus, A. J., Blockley, E. W., Yool, A., Walton, J., Hardacre, C., Andrews, T., Bodas-Salcedo, A., Stringer, M., de Mora, L., Harris, P., Hill, R., Kelley, D., Robertson, E., and Tang, Y.: UKESM1.1: development and evaluation of an updated configuration of the UK Earth System Model, Geosci. Model Dev., 16, 1569-1600, 10.5194/gmd-16-1569-2023, 2023.

Mulcahy, J. P., Johnson, C., Jones, C. G., Povey, A. C., Scott, C. E., Sellar, A., Turnock, S. T., Woodhouse, M. T., Abraham, N. L., Andrews, M. B., Bellouin, N., Browse, J., Carslaw, K. S., Dalvi, M., Folberth, G. A., Glover, M., Grosvenor, D. P., Hardacre, C., Hill, R., Johnson, B., Jones, A., Kipling, Z., Mann, G., Mollard, J., O'Connor, F. M., Palmieri, J., Reddington, C., Rumbold, S. T., Richardson, M., Schutgens, N. A. J., Stier, P., Stringer, M., Tang, Y., Walton, J., Woodward, S., and Yool, A.: Description and evaluation of aerosol in UKESM1 and HadGEM3-

GC3.1 CMIP6 historical simulations, Geoscientific Model Development, 13, 6383-6423, 10.5194/gmd-13-6383-2020, 2020.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461-3482, 10.5194/gmd-9-3461-2016, 2016.

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980-2020, Atmospheric Chemistry and Physics, 7, 4419-4444, 10.5194/acp-7-4419-2007, 2007.

Samset, B. H., Sand, M., Smith, C. J., Bauer, S. E., Forster, P. M., Fuglestvedt, J. S., Osprey, S., and Schleussner, C.-F.: Climate Impacts From a Removal of Anthropogenic Aerosol Emissions, Geophysical Research Letters, 45, 1020-1029, https://doi.org/10.1002/2017GL076079, 2018.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y. C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geosci. Model Dev., 13, 6165-6200, 10.5194/gmd-13-6165-2020, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the UK Earth System Model, J. Adv. Model. Earth Syst., 11, 4513-4558, 10.1029/2019ms001739, 2019.

Smith, C. J. and Forster, P. M.: Suppressed Late-20th Century Warming in CMIP6 Models Explained by Forcing and Feedbacks, Geophysical Research Letters, 48, 10.1029/2021gl094948, 2021.

Steffen, W., Crutzen, P. J., and McNeill, J. R.: The Anthropocene: Are humans now overwhelming the great forces of nature, Ambio, 36, 614-621, 10.1579/0044-7447(2007)36[614:taahno]2.0.co;2, 2007.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'Ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geoscientific Model Development, 12, 2727-2765, 10.5194/gmd-12-2727-2019, 2019.

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., Rast, S., Kokkola, H., Schultz, M., Schroeder, S., Daskalakis, N., Barthel, S., Heinold, B. and Lohmann, U.: The global aerosol-climate model ECHAM6.3-HAM2.3-Part 1: Aerosol evaluation, Geoscientific Model Development, 12, 1643-1677. 2019.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T. and Tie, X.: Analysis and quantification of the diversities of aerosol life cycles within AeroCom, Atmospheric Chemistry and Physics, 6, 1777-1813. 2006.

Vestreng, V., Myhre, G., Fagerli, H., Reis, S., and Tarrasón, L.: Twenty-five years of continuous sulphur dioxide emission reduction in Europe, Atmos. Chem. Phys., 7, 3663-3681, 10.5194/acp-7-3663-2007, 2007.

Wang, Z., Lin, L., Xu, Y., Che, H., Zhang, X., Zhang, H., Dong, W., Wang, C., Gui, K., and Xie, B.: Incorrect Asian aerosols affecting the attribution and projection of regional climate change in CMIP6 models, Npj Climate and Atmospheric Science, 4, 10.1038/s41612-020-00159-2, 2021.

Watterson, I. G. and Dix, M. R.: Effective sensitivity and heat capacity in the response of climate models to greenhouse gas and aerosol forcings, Quarterly Journal of the Royal Meteorological Society, 131, 259-279. 2005.

Wilcox, L. J., Highwood, E. J., and Dunstone, N. J.: The influence of anthropogenic aerosol on multi-decadal variations of historical global climate, Environmental Research Letters, 8, 10.1088/1748-9326/8/2/024033, 2013.

Wu, T., Zhang, F., Zhang, J., Jie, W., Zhang, Y., Wu, F., Li, L., Yan, J., Liu, X., Lu, X., Tan, H., Zhang, L., Wang, J., and Hu, A.: Beijing Climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations, Geosci. Model Dev., 13, 977-1005, 10.5194/gmd-13-977-2020, 2020.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, Journal of the Meteorological Society of Japan, 97, 931-965, 10.2151/jmsj.2019-051, 2019.

Zhang, J., Furtado, K., Turnock, S. T., Mulcahy, J. P., Wilcox, L. J., Booth, B. B., Sexton, D., Wu, T., Zhang, F., and Liu, Q.: The role of anthropogenic aerosols in the anomalous cooling from 1960 to 1990 in the CMIP6 Earth System Models, Atmos. Chem. Phys. Discuss., 1-39, 10.5194/acp-2021-570, 2021a.

Zhang, J., Wu, T., Zhang, F., Furtado, K., Xin, X., Shi, X., Li, J., Chu, M., Zhang, L., Liu, Q., Yan, J., Wei, M., and Ma, Q.: BCC-ESM1 Model Datasets for the CMIP6 Aerosol Chemistry Model Intercomparison Project (AerChemMIP), Advances in Atmospheric Sciences, 38, 317-328, 10.1007/s00376-020-0151-2, 2021b.