The dim and distant past: Constraining aerosol forcing history in the 20th century

Kine Onsum Moseid

Dissertation for the degree of Philosophiae Doctor (PhD)

Section for Meteorology and Oceanography

Department of Geosciences

University of Oslo

December 2021

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2490

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

:Preface

This synthesis and collection of papers are submitted for the degree of pilosophiae doctor (PhD) in atmospheric physics and chemistry at the Section for Meteorology and Oceanography (MetOs), Department of Geosciences, University of Oslo. The work has been performed from September 2018 until December 2021 at The Norwegian Meteorological Institute (MetNo) which also provided the funding. The research has been supervised by Michael Schulz (MetNo) and Trude Storelvmo (MetOs). The thesis consists of an introduction part and the following papers. Summary of all three papers, including author contributions, are specified in Chapter 4 of the introduction part.

Paper I: **Kine Onsum Moseid**, Michael Schulz, Trude Storelvmo, Ingeborg Rian Julsrud, Dirk Olivié, Pierre Nabat, Martin Wild, Jason N.S. Cole, Toshihiko Takemura, Naga Oshima, Susanne E. Bauer, Guillaume Gastineau, (2020), "Bias in CMIP6 models as compared to observed regional dimming and brightening", Atmospheric Chemistry and Physics, doi:10.5194/acp-20-16023-2020

Paper II: **Kine Onsum Moseid**, Michael Schulz, Anja Eichler, Margit Schwikowski, Joseph R. McConnell, Dirk Olivié, Allison S. Criscitiello, Karl J. Kreutz, Michel Legrand, (2021), "Using ice cores to evaluate CMIP6 aerosol concentrations over the historical era", Journal of Geophysical Research, *in review*

Paper III: **Kine Onsum Moseid**, Michael Schulz, Trude Storelvmo, (2021), "Importance of BC lifetime for radiative effects in CMIP6 models", *In prep, planned for Geophysical Research Letters*

Other publications from the PhD period that are not included in the thesis:

I Øyvind Seland, Mats Bentsen, Dirk Olivié, Thomas Toniazzo, Ada Gjermundsen, Lise Seland Graff, Jens Boldingh Debernard, Alok Kumar Gupta, Yan-Chun He, Alf Kirkevåg, Jörg Schwinger, Jerry Tjiputra, Kjetil Schanke Aas, Ingo Bethke, Yuanchao Fan, Jan Griesfeller, Alf Grini, Chuncheng Guo, Mehmet Ilicak, Inger Helene Hafsahl Karset, Oskar Landgren, Johan Liakka, **Kine Onsum Moseid**, Aleksi Nummelin, Clemens Spensberger, Hui Tang, Zhongshi Zhang, Christoph Heinze, Trond Iversen, and Michael Schulz, (2020), "Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6

DECK, historical, and scenario simulations", Geoscientific Model Development, doi:10.5194/gmd-13-6165-2020

II Ingeborg Rian Julsrud, Trude Storelvmo, Michael Schulz, **Kine Onsum Moseid**, Martin Wild, (2021), "Disentangling Aerosol and Cloud effects on Dimming and Brightening in Observations and CMIP6", Journal of Geophysical Research, *in review*

III **Kine Onsum Moseid**, Stefan Hofer, Michael Schulz, (2021), "What might the future climate look like?", Frontiers for Young Minds, *Accepted*

:Acknowledgements

First and foremost, I would like to thank my supervisors, Michael Schulz and Trude Storelvmo, for guidance and support. I dare say I have been very lucky to have the two of you. A special thanks to Michael for prioritizing me, encouraging me to trust my scientific intuition, and for allowing me to be free to do the work I wanted to do. I have learned a lot that I will take with me in the rest of my career.
Thank you also to my colleagues at MET, especially Ada, for making my time here joyful and educational.
I would then like to thank the FullBullpen 3.0 - Evelien, Rob, Franziska and Tim - for all the power hours, karaoke nights, trips, general support and encouragement. An extra special mention to Tim for our pomodoro sessions, rants, and for sacrificing your time and brain to give me valuable feedback on my work. Thank you so much to my friends Stefan and Marianne, for our fishing trips and for reading parts of my thesis (respectively), and to Finn for the love.
A great big thank you to Karoline, my sealmate (not a typo), for keeping tabs on me, being your supportive self, and being a good friend.
Thank you to my særp vegas girls that have supported me through the ups and downs in this process, a non-complete list of your contributions follows: Kristine for the morning coffees, Ida for the extravagant dinner parties, Amanda for the RH-discussions, Anette for your compassionate check-ins and Julie for knowing my schedule and being beyond impressed by everything I do.
To the rest of my friends, thank you for being patient with me and allowing me to retract into work in the past year, I am ready to come back out now.
Thank you so much to my mom and dad with partners, my sister Kaja and Ole-M, and Santi and Ulvang, for being supportive, happy to see me, and proud of me.

Thank you for the love to my partner in life and fellow bon vivant, Jonathan.

A final thank you to my nephews to which this thesis is dedicated to,
Karsten and Lavrans.

Oslo, February 17, 2022
Kine Onsum Moseid

Contents

Part I

Thesis

Chapter 1

:Introduction

1.1 Motivation

"Each of the last four decades has been successively warmer than any decade that preceded it since 1850" *IPCC* (2021).

Climate change is already affecting large parts of the world, and for us to prepare for the implications and understand the climate of our future, it is essential to understand the climate of the dim and distant past. Out of the multiple natural and anthropogenic factors affecting climate, greenhouse gases (GHGs) and aerosols from human activity dominate over the industrial era. The warming effect of the accumulating GHGs in the atmosphere is partly compensated by aerosol cooling, and the historical disentanglement of these two effects can be used to estimate the future warming potential. The effects of GHGs and aerosols are frequently expressed by their radiative impact on the Earth's energy balance, also called *forcings*. The decades following the 1980s is referred to as the instrumental period of the industrial era, because this is when measurements of the state of the climate became frequent and precise enough to be used for climate monitoring. Changes to the top of atmosphere (TOA) energy balance are measured using satellites, where sophisticated methods can identify aerosol radiative effects amid increased outgoing radiation connected to the observed surface temperature increase (*Loeb et al.*, 2018). However, human activity has emitted GHGs and aerosols since the industrial revolution in the 1800s, and observations are sparse before the instrumental period. Earth system models (ESMs) are used to fill this gap and estimate the effective radiative forcing of GHGs, tropospheric aerosols, and other climate forcers.

Figure 1.1 illustrates the temporal evolution of these estimates between the onset of the industrial revolution and the present day. The data shown in Figure 1.1 originates from a multitude of current generation climate models, all participating in the Coupled Model Intercomparison Project phase 6 (CMIP6). The tropospheric aerosol forcing compensates for a large part of the GHG forcing. Only after 1970 did the total anthropogenic forcing become positive and it has increased steadily since then. Whether the recently reduced negative aerosol forcing has contributed significantly to the overall positive forcing of climate remains an open question. This is reflected in the is large historical uncertainty connected to the tropospheric aerosol forcing in Figure 1.1 (gray shading)

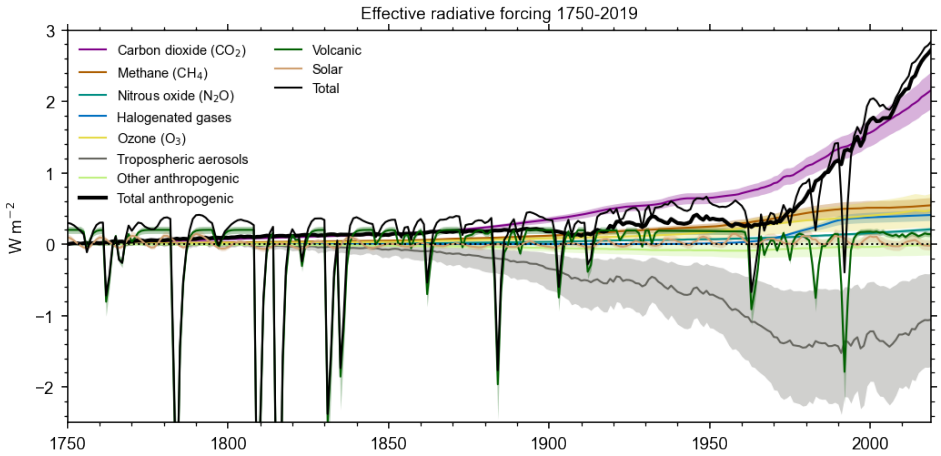

Figure 1.1: *Temporal evolution of effective radiative forcing related to different climate forcers from 1750 to 2019. This figure is made by Chris Smith[a], attained from personal correspondence, and is made using the same data as found in Figure 2.10 in Gulev et al. (2021).*
[a] ORCiD: 0000-0003-0599-4633

which represents the 5-95% uncertainty range according to these models. This uncertainty is predominantly attributed to a poor understanding (and thereby large model spread) regarding aerosol-cloud and aerosol-radiation interactions. As aerosol processes occur on small scales in both time (seconds-minutes) and space (μm-m), their representation in ESMs is inevitably parameterized. These simplified parameterizations, also called aerosol treatments, cover the entire life cycle of aerosol species from emission to deposition, and are the cause for the large inter-model differences, and a subsequent large model spread in the diagnosed aerosol forcings.

Many efforts have been made to constrain the tropospheric aerosol forcing through enhancing the confidence in aerosol-cloud forcing estimates (*Diamond et al.*, 2020; *Fiedler et al.*, 2019; *Gettelman*, 2015; *Karset et al.*, 2018), but even as a parametrization is evaluated and validated at a certain time and place, the extrapolated spatial and temporal representativeness remains an issue. Although direct aerosol observations over the industrial era are scarce, ground-based stations have recorded the downwelling shortwave radiation since the mid-1900s, which can work as a proxy for the aerosol forcing (*Cherian et al.*, 2014). Studies comparing the previous generations of models (CMIP5) to observations showed that the models underestimated this aerosol forcing proxy (*Allen et al.*, 2013; *Storelvmo et al.*, 2018), and work in this thesis assess whether this is still the case for state-of-the-art ESMs (CMIP6).

While the uncertainty as shown in Figure 1.1 reflects inter-model spread only (attained with identical emissions), work in this thesis focuses on an additional source of uncertainty which is not yet accounted for: the aerosol and aerosol precursor emission inventories. Aerosol and aerosol precursor emission inventories are particularly uncertain for the pre-satellite era, meaning before 1980.

Estimating historical burdens and emissions is a much more tedious task for anthropogenic aerosols than for GHGs, as aerosols are short lived climate forcers with lifetimes averaging on about one week (*Naik et al.*, 2021), that spread heterogeneously around the globe.

The uncertainty in emission estimates is seldom discussed due to emission inventory datasets having so far been presented without uncertainties in large model intercomparison projects (CMIP6, *Hoesly et al.* (2018)). Until these uncertainties are published, it is important to keep in mind that emission inventories are a potential source of bias in research related to radiative forcing caused by aerosols when investigating the pre-satellite industrial era.

Whether a bias between models and observations is caused by erroneous aerosol emission inventories or the treatment of aerosols within the model can be difficult to disentangle. This thesis proposes to use ice core archives of the two aerosol species black carbon and sulfate both close and far from anthropogenic emission sources to aid the disentanglement. The method assumes that inter-model differences in aerosol parametrizations become more apparent with increasing distance from the aerosols' emission source. This assumption is based on the fact that an atmospheric model in an ESM simulates the aerosol concentration starting at its emission source and transporting it with each time step. Close to an aerosol emission source, the inter-model spread is small as fewer time steps had to be simulated. However, as the time steps within a model progress with increasing distance from the emission source, the aerosol treatment in the individual models becomes more apparent, leading to a larger inter-model spread. Thus, the biases in the black carbon/sulfate concentration compared to observations in liquid ice are either dominated by biases in emission (close to the source) or are a combination of errors in emission and inter-model spread (far from the source).
Lifetimes of tropospheric aerosols depend on the ESM and their aerosol treatment. These model differences can highly affect the black carbon/sulfate concentration depending on the distance from the aerosol emission source at which the concentrations are investigated. Therefore, a lifetime study of the aerosol black carbon is added to this thesis complementing the model-ice core comparison. This, together with the model-ground station comparison regarding incoming solar radiation provides insights to our aerosol forcing history from 1850 until present day, and the objectives of this thesis are outlined in the next section.

1.2 Objectives

The overall objective of this thesis is to **understand the aerosol forcing history through analyses of state-of-the-art Earth system models, and investigate whether a bias to observations can be attributed to model-errors or flawed emission inventories**. This is done through comparing Earth system model output to observations. Two observational data sets have been used, one is based on surface radiation measurements as recorded by ground stations and the other is an ice core aerosol concentration dataset that has been compiled for this thesis, including previously unpublished records. The third study in this thesis is a pure Earth system model study, inspired by findings in the first two studies. An overview of the Earth system model data in addition to the observational data sets are presented in Chapter 3.
The overall objective is met by a series of specific objectives in each of the studies enclosed in this thesis:

Paper I
To compare gap-filled observational surface shortwave radiation data to that of CMIP6 models, and evaluate model performance of surface energy balance previously shown to be directly related to anthropogenically emitted aerosols (*Cherian et al.*, 2014). This comparison will provide new knowledge in the performance of the newest generation of ESMs regarding aerosol forcing

Paper II
To make use of the under-utilized aerosol archive in ice cores in the evaluation on Earth system models' ability to reproduce aerosol concentrations over the industrial era. This evaluation provides a useful assessment of a modern ensemble of ESMs and their emission inventories which form the basis of aerosol forcing estimates in recent evaluations (*Forster et al.*, 2021)

Paper III
To investigate BC lifetimes in CMIP6 over the historical era, and determine its importance for radiative effects

The background information used as basis for the work in this thesis is presented in Chapter 2, and a summary of findings and answers to specific objectives are found in Chapter 4. Consequences for aerosol forcing history from all three papers are discussed and point to a larger contribution of black carbon to aerosol forcing in the early and mid 20th century than what is presented by current CMIP6 emission inventories (Chapter 5).

Chapter 2

:Background

This chapter will introduce the reader to the basic scientific concepts and background necessary for understanding the work carried out in this thesis. New findings within the research field and relevant scientific introductions are presented individually within each paper in Part II.

2.1 Earth's energy balance

In a stable climate, the incoming solar radiation at TOA is balanced by the outgoing longwave radiation from Earth plus the reflected shortwave radiation (Fig. 2.1). This TOA balance is exposed to changes through either human or natural perturbations, and the perturbation induced energy imbalance (measured in Wm^{-2}) is called Effective Radiative Forcing (ERF, *Forster et al.*, 2021). A prime example of such a perturbation is the anthropogenically added greenhouse gas CO_2, which results in a positive ERF. The warming potential of CO_2 was first discovered and published by the climate science pioneer Eunice Foote in 1856 (*Foote*, 1856), even though the scientist John Tyndall is often credited the discovery from his 1859 publication[1]. A positive ERF means more energy is entering the Earth system than leaving it, which will inherently increase the temperature within the Earth system, until ultimately, following Stefan Boltzmanns law, the outgoing longwave radiation increases such that the system is in balance again (*Charney et al.*, 1979). The radiative effect of the temperature change and the perturbation induced ERF result in a net TOA radiative imbalance ΔN, a relation which can be expressed as follows:

$$\Delta N = \Delta F + \alpha \Delta T \tag{2.1}$$

where ΔF represents ERF, ΔT is the temperature change, and α is the climate feedback parameter. The increased temperature causes increased outgoing thermal radiation which dampens the initial temperature change with a feedback mechanism called the Planck feedback which represents one out of many feedbacks in our climate system. The feedback parameter α assumes that changes in the radiative fluxes are proportional to surface temperature changes, and would be a negative number if representing just the Planck feedback. The details of feedback mechanisms are not central to this thesis,

[1] https://www.nature.com/articles/d41586-019-02117-2, Accessed: 28.11.21

it rather focuses instead on ΔF in Equation 2.1 and specific aspects of the surface energy balance, as well as their temporal evolution.

Present day energy budgets at TOA, the surface, and within the atmosphere can be seen in detail in Figure 2.1. The incoming shortwave radiation is roughly 340 Wm^{-2}, and although reflection is dependent on cloud cover and other atmospheric components the averaged reflected shortwave radiation is about 100 Wm^{-2} at TOA. Clouds interact considerably with radiation, and it is therefore useful to divide the energy balance in two categories, one for all sky conditions, and one for cloud free (clear sky in Figure 2.1) conditions. In the top panel of Figure 2.1 on the left we can identify an imbalance of 0.6 Wm^{-2}, which represents ΔN as expressed in Equation 2.1. This imbalance is an estimation of the residual TOA energy, or in other words how much the climate system absorbs at present day. Most of the residual energy is stored in the oceans (*von Schuckmann et al.*, 2020). The present day TOA energy balance is provided by satellite measurements of shortwave and longwave fluxes, but fewer observations are available for surface fluxes. In addition to the radiative fluxes, latent heat and sensible heat fluxes need to be taken into account when calculating the surface energy budget. These two surface fluxes are unquantified for clear sky conditions as seen in Figure 2.1, as clear sky is a fictive state of what the energy balance would be given there was no clouds. For this fictive state we see the TOA imbalance is 20 Wm^{-2}. Out of the incoming solar radiation only $\sim$ 54 % reaches the surface in an average present day with all sky conditions, and processes affecting this metric are explained in the next section.

2.1.1 Downwelling shortwave radiation at the surface

While the incoming solar radiation at TOA is equal for both schematics in Figure 2.1, the shortwave radiation reaching the surface (*solar down surface*) differs between the two. Downwelling shortwave radiation at the surface is affected by clouds, aerosols, water vapor, and radiatively active gases (*Wild et al.*, 2019) either by reflection, scattering, or absorption. When combining both the scattered downwelling solar radiation in addition to the direct downwelling solar radiation we use the term *global radiation*, and note that *global* does not refer to a spacial distribution here. Global shortwave radiation is typically measured by ground stations (pyranometers, explained more in Section 3.3.1), and the oldest continuous record is from Stockholm (see Figure 2.2).
The multidecadal overall decrease in global shortwave radiation from the 1950s to the 1980s seen in Figure 2.2 has been identified in many measurement sites across large spacial distances, and is referred to as *global dimming* (*Liepert*, 2002; *Wild*, 2009). The positive trend in global shortwave radiation following global dimming is called *brightening*, which has also been recorded in multiple measurement sites. These measured variations in surface solar radiation far exceed natural variations in insolation, which are typically around 0.1 Wm^{-2} in the 11-year-cycle (*Fröhlich*, 2006). Given the modest variations in TOA insolation, the cause of multidecadal variations in surface solar radiation has to be within the climate system. The timing of dimming and brightening differs between continents, and while the trends shown for Stockholm in Figure 2.2 are representative of European timing in trends, the onset of brightening in China happened roughly twenty years later, well into the 2000s (*Sanchez-Lorenzo et al.*, 2015; *Schwarz et al.*, 2020), and in India it is yet to come. *Kvalevåg and Myhre* (2007) estimated that

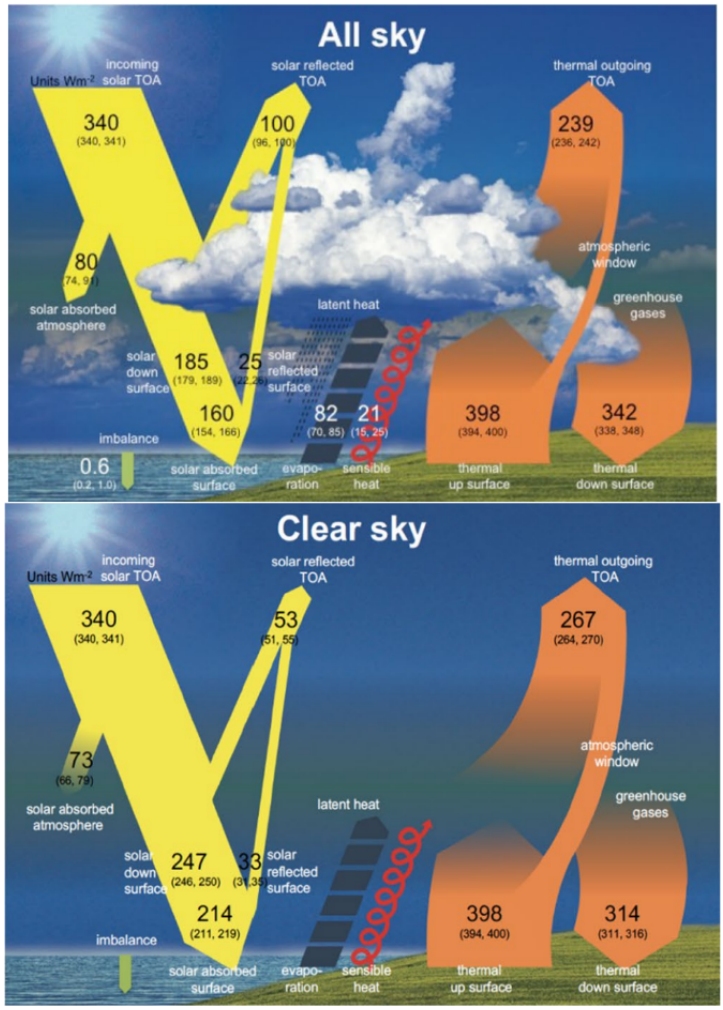

Figure 2.1: *Diagrams of the global mean energy balance of the Earth under all-sky (top) and clear-sky conditions (bottom), representing present day climate at the beginning of the 21th century. Units* Wm^{-2}. *Figure adapted from Wild et al. (2015) and Wild et al. (2019)*

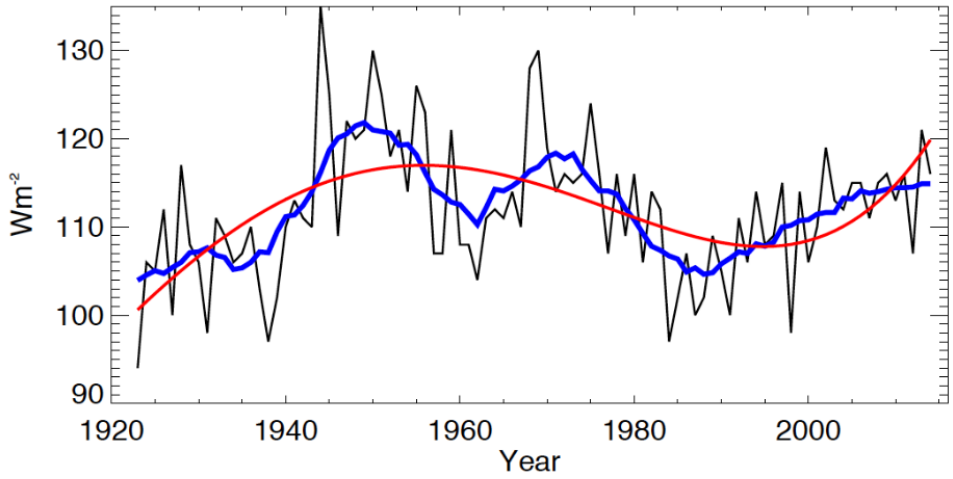

Figure 2.2: *Surface downward shortwave radiation (Wm^{-2}) measured in Stockholm since 1922. The blue line shows the 5-year moving average; the red line shows the fourth-order polynomial fit. Figure from Wild et al. (2017)*

the effect of radiatively active gases' effect on downwelling shortwave surface radiation was -0.31 Wm^{-2}, and that water vapor and ozone were the largest contributors (-0.29 Wm^{-2} and +0.33 Wm^{-2}, respectively). The effect of water vapor changes in the atmospheric column as a result of global warming has been found to be on the order of less than 0.5 Wm^{-2} since the 1960s (*Wild*, 2009). These orders of magnitude are too small to explain the measured dimming and brightening. This leaves clouds and aerosols, both separately and interactively, as potential candidates for the explanation of regional multidecadal changes in downwelling shortwave radiation at the surface.

In this thesis, aerosol effects as presented by ESMs are evaluated and compared against surface measurements of downwelling shortwave radiation. The next section will give the aerosol background necessary for understanding the results presented within enclosed work.

2.2 Aerosols

Aerosols are small solid or liquid particles suspended in air, which originate from either anthropogenic or natural activity. Examples of naturally emitted aerosols are dust from deserts, sea salt from wave breaking in the ocean, and black carbon from biomass burning which all can be found in the schematic in Figure 2.3. However, biomass burning can also be caused by human activity. Black carbon from biomass burning is therefore an example of an aerosol with both natural and anthropogenic origin, though anthropogenic black carbon emissions mostly stem from fossil fuel burning. Fossil fuel burning is the main human activity across sectors resulting in aerosol and aerosol precursor emissions. An aerosol precursor is a gas that once emitted can oxidize to condensable species and thereby form an aerosol in the atmosphere. Contrasting primary aerosols which are directly emitted into the atmosphere, in situ formed aerosols are called secondary aerosols, and the most important one for work in this thesis is sulfate (SO_4), which originates from the aerosol precursor sulfur dioxide (SO_2). Sul-

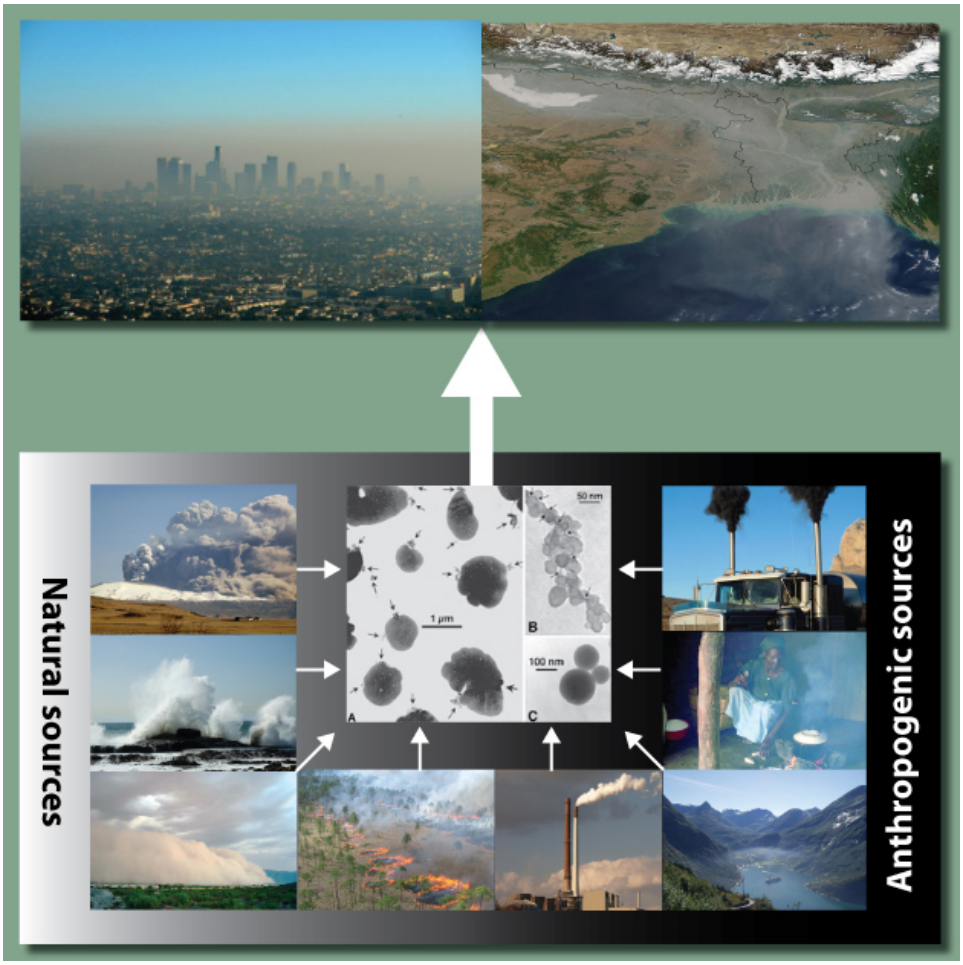

Figure 2.3: *Top: local and large scale air pollution. Sources include (bottom, counterclockwise) volcanic eruptions (producing volcanic ash and sulphate), sea spray (sea salt and sulphate aerosols), desert storms (mineral dust), savannah biomass burning (BC and OC), coal power plants (fossil fuel BC and OC, sulphate, nitrate), ships (BC, OC, sulphates, nitrate), cooking* (domestic BC and OC), road transport (sulphate, BC, VOCs yielding OC). Center: Electron microscope images of (A) sulphates, (B) soot, (C) fly ash, a product of coal combustion (Posfai et al., 1999). Figure and caption from* https://www.nature.com/scitable/knowledge/library/aerosols-and-their-relation-to-global-climate-102215345/*, accessed: 30.11.21*

fur dioxide can be emitted naturally by volcanic eruptions and from marine outgassing DMS, or anthropogenically through fossil fuel burning in the energy or industry sector. Anthropogenic emissions of aerosols and aerosol precursors have varied over time and space during the historical era, and human induced changes to aerosol burdens in the atmosphere affect the radiative budget through aerosol-radiative and aerosol-cloud interactions.

2.2.1 Aerosol effects on the radiative balance

Aerosols affect the radiative balance directly by scattering or absorbing shortwave radiation (*aerosol-radiation interactions*). In addition, aerosols indirectly interact with radiation by changing cloud properties (*aerosol-cloud interactions*). For example, sulfate can act as a cloud condensation nuclei (CCN) and thereby increase the cloud droplet number concentration within a cloud which makes the cloud more reflective, famously known as the Twomey effect (*Twomey*, 1977). The Twomey effect can be seen with some imagination in the top panel of Figure 2.1 on the left, where an increased atmospheric burden of sulfate and thereby an increase of CCN would increase the flux read in the yellow arrow representing *solar reflected TOA*. Following this schematic, more energy leaves the system with anthropogenically added sulfate than without, which results in a negative ERF (see Section 2.1). Sulfate is one example of an aerosol which interacts with clouds, others include organic carbon, nitrate, and the natural source sea salt. As mentioned above, a negative ERF means more energy is leaving the system than entering it, therefore cooling the Earth system.

A sophisticated emission component ERF analysis is performed by *Naik et al.* (2021) and is shown in Figure 2.4. The authors find that emissions of sulfur dioxide have had a strong negative ERF (-0.90 [-0.24 to -1.56] Wm^{-2}) from 1750 to present day, and that most of this negative ERF (-0.68 Wm^{-2}) comes from sulfate-cloud interactions.
A more radiatively complicated aerosol species is black carbon (BC), which overall has a slightly positive ERF where BC-cloud interactions contribute to a negative ERF, and BC-radiation interactions contribute to a positive ERF. Black carbon is considered to be an *absorbing aerosol*, which means it heats up the atmosphere by absorbing incoming solar radiation. It also inhibits the radiation from reaching the surface and being reflected back to space (*Bond et al.*, 2013). In addition, BC can be deposited on reflective surfaces such as ice and snow, thus lowering the albedo of the surface, and warm up its surroundings. The last example explains how BC can contribute to the so-called ice-albedo feedback, which is a positive feedback. In contrast to the BC-radiation interactions which warm the climate system, BC-cloud interactions contribute to a negative ERF according to Figure 2.4. BC-cloud interactions can be separated into two categories, the indirect effect and the semi-indirect effect. BC can act as an ice nucleating particle, which is suggested to have a modest cooling effect (*McGraw et al.*, 2020), and if BC is coated by hygroscopic material (such as sulfate or organic matter (*Liu et al.*, 2011)) it can also act as a CCN and affect clouds in the same way as described above for sulfate. This is defined as the indirect effect (*Cherian et al.*, 2017). The semi-direct effect describes the process where the absorption of solar radiation by BC leads to local warming and a shift in the atmospheric temperature structure, ultimately altering

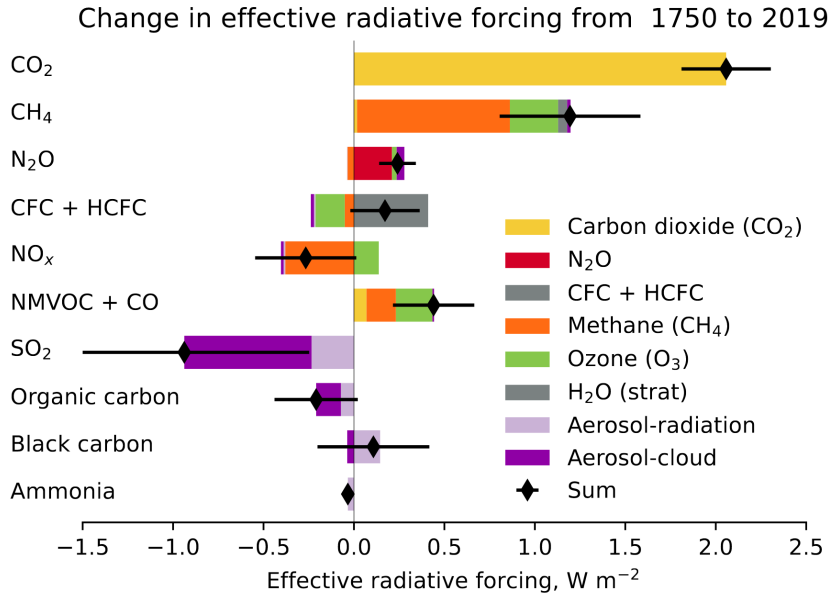

Figure 2.4: *Contributions to effective radiative forcing (ERF) from component emissions between 1750 to 2019 based on CMIP6 models (Thornhill et al., 2021). Error bars are 5-95% and for the ERF account for uncertainty in radiative efficiencies and multi-model error in the means. ERF due to aerosol radiation (ERFari) and cloud effects are calculated from separate radiation calls for clear-sky and aerosol free conditions (Ghan, 2013; Thornhill et al., 2021). "Cloud" includes cloud adjustments (semi-direct effect) and ERF from indirect aerosol-cloud to -0.22 Wm^{-2} for ERFari and -0.84 m^{-2} interactions(ERFaci). The aerosol components (SO_2, organic carbon, black carbon) are scaled to sum to -0.22 Wm^{-2} for ERFari and -0.84 Wm^{-2} for "cloud". Adapted from Figure 6.12 in Naik et al. (2021)*

the distribution of clouds (*Bond et al.*, 2013).
The error bar shown for black carbon in Figure 2.4 is a visualisation of intermodel disagreement in BC-related processes, such as BC interactions with the cryosphere, clouds, and hygroscopic material. The latter is the topic of Paper III in this thesis.

2.2.2 Observations of aerosols

In contrast to well-mixed greenhouse gases, aerosols are heterogeneously distributed around the globe, making point measurements from ground stations non-representative of global values. Therefore, satellites are currently the only global experimental technique to observed aerosol loadings. However, satellite-observations of aerosols only date back to 1979 (*Herman et al.*, 1997) whereas anthropogenic aerosol emissions started long before then. Evidence for aerosol concentrations in the pre-satellite era are sparse, and work in this thesis focuses in investigating the history of anthropogenic aerosols using non-satellite observations.
Aerosols can either be measured directly or through proxies. An example of a proxy is downwelling shortwave radiation (*Cherian et al.*, 2014). As explained in the previous section (and seen in Figure 2.1), aerosols affect how much incoming sunlight reaches the surface, and can thereby be used to evaluate long term changes in atmospheric aerosol abundance. This type of evaluation is precisely what is done in Paper I.

Another proxy for aerosol concentrations is found in ice cores. Aerosols can be deposited either directly as dry particles or incorporated in precipitation as wet deposition. Ice cores are usually sampled in the accumulation zone of an ice cap or glacier which is situated in a location where surface melting is believed to not penetrate annual ice layers. Such locations are usually found in high alpine sites or ice sheets such as Greenland and Antarctica. After an ice core is drilled the age of all layers is determined through existing chronologies using several age markers, such as fallout from volcanic eruptions or high radiation layers from thermonuclear testing. Although there are some uncertainties connected with the age markers in an ice core, they prove very useful for evaluating long term trends in aerosol concentrations in ice, which is done for the industrial era in Paper II.
Ground stations and ice cores are spatially sparse, and the former does not cover the entirety of the industrial era. Computational tools are needed to fill the gaps in the anthropogenic aerosol history.

2.3 Earth system modelling

To handle the problem of sparse observations of aerosols and aerosol effects in the pre-satellite era, ESMs are used. Figure 2.5 shows the evolution of how individual component models have merged gradually to the global coupled system. For example, interactive vegetation has been a standalone model from its creation in the late 1980s until being incorporated in the coupled climate models in the mid 2010s. ESMs are a subset of coupled climate models, and have in common that they include interactions between biogeochemical processes and feedbacks (*Eyring et al.*, 2016). ESMs use a lattice approach to represent the atmosphere, ocean, cryosphere, biosphere, and physical equations describe the processes within each grid box and their interactions. The resolution of the grid boxes in the atmospheric component is generally between 100 and 200 km, and any physical processes occurring on scales smaller than the ESM grid boxes needs to be parameterized (see Section 2.3.2).

ESMs are used for simulating the far past (paleo), the recent past (historical era), present, and future. They differ from forecasting models in that they focus on long term changes rather than hourly changes in the Earth system, and they span globally instead of regionally.
Simulations of the recent past are compared to observations and evaluated based on ESM output, which is however only as good as the input, which is the subject of the next section.

2.3.1 Emission inventories in Earth system modelling

The newest, best-guess estimate of anthropogenic aerosol and aerosol precursor emissions is presented as a gridded data set in *Hoesly et al.* (2018) and is developed with the Community Emissions Data System (CEDS)[2]. This data set is the basis for all emission experiments performed in this thesis, and the basis for experiments under the CMIP6

[2]http://www.globalchange.umd.edu/ceds/

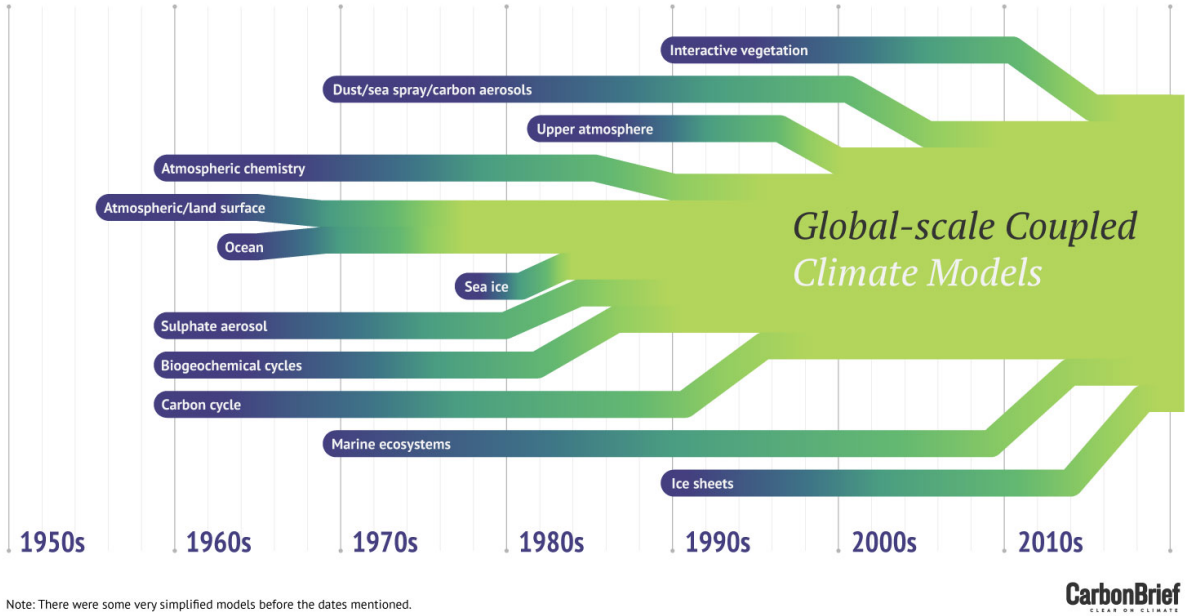

Figure 2.5: *Overview of when mathematical component models were developed (blue) and at what time they were individually merged to a coupled system (green). Adapted from Graphic by Rosamund Pearce; based on the work of Dr Gavin Schmidt* [a]

[a] https://www.carbonbrief.org/qa-how-do-climate-models-work, Accessed: 09.11.21

framework which is further explained in Chapter 3. Work in this thesis focuses specifically on the aerosols sulfate and black carbon, and emission inventories relevant to these aerosols are shown in Figure 2.6 sorted by emission sectors and emission regions.

Sulfur dioxide is predominantly emitted by the energy and industrial sector, and the dominating emission region depends on the decade. The prime time for sulfur dioxide emissions in North America, Europe, and the former Soviet Union have come and gone, and Asia is currently the main emission contributor (Fig. 2.6). The timeseries of sulfur dioxide emissions have been made using reliable existing emission inventories and matching them to default estimates. Existing emission inventories are only available in the recent past, so the data is extended further back in time by using activity drivers (population/energy consumption) and emission factors. The activity drivers from before 1970 are found by estimating total fuel use for each country using CO_2 trends (from CDIAC, *Andres et al.*, 1999; *Boden et al.*, 2016), and disaggregating the total fuel use into sectors (*Bond et al.*, 2004, 2007), where sulfur dioxide is largely emitted in the industrial sector "hard coal" (*Hoesly et al.*, 2018). The emission factor is extended back in time by converging to a given value in a given year - often 0 in 1850.

Black carbon as prescribed in *Hoesly et al.* (2018) is largely based on the emission inventory estimate presented in *Bond et al.* (2007), which is developed by assuming that emissions of particulate matter depend on the choice of combustion technology and the amount of fuel consumed. Total fossil fuel consumption before 1950 is estimated in *Andres et al.* (1999), and a great effort was made by *Bond et al.* (2007) to estimate sectorial activity drivers such as domestic fuel, railroads and steel industry to name a few. These sectorial divisions have been reaggregated to match sectors defined

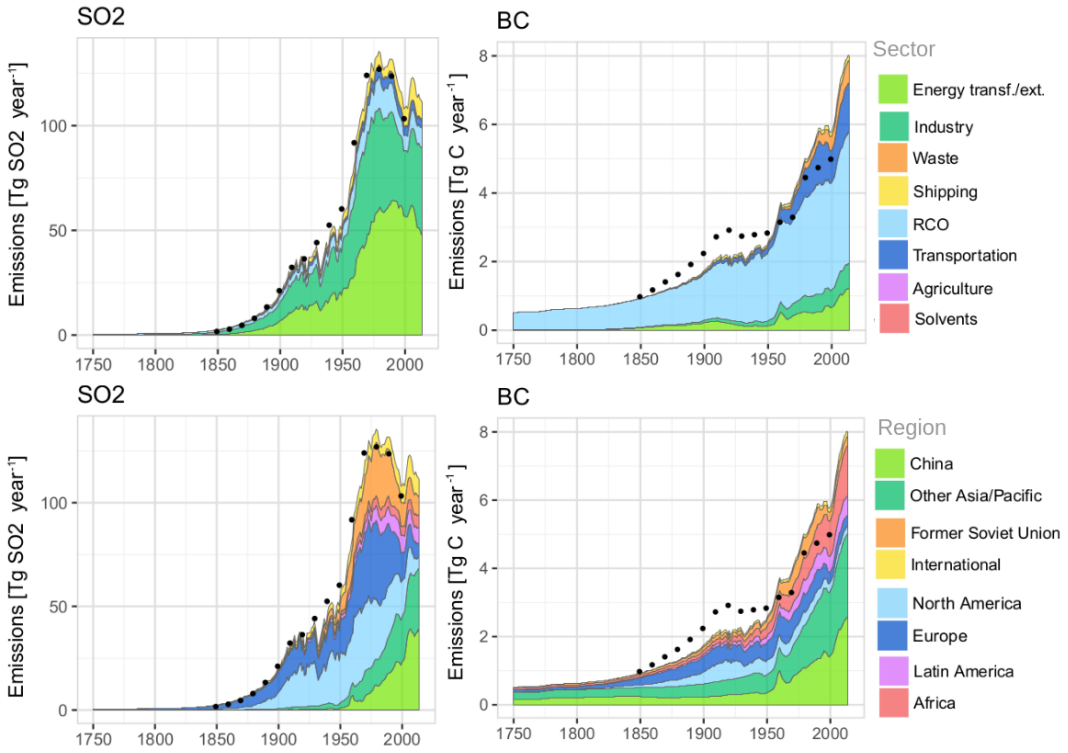

Figure 2.6: *Emissions estimates for sulfur dioxide and black carbon sorted by sector (top row) and region (bottom row). Dotted lines are emissions according to Lamarque et al. (2010). "RCO" stands for residential, commercial, and other. Adapted from Hoesly et al. (2018).*

in *Hoesly et al.* (2018) which can be found in Figure 2.6, where *residential, commercial, and other (RCO)* is found as the main source of BC emissions. Emission factors for BC differ between sources, and although *Bond et al.* (2007) admit these estimates are obtained with methods that come with uncertainties (e.g. observer-based opacity measurements for coal stokers), they remain the best-guess estimates at present day.
As with sulfur dioxide, combining activity drivers and emission factors for the time period before 1970 result in the BC emission estimates shown in Figure 2.6. Some of the emitting regions investigated in this thesis experienced major socioeconomic events in the time period covered, such as wars and depression. The resulting non-linearities in emission trends are not captured in the methods used in *Bond et al.* (2007), as they focus on long term growth in BC emissions.

Limitations to the dataset provided by *Hoesly et al.* (2018) include variation in the methodology used for estimating emissions from specific countries/regions, sectorial coverage, level of detail, and consistency over time and space. In addition, there is no uncertainty estimate available for this data set.

2.3.2 Aerosols in Earth system modelling

Aerosol processes fall within the subgrid scale of ESMs, and the parameterizations of aerosol processes are found within the aerosol schemes in the atmospheric component of an ESM. Aerosol schemes have large variations across development groups,

including the handling of size distributions, what aerosol types are included and so on. Using modal schemes to represent aerosol size distribution is most common, the model CESM2 uses a four mode scheme (MAM4) which includes a total of 18 transported aerosol tracers and 5 precursor gases (*Liu et al.*, 2016). In NorESM2-LM, which is the most used model in this thesis, a "production-tagged" scheme OsloAer6 is used. OsloAer6 includes 21 transported aerosols and 8 gas tracers (*Kirkevåg et al.*, 2018). Typically included aerosols include black carbon, sea salt, dust, primary organic matter, secondary organic matter, sulfate, and soil dust (clay). Aerosol size distribution can also be represented in sectorial bins which believed to be closer to first principles (*Blichner et al.*, 2021), however this is more computationally costly than the modal approach and is therefore often shelved. Work in this thesis will present results from a total of 13 separate climate models, which all have individual aerosol schemes.
A particular aspect of aerosol scheme differences is relevant to work in this thesis and concerns the treatment of black carbon's interaction with hygroscopic materials. The potential condensation of hygroscopic material on BC-containing particles changes the atmospheric lifetime of BC. Some ESMs require eight monolayers (a single, closely packed layer of molecules) of sulfate equivalent condensate on a black carbon particle, for it to convert from a hydrophobic to a hydrophilic state, while for other ESMs this conversion is only dependent on the available concentration of hydroxyl radical (OH), and some models have even more simple ageing schemes, which prescribe an exponential decay with time and consequential transformation of insoluble to soluble BC. The aerosol specific error bars shown in Figure 2.4 are partly a result of differences in aerosol schemes, as the data shown in this figure is based on simulations performed by a great number of ESMs.

Chapter 3

:Research tools

In this chapter I will first give an overview of the framework allowing for the multi-model analysis performed in Paper I, II, and III. Additional experiments outside this framework have been performed using the model NorESM2-LM for Paper II, and this model is further compared with the model CESM2 in Paper III, so I present these models along with details of the experiments performed in NorESM2-LM before finally presenting the two observational data sets used for Paper I and II.

3.1 CMIP6

The Coupled Model Intercomparison Project (CMIP) is initiated by the World Climate Research Programme (WCRP) and aims to better understand past, present and future climate changes through analysing the outcomes of multiple global climate models. The idea behind CMIP is to provide basis for climate assessments such as the recently published sixth assessment report (AR6) by the Intergovernmental Panel on Climate Change (IPCC).
To provide this basis a multitude of coupled climate models perform a set of common experiments before the multi-model output is standardized and made publicly available[1]. The CMIP effort makes valuable model data available to scientists beyond those who run the models, and provides a framework for delivering high quality climate information for IPCC ARs, political negotiations, and climate science in general. The project was first started in 1995, and has since gone through several phases as new generations of climate models have evolved. The model data used in this thesis are all from the newest generation of model development, which is from phase 6 of CMIP (CMIP6) (*Eyring et al.*, 2016).
Within CMIP6 there exists 21 smaller model intercomparison efforts, including AerChemMIP, RFMIP, and DAMIP, which are relevant to the work in this thesis and will be presented below. But first I will explain the historical experiment, which is one of the entry-card experiments of CMIP6.

[1] https://www.wcrp-climate.org/wgcm-cmip, Accessed: 25.11.21

3.1.1 Historical

Relevant for Paper I, II and III

To participate in any of the smaller model intercomparison efforts within CMIP6 a model needs to perform a set of basic experiments, and the *historical* experiment is one of them. The historical time period is defined as 1850 through 2014, and each model uses the same emission inventories for aerosols, aerosol precursors, and greenhouse gases as prescribed in *Hoesly et al.* (2018) and *van Marle et al.* (2017), see Chapter 2.3.1.
It is easier to estimate the uncertainty in model response to a given forcing when all models use the same input data sets as opposed to induvidual ones (*Eyring et al.*, 2016). The major purpose of the historical simulation is to evaluate model performance, both as compared to observations but also in intermodel comparison studies. In Paper I and II output from the historical simulation from several models have been compared to observations to evaluate model performance, and the comparison methods are described in the individual Papers.

3.1.2 AerChemMIP

Relevant for Paper I and III

The Aerosol Chemistry Model Intercomparison Project (AerChemMIP) is designed to quantify air quality and climate impacts of aerosols and chemically reactive gases (*Collins et al.*, 2017). The experiments within AerChemMIP target four scientific questions presented in *Collins et al.* (2017), and the first of these four questions is relevant to this thesis: How have anthropogenic emissions contributed to global radiative forcing and affected regional climate over the historical period?
In this question "anthropogenic emissions" refer to non-CO_2 emissions, such as near term climate forcers (NTCFs), halocarbons and nitrous oxide. Multiple AerChemMIP experiments have been used in work related to this thesis, and they are explained as follows. The first experiment is called *hist-piAer*, and is an ocean-coupled experiment. When experiment names within the CMIP6 framework contains the word *hist* they have been run for the same time period at the *historical* period unless otherwise is specified. The remaining part of the experiment name generally explains what differentiates this experiment from the *historical* one, so for the first AerChemMIP experiment *hist-piAer* a historical simulation has been run but all aerosols (Aer) have been kept at pre-industrial (pi) levels. Two additional experiments *hist-piNTCF* and *histSST* was also included in Paper I, and while *hist-piNTCF* is self explanatory following the above description of MIP naming convention, *histSST* only differs from the *historical* simulation in that it contains prescribed historical sea surface temperatures (SSTs) instead of a coupled ocean configuration. These experiments were used to evaluate how anthropogenic aerosol emissions and a coupled ocean would affect downwelling shortwave radiation at the surface in Paper I.

The quantification of species-specific ERF was presented in Figure 2.4 in Chapter 2.2.1, and the aerosol-related ERFs in this figure have been calculated from the so-called pi-

Clim experiments within the AerChemMIP framework. All experiments which names contain *piClim* has prescribed pre-industrial climatology of SSTs and sea ice. This is done to exclude ocean temperature responses and following feedbacks to ensure consistency in ERF estimates (see Eq. 2.1). Several *piClim* experiments from AerChemMIP have been used in Paper III, both for investigating species specific present day effects on the radiative TOA and surface balance, and for investigating variations in black carbon lifetime following changes in other soluble species within different global climate models. AerChemMIP provides species specific radiative forcing experiments by building on the *piClim*-protocol defined and initiated by the Radiative Forcing Model Intercomparison Project (RFMIP).

3.1.3 RFMIP

Relevant for Paper I and III

The main objective of RFMIP is to investigate ΔF in Equation 2.1, also called ERF throughout this thesis. Experiments in RFMIP are designed to diagnose both present day and time-evolving ERF from land use, greenhouse gases, aerosols (non-species-specific), either individually or combined (*Pincus et al.*, 2016).
In Paper I we have used model data from two time-evolving experiments in RFMIP, *piClim-histaer* and *piClim-histall*. The difference between the two is that one only contains historical evolution in aerosol emissions without greenhouse gas emissions or other forcing agents. These experiments are used to compare modelled surface solar energy changes between an all-aerosol historical development and an all-agent historical development, which is further discussed and compared to observations in Paper I.
In Paper III we use model output from the two experiments *piClim-control* and *piClim-aer*. None of these have time-evolving changes in forcing agents, but rather 30 year simulations with constant pre-industrial and present day aerosol forcing agents, respectively. Model data from the experiment *piClim-control* is used as pre-industrial reference run for all ERF calculations performed in this thesis, both when using RFMIP and AerChemMIP perturbation experiments.

3.1.4 DAMIP

Relevant for Paper I

The Detection and Attribution Model Intercomparison Project (DAMIP) has a primary objective of estimating the contributions of natural and anthropogenic forcing changes to both regional and global observed changes (*Gillett et al.*, 2016). All DAMIP experiments are fully coupled, meaning we cannot separate ERF from the total TOA energy shift caused by surface temperature responses and feedbacks, and they are run extended historically from 1850 to 2020. This makes experiment output from DAMIP an excellent tool for work done in Paper I of this thesis, as this work focuses on historical changes in surface energy, not TOA ERF. The three experiments from DAMIP used in this thesis are *hist-nat*, *hist-GHG*, and *hist-aer*. The first experiment only contains natural historical forcings such as solar irradiance and stratospheric aerosols, the second has no natural forcings, but a historical evolution of well mixed greenhouse gases. The

third experiment contains anthropogenic aerosols only (BC, OC, SO_2, SO_4, NO_x, NH_3, CO, NMVOC).

Further details on which models participating in AerChemMIP, RFMIP and DAMIP was used, what variables were downloaded, and how they have been treated can be found in the individual papers.

3.2 NorESM2-LM and CESM2

Relevant for Paper II, III

NorESM2-LM (*Kirkevåg et al.*, 2018; *Seland et al.*, 2020) is one of the CMIP6 released versions of the Norwegian Earth System Model (NorESM). NorESM is based on the Community Earth System Model (CESM) which is a widely used ESM that participates in many of the aerosol relevant efforts within CMIP6 (*Danabasoglu et al.*, 2020). CESM2 is developed mainly at the National Center for Atmospheric Research (NCAR), and while one is based on the other there are many important differences between the two ESMs. For one NorESM2 uses a completely different ocean model (Bergen Layered Ocean Model: BLOM; Bentsen et al., 2021, in prep.) than CESM2, which contributes majorly to the different climate sensitivities found in NorESM2-LM and CESM2 (*Gjermundsen et al.*, 2021). The atmospheric model CAM6-Nor in NorESM2-LM is based on CAM6, the atmospheric model from CESM version 2, but the aerosol scheme differs between the two, which is a great opportunity to disentangle host model effects from aerosol scheme differences in this thesis.

3.2.1 BC solubility

As mentioned above, the main difference between the atmospheric model in CESM2 (CAM6) and the one in NorESM2-LM (CAM6-Nor) is the aerosol scheme. Although these aerosol schemes differ greatly in a variety of ways, for Paper III one difference is central: how they treat black carbon solubility. The atmospheric model CAM6-Nor uses the aerosol module OsloAero6 together with calculations provided in *Liu et al.* (2012) to calculate BC ageing, which is relevant to BC lifetime. The requirements for BC to move from its primary emitted state (hydrophobic) to an aged state (hydrophilic) is a condensate thickness exceeding three monolayers of sulfate equivalents (*Kirkevåg et al.*, 2018). CAM6 uses the aerosol module MAM4 (*Liu et al.*, 2016) which requires eight monolayers of sulfate equivalent condensates to move a BC particle from its primary state to an aged state.

3.2.2 Experiments using NorESM2-LM

Paper II contains an investigation and comparison of aerosol concentrations in ice in CMIP6 models as compared to those archived in ice cores. The location of the ice cores is often very remote from the anthropogenic sources and the comparison quickly opened the question *where does the recorded aerosol come from?* To answer this question, source attribution experiments were performed using NorESM2-LM. We investi-

gated the source for sulfate and black carbon concentrations in ice by perturbing emissions of the sulfate precursor sulfur dioxide and black carbon in different continental regions. The regions were chosen as they are each believed to be important emission regions at some point over the historical era. The perturbation experiments are listed together with their reference simulations (*historical* from CMIP6 and *histSST* from AerChemMIP) in Table 3.1.

Table 3.1: *Emissions for CMIP6 are described in Hoesly et al. (2018) for anthropogenic emissions and in van Marle et al. (2017) for biomass burning emissions. Asia refers to South Asia + Eastern Asia + Central Asia (excluding Russia, Southeast Asia and Middle East) according to HTAP2 regions as defined in Galmarini et al. (2017). SST/SIC : prescribed sea surface temperature and sea-ice cover from the historical simulation.*

Experiment	Emission perturbation				Ocean model
	Species	Region	Sector	Size	
historical	–	–	–	–	Full ocean
histSST	–	–	–	–	SST/SIC
histSST-so2x2nam	SO_2	North-Am.	Anthrop.	+100%	SST/SIC
histSST-so2x2eur	SO_2	Europe	Anthrop.	+100%	SST/SIC
histSST-so2x2asi	SO_2	Asia	Anthrop.	+100%	SST/SIC
histSST-bcx2nam	BC	North-Am.	Anthrop.	+100%	SST/SIC
histSST-bcx2eur	BC	Europe	Anthrop.	+100%	SST/SIC
histSST-bcx2asi	BC	Asia	Anthrop.	+100%	SST/SIC
histSST-biox2	SO_2, BC	Global	Biom. burn.	+100%	SST/SIC

Note that while these experiments were designed and analysed by me, they were performed by coauthor Dirk Olivié

3.3 Observations

The instrumental period, the period when aerosol observations became global and credible, began long after human activity started emitting aerosols. Work in this thesis focuses on the 20th century, and given how rare long term records of aerosol related observations are, I am proud to present the two observational data sets used for Paper I and II.

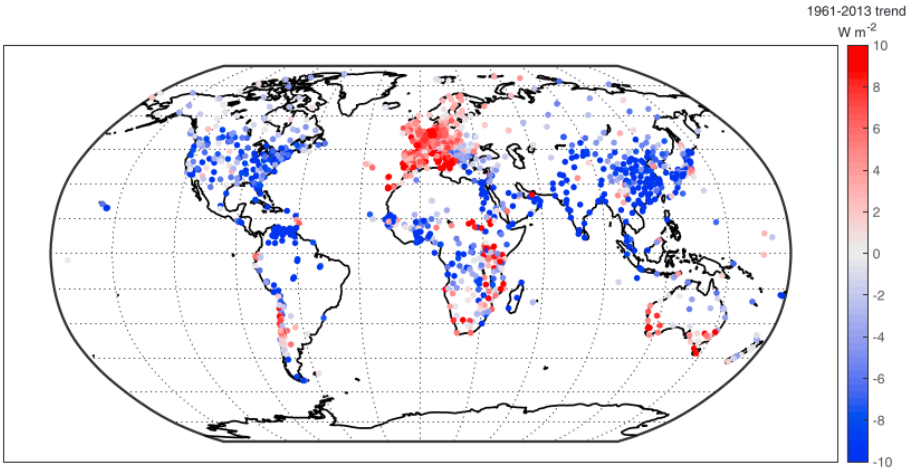

Figure 3.1: *Global overview of location of Global Energy Balance Archive stations included in Paper I, along with the sign and magnitude of their downward flux of solar radiation at the surface trend, calculated based on 5-year running means from 1961 to 2013 (red corresponds to brightening and blue to dimming). Figure and parts of caption from Storelvmo et al. (2018).*

3.3.1 Gap filled GEBA

Relevant for Paper I

The Global Energy Balance Archive (GEBA) is an ETH Zurich maintained database for storage of surface energy fluxes measured worldwide (*Wild et al.*, 2017). Thanks to GEBA we have access to monthly mean data of global radiation (see Chap. 2.1.1) at more than 2200 locations, mostly measured using pyranometer instruments. However, many stations were installed in more recent times, and continuous long term records are few, which is why we use a dataset we call gap filled GEBA in the work in this thesis. Gap filled GEBA uses 1487 stations from GEBA and applied the machine learning technique random forest (*Breiman*, 2001) to temporally fill missing monthly values at each station from 1961-2014. The method is evaluated and presented in *Leirvik and Yuan* (2021). An overview of the spatial distribution of the stations in gap filled GEBA can be found in Figure 3.1, which also includes the surface downwelling shortwave radiation trend per station.
A benefit of using this dataset is that we can compare our results using the new CMIP6 model data to that of *Storelvmo et al.* (2018) which was using CMIP5 data, and of which Figure 3.1 is attained from. We divided the GEBA data into regions based on countries and continents, and co-located the model output to station locations (see Paper I).

3.3.2 Ice cores

Relevant for Paper II

The most direct tracer of aerosols in the pre-instrumentation era are aerosol proxy records from glaciers and ice sheets. The ice-preserved aerosol record can also be used for the more recent historical period, such as during the post-industrial revolution and pre-satellite era when anthropogenic aerosol emissions were high, yet aerosol observa-

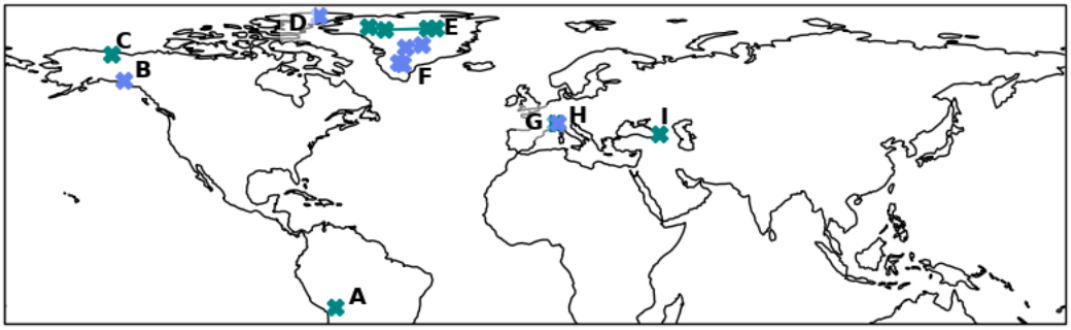

Figure 3.2: *Global overview of location of ice cores gathered and used in Paper II. 15 locations are used, and the multiple ice cores in Greenland are averaged together into a "southern" and "northern" part. This figure is adapted from Figure 1 in Paper II*

tions were sparse. Annual average sulfate and black carbon concentration records from 15 ice cores were gathered (see Table 3.2), and decadally averaged to better represent aerosol trend evolution over the historical era. Figure 3.2 shows the spatial distribution of the ice cores, and several aerosol records from these cores have not previously been published, and is denoted in Table 3.2 as *This study*, which refers to Paper II.

The ice cores were selected based on a set of requirements. We required the ice core to have at least annual resolution, and both black carbon and sulfate records. In addition, we excluded some ice cores which met these criteria due to them being located in regions that were believed to have a strong melting. By strong melting we mean that the location is prone to such warming events so that snow melt exceed the annual snow layers, which impose high uncertainties in the timing of the aerosol record. Examples of regions we excluded based on this was Lomonosovfonna in Svalbard and Upper Freemont Glacier in Wyoming, USA.
Model outputs were colocated to the location of an ice core and then a 3x3 grid surround the ice core was selected when comparing model outputs to ice core data. This was done as ice cores are often situated in alpine areas the models cannot resolve with their spatial resolution, so the aerosol concentrations from models are presented in an "ice core area" rather than at the ice core site.

Table 3.2: *Overview of the ice cores gathered and used in Paper II along with respective references where BC and sulfate data can be found. Previously unpublished data is referenced as This study. This table is also found in Paper II.*

Site	Lat	Lon	BC	sulfate
Illimani	-16.62	-67.76	*Osmont et al.* (2019)	*Kellerhals et al.* (2010)
Eclipse	60.5	-139.5	*This study*	*This study*
McCall Glacier	69.3	-143.8	*This study*	*This study*
Mt Oxford	82.2	-73.2	*This study*	*This study*
Greenland				
Northern				
NGT_B19	78.0	-36.4	*This study*	*This study*
Tunu2013	78.0	-33.9	doi:10.18739/A2ZQ1G	*Sigl et al.* (2015)
NEEM_2011_S1	77.5	-51.1	*Zennaro et al.* (2014)	*Sigl et al.* (2013)
Humboldt	78.5	-56.8	*McConnell* (2010)	*Sigl et al.* (2013)
Southern				
Summit2010	72.6	-38.3	doi:10.18739/A2XV7T	doi:10.18739/A2XV7T
D4	71.4	-43.9	*McConnell* (2007)	*McConnell* (2007)
ACT11d	66.5	-46.3	*This study*	doi:10.18739/A2Z933
ACT2	66.0	-45.2	*McConnell and Edwards* (2008)	*McConnell and Edwards* (2008)
Col Du Dme	45.8	6.9	*This study*	*Preunkert et al.* (2001) *Legrand et al.* (2013)
Colle Gnifetti	45.9	7.85	*Sigl et al.* (2018)	*Engardt et al.* (2017)
Mt Elbrus	42.4	42.4	*Lim et al.* (2017)	*Preunkert et al.* (2019)

Chapter 4

:Presentation of findings

This chapter presents summaries and main findings of the three studies enclosed in this thesis. Each paper presents its corresponding objective as presented in the introduction, and a main conclusion.
The three papers follow a storyline which can be explained as follows: A first look into the newly released CMIP6 model results reveal that models underestimate the historical aerosol effects believed to be connected to anthropogenic aerosol emissions (Paper I). A further investigation (Paper II) confirms that indeed there are discrepancies between aerosol concentration trends found in ice and as simulated by models, pointing to errors in European BC emissions. The implications of different atmospheric loads of BC is further explored with the same models as used in the ice core study (Paper III), but this time large intermodel differences in BC lifetime is revealed. The BC lifetime differences found in Paper III may contribute to biased atmospheric BC loads found in Paper II side-by-side to erroneous emission inventories.

4.1 Paper summary

4.1.1 Paper I: Bias in CMIP6 models as compared to observed regional dimming and brightening

Kine Onsum Moseid, Michael Schulz, Trude Storelvmo, Ingeborg Rian Julsrud, Dirk Olivié, Pierre Nabat, Martin Wild, Jason N.S. Cole, Toshihiko Takemura, Naga Oshima, Susanne E. Bauer, Guillaume Gastineau

Manuscript published in Atmospheric Chemistry and Physics, December 2020

Objective

To compare gap filled observational surface shortwave radiation data to that of CMIP6 models, and evaluate model performance of surface energy balance believed to be directly related to anthropogenically emitted aerosols.

Summary

In this paper, we investigated global and regional aerosol effects over the time period 1961-2014 by looking at surface downwelling shortwave radiation (SDSR). We used observations from ground stations as well as multiple experiments from eight Earth System Models (ESMs) participating in the Coupled Model Intercomparison Project Version 6 (CMIP6). We find that model experiments without anthropogenic emission of aerosols do not show changes in SDSR (dimming), while experiments that do include anthropogenic aerosol emissions shows a reduction of SDSR. The modelled SDSR in experiments including anthropogenic emissions represent observed SDSR evolution well in Europe, but poorly in China, and we suggest this is connected to underestimated aerosol emissions estimates there.

Main findings

- In the historical experiments, CMIP6 models underestimate global dimming compared to what was measured by the ground stations. The models reproduce the transient development of SDSR well in Europe but poorly in Asia, specifically China
- When investigating simulations performed in the DAMIP, RFMIP and AerChemMIP we find that only those simulations containing anthropogenic aerosol emissions show any dimming at all, yet the dimming is underestimated by most models. This further underlines that the phenomena global dimming and brightening is connected to anthropogenic aerosol emissions, which is widely thought but not proven
- We suggest that the continuous decrease in simulated SDSR is related to the continuous increase in atmospheric sulfate burden in the historical simulations over China, and that the cause of the discrepancy between model and observations in transient SDSR in China is partly in erroneous emission inventories

Main conclusion

According to our study, global dimming is likely a results of anthropogenic aerosol emissions, however, CMIP6 models underestimate the dimming which raises question to whether they reproduce historical climate evolution for the right or wrong reason.

Author contribution: *I helped design the study, did all analysis of the CMIP6 data and the comparison to the observational data. I also wrote the main text for the paper with guidance from supervisors, and input from co-authors.*

4.1.2 Paper II: Using ice cores to evaluate CMIP6 aerosol concentrations over the historical era

Kine Onsum Moseid, Michael Schulz, Anja Eichler, Margit Schwikowski, Joseph R. McConnell, Dirk Olivié, Allison S. Criscitiello, Karl J. Kreutz, Michel Legrand

Manuscript submitted to Journal of Geophysical Research

Objective

To make use of the under-utilized aerosol archive in ice cores in the evaluation on Earth system models ability to reproduce aerosol concentrations over the industrial era.

Summary

We have gathered sulfate and black carbon (BC) data from 15 ice cores, including some previously unpublished, and compared the respective aerosol concentrations in ice cores to that of 11 Earth System Models (ESMs). The relative temporal trend in concentrations of sulfate is largely represented well in models, while the relative temporal trend in BC concentration is not. The aerosol concentration magnitudes do not compare well between ice cores and models for both sulfate and BC, for most regions, and the reason for magnitude differences is not known or further investigated in this study. We performed source attribution experiments using NorESM2-LM, and found that European emissions of BC contribute to BC concentration found in Northern Greenland. Ice cores from the European Alps shows an early 20th century maximum in BC concentration, while all models agree on a late 20th maximum BC concentration in the same region, and we suggest this is caused by errors in European emission inventories. The model spread in BC concentration increase in time the further away from emission sources an investigated ice core location is, which we interpret as signals of differences in aerosol treatment leading to deposition within the models.

Main findings

- We find that emission changes of sulfate precursors in the CMIP6 emission inventories are consistent with the observations presented in this paper, and note that while trends are represent well, absolute magnitudes of concentration is not
- There is high model agreement in the model to observation bias in black carbon concentration relative trend in Europe, which suggest the bias is sourced to what the models all have in common: emission inventories for BC.
- Areas far from its emission source region (according to NorESM2-LM experiments) have a diverging model range in time, which we believe is connected to intermodel differences in BC residence times
- Source attribution experiments reveal relatively large European contributions to Northern Greenland BC concentrations, as European emission sources of BC are suggested to be erroneous in this study this partly explains the model to observation bias in this region

Main conclusion

While sulfate concentration trends are well represented in CMIP6 models, we have found evidence suggesting that European emission inventories for BC as prescribed in CMIP6 are erroneous. The BC concentration bias between models and ice cores found in Europe is also found in other regions where European emissions contribute, but biases here are also a result of differences in aerosol treatment within models.

Author contribution:*I helped design the study, designed the NorESM2-LM contribution experiments, performed all analysis and led the writing.*

4.1.3 Paper III: Importance of BC lifetime for radiative effects in CMIP6 models

Kine Onsum Moseid, Michael Schulz, Trude Storelvmo

Manuscript in preparation, planned for Geophysical Research Letters

Objective

To investigate BC lifetimes in CMIP6 over the historical era, and determine its importance for radiative effects.

Summary

We have used multiple CMIP6 efforts to map BC lifetimes both globally and in the northern hemisphere within several models. We first find that 11 models have a large spread regarding BC lifetime in pre industrial times, but this spread is smaller in present day. This is because some models have large temporal changes in BC lifetime over the historical period, mainly a reduction. We investigate two models which have very similar atmospheric code, except their aerosol schemes, which are situated on opposite sides of the temporal BC lifetime change-story. NorESM2-LM does not exhibit large changes in lifetime across the historical era, but CESM2 does, and the BC lifetime reduction observed in CESM2 is anti-correlated to atmospheric load of sulfate. When investigating results from RFMIP and AerChemMIP we find that the models with a long BC lifetime also has the strongest negative surface ERF out of the group, and the largest atmospheric absorption.

Main findings

- CMIP6 present-day black carbon lifetimes have a range of 3.8 days to 8.5 days
- Some Earth system models have a significant historical reduction in black carbon lifetime of 1-4 days, which we explain by intermodel differences in aerosol code for coating black carbon with soluble material
- Both CESM2 and NorESM2-LM include BC coating in their aerosol schemes, but CESM2 requires more sulfate equivalents in order for a BC particle to coat compared to NorESM2-LM. This means BC is sufficiently coated in pre industrial times in NorESM2-LM, and result in NorESM2-LM having a shorter pre industrial BC lifetime than CESM2. As anthropogenically added sulfate increases, BC is coated also in CESM2, and the BC lifetime in CESM2 decreases
- A long BC lifetime leads to large atmospheric absorption, and a strong negative surface ERF, according to experiments within RFMIP and AerChemMIP

Main conclusion

Models of CMIP6 still have a large spread in BC lifetimes. Some models have a large decrease in BC lifetime across the historical era which we suggest is a result of BC

coating being dependent on sulfate load in these models. BC lifetimes determine atmospheric absorption and contribute to the total aerosol surface ERF.

Author contribution: *I designed the study, performed all analysis and wrote the paper.*

4.2 Summary of findings in the three papers

The findings from Paper I, II and III highlight two areas with potential for improvement in future projects like CMIP, aerosol treatment within the ESMs and in the input aerosol emission inventories. The bias between models and observations in Paper I is suggested to be caused by biases in emission inventories, while in Paper II observations-to-model bias seems to be a combination of biases in emission inventories and in model aerosol treatments, depending on the region investigated. While work in this thesis only focused on the aerosols sulfate and BC, other aerosol species with radiative impacts (see Figure 2.4) should also be investigated in regards to emission inventories over the historical period. Improving the accuracy of emission input inventories is an important step towards reducing the uncertainty in estimated aerosol forcing history as shown in Chapter 1. Our findings from Paper I and II show that observations-model biases are regionally very different, and a region-based approach for emission inventory investigations is recommended based on these findings.
The pure ESM comparison done in Paper III shows that there are large intermodel differences in absolute and temporal development of BC lifetime which result in a large intermodel difference in atmospheric absorption. However it is difficult to reduce the intermodel spread when observations of BC lifetime does not exist, so we recommend more work to be done towards understanding the life cycle of BC in the atmosphere, for then to determine which aerosol treatments within the models are most appropriate.

The next chapter will discuss the results of the three papers in more detail, together with limitations of the work and future outlook.

Chapter 5

:Conclusions, discussion and future outlook

The overall objective of this thesis is to **understand the aerosol forcing history through analyses of state-of-the-art Earth system models, and investigate whether a bias to observations can be attributed to model-errors or flawed emission inventories**. The enclosed studies contribute to this understanding by using two observational data sets to identify model biases, and a multitude of simulations from a total of 13 models from the CMIP6 archive (as explained in Chapter 3). Our findings show that the aerosol forcing history as estimated by Earth system models contains uncertainties both connected to aerosol processes within the models and to their emission inputs. This means that we have uncovered biases connected to uncertainties that are not frequently investigated, such as in emission inventories. The current chapter presents the conclusions and implications of the enclosed studies in connection with the specific research objectives presented in Chapter 1.

5.1 Concluding remarks and implications

Paper I of this thesis aims at *comparing gap filled observational surface shortwave radiation data to that of CMIP6 models, and evaluating model performance of surface energy balance believed to be directly related to anthropogenically emitted aerosols*. Investigating an aerosol forcing proxy such as downwelling shortwave surface radiation in ESMs and observations revealed an important bias, and these findings serve as excellent basis for the rest of the studies enclosed. Paper I concludes that models underestimate the aerosol forcing proxy, especially in the 1970-1990s, meaning more downwelling shortwave radiation reaches the surface in CMIP6 models globally than is measured by instruments. These results compare well to findings in *Storelvmo et al.* (2018) which investigated the same metric using an earlier generation of models (CMIP5). This means state-of-the-art Earth system models have not improved in regards to this global aerosol forcing proxy since the previous generation, despite the many developments in the field of aerosol research. In order to reduce this bias in future generation of models one needs to understand the cause of the bias. For this Paper I suggests that part of the bias can be sourced to underestimations in Asian sulfur dioxide emissions. The conclusions of Paper I differ from the ones in *Storelvmo et al.* (2018), and the reason for that is found in the regional investigations, especially in Asia.

While the CMIP5 ensemble produced practically no surface radiation trends in Asia, the subset of CMIP6 models used in Paper I did produce trends, although weaker than observed. When the CMIP5 aerosol emissions anticorrelates well with downwelling shortwave radiation yet still produce no CMIP5 radiative trends, the logical conclusion is that model deficiencies are to blame. As the CMIP6 model subset in Paper I did produce trends it is logical to call in emission inventory as a potential cause of bias.
Downwelling shortwave radiation plays an important part in our climate's energy balance (see Figure 2.1 in Chapter 2), and the global anomaly model-observation bias for this metric is found to be around 5 Wm^{-2} in the late 1980s in Paper I. Implications of our findings from Paper I include: if models represent the historical surface temperature development as observed, but allow more shortwave radiation to reach the surface than observed over the same period, the models are most likely simulating the right temperature for the wrong reason. In addition, we find that the ESM bias is of the same magnitude in CMIP5 as in CMIP6, despite almost ten years of model development from one to the other. This, together with the fact that the aerosol emission inventories of BC and sulfur dioxide have not changed greatly in magnitude from one generation to the other points to a need for investigations of the uncertainty in aerosol emission inventories.

The need for investigating aerosol emission inventories is then the basis for Paper II, of which the objective is *to make use of the under-utilized aerosol archive in ice cores in the evaluation on Earth system models' ability to reproduce aerosol concentrations over the industrial era*.
By comparing sulfate and BC concentrations in nine regions close and far from emission sources we assessed which biases could be attributed to errors in emission inventories and which could be due to aerosol process representation within the models. The 11 models of the study all agree in their mistake in BC concentration timing in Europe, which is confirmed using two Alpine ice cores. According to our observations, the maximum BC emissions in at least Western Europe must have occurred before 1950, not after, as is simulated in the models (see Chapter 2 and *Hoesly et al.* (2018)). While sulfate deposition evolution agrees in models and ice cores, the timing bias for BC points to a severe flaw in the emission inventory. Since BC deposition reached its maximum between 1910 and 1940 the atmospheric burden must also have been at its maximum at this time, unless BC lifetime in Europe was much longer in these decades. As explained in Chapter 2, BC is a complex aerosol which can absorb or scatter solar radiation depending on its interactions in the atmosphere. These two effects give diverging TOA ERFs, but both yield in a negative surface ERF. Therefore a higher pre-1950 load of BC than what is believed at present will imply a stronger negative BC contribution to the surface aerosol forcing, while the total climate effect of BC is more uncertain at present. Unfortunately the time period of the emission flaw in European BC as revealed by ice core data is not covered in Paper I, where shortwave downwelling radiation at the surface was well represented in Europe between 1961 and 2014.

Ice core concentration data located far from emission sources reveal large model differences in processes leading to aerosol concentrations in ice. In Greenland the models even disagreed on whether BC concentration exhibit a positive or negative trend between 1950 and present day. Since the temporal evolution of sulfate in ice cores was

generally simulated as observed in Greenland we assume that transport is not the cause of intermodel differences in BC concentration trend, but that this larger inconsistency for BC is rather due to an inter-model difference in BC lifetime and its evolution over time. Some models exhibit a decreasing BC lifetime over the historical period, and these are the models that show a negative BC concentration trend in Greenland ice. The ice cores in Greenland also show a negative BC concentration trend, but the temporal evolution is quite different form any of the models in the study. As source-attribution studies using NorESM2-LM show that much of the BC found in Greenland originate in Europe, the model-ice core biases there could be due to a mix of emission inventory flaws and inter-model differences. The climate implications of this study depend on the climate impacts of BC, which raises new questions related to intermodel differences in BC lifetime, and why the lifetime changes temporally in some models while not in others.

Paper III aims to aid in answering the questions raised by results in Paper II through the objective
to investigate BC lifetimes in CMIP6 over the historical era, and determine its importance for radiative effects.
Present day BC lifetime has an inter-model spread of 3.8 to 8.5 days, and the spread is even larger when investigating pre-industrial lifetimes (3.9 to 9.8 days). This reduction in spread is due to some models having a decreasing BC lifetime over the historical era, which we attribute to a considerable sensitivity of models and their BC ageing to anthropogenically added sulfate. Most models in this study have sulfate or sulfate equivalent sensitivities regarding the coating of BC, but the models differ in how much sulfate is needed to coat BC and make it soluble and be removed more quickly by precipitation. Some models which include BC sensitivities to soluble materials in their BC aerosol formulations show no historical BC lifetime reduction, probably due to that BC particles are quickly and sufficiently coated already at pre-industrial levels of soluble material to be removed by wet deposition. The three models in our study with a significant reduction of BC lifetime require more sulfate, i.e. anthropogenic emissions, to coat BC to become soluble than what was available in pre-industrial times. A BC particle which is coated by sulfate will act as a CCN and be wet deposited at the same rate as sulfate, and since sulfate in general has a lower lifetime than BC, coating leads to BC lifetime reductions.
The few studies that try to constrain BC lifetime with observations conclude that present day lifetime should be less than 5 days, but they do not discuss if BC lifetime has changed over the historical era or not. The results of Paper III then have a selection of implications, depending on which group of models represent reality. If pre-industrial levels of soluble materials are sufficient to coat BC particles we can also assume BC lifetime has remained semi-constant throughout the industrial era. Out of the models with semi-constant lifetimes, five models in Paper III have BC lifetimes over 5 days, and consequentially have a higher atmospheric absorption and surface ERF than their low-lifetime counterparts. If BC lifetime in reality is less than 5 days, the surface energy balance in long-lifetime models is either biased high, or right for the wrong reasons.

All the models in which BC lifetime is reduced over the historical era exhibit a present

day BC lifetime of over 5 days. However, if we assume the historical BC lifetime variations in these models to be true, the implications are slightly different than for the semi-constant-lifetime-scenario. The reduction in BC lifetime in these three models are on the order of 1-4 days, which are all significant lifetime changes with or without correcting for the general lifetime bias (less than 5 days). Since a reduction in lifetime implies a smaller surface forcing effect, the models which reduce their BC lifetime following sulfur dioxide emissions would also necessarily reduce their dimming efficiency by BC during this time, contributing to a weaker surface downwelling shortwave radiation anomaly for these models. This can be compared to results in Paper I, where models in general indeed underestimated the dimming effect, but we note that the underestimation in Paper I is found for models with and without historical BC lifetime reductions.
Given the three models in Paper III represent reality, a reduction of sulfur dioxide emissions will prolong BC lifetimes. In the last decades, air pollution policies have been successful in reducing global emissions of sulfur dioxide, and will continue to do so in the future. If future BC lifetime will be prolonged compared to present day we will have a stronger atmospheric absorption and a more negative surface ERF in the future than at present day, given BC emissions increase or remain unchanged.

In summary these studies have enhanced understanding of the aerosol forcing history by identifying where the CMIP6 models appear to perform well compared to long-term observations (Paper I: Surface downwelling shortwave radiation anomalies in Europe, and partly India. Paper II: long term sulfate concentration trends in ice across several Northern Hemisphere regions.). A subset of Models in CMIP6 are found to underestimate historical aerosol effects, and this may be due to underestimated aerosol emission in China in the late 20th century (Paper I) and missing European BC emissions in the early 20th century (Paper II). However, some models probably also overestimate historical BC effects, due to long BC lifetimes (Paper III).

5.2 Research limitations and future outlook

A central limitation to aerosol forcing history investigations is the few long term data sets available for aerosols and aerosol proxies. However, more data sets exist beyond the ones presented and used in work in this thesis. One such data set is sunshine duration measurements, which can be used to estimate AOD on cloud-free days (*Sanchez-Romero et al.*, 2016; *Wandji Nyamsi et al.*, 2020), and which extend all the way back to the invention of the Campbell-Stokes sunshine recorder in the late 19th century (*Sanchez-Romero et al.*, 2014). Sunshine duration measurements have previously shown to be consistent with surface solar downwelling radiation (*Allen et al.*, 2013), and would therefore complement our Paper I nicely. In addition, pan evaporation measurements and daily temperature range are also valuable metrics which can be used to investigate aerosol effects. Pan evaporation is a proxy for direct sunshine, and daily temperature range can divide between greenhouse gas warming and aerosol cooling as greenhouse gas warming happens 24 hours a day, while aerosol mainly interact with the solar radiation (*Wild*, 2009).

There are limitations to investigating all-sky downwelling shortwave surface radiation measurements as an aerosol forcing proxy, as these measurements are not only affected by aerosols. While previous studies have found aerosol-radiation interactions as the dominant cause of the observed dimming and brightening in Europe and East Asia (*Allen et al.* (2013) and references within), other studies have pointed to aerosol-radiation interactions and aerosol-cloud interactions playing an equal part (*Ohmura*, 2009). The disentangling of aerosol and cloud effects in CMIP6 and the radiative measurements used in Paper I is the topic of Julsrud et al (in review), which is listed in the Preface of this thesis as the second publication from the PhD period which is not included in this thesis.

In Paper I sulfate is used as a proxy for anthropogenic aerosols in Europe and China. However from the experience gained during this thesis (see Paper II and III) Paper I should have put similar attention to BC as sulfate. BC trends influence surface radiation efficiently and the temporal evolution of BC emissions is less understood, as our comparison to ice core derived trends clearly show.

The large intermodel spread in BC concentrations found in Greenland in Paper II were attributed to intermodel differences in BC lifetimes. However, models had large intermodel differences in sulfate concentrations as well, even in pre-industrial times. This needs to be further investigated, and a similar lifetime study as for BC in Paper III should be performed for sulfate.
While Paper III explains the historical reductions in lifetime as an effect of anthropogenically added soluble material, some investigations should be made to determine the climate effects of the different BC emission regions. Where the dominant regions for BC emissions are have changed over the historical period, with a shift from North America and Europe, to South- and South East Asia. This means BC is now emitted in drier climates, which should result in prolonged BC lifetimes.
Further efforts to constrain BC lifetime would benefit our understanding of the energy balance, both at the top of the atmosphere and at the surface. We recognize that observing BC lifetime and its sensitivity to sulfate remains a difficult task. So far such constraining efforts are made only by using observations from aircraft looking at vertical distribution of BC in remote regions.

Uncertainty regarding aerosol emission has been mentioned in work outside this thesis before (*Bond et al.*, 2013; *Hodnebrog et al.*, 2014). While Paper II suggests that the temporal evolution of sulfate emission *trends* from Europe and North America were correct, it could not present conclusions regarding emission magnitude. Papers I and II give some indications that aerosol emissions are underestimated in Asia and that BC emissions are underestimated in Europe in earlier decades of the 20th century. If major anthropogenic aerosol emission estimations have been underestimated over the historical era, the current model simulations are missing a cooling effect. If models accurately represent the observed temperature record without such an additional cooling component, they are also missing a compensating warming component. We can only speculate whether the warming potential of CO_2 is greater than previously thought, or that additional BC has added a warming component in the historical period. Further work on constraining the overall aerosol cooling is needed to enlighten the field. I sug-

gest that the making of uncertainty estimates connected to aerosol emissions should be the prioritized next step for improving confidence in temporal aerosol forcing estimates between present day and the dim and distant past.

Bibliography

Allen, R. J., J. R. Norris, and M. Wild (2013), Evaluation of multidecadal variability in CMIP5 surface solar radiation and inferred underestimation of aerosol direct effects over Europe, China, Japan, and India, *Journal of Geophysical Research: Atmospheres*, *118*(12), 6311–6336, doi:10.1002/jgrd.50426. 1.1, 5.2

Andres, R. J., D. J. Fielding, G. Marland, T. A. Boden, N. Kumar, and A. T. Kearney (1999), Carbon dioxide emissions from fossil-fuel use, 1751-1950, *Tellus B: Chemical and Physical Meteorology*, *51*(4), 759–765, doi:10.3402/tellusb.v51i4.16483, publisher: Taylor & Francis _eprint: https://doi.org/10.3402/tellusb.v51i4.16483. 2.3.1

Blichner, S. M., M. K. Sporre, R. Makkonen, and T. K. Berntsen (2021), Implementing a sectional scheme for early aerosol growth from new particle formation in the Norwegian Earth System Model v2: comparison to observations and climate impacts, *Geoscientific Model Development*, *14*(6), 3335–3359, doi:10.5194/gmd-14-3335-2021, publisher: Copernicus GmbH. 2.3.2

Boden, T. A., R. J. Andres, and G. Marland (2016), Global, Regional, and National Fossil-Fuel CO2 Emissions (1751 - 2013) (V. 2016), *Tech. rep.*, Environmental System Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE) (United States); Carbon Dioxide Information Analysis Center (CDIAC), Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States), doi:10.3334/CDIAC/00001_V2016. 2.3.1

Bond, T. C., D. G. Streets, K. F. Yarber, S. M. Nelson, J.-H. Woo, and Z. Klimont (2004), A technology-based global inventory of black and organic carbon emissions from combustion, *Journal of Geophysical Research: Atmospheres*, *109*(D14), doi:10.1029/2003JD003697, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2003JD003697. 2.3.1

Bond, T. C., E. Bhardwaj, R. Dong, R. Jogani, S. Jung, C. Roden, D. G. Streets, and N. M. Trautmann (2007), Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850âĂŞ2000, *Global Biogeochemical Cycles*, *21*(2), doi:10.1029/2006GB002840, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006GB002840. 2.3.1

Bond, T. C., S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman,

H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender (2013), Bounding the role of black carbon in the climate system: A scientific assessment, *Journal of Geophysical Research: Atmospheres*, *118*(11), 5380–5552, doi:10.1002/jgrd.50171, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50171. 2.2.1, 5.2

Breiman, L. (2001), Random Forests, *Machine Learning*, *45*(1), 5–32, doi:10.1023/A:1010933404324. 3.3.1

Charney, J. G., A. Arakawa, D. J. Baker, B. Bolin, R. E. Dickinson, R. M. Goody, C. E. Leith, H. M. Stommel, and C. I. Wunsch (1979), *Carbon Dioxide and Climate: A Scientific Assessment*, National Academy of Sciences, Washington, DC. 2.1

Cherian, R., J. Quaas, M. Salzmann, and M. Wild (2014), Pollution trends over Europe constrain global aerosol forcing as simulated by climate models, *Geophysical Research Letters*, *41*(6), 2176–2181, doi:10.1002/2013GL058715, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013GL058715. 1.1, 1.2, 2.2.2

Cherian, R., J. Quaas, M. Salzmann, and L. Tomassini (2017), Black carbon indirect radiative effects in a climate model, *Tellus B: Chemical and Physical Meteorology*, *69*(1), 1369,342, doi:10.1080/16000889.2017.1369342, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/16000889.2017.1369342. 2.2.1

Collins, W. J., J.-F. Lamarque, M. Schulz, O. Boucher, V. Eyring, M. I. Hegglin, A. Maycock, G. Myhre, M. Prather, D. Shindell, and S. J. Smith (2017), AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, *Geosci. Model Dev.*, *10*(2), 585–607, doi:10.5194/gmd-10-585-2017. 3.1.2

Danabasoglu, G., J.-F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, C. Hannay, M. M. Holland, W. G. Large, P. H. Lauritzen, D. M. Lawrence, J. T. M. Lenaerts, K. Lindsay, W. H. Lipscomb, M. J. Mills, R. Neale, K. W. Oleson, B. Otto-Bliesner, A. S. Phillips, W. Sacks, S. Tilmes, L. v. Kampenhout, M. Vertenstein, A. Bertini, J. Dennis, C. Deser, C. Fischer, B. Fox-Kemper, J. E. Kay, D. Kinnison, P. J. Kushner, V. E. Larson, M. C. Long, S. Mickelson, J. K. Moore, E. Nienhouse, L. Polvani, P. J. Rasch, and W. G. Strand (2020), The Community Earth System Model Version 2 (CESM2), *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001,916, doi:10.1029/2019MS001916, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001916. 3.2

Diamond, M. S., H. M. Director, R. Eastman, A. Possner, and R. Wood (2020), Substantial Cloud Brightening From Shipping in Subtropical Low Clouds, *AGU Advances*, *1*(1), e2019AV000,111, doi:10.1029/2019AV000111, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019AV000111. 1.1

Engardt, M., D. Simpson, M. Schwikowski, and L. Granat (2017), Deposition of sulphur and nitrogen in Europe 1900-2050. Model calculations and comparison to historical observations, *Tellus B: Chemical and Physical Meteorology*, *69*(1), 1328,945, doi:10.1080/16000889.2017.1328945. 3.2

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016), Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, *9*(5), 1937–1958, doi:https://doi.org/10.5194/gmd-9-1937-2016. 2.3, 3.1, 3.1.1

Fiedler, S., S. Kinne, W. T. K. Huang, P. Räisänen, D. O'Donnell, N. Bellouin, P. Stier, J. Merikanto, T. van Noije, R. Makkonen, and U. Lohmann (2019), Anthropogenic aerosol forcing - insights from multiple estimates from aerosol-climate models with reduced complexity, *Atmospheric Chemistry and Physics*, *19*(10), 6821–6841, doi: 10.5194/acp-19-6821-2019, publisher: Copernicus GmbH. 1.1

Foote, E. (1856), Circumstances affecting the heat of the sun's rays, *The American Journal of Science and Arts*, *XXII*(2), 2nd Series, v. XXII/no. LXVI,. 2.1

Forster, P., T. Storelvmo, K. Armour, W. Collins, J.-L. Dufresne, D. Frame, D. J. Lunt, T. Mauritsen, M. D. Palmer, M. Watanabe, M. Wild, and X. Zhang (2021), The Earth's energy budget, climate feedbacks, and climate sensitivity, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, ö. Yelekçi, R. Yu, and B. Zhou, Cambridge University Press. 1.2, 2.1

Fröhlich, C. (2006), Solar Irradiance Variability Since 1978, in *Solar Variability and Planetary Climates*, edited by Y. Calisesi, R. M. Bonnet, L. Gray, J. Langen, and M. Lockwood, Space Sciences Series of ISSI, pp. 53–65, Springer, New York, NY, doi:10.1007/978-0-387-48341-2_5. 2.1.1

Galmarini, S., B. Koffi, E. Solazzo, T. Keating, C. Hogrefe, M. Schulz, A. Benedictow, J. J. Griesfeller, G. Janssens-Maenhout, G. Carmichael, J. Fu, and F. Dentener (2017), Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, *Atmospheric Chemistry and Physics*, *17*(2), 1543–1555, doi:10.5194/acp-17-1543-2017, publisher: Copernicus GmbH. 3.1

Gettelman, A. (2015), Putting the clouds back in aerosolâĂŞcloud interactions, *Atmospheric Chemistry and Physics*, *15*(21), 12,397–12,411, doi:https://doi.org/10.5194/acp-15-12397-2015. 1.1

Ghan, S. J. (2013), Technical Note: Estimating aerosol effects on cloud radiative forcing, *Atmospheric Chemistry and Physics*, *13*(19), 9971–9974, doi:https://doi.org/10.5194/acp-13-9971-2013. 2.4

Gillett, N. P., H. Shiogama, B. Funke, G. Hegerl, R. Knutti, K. Matthes, B. D. Santer, D. Stone, and C. Tebaldi (2016), The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6, *Geoscientific Model Development*, *9*(10), 3685–3697, doi:https://doi.org/10.5194/gmd-9-3685-2016. 3.1.4

Gjermundsen, A., A. Nummelin, D. Olivié, M. Bentsen, Ø. Seland, and M. Schulz (2021), Shutdown of Southern Ocean convection controls long-term greenhouse gas-induced warming, *Nat. Geosci.*, *14*(10), 724–731, doi:10.1038/s41561-021-00825-x, bandiera_abtest: a Cg_type: Nature Research Journals Number: 10 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Climate and Earth system modelling;Climate change;Ocean sciences;Projection and prediction Subject_term_id: climate-and-earth-system-modelling;climate-change;ocean-sciences;projection-and-prediction. 3.2

Gulev, S. K., P. W. Thorne, J. Ahn, F. J. Dentener, C. M. Domingues, S. Gerland, D. Gong, D. S. Kaufman, H. C. Nnamchi, J. Quaas, J. A. Rivera, S. Sathyendranath, S. L. Smith, B. Trewin, K. von Shuckmann, and R. S. Vose (2021), Changing state of the climate system, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, ö. Yelekçi, R. Yu, and B. Zhou, Cambridge University Press. 1.1

Herman, J. R., P. K. Bhartia, O. Torres, C. Hsu, C. Seftor, and E. Celarier (1997), Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data, *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16,911–16,922, doi:10.1029/96JD03680, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/96JD03680. 2.2.2

Hodnebrog, O., G. Myhre, and B. H. Samset (2014), How shorter black carbon lifetime alters its climate effect, *Nat Commun*, *5*(1), 5065, doi:10.1038/ncomms6065, bandiera_abtest: a Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Atmospheric science Subject_term_id: atmospheric-science. 5.2

Hoesly, R. M., S. J. Smith, L. Feng, Z. Klimont, G. Janssens-Maenhout, T. Pitkanen, J. J. Seibert, L. Vu, R. J. Andres, R. M. Bolt, T. C. Bond, L. Dawidowski, N. Kholod, J.-i. Kurokawa, M. Li, L. Liu, Z. Lu, M. C. P. Moura, P. R. O'Rourke, and Q. Zhang (2018), Historical (1750âĂŞ2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geoscientific Model Development*, *11*(1), 369–408, doi:https://doi.org/10.5194/gmd-11-369-2018. 1.1, 2.3.1, 2.6, 3.1.1, 3.1, 5.1

IPCC (2021), Summary for policymakers, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, ö. Yelekçi, R. Yu, and B. Zhou, Cambridge University Press. 1.1

Karset, I. H. H., T. K. Berntsen, T. Storelvmo, K. Alterskjaer, A. Grini, D. Olivié, A. Kirkevåg, Ø. Seland, T. Iversen, and M. Schulz (2018), Strong impacts on aerosol indirect effects from historical oxidant changes, *Atmospheric Chemistry and Physics*,

18(10), 7669–7690, doi:10.5194/acp-18-7669-2018, publisher: Copernicus GmbH. 1.1

Kellerhals, T., S. Brütsch, M. Sigl, S. Knüsel, H. W. Gäggeler, and M. Schwikowski (2010), Ammonium concentration in ice cores: A new proxy for regional temperature reconstruction?, *Journal of Geophysical Research: Atmospheres*, *115*(D16), doi:10.1029/2009JD012603, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD012603. 3.2

Kirkevåg, A., A. Grini, D. Olivié, Ø. Seland, K. Alterskjaer, M. Hummel, I. H. H. Karset, A. Lewinschal, X. Liu, R. Makkonen, I. Bethke, J. Griesfeller, M. Schulz, and T. Iversen (2018), A production-tagged aerosol module for Earth system models, OsloAero5.3 - extensions and updates for CAM5.3-Oslo, *Geoscientific Model Development*, *11*(10), 3945–3982, doi:10.5194/gmd-11-3945-2018, publisher: Copernicus GmbH. 2.3.2, 3.2, 3.2.1

Kvalevåg, M. M., and G. Myhre (2007), Human Impact on Direct and Diffuse Solar Radiation during the Industrial Era, *Journal of Climate*, *20*(19), 4874–4883, doi:10.1175/JCLI4277.1, publisher: American Meteorological Society Section: Journal of Climate. 2.1.1

Lamarque, J.-F., T. C. Bond, V. Eyring, C. Granier, A. Heil, Z. Klimont, D. Lee, C. Liousse, A. Mieville, B. Owen, M. G. Schultz, D. Shindell, S. J. Smith, E. Stehfest, J. V. Aardenne, O. R. Cooper, M. Kainuma, N. Mahowald, J. R. McConnell, V. Naik, K. Riahi, and D. P. v. Vuuren (2010), Historical (1850âĂŞ2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmospheric Chemistry and Physics*, *10*(15), 7017–7039, doi:https://doi.org/10.5194/acp-10-7017-2010. 2.6

Legrand, M., S. Preunkert, B. May, J. Guilhermet, H. Hoffman, and D. Wagenbach (2013), Major 20th century changes of the content and chemical speciation of organic carbon archived in Alpine ice cores: Implications for the long-term change of organic aerosol over Europe, *Journal of Geophysical Research: Atmospheres*, *118*(9), 3879–3890, doi:https://doi.org/10.1002/jgrd.50202, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50202. 3.2

Leirvik, T., and M. Yuan (2021), A Machine Learning Technique for Spatial Interpolation of Solar Radiation Observations, *Earth and Space Science*, *8*(4), e2020EA001,527, doi:10.1029/2020EA001527, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020EA001527. 3.3.1

Liepert, B. G. (2002), Observed reductions of surface solar radiation at sites in the United States and worldwide from 1961 to 1990, *Geophysical Research Letters*, *29*(10), 61–1–61–4, doi:10.1029/2002GL014910. 2.1.1

Lim, S., X. Faïn, P. Ginot, V. Mikhalenko, S. Kutuzov, J.-D. Paris, A. Kozachek, and P. Laj (2017), Black carbon variability since preindustrial times in the eastern part of Europe reconstructed from Mt. Elbrus, Caucasus, ice cores, *Atmospheric Chemistry and Physics*, *17*(5), 3489–3505, doi:10.5194/acp-17-3489-2017, publisher: Copernicus GmbH. 3.2

Liu, J., S. Fan, L. W. Horowitz, and H. Levy II (2011), Evaluation of factors controlling long-range transport of black carbon to the Arctic, *Journal of Geophysical Research: Atmospheres*, *116*(D4), doi:10.1029/2010JD015145, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2010JD015145. 2.2.1

Liu, X., R. C. Easter, S. J. Ghan, R. Zaveri, P. Rasch, X. Shi, J.-F. Lamarque, A. Gettelman, H. Morrison, F. Vitt, A. Conley, S. Park, R. Neale, C. Hannay, A. M. L. Ekman, P. Hess, N. Mahowald, W. Collins, M. J. Iacono, C. S. Bretherton, M. G. Flanner, and D. Mitchell (2012), Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, *Geoscientific Model Development*, *5*(3), 709–739, doi:10.5194/gmd-5-709-2012, publisher: Copernicus GmbH. 3.2.1

Liu, X., P.-L. Ma, H. Wang, S. Tilmes, B. Singh, R. C. Easter, S. J. Ghan, and P. J. Rasch (2016), Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model, *Geoscientific Model Development*, *9*(2), 505–522, doi:10.5194/gmd-9-505-2016, publisher: Copernicus GmbH. 2.3.2, 3.2.1

Loeb, N. G., D. R. Doelling, H. Wang, W. Su, C. Nguyen, J. G. Corbett, L. Liang, C. Mitrescu, F. G. Rose, and S. Kato (2018), Clouds and the earth's radiant energy system (ceres) energy balanced and filled (ebaf) top-of-atmosphere (toa) edition-4.0 data product, *Journal of Climate*, *31*(2), 895 – 918, doi:10.1175/JCLI-D-17-0208.1. 1.1

McConnell, J. R. (2007), 20th century industrial black carbon emissions altered and climate forcing, *SCIENCE*, *317*. 3.2

McConnell, J. R. (2010), New Directions: Historical black carbon and other ice core aerosol records in the Arctic for GCM evaluation, *Atmospheric Environment*, *44*(21), 2665–2666, doi:10.1016/j.atmosenv.2010.04.004. 3.2

McConnell, J. R., and R. Edwards (2008), Coal burning leaves toxic heavy metal legacy in the Arctic, *PNAS*, *105*(34), 12,140–12,144, doi:10.1073/pnas.0803564105. 3.2

McGraw, Z., T. Storelvmo, B. H. Samset, and C. W. Stjern (2020), Global Radiative Impacts of Black Carbon Acting as Ice Nucleating Particles, *Geophysical Research Letters*, *47*(20), e2020GL089,056, doi:10.1029/2020GL089056, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL089056. 2.2.1

Naik, V., S. Szopa, B. Adhikary, P. E. Artaxo Netto, T. Berntsen, W. D. Collins, S. Fuzzi, L. Gallardo, A. Kiendler-Scharr, Z. Klimont, H. Liao, N. Unger, and P. Zanis (2021), Short-lived climate forcers, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, ö. Yelekçi, R. Yu, and B. Zhou, Cambridge University Press. 1.1, 2.2.1, 2.4

Ohmura, A. (2009), Observed decadal variations in surface solar radiation and their causes, *Journal of Geophysical Research: Atmospheres*, *114*(D10), doi:10.1029/2008JD011290, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD011290. 5.2

Osmont, D., M. Sigl, A. Eichler, T. M. Jenk, and M. Schwikowski (2019), A Holocene black carbon ice-core record of biomass burning in the Amazon Basin from Illimani, Bolivia, *Climate of the Past*, *15*(2), 579–592, doi:10.5194/cp-15-579-2019, publisher: Copernicus GmbH. 3.2

Pincus, R., P. M. Forster, and B. Stevens (2016), The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, *Geoscientific Model Development*, *9*(9), 3447–3460, doi:10.5194/gmd-9-3447-2016, publisher: Copernicus GmbH. 3.1.3

Preunkert, S., M. Legrand, and D. Wagenbach (2001), Sulfate trends in a Col du Dôme (French Alps) ice core: A record of anthropogenic sulfate levels in the European midtroposphere over the twentieth century, *J. Geophys. Res.*, *106*(D23), 31,991–32,004, doi:10.1029/2001JD000792. 3.2

Preunkert, S., M. Legrand, S. Kutuzov, P. Ginot, V. Mikhalenko, and R. Friedrich (2019), The Elbrus (Caucasus, Russia) ice core record - Part 1: reconstruction of past anthropogenic sulfur emissions in south-eastern Europe, *Atmospheric Chemistry and Physics*, *19*(22), 14,119–14,132, doi:10.5194/acp-19-14119-2019, publisher: Copernicus GmbH. 3.2

Sanchez-Lorenzo, A., M. Wild, M. Brunetti, J. A. Guijarro, M. Z. Hakuba, J. Calbó, S. Mystakidis, and B. Bartok (2015), Reassessment and update of long-term trends in downward surface shortwave radiation over Europe (1939-2012), *Journal of Geophysical Research: Atmospheres*, *120*(18), 9555–9569, doi:10.1002/2015JD023321. 2.1.1

Sanchez-Romero, A., A. Sanchez-Lorenzo, J. Calbó, J. A. González, and C. Azorin-Molina (2014), The signal of aerosol-induced changes in sunshine duration records: A review of the evidence, *Journal of Geophysical Research: Atmospheres*, *119*(8), 4657–4673, doi:10.1002/2013JD021393, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD021393. 5.2

Sanchez-Romero, A., A. Sanchez-Lorenzo, J. A. González, and J. Calbó (2016), Reconstruction of long-term aerosol optical depth series with sunshine duration records, *Geophysical Research Letters*, *43*(3), 1296–1305, doi:10.1002/2015GL067543, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL067543. 5.2

Schwarz, M., D. Folini, S. Yang, R. P. Allan, and M. Wild (2020), Changes in atmospheric shortwave absorption as important driver of dimming and brightening, *Nat. Geosci.*, *13*(2), 110–115, doi:10.1038/s41561-019-0528-y, bandiera_abtest: a Cg_type: Nature Research Journals Number: 2 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Atmospheric science;Climate change;Climate sciences;Environmental sciences Subject_term_id: atmospheric-science;climate-change;climate-sciences;environmental-sciences. 2.1.1

Seland, O., M. Bentsen, L. Seland Graff, D. Olivié, T. Toniazzo, A. Gjermundsen, J. B. Debernard, A. K. Gupta, Y. He, A. Kirkevåg, J. Schwinger, J. Tjiputra, K. Schancke Aas, I. Bethke, Y. Fan, J. Griesfeller, A. Grini, C. Guo, M. Ilicak, I. H. Hafsahl Karset, O. Landgren, J. Liakka, K. Onsum Moseid, A. Nummelin, C. Spensberger, H. Tang, Z. Zhang, C. Heinze, T. Iverson, and M. Schulz (2020), The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations, *Geoscientific Model Development Discussions*, pp. 1–68, doi:https://doi.org/10.5194/gmd-2019-378, publisher: Copernicus GmbH. 3.2

Sigl, M., J. R. McConnell, L. Layman, O. Maselli, K. McGwire, D. Pasteris, D. Dahl-Jensen, J. P. Steffensen, B. Vinther, R. Edwards, R. Mulvaney, and S. Kipfstuhl (2013), A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years, *Journal of Geophysical Research: Atmospheres*, *118*(3), 1151–1169, doi:10.1029/2012JD018603, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2012JD018603. 3.2

Sigl, M., M. Winstrup, J. R. McConnell, K. C. Welten, G. Plunkett, F. Ludlow, U. Büntgen, M. Caffee, N. Chellman, D. Dahl-Jensen, H. Fischer, S. Kipfstuhl, C. Kostick, O. J. Maselli, F. Mekhaldi, R. Mulvaney, R. Muscheler, D. R. Pasteris, J. R. Pilcher, M. Salzer, S. Schüpbach, J. P. Steffensen, B. M. Vinther, and T. E. Woodruff (2015), Timing and climate forcing of volcanic eruptions for the past 2,500 years, *Nature*, *523*(7562), 543–549, doi:10.1038/nature14565. 3.2

Sigl, M., N. J. Abram, J. Gabrieli, T. M. Jenk, D. Osmont, and M. Schwikowski (2018), 19th century glacier retreat in the Alps preceded the emergence of industrial black carbon deposition on high-alpine glaciers, *The Cryosphere*, *12*(10), 3311–3331, doi:10.5194/tc-12-3311-2018, publisher: Copernicus GmbH. 3.2

Storelvmo, T., U. K. Heede, T. Leirvik, P. C. B. Phillips, P. Arndt, and M. Wild (2018), Lethargic Response to Aerosol Emissions in Current Climate Models, *Geophysical Research Letters*, *0*(0), doi:10.1029/2018GL078298. 1.1, 3.1, 3.3.1, 5.1

Thornhill, G., W. Collins, D. Olivié, R. B. Skeie, A. Archibald, S. Bauer, R. Checa-Garcia, S. Fiedler, G. Folberth, A. Gjermundsen, L. Horowitz, J.-F. Lamarque, M. Michou, J. Mulcahy, P. Nabat, V. Naik, F. M. O'Connor, F. Paulot, M. Schulz, C. E. Scott, R. Séférian, C. Smith, T. Takemura, S. Tilmes, K. Tsigaridis, and J. Weber (2021), Climate-driven chemistry and aerosol feedbacks in CMIP6 Earth system models, *Atmospheric Chemistry and Physics*, *21*(2), 1105–1126, doi:https://doi.org/10.5194/acp-21-1105-2021, publisher: Copernicus GmbH. 2.4

Twomey, S. (1977), The Influence of Pollution on the Shortwave Albedo of Clouds, *Journal of the Atmospheric Sciences*, *34*(7), 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences. 2.2.1

van Marle, M. J. E., S. Kloster, B. I. Magi, J. R. Marlon, A.-L. Daniau, R. D. Field, A. Arneth, M. Forrest, S. Hantson, N. M. Kehrwald, W. Knorr, G. Lasslop, F. Li, S. Mangeon, C. Yue, J. W. Kaiser, and G. R. van der Werf (2017), Historic global

biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750âĂŞ2015), *Geoscientific Model Development*, *10*(9), 3329–3357, doi:10.5194/gmd-10-3329-2017, publisher: Copernicus GmbH. 3.1.1, 3.1

von Schuckmann, K., L. Cheng, M. D. Palmer, J. Hansen, C. Tassone, V. Aich, S. Adusumilli, H. Beltrami, T. Boyer, F. J. Cuesta-Valero, D. DesbruyÃĺres, C. Domingues, A. García-García, P. Gentine, J. Gilson, M. Gorfer, L. Haimberger, M. Ishii, G. C. Johnson, R. Killick, B. A. King, G. Kirchengast, N. Kolodziejczyk, J. Lyman, B. Marzeion, M. Mayer, M. Monier, D. P. Monselesan, S. Purkey, D. Roemmich, A. Schweiger, S. I. Seneviratne, A. Shepherd, D. A. Slater, A. K. Steiner, F. Straneo, M.-L. Timmermans, and S. E. Wijffels (2020), Heat stored in the Earth system: where does the energy go?, *Earth System Science Data*, *12*(3), 2013–2041, doi:10.5194/essd-12-2013-2020, publisher: Copernicus GmbH. 2.1

Wandji Nyamsi, W., A. Lipponen, A. Sanchez-Lorenzo, M. Wild, and A. Arola (2020), A hybrid method for reconstructing the historical evolution of aerosol optical depth from sunshine duration measurements, *Atmospheric Measurement Techniques*, *13*(6), 3061–3079, doi:10.5194/amt-13-3061-2020, publisher: Copernicus GmbH. 5.2

Wild, M. (2009), Global dimming and brightening: A review, *Journal of Geophysical Research*, *114*, doi:10.1029/2008JD011470. 2.1.1, 5.2

Wild, M., D. Folini, M. Z. Hakuba, C. Schär, S. I. Seneviratne, S. Kato, D. Rutan, C. Ammann, E. F. Wood, and G. König-Langlo (2015), The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models, *Clim Dyn*, *44*(11), 3393–3429, doi:10.1007/s00382-014-2430-z. 2.1

Wild, M., A. Ohmura, C. Schär, G. Müller, D. Folini, M. Schwarz, M. Z. Hakuba, and A. Sanchez-Lorenzo (2017), The Global Energy Balance Archive (GEBA) version 2017: a database for worldwide measured surface energy fluxes, *Earth System Science Data*, *9*(2), 601–613, doi:https://doi.org/10.5194/essd-9-601-2017. 2.2, 3.3.1

Wild, M., M. Z. Hakuba, D. Folini, P. Dörig-Ott, C. Schär, S. Kato, and C. N. Long (2019), The cloud-free global energy balance and inferred cloud radiative effects: an assessment based on direct observations and climate models, *Clim Dyn*, *52*(7), 4787–4812, doi:10.1007/s00382-018-4413-y. 2.1.1, 2.1

Zennaro, P., N. Kehrwald, J. R. McConnell, S. Schüpbach, O. J. Maselli, J. Marlon, P. Vallelonga, D. Leuenberger, R. Zangrando, A. Spolaor, M. Borrotti, E. Barbaro, A. Gambaro, and C. Barbante (2014), Fire in ice: two millennia of boreal forest fire history from the Greenland NEEM ice core, *Climate of the Past*, *10*(5), 1905–1924, doi:10.5194/cp-10-1905-2014, publisher: Copernicus GmbH. 3.2

Part II

Papers

I

Paper I

Bias in CMIP6 models as compared to observed regional dimming and brightening

Kine Onsum Moseid, Michael Schulz, Trude Storelvmo, Ingeborg Rian Julsrud, Dirk Olivié, Pierre Nabat, Martin Wild, Jason N.S. Cole, Toshihiko Takemura, Naga Oshima, Susanne E. Bauer, Guillaume Gastineau
Atmospheric Chemistry and Physics, 2020
doi:10.5194/acp-20-16023-2020

Atmos. Chem. Phys., 20, 16023–16040, 2020
https://doi.org/10.5194/acp-20-16023-2020

Atmospheric Chemistry and Physics
Open Access
EGU

Bias in CMIP6 models as compared to observed regional dimming and brightening

Kine Onsum Moseid[1], Michael Schulz[1,2], Trude Storelvmo[2], Ingeborg Rian Julsrud[2,1], Dirk Olivié[1], Pierre Nabat[3], Martin Wild[4], Jason N. S. Cole[5], Toshihiko Takemura[6], Naga Oshima[7], Susanne E. Bauer[8], and Guillaume Gastineau[9]

[1]Norwegian Meteorological Institute, Research Department, Oslo, Norway
[2]University of Oslo,Department of Geosciences, Section for Meteorology and Oceanography, Oslo, Norway
[3]Centre National de Recherches Meteorologiques (CNRM), Universite de Toulouse, Météo-France, CNRS, Toulouse, France
[4]Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland
[5]Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada
[6]Climate Change Science Section, Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan
[7]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Ibaraki, Japan
[8]Center for Climate Systems Research, Columbia University, NASA Goddard Institute for Space Studies, New York, NY, USA
[9]LOCEAN, IPSL, Sorbonne Université, IRD, MNHN, CNRS, Paris, France

Correspondence: Kine Onsum Moseid (kristineom@met.no)

Received: 30 December 2019 – Discussion started: 4 February 2020
Revised: 9 November 2020 – Accepted: 9 November 2020 – Published: 22 December 2020

Abstract. Anthropogenic aerosol emissions have increased considerably over the last century, but climate effects and quantification of the emissions are highly uncertain as one goes back in time. This uncertainty is partly due to a lack of observations in the pre-satellite era, making the observations we do have before 1990 additionally valuable. Aerosols suspended in the atmosphere scatter and absorb incoming solar radiation and thereby alter the Earth's surface energy balance. Previous studies show that Earth system models (ESMs) do not adequately represent surface energy fluxes over the historical era. We investigated global and regional aerosol effects over the time period 1961–2014 by looking at surface downwelling shortwave radiation (SDSR). We used observations from ground stations as well as multiple experiments from eight ESMs participating in the Coupled Model Intercomparison Project Version 6 (CMIP6). Our results show that this subset of models reproduces the observed transient SDSR well in Europe but poorly in China. We suggest that this may be attributed to missing emissions of sulfur dioxide in China, sulfur dioxide being a precursor to sulfate, which is a highly reflective aerosol and responsible for more reflective clouds. The emissions of sulfur dioxide used in the models do not show a temporal pattern that could explain observed SDSR evolution over China. The results from various aerosol emission perturbation experiments from DAMIP, RFMIP and AerChemMIP show that only simulations containing anthropogenic aerosol emissions show dimming, even if the dimming is underestimated. Simulated clear-sky and all-sky SDSR do not differ greatly, suggesting that cloud cover changes are not a dominant cause of the biased SDSR evolution in the simulations. Therefore we suggest that the discrepancy between modeled and observed SDSR evolution is partly caused by erroneous aerosol and aerosol precursor emission inventories. This is an important finding as it may help interpret whether ESMs reproduce the historical climate evolution for the right or wrong reason.

1 Introduction

Aerosol particles scatter and absorb radiation and change the radiative properties of clouds, thereby altering Earth's energy balance (Boucher et al., 2013). Anthropogenic aerosol emissions have substantially increased over the last century, but the quantification of the effect has been characterized by large uncertainties. Earth system models (ESMs) are evalu-

ated based on their ability to reproduce the climate evolution of the past 165 years, and the sparsity of aerosol-related observations in the pre-satellite era plays a dominant role in the uncertainty connected to these historical experiments. An improved understanding of the historical aerosol effect would increase the accuracy and credibility of ESM future climate projections.

Aerosol particles cause changes in the amount of sunlight reaching the surface together with changes in insolation, cloud cover, water vapor and other radiatively active gases (Wild et al., 2018). Extraterrestrial influences like the 11-year cycle of the Sun have not created any important trends on decadal timescales in Earth's surface solar radiation in the past century (Eddy et al., 1982; Wild, 2009). Water vapor amount has not changed sufficiently in recent decades to have an effect on decadal fluctuations of incoming sunlight at the surface (Wild, 2009; Wang and Yang, 2014; Yang et al., 2019; Hoyt and Schatten, 1993; Ramanathan and Vogelmann, 1997; Solomon et al., 2010), and radiatively active gases dominate in the longwave spectrum (Ramanathan et al., 1989).

The relative roles of clouds, aerosols and their interactions in historical variations of surface downwelling shortwave radiation (SDSR) are still disputed, but previous studies have found that aerosol effects dominate on multidecadal timescales, while cloud effects are relevant on shorter timescales (Wild, 2016; Romanou et al., 2007). Aerosol effects can be divided into the direct and indirect effects. The direct effect is the scatter or absorption directly caused by a dry aerosol, also called the aerosol–radiation interaction (ari) (Boucher et al., 2013), and the indirect effect is how aerosols change properties in clouds, also called aerosol–cloud interactions (aci). Aci includes both a change in cloud lifetime and most importantly a change in cloud albedo, making the cloud appear brighter (Boucher et al., 2013).

Assuming aerosol effects dominate the multidecadal timescales, SDSR can serve as a proxy for aerosol effects. The Global Energy Balance Archive (GEBA) dataset contains measurements of SDSR as far back as in 1922 (Wild et al., 2017) and as such represents a unique and valuable dataset for evaluation of simulated aerosol effects prior to the satellite era.

Observed SDSR from the GEBA dataset reveals a widespread negative trend from the 1950s to the late 1980s, commonly referred to as "global dimming" (Liepert, 2002; Wild, 2016). The magnitude of this dimming differs vastly between regions, which is expected if the cause of dimming were regionally different increases in aerosol emissions, as has been proposed by Wild et al. (2007), Sanchez-Romero et al. (2014), and Wild (2016). In some areas a positive trend in SDSR follows the dimming, and this SDSR increase has been termed "brightening" (Wild et al., 2005). Brightening is connected to the reduction in anthropogenic aerosol emission (Nabat et al., 2014). Fewer particles suspended in the air allow for more sunlight to reach the surface and thus an increase in the measured SDSR. Previous studies show that historical simulations from ESMs do not reproduce the observed global transient development of SDSR (Storelvmo et al., 2018; Wild, 2009; Allen et al., 2013; Wild and Schmucki, 2011). The cause of this discrepancy is not known but may be connected to uncertainties in aerosol emission inventories of the past, or, as Storelvmo et al. (2018) suggested, other uncertainties concern how models treat processes that translate aerosol emissions into radiative forcing.

In this study we use gap-filled data based on the GEBA dataset, together with several recent CMIP6 historical model experiments from eight climate models to investigate the aerosol effect in the time period 1961–2014, globally and regionally. In the middle of this time period (around the late 1990s), the main region of high anthropogenic aerosol emissions shifted from Europe and North America to Asia. We have chosen to focus on the regions of Europe and Asia in this study, as the models exhibit diverging abilities to reproduce the observed SDSR in these regions. We also use observational cloud cover data to briefly assess the role of cloud cover in the historical development of SDSR. We explore the relation between regional SDSR and aerosol emissions using a set of ESM experiments with differing aerosol emissions; some have pre-industrial aerosol emissions, while others use the most recent and best available historical aerosol emission inventory (Hoesly et al., 2018). This paper thereby provides new insights into the question of whether state-of-the-art ESMs can adequately reproduce a part of the changes in the surface energy budget over the historical era. This is in turn an important indication of whether the ESMs reproduce the dominant processes governing the historical climate evolution.

The paper is structured as follows. In Sect. 2 we begin by presenting the two observational datasets used, followed by a detailed description of the experiments simulated by the eight models chosen to be part of this study. The methods used to obtain and analyze the data finalize Sect. 2. The results are presented in Sect. 3, starting with a global view of dimming and brightening before focusing on regional assessments of SDSR, clear-sky SDSR, and cloud cover. Section 4 discusses the implications of our results and how they compare to previous studies, before final conclusions are presented in Sect. 5.

2 Data and methods

2.1 Observations

The GEBA holds data from ground-based stations measuring energy fluxes at the Earth's surface around the globe (Wild et al., 2017). Pyranometers were used in most of the measurement sites, which have an accuracy limitation of 3 %–5 % of the full signal (Michalsky et al., 1999; Wild et al., 2013). We use the monthly mean data from 1487 stations

in the time period 1961–2014 measuring downwelling shortwave radiation. The GEBA dataset has been complemented by a machine learning technique (random forests, Breiman, 2001) as explained in Storelvmo et al. (2018) to cover time periods of missing observations in the measurements and facilitate comparison to the gridded model data. This allows for all 1487 stations to have data on each time step, so that all regions have a complete record and the same number of stations throughout the entire time period in question.

Monthly mean cloud cover data are provided by the Climatic Research Unit (University of East Anglia) and NCAS, and we are using version 4.02 of this dataset (CRU). CRU covers the period 1901–2017 (Harris et al., 2020) and consists of a climatology made from measurements at meteorological stations around the globe, interpolated to a 0.5° latitude–longitude resolution grid covering continental areas. Information on interpolation methods and procedures used to create the gridded dataset is given in Harris et al. (2020) and references therein. In short, CRU has its foundation in station data but is interpolated to a grid using angular-distance weighting. The cloud cover variable is largely derived as a secondary variable, based on measurements of other parameters such as sunshine hours and diurnal temperature range.

2.2 Models and CMIP6

Eight climate models (NorESM2, CanESM5, MIROC6, CESM2, CNRM-ESM2-1, GISS-E2-1-G, IPSL-CM6A-LR, MRI-ESM2-0) were chosen for this study, based on available data and their involvement in relevant model intercomparison projects within the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016). As this study focuses on dimming and brightening, we have chosen experiments from model intercomparison projects (MIPs) that include perturbed historical simulations with which one can single out the effect of anthropogenic aerosol emissions in our diagnostic variables. An overview of models and experiments can be found in Table 1. This section will give a more detailed description of the experiments in Table 1 and explain why they were chosen.

Every model that takes part in CMIP6 has to deliver a set of common experiments; among these is the *historical* simulation. As can be seen in Table 1, all the models have provided historical simulation results. All other experiments listed in Table 1 are simulations covering the historical period (1850–2014) but with specific alterations dependent on what model intercomparison project they are a part of.

The Detection and Attribution Model Intercomparison Project (DAMIP) has the goal of improving estimations of the climate response to individual forcings (Gillett et al., 2016) and includes three relevant experiments. One experiment traces exclusively the impact of anthropogenically emitted aerosols as forcing agents over the historical period and is called *hist-aer*. This means no anthropogenic greenhouse gas emissions or natural climate forcings are used in this simulation. The *hist-nat* experiment consists of only the perturbations due to the evolution of the natural forcing, e.g., from stratospheric aerosols of volcanic origin and solar irradiance variations. Finally, the *hist-GHG* experiment has only forcings from changes in the well-mixed greenhouse gases. These experiments were chosen as they give a unique insight into how a fully coupled climate model attributes responses over the historical period to the different climate forcings.

While DAMIP provides a good framework for one of the main questions in CMIP6, namely how the Earth system responds to forcing, RFMIP, the Radiative Forcing Model Intercomparison Project, focuses on understanding the forcing itself. RFMIP contains a large set of experiments to further understand the radiative forcing of the past and the present (Pincus et al., 2016). We use two experiments from RFMIP, both with sea surface temperatures prescribed to pre-industrial values. One experiment includes both anthropogenic and natural aerosol emissions (*piClim-histall*), while the other only includes anthropogenic aerosol emissions (*piClim-histaer*). When sea surface temperatures are kept to pre-industrial values, the global surface temperature development stalls, and the simulation will keep to first order a pre-industrial climate. Sea surface temperatures changes would have an effect on cloud cover, which in turn can affect SDSR. These *piClim* experiments will show the direct atmospheric forcing on SDSR due to greenhouse gases and aerosols, alone or in combination, without including cloud cover changes induced by global warming.

The third MIP included in this study is the Aerosol Chemistry Model Intercomparison Project (AerChemMIP), which is designed to answer questions regarding the specific effect of aerosols and other near-term climate forcers (NTCF) on climate. NTCFs include methane, tropospheric ozone, aerosols and their precursors (Collins et al., 2017). Three experiments have been selected from AerChemMIP, *histSST*, with all forcing agents included, and two perturbations which have pre-industrial aerosol emissions: *hist-piAer* and *hist-piNTCF*. The hist-piNTCF experiment has in addition pre-industrial NTCF levels for ozone. A difference in these two simulations would only appear if ozone concentrations were computed in an interactive chemistry scheme. These two simulations are coupled and are comparable to the historical experiment. The experiment *histSST* uses all forcing agents and the sea surface temperatures derived from the historical simulation so that the temperature evolution, and hence its effect on SDSR, should be similar to the historical experiment but removes responses involving a coupled ocean. These experiments together with the historical experiment were chosen to differentiate between historical changes in aerosol and tropospheric ozone or whether a mixing layer in the ocean may have had an effect on dimming.

Data from all experiment ensembles from each of the MIPs listed above provide useful information on the role of anthropogenic aerosol emission in dimming and/or brightening.

Table 1. Model participation, as used in this study, in CMIP6 model intercomparison projects (MIPs) and their experiments.

Experiment	NorESM2	CanESM5	MIROC6	CESM2	CNRM-ESM2-1	GISS-E2-1-G	IPSL-CM6A-LR	MRI-ESM2-0
historical	x	x	x	x	x	x	x	x
hist-aer	x	x	x			x	x	x
hist-GHG	x	x	x	x		x	x	x
hist-nat	x	x		x		x	x	x
piClim-histaer	x	x	x			x		
piClim-histall	x	x	x			x	x	
hist-piAer	x		x					x
hist-piNTCF	x		x		x			x
histSST	x		x		x			x

2.3 Methods

The GEBA stations have been divided into regions based on the country and continent. The number of stations in a region is presented together with the first results in the caption of Fig. 2. The number of stations per region remains constant throughout the time period because of our gap-filling approach. A figure with the spatial distributions and trend of SDSR per station in GEBA used in this study is found in Fig. 1 in Storelvmo et al. (2018).

All model output and CRU results have been co-located to GEBA station locations using the nearest neighbor method. This entails that if two GEBA stations are within one grid box of a model, data from that grid box will be retrieved twice by nearest neighbor interpolation, as every station has been weighted equally. A global mean is defined here as the mean of a variable across all GEBA station locations. A regional mean is a mean of a variable across the GEBA station locations registered to that same region in the GEBA data. When a result is shown as an anomaly, as opposed to an absolute value, the general formula has been to subtract the baseline value, defined as the mean of the first 5 years of the investigated time period (1961–2014), from the time series in question. To clarify – first an average value per year per region is calculated, and then a new mean is created from the first 5 years of this time series. This 5-year mean is then subtracted from each year in the time series for the region in question and presented as an anomaly. We will often present data as 6-year averages, as yearly variabilities are not the focus of this study. These 6-year averages are simply made by dividing the time series over 54 years (1961–2014) into nine equal intervals and averaging these intervals together. When the atmospheric burdens of SO_4 are shown together with observed SDSR from GEBA, the time series have been smoothed using a 10-year running mean, and this is the only data in the paper shown using this smoothing technique.

The "baseline" values for global SDSR and cloud cover in the models and observations of this study can be found in the Appendix in Table A1.

The model data have been retrieved from the Earth System Grid Federation (ESGF) (Cinquini et al., 2014). ESGF is a data management system consisting of multiple geographically distributed nodes that coordinate through a peer-to-peer (P2P) protocol (Fan et al., 2015). We have used three ensemble members for the *historical* experiment to present internal variability in the models and one ensemble member from the rest of the experiments shown, as not every experiment had requested more than one simulation. Table A2 in the Appendix shows the resolutions, aerosol schemes, and aerosol complexity of the models in this study, and Sect. A2 explains the variables and variant labels downloaded.

3 Results

3.1 Model variability

Figure 1 shows the SDSR anomaly for each model of the study co-located to all GEBA stations, 1487 in total as compared to the observed SDSR anomaly. The aerosol effective radiative forcing (aerosol ERF) corresponding to each model is obtained from Smith et al. (2020) and is listed in each panel to illustrate the strength of the aerosol radiative effect in the model.

Each climate model has its own internal variability and thereby represents its separate climate systems. SDSR is a highly variable metric on a year-to-year basis, which can be seen both in the GEBA data in black in Fig. 1 and in following a single ensemble member per model. Within each model ensemble one can see that no member is equal to another, which is a clear signal of the internal variability of each model. The spread of all three ensemble members in a 6-year period can be read from the height (interquartile range) of the boxes in the 6-year intervals; note that this spread is dominated by large inter-annual variabilities within each member. One example is GISS-E2-1-G, where each ensemble member has large interannual variabilities: the boxes present long whiskers and large interquartile ranges, but when comparing the ensemble member 6-year means one by one they mostly

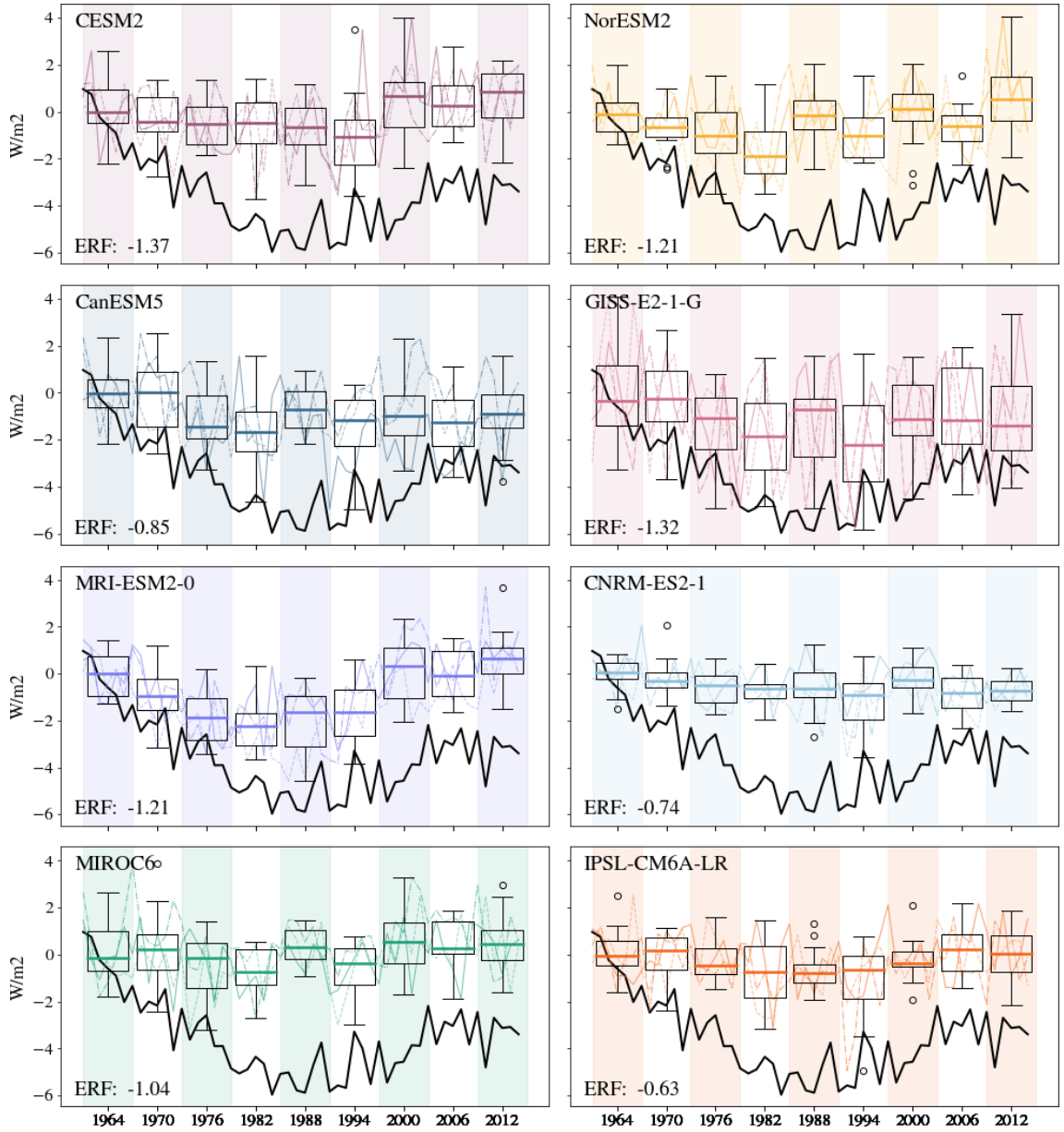

Figure 1. Global surface downwelling shortwave radiation (SDSR) anomaly at the surface for GEBA (black) and three ensemble members for the historical simulation of the eight models in this study. The boxes are made for 6-year intervals (shaded in background) based on 6-yearly means and three ensemble members per model. Colored lines behind boxes show yearly values of SDSR anomaly per ensemble member. The height of each box represents the interquartile range of the data, and the thick colored line within each box is the median. The whiskers show the minimum and maximum values of the selection of data, and the outliers are shown as a hollow dot. Results are co-located to all GEBA stations (1487) throughout the time period. The aerosol ERF as found in Smith et al. (2020) per model is shown in the bottom left of each panel.

agree on their magnitudes of SDSR anomaly, so the intra-ensemble spread is not large for GISS-E2-1-G. We find (not shown here) that the model with the least interannual variabilities is CNRM-ES2-1, while the model with the largest inter-ensemble disagreements is CanESM5.

Figure 1 also shows that the models in general do not agree with the observed global SDSR anomaly shown in black. Dimming and brightening are tendencies in surface radiation that are observed on longer than interannual timescales; with this in mind, SDSR from models will in general be presented as 6-year means for the remainder of this paper. The model MRI-ESM2-0 shows the most similar SDSR evolution compared to the observed data according to Fig. 1.

The model with the strongest aerosol ERF is CESM2, while the weakest aerosol ERF is presented by IPSL-CM6A-LR.

3.2 Dimming and brightening

The change in SDSR in the *historical* simulations from the eight models is presented together with GEBA data in Fig. 2. Panel (a) of this figure corresponds to the results shown in Fig. 1. Each model graph in Fig. 2 represents the ensemble mean of the model in question averaged over 6 years, based on three ensemble members. GEBA data are shown in black, also as 6-year averages, but with the yearly time series shown in grey in the background. Model simulations show small changes in global SDSR compared to observations (Fig. 2a). Global SDSR is observed to decrease over the 1487 stations until the late 1980s before increasing again, clearly showing the global "dimming" and "brightening" as found in previous studies listed in the introduction.

None of the models outperform one another globally, and there is a discrepancy of about 2–3 $\mathrm{W\,m^{-2}}$ between models and observations. To further identify from where this discrepancy originates, we consider some geographical regions separately. Asia and Europe are relevant regions in regards to anthropogenic aerosol emissions (as explained in Sect. 1) and thereby also relevant to global dimming and brightening. The historical SDSR evolutions in Europe and Asia are presented in Fig. 2b and c, respectively. European SDSR is relatively well represented by the model simulations. The yearly GEBA time series has values within the shaded area that shows the standard deviation of the total of 24 model ensemble values in almost every 6-year period in Europe. The dimming in Europe is believed to have started before 1961 (Wild, 2009), which partly explains why the initial European dimming in Fig. 2b is weak. GEBA shows a short-term positive anomaly between 1970 and 1980, which is not caught by the models. This peak is currently unexplained, but a short assessment of its possible association with changes in cloud cover is found in Sect. A1 in the Appendix.

There is generally a large discrepancy between model simulations and observations of SDSR in Asia, as seen in Fig. 2c. The ground stations in Asia show a noticeable trend change in SDSR in the transition from the 1980s to 1990s that is not apparent in the model simulations. The historical model simulations show a consistent negative trend during the entire historical period in question in Asia. Historically, countries with relatively high emissions in Asia include India, Japan, and China (Hoesly et al., 2018), and the SDSR evolution for each of these countries is shown in Fig. 2d, e, and f, respectively.

Figure 2d shows that the models capture a relatively strong negative trend of SDSR in India, with MIROC6 being the model with the most modest trend. There are evident differences between observations and simulations in both Japan and China. Ground stations in Japan show a sharp decrease in SDSR until the early 1970s followed by some variations until a new minimum value is reached around 1990 before an increase in SDSR is measured. The minimum value around 1990 and the following positive trend is similar to that of China. Japan is downwind of the Asian continent and thus believed to be influenced by aerosol emissions from China. Model simulations do not capture the magnitude of dimming in Japan or the apparent brightening in the 1990s. The timing of minimum SDSR occurs differently in models, which was also seen in Fig. 2a.

Observations from China (Fig. 2f) show a trend change in SDSR similar to the one identified in Fig. 2c for Asia as a whole, with the minimum value found in 1989. We note that China consists of 119 GEBA stations, while Asia as a whole consists of 311 stations; thus, the Asian average is largely impacted by SDSR as measured in Chinese stations. In general the historical model simulations show dimming throughout the historical period in China, meaning none of them shows a similar trend change to the one from the observational dataset. This post-1990 trend change is a source of discussion within the field, and a thorough assessment, relevant to the conclusions from this study, is found in Sect. 4.1.

3.3 Dimming and brightening over China in various CMIP6 experiments

In order to understand which forcing agents are responsible for the overall trends in SDSR in the models, we now investigate China for the experiments listed in Table 1. Figure 3a shows perturbed historical simulations as performed in DAMIP together with observations of SDSR. DAMIP has two experiments without historical anthropogenic aerosol emission (dashed/*hist-nat* and stippled/*hist-GHG* lines) and one experiment with historical anthropogenic aerosol emissions (solid lines/*hist-aer*). The experiment *hist-aer* is the only experiment in DAMIP exhibiting a distinguishable dimming signal. SDSR from *hist-aer* shows patterns similar to the *historical* simulations with continuous dimming throughout the period, unlike the observed SDSR. SDSR in the experiments *hist-nat* and *hist-GHG* do not show signs of dimming or brightening over the investigated period in China, which confirms that water vapor or stratospheric aerosols are not the dominant cause of multidecadal dimming signals in the fully coupled historical model simulations. This is supported by previous work, as mentioned in the introduction.

Out of the three experiments from AerChemMIP only, *histSST*, has prescribed sea surface temperatures and contains changes in anthropogenic aerosol emissions. This is consistent with the time evolution of SDSR in *histSST* as the simulations diverge from the other simulations as time progresses (Fig. 3b). Keeping in mind that *histSST* also has anthropogenic greenhouse gas (GHG) emissions in addition to natural forcers, the only difference from *histSST* to the *historical* experiment is the absence of a coupled ocean and the use of prescribed sea surface temperatures. The model MRI-

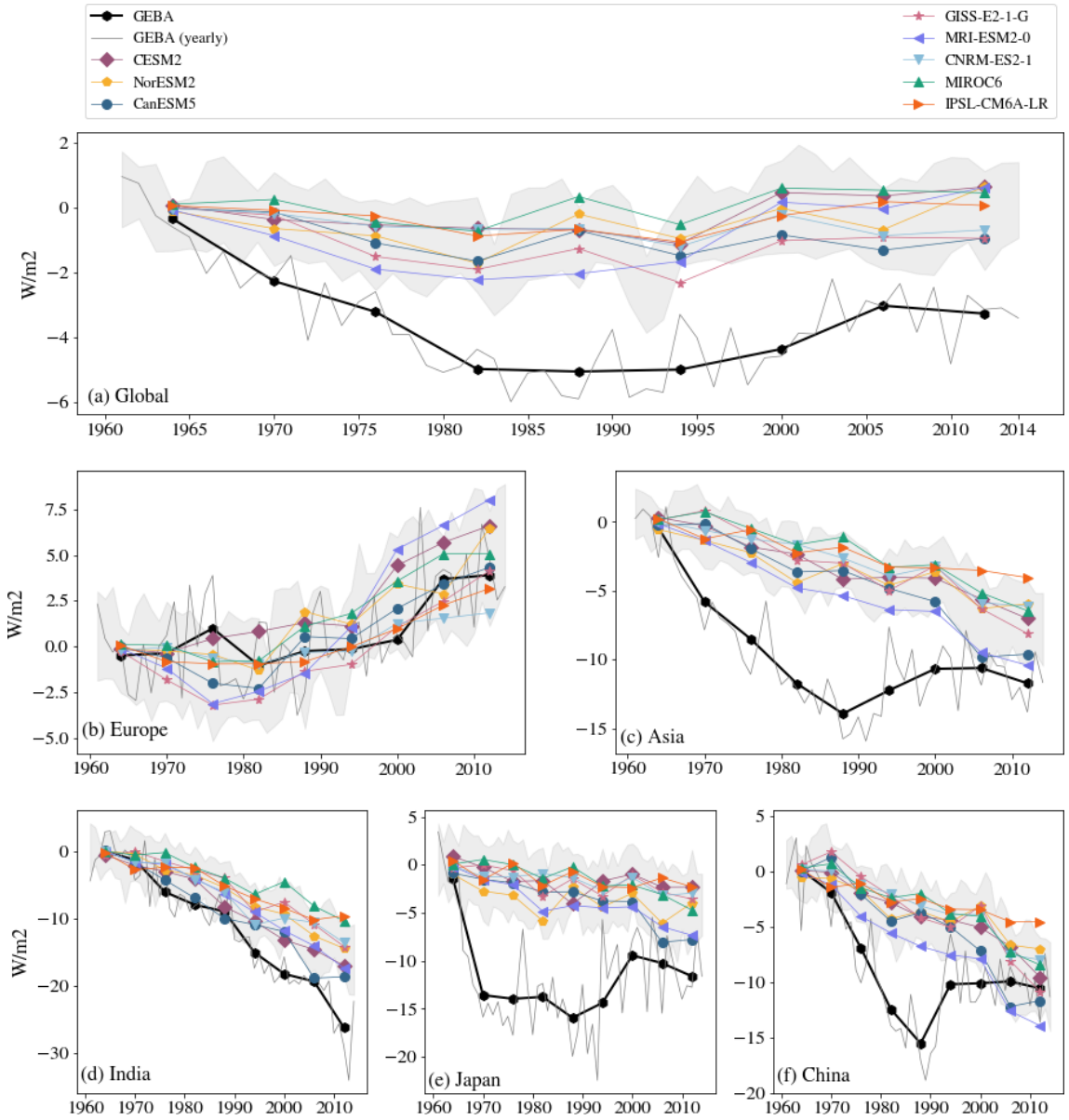

Figure 2. Six-year averages of the SDSR anomaly at the surface for GEBA and eight Earth system models. Results are co-located at **(a)** all GEBA stations (1487), **(b)** European (503), **(c)** Asian (311), **(d)** Indian (15), **(e)** Japanese (100), and **(f)** Chinese (119) stations. Numbers in parentheses are the number of ground stations in the respective region. The entire 54-year period has been divided into intervals of 6 years and then averaged together to make nine data points as shown by the markers. The grey shading represents 1 standard deviation from the yearly total ensemble mean.

ESM2-0 presents the strongest dimming in both DAMIP and AerChemMIP. The simulations with pre-industrial aerosols (*hist-piAer*) and pre-industial near-term climate forcers, including aerosols and ozone (*hist-piNTCF*), show very small or negligible changes in the SDSR over the time period considered.

Recall that the experiments of RFMIP utilize pre-industrial SSTs, meaning essentially there is no global warming in these experiments. In the RFMIP experiments shown in Fig. 3c both *piClim-histaer* and *piClim-histall* contain anthropogenic aerosol emissions, and all simulations show a continuous dimming throughout the period. There is no clear distinction between experiments containing GHG emissions in addition to anthropogenic aerosol emissions (solid lines/*piClim-histall*) and the experiments only containing anthropogenic aerosol emissions (stippled lines/*piClim-histaer*). This implies that greenhouse gases without their global warming effect do not affect multidecadal all-sky SDSR in a significant way over China throughout the period,

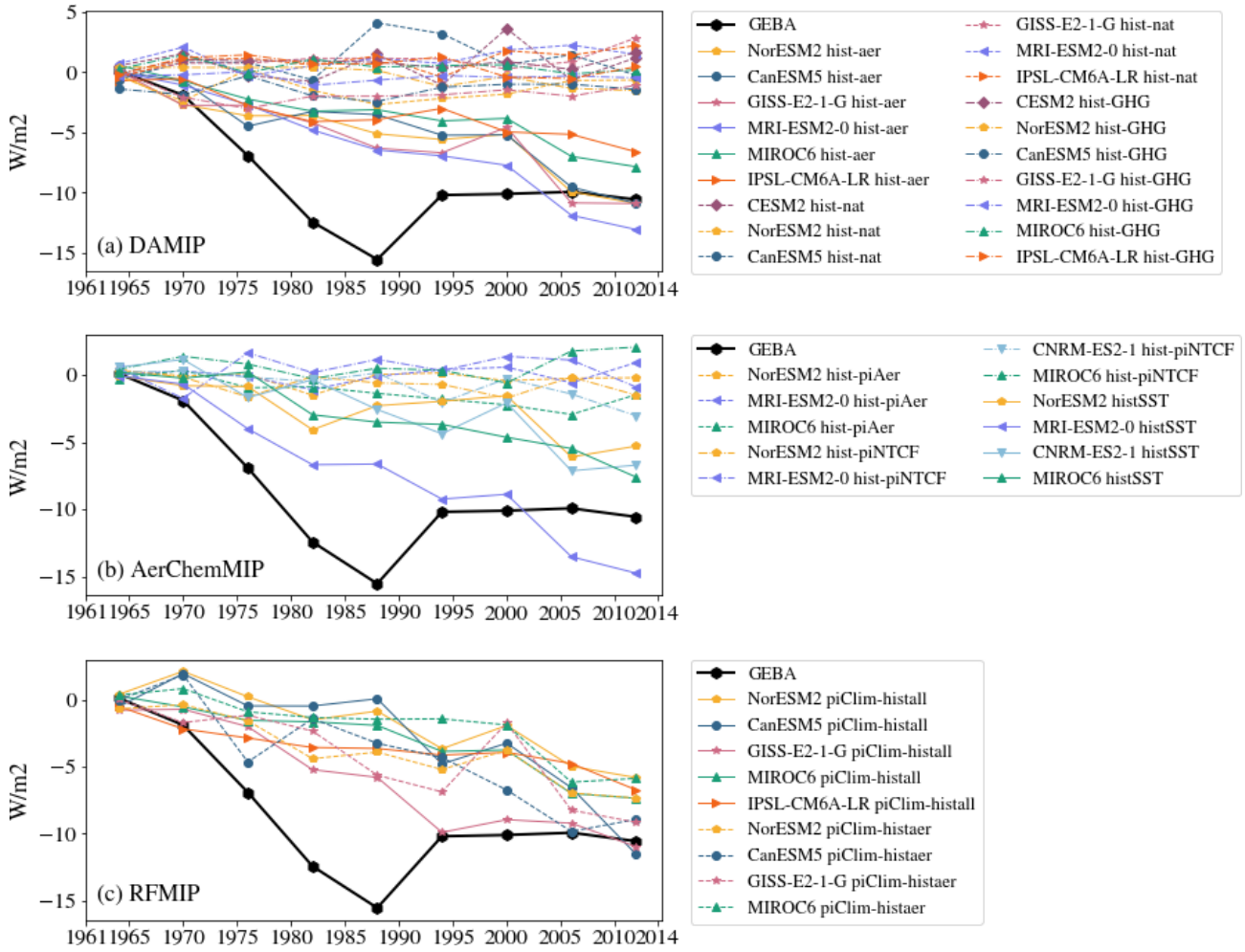

Figure 3. SDSR anomaly in China for all the CMIP6 simulations as listed in Table 1. All model results are co-located at GEBA station locations registered to China (119 stations). The entire 54-year period has been divided into intervals of 6 years and then averaged together to make nine data points as shown by the markers.

again supported by previous work mentioned in the introduction.

Overall there is a clear difference in SDSR between experiments that include anthropogenic aerosol emissions and experiments that do not. Dimming is apparent in every simulation containing anthropogenic aerosol emissions but absent in the simulations using aerosols maintained at constant pre-industrial levels. This points to anthropogenic aerosol emissions playing a key role in dimming. Whether the sea surface temperature is pre-industrial, prescribed historical, or decided by a coupled ocean model seems to be unimportant for the SDSR temporal evolution in China in most models.

No distinct flattening or brightening is identified in any of the simulations in which dimming is identified, and therefore none of the model simulations shows a temporal evolution of SDSR close to the one seen in observations over China.

All-sky SDSR changes can be further decomposed by the models into a clear-sky contribution as well as a contribution from changes in cloud cover or other cloud properties. In the next section we present the decomposed contributions to all-sky SDSR in China to further understand the discrepancy seen in Fig. 3.

3.4 Clear-sky SDSR and cloud cover in China

So far we have only evaluated all-sky SDSR, which is influenced by clouds and any aerosol radiative effects. Table 2 shows changes in cloud cover, all-sky SDSR, and clear-sky SDSR within three different time periods for the models and observational datasets of this study. Between the years 1961 and 1989 GEBA shows a strong negative change in all-sky SDSR in Fig. 2f. In Table 2 we thus show changes in this time period by making two 3-year means and subtracting them from one another. This is done to avoid extreme values as we are working with metrics exhibiting large year-to-year variations. This has been done for two additional time periods

which have been chosen based on the temporal development in the all-sky SDSR as measured by GEBA in China (see Fig. 2f), summarized in the second lowest row in Table 2.

In the first time period the models do not agree on the sign of cloud cover change, and the simulated all-sky SDSR is weaker than the observed one, which was already established in the previous section. Clear-sky SDSR does not differ largely from all-sky SDSR within the models. For some models the negative change in clear-sky SDSR is stronger than in all-sky SDSR, meaning that the aerosol direct effect may contribute significantly to dimming for these models. The aerosol indirect effect changes the radiative properties of clouds in two ways: by making them appear brighter and by altering their lifetime (Boucher et al., 2013). Therefore, a weak change in cloud cover followed by a strong change in all-sky and clear-sky SDSR points to both the direct and brightening indirect aerosol effects being the primary cause of SDSR change, as an altered lifetime of clouds would imply cloud cover changes.

In the second time period GEBA shows a positive change (which will be further discussed in Sect. 4.1), and CRU shows a cloud cover change of $+3.0\,\%$. Intuitively, an increase in cloud cover would not create a brightening at surface level. The observations are thus not consistent in this time period if only cloud cover effects were important. The models disagree in their sign of cloud cover changes and all-sky and clear-sky SDSR. In the final time period where GEBA shows a weak slightly positive change in all-sky SDSR, every model in this study shows a dimming. All models apart from MIROC6 show simulated clear-sky SDSR changes that are stronger than the changes found in all-sky SDSR. Together with the inconsistent simulated cloud cover and all-sky SDSR changes for this time period, we suggest that both direct and indirect aerosol effects are responsible for the changes in SDSR found in the model simulations.

All models show dimming in clear-sky and all-sky SDSR in the first and last time periods. Some models show a weak positive change in all-sky SDSR in the same period as GEBA presents a strong brightening. Both observed and simulated changes in cloud cover neither act as a brightening mask for clear-sky dimming nor are convincingly a cause of dimming/brightening in either observed or simulated all-sky SDSR. A rough calculation of the effect of $1\,\%$ cloud cover increase on SDSR in China is found in Sect. A3 in the Appendix, indicating that such an increase could result in a dimming of 1–$3.5\,\mathrm{W\,m^{-2}}$. As such it shows that observed and modeled changes in cloud cover, as reported in Table 2, can lead to important contributions to the dimming and brightening signals in SDSR. However, this calculation is idealized, does not isolate the cloud cover change effect in the model results and does not explain the inconclusive data reported in Table 2. It is important to note that the robustness of observed cloud cover changes must be verified by satellite observations, which goes beyond the scope of this study.

In Sect. 3.3 we showed that dimming was only apparent in simulations that included anthropogenic aerosol emissions. In this section we found the clear-sky SDSR to be close in value to or even stronger than all-sky SDSR, indicating the simulated dimming is primarily caused by aerosol effects. Table 2 underlines previous findings: dimming in models is overall weaker than in observations. The next section will then show how the simulated aerosol burdens are connected to SDSR.

3.5 Atmospheric burden of SO_4

In the atmosphere, the *presence* of a reflective aerosol is the cause of scattered shortwave radiation, and the emission of its precursor is only an indirect indicator of its presence. All CMIP6 simulations mentioned above have utilized the same anthropogenic sulfur dioxide gas emissions; however, the resultant dimming differed considerably. SO_4 aerosol burdens should be more closely linked to the radiative effect. Therefore, we present here also the simulated anomalies in burden of SO_4 in the various models over Europe, a location where dimming and brightening are fairly well represented in simulations, and over China, where dimming and brightening are poorly represented in simulations (Fig. 4a and b, respectively). The sulfate burdens are co-located to GEBA station locations in the respective regions. As expected, sulfate aerosols have an important role in European dimming and brightening, as the simulated burdens of SO_4 show a strikingly similar pattern (but with opposite sign) to the observed SDSR over Europe for all the models. The maximum burdens are found in the early to mid 1980s depending on the model, and the minimum SDSR around the same time. The various models differ in the magnitude of change in SO_4 burden over Europe, but all show similar tendencies. MRI-ESM2-0 is the model with the largest changes, and GISS-E2-1-G is the model with the smallest changes in SO_4 burden. The same is observed over China, where MRI-ESM2-0 has an SO_4 burden at the end of the time period which is more than double the burden of the other models (except NorESM2). In contrast to Europe, the observed SDSR evolution does not mirror well the simulated SO_4 burden time series over the GEBA stations in China. In order for the SO_4 burden to be the main cause of the observed changes in SDSR, the Asian SO_4 burden would have to peak around the late 1980s, which is not seen in the models in Fig. 4b. All the simulated historical SO_4 burdens increase until 2010, showing no signs of either a trend change or a flattening of aerosol-induced dimming. Assuming GEBA data provide a reasonable representation – within uncertainty bounds as discussed in Sect. 4.1 – of the historical development of SDSR and implicitly sulfur burdens in China, the problem in SO_4 burden must come from either the emissions, aerosol formation, transport or the removal processes of SO_4.

It appears, however, that the simulated burdens of SO_4 co-located to GEBA stations in China follow quite closely

Table 2. Changes in Chinese cloud cover (%), all-sky SDSR AS (W m^{-2}), and clear-sky SDSR CS (W m^{-2}) between two 3-year means for three time periods. All model results are means made from three ensemble members of the historical simulation, co-located and extracted at Chinese GEBA station locations. Changes in cloud cover are from CRU gridded data and represent means co-located to Chinese GEBA stations.

	[1961–1963]–[1987–1989]			[1990–1992]–[1997–1999]			[2000–2002]–[2012–2014]		
Data	(%)	AS (W m^{-2})	CS (W m^{-2})	(%)	AS (W m^{-2})	CS (W m^{-2})	(%)	AS (W m^{-2})	CS (W m^{-2})
NorESM2	−1.0	−4.6	−4.0	0.6	−1.0	−0.4	0.3	−3.9	−5.0
CanESM5	−0.4	−3.5	−4.6	−0.1	0.8	0.6	−1.7	−2.4	−5.7
MIROC6	0.4	−4.4	−3.6	1.2	−1.3	−0.4	0.5	−5.5	−5.3
CESM2	−1.0	−2.6	−3.6	−0.2	0.0	−0.2	0.0	−5.3	−6.7
CNRM-ESM2-1	−0.4	−3.3	−5.2	−0.6	1.1	−1.0	−0.9	−3.5	-6.5
GISS-E2-1-G	1.3	−3.7	−6.4	−0.2	−0.7	−1.2	2.5	−8.7	−9.9
IPSL-CM6A-LR	−1.2	−3.3	−5.0	0.5	−0.6	−0.1	−1.6	−0.4	-1.9
MRI-ESM2-0	−0.1	−7.1	−6.9	−0.3	−0.8	−0.9	−1.1	−4.9	-8.8
MODELMEAN	−0.3	−4.1	−4.9	0.1	−0.3	−0.4	−0.2	−4.3	−6.2
GEBA		−15.4			6.6			0.9	
CRU	0.1			3.0			−1.0		

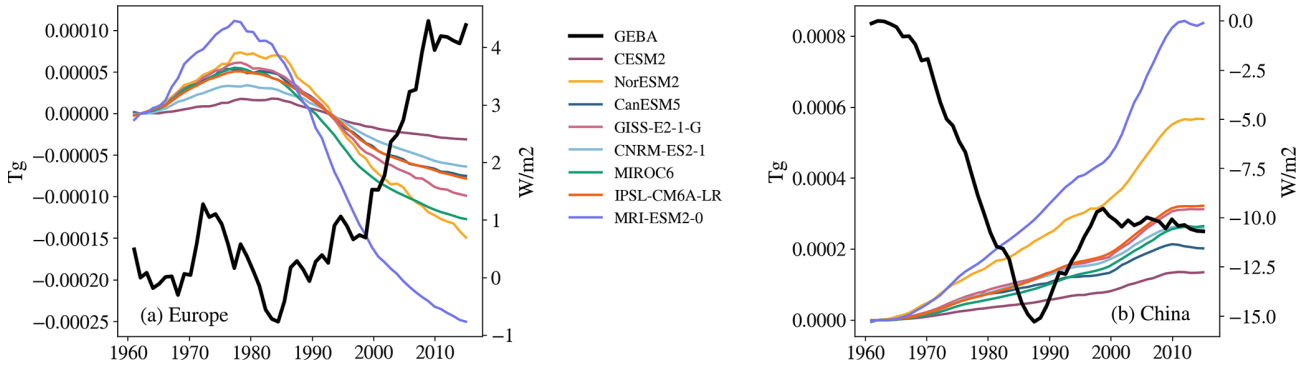

Figure 4. Anomaly of simulated atmospheric load of sulfate per model together with observed all-sky SDSR anomaly in **(a)** Europe and **(b)** China. The GEBA data are shown as yearly anomalies, while the atmospheric loads have been smoothed using a 10-year running mean technique as explained in Sect. 2.3.

the time series of emitted SO_2 in the climate models over China (shown in Appendix Fig. A2), which indicates that SO_4 formation and export of sulfur from the Chinese region remains rather similar in the period investigated. Following the logic that emission correlates with burden, which again anti-correlates with SDSR changes, the temporal development of SDSR seen in GEBA cannot be explained from the current emission inventories, given that sulfate aerosols play an important role in SDSR in China.

4 Discussion

The climate effect of aerosol emissions over the industrial era is poorly constrained, in part due to lack of observations and uncertainty in emissions. The uncertainty in past aerosol climate effects is an important reason for the large spread in climate projections for the future. Here, we investigate the effect of aerosols in GEBA which provides valuable observations of historical shortwave radiation at the surface.

We have shown that a subset of models participating in CMIP6 does not accurately represent the observed dimming and brightening trends globally and regionally in their *historical* simulation. This is comparable to that of Storelvmo et al. (2018) and Wild and Schmucki (2011), who showed that the CMIP5 and CMIP3 ensemble mean SDSR globally co-located to GEBA stations does not represent dimming or brightening. Our findings show that reproducibility of SDSR has not improved from CMIP5 to CMIP6. We find that most models show an underestimation of changes in SDSR compared to observations, and the development over time greatly differs between model and observations, especially in China. This is in agreement with Allen et al. (2013), who studied the CMIP5 ensemble mean and found a continuous dimming trend over China but with a severely underestimated magni-

tude of modeled clear-sky SDSR during the dimming period compared to a clear-sky proxy based on GEBA data.

The simulated SDSR on decadal timescales over China does not differ significantly when comparing the RFMIP experiments (Fig. 3) to the historical experiment. RFMIP experiments have pre-industrial sea surface temperatures and thus do not include global-warming-induced cloud cover changes. When experiments with historical cloud cover changes show dimming in the same magnitude as experiments without historical cloud cover changes, the dimming can be assumed to be dominated by aerosol effects in China. This complements the findings by Folini and Wild (2015) where sea surface temperatures correlate with cloud cover, not aerosol effects. Table 2 showed inconclusive connections between modeled and observed cloud cover but clear connections between clear-sky SDSR and all-sky SDSR, again pointing to aerosol effects dominating SDSR time evolution in China.

The climate models strongly underestimate the dimming observed in China in addition to not representing the post-1990 trend change. This trend change is the topic of discussion in the next section.

4.1 The post-1990 trend change in China

Several studies have tried to explain the trend change as presented here by GEBA in China in the transition from the 1980s to the 1990s. Streets et al. (2006) proposed a peak in combined emissions of SO_2 and black carbon in 1988–1989 as a possible explanation. A later study questions the quality of the observational data showing the trend change (Tang et al., 2011), while recent studies propose the post-1990 initial, strong brightening is to a considerable extent an artifact of a nationwide change in SDSR measurements (Wang and Wild, 2016; Yang et al., 2018). The change in SDSR measurements includes a replacement of SDSR instrumentation, an increase in measurement frequency and in addition an update in the classification of SDSR stations, and Wang and Wild (2016) conclude that the upward trend ("jump" between 1990 and 1999) should be considerably weaker and that only 20 % of the "jump" has actual physical causes. Yang et al. (2018) homogenized the data from Wang and Wild (2016) and Wang et al. (2012) and presented a new SDSR evolution (results can be seen in Yang et al., 2018, Fig. 10). The newly homogenized data exhibit a significant dimming trend ($-6.13 \pm 0.47\,\mathrm{W\,m^{-2}}$ per decade) between 1958 and 1990, a flattening of the curve in 1991–2005, followed by a brightening trend ($6.13 \pm 1.77\,\mathrm{W\,m^{-2}}$ per decade) between the years 2005 and 2016. We can use Fig. 2f to compare our model data to these homogenized data and see that even without a larger "jump" in the data around 1990 there are still large discrepancies between model and observation, both in the form and magnitude of the brightening period after 1990. All the models show dimming in the flattening period of the newly homogenized data. All the models apart from CanESM5 show an averaged negative trend between the 6-year means of 2003–2008 and 2009–2014, where the newly homogenized data show a brightening. A similar "jump" to the one seen in China can be identified slightly later in Japan (Fig. 2e). To our knowledge, we have no information on either a replacement of instruments or an update in the classification of SDSR stations in Japan. Norris and Wild (2009) investigated the role of clouds for historical SDSR observations in China and Japan and found the post-1990 brightening in Japan to be statistically significant, while the Chinese brightening was found to be insignificant. In this paper (published before Wang and Wild, 2016) half of the post-1990 brightening in China and one-third for Japan were attributed to a reduction in cloud cover. These results point to a need for more studies assessing and evaluating available observational SDSR data. However, models do not accurately represent the strength of dimming throughout the whole period or the change in trend after 1990 and thus the time evolution of SDSR observed in China, with or without the early 1990s "jump" in brightening.

4.2 Aerosol effect on dimming

Out of all the experiments presented in Table 1 and Fig. 3, only those containing anthropogenic aerosol emissions showed dimming. This is expected as aerosols have been presented as the main cause of reduction in SDSR in China by previous studies (Wild, 2009; Yunfeng et al., 2001; Kaiser and Qian, 2002).

Storelvmo et al. (2018) argue that the discrepancy seen between observed and modeled CMIP5 model mean global SDSR can be attributed to errors in the treatment of processes that translate aerosol emissions into clear-sky and all-sky radiative forcings. Here, we can see an anti-correlation between simulated SO_4 burdens from Fig. 4a and b and simulated SDSR from Fig. 2b and f, respectively. Therefore we suggest that the simulated SDSR is dominantly a result of simulated SO_4 burdens. Simulated SDSR agrees relatively well with observed SDSR in Europe (Fig. 2b) along with simulated SO_4 burden anti-correlating relatively well with observed SDSR in Europe (Fig. 4a). This means that the model code translating burdens into SDSR in Europe can simulate changes in SDSR as a consequence of changes in aerosol emissions. If models translate burden into SDSR correctly in Europe, this does not necessarily mean that they translate burden into SDSR correctly in other regions. However, we suggest that the code translating burdens into SDSR should also work correctly in China, since also in China we find that aerosols are the main cause of dimming, in agreement with Wild (2009), Yunfeng et al. (2001), and Kaiser and Qian (2002). Note also that we find no consistency between observed cloud cover changes, GEBA data and simulated cloud cover and SDSR anomalies in China (Table 2). By suggesting the translation process from burden to SDSR is behaving correctly in both regions, the potential source of

error causing discrepancies between observed and simulated SDSR can be traced to the causes of the simulated atmospheric burdens in the first place.

The sulfur dioxide emission inventory used as input for historical model simulations in CMIP6 is shown in Hoesly et al. (2018) (Fig. 3), and the emissions as translated in four of the models of this study is shown in Fig. A2.

Hoesly et al. (2018) have pointed to the need for emission uncertainties, but this has not been done for these emissions. Aas et al. (2019) have studied global and regional trends in atmospheric sulfur and found that uncertainties in emissions were largest in Asia, even if their study only went back to 1990.

Previous studies estimating SO_2 emissions include Lu et al. (2010), who found that sulfur dioxide emissions in China increased by 53 % between 2000 and 2006 using technology-based methodology and thereby found similar results to that of Hoesly et al. (2018). Lu et al. (2010) also compared AOD-derived SDSR to GEBA-based SDSR data as shown in Streets et al. (2006) and found the GEBA-based SDSR data to not accurately represent SDSR development in East Asia; this further underlines the need for more studies evaluating SDSR observations. Other studies such as Koukouli et al. (2018) have used satellite observations to estimate a new emission inventory for SO_2 between 2005 and 2015 in China. We note that the year 2005 in Fig. A2 is directly after the sharp increase in SO_2 emissions, and the biggest differences between the estimation made by Koukouli et al. (2018) and the SO_2 emission inventories in CMIP6 are a decrease in emissions after the year of 2011. This decrease in SO_2 emissions would intuitively result in a brightening, which is identified over the same time period in the homogenized data by Yang et al. (2018) (Fig. 10 therein).

The modeled emissions of SO_2 as shown in Fig. A2 over China showed no trace of a significant change in trend after 1990 in our observed SDSR time series as discussed in the previous section. Assuming sulfate burden is responsible for the observed multiyear trends of SDSR, we argue that errors in emissions inventories in China could be part of the problem.

5 Conclusions

Earlier studies have shown that previous generations of Earth system models have not been able to reproduce the transient development of surface downwelling shortwave radiation (SDSR) in the last decades since 1960 when observations became available (Storelvmo et al., 2018; Allen et al., 2013). This discrepancy is hypothesized to be related to increasing and then partially decreasing trends in global aerosol emissions and subsequent aerosol radiative effects, but the exact cause is unknown.

In this paper, we compared observations to model-simulated surface downwelling shortwave radiation and cloud cover in specific regions for the time period 1961 to 2014. We found that in the *historical* experiments, CMIP6 models reproduce the transient development of SDSR well in Europe but poorly in Asia. The multiple historical and associated perturbation experiments performed under CMIP6 reveal that only those simulations containing anthropogenic aerosol emissions show dimming, and the dimming is underestimated by most models. China exhibits a sharp positive trend in observed SDSR in the 1990s that is not found in historical model simulations. This "jump" has been suggested to be an artifact, but historical simulations also do not accurately represent the homogenized observed SDSR as proposed by Yang et al. (2018). We suggest that the continuous decrease in simulated SDSR is related to the continuous increase in atmospheric sulfate burden in the *historical* simulations over China. Following this logic, the observed transient development of SDSR points to the evolution of the sulfate burden in the models being wrong in this region. The sulfate burden is a result of sulfur dioxide emissions, gas-to-particle conversion and wet deposition. Sulfur dioxide emissions over China show neither sign of the observed trend change from gap-filled GEBA data nor of the brightening-followed flattening from the homogenized data as proposed by Yang et al. (2018). Sulfur dioxide emissions used in the models over China have a strong increase in the early 2000s, which can be observed as a sharp dimming at the same time in Fig. 2f. We suggest that the cause of the discrepancy between model and observations in transient SDSR in China is partly in erroneous emission inventories.

As the observed climate change is the result of warming from greenhouse gases and simultaneous cooling from aerosol radiative effects, getting aerosol emissions correct is important in Earth system models.

Since the SDSR measurements are not only sensitive to aerosol effects, they might not be the most accurate way to infer historic aerosol loads and forcing. Further studies could include other observations and proxies for aerosol effects in the historical era, such as long-term satellite-retrieved aerosol optical depth, deposition of anthropogenic sulfur, organic carbon and nitrate in ice cores, as well as daily temperature range records.

Appendix A: Additional information

A1 The European SDSR evolution

Figure A1 suggests cloud cover variation as a possible explanation of the local maximum in observed European SDSR during the period 1967–1978. Cloud cover exhibited a substantial minimum simultaneous with the maximum in SDSR. The peak is not reproduced in the historical runs of Earth system models studied herein (see Fig. 2b). Cloud cover variations that are not externally forced but are rather a result of internal variability cannot be expected to be reproduced in fully coupled Earth system models. This might serve as an explanation why the substantial peak in SDSR between 1967 and 1978 is not reproduced in the Earth system models.

A2 The data downloaded from ESGF

Table A2 shows an overview of the eight models used in this study. For the historical simulations three ensemble members per model were downloaded, with the variant labels r[1,2,3]i1p1f[1,2] for the variables rsds, rsdscs and clt. In addition, the variables loadso4 and areacella were downloaded for one ensemble member per model in the historical simulation per model. In the remaining experiments listed in Table 1 only one ensemble member per model was downloaded for the variable rsds; this was done as not every model provides more than one simulation per experiment.

A3 Effects of cloud cover change on all-sky SDSR

If we assume that $E_{\text{clear sky}}$ is the diurnal average clear-sky SDSR in a region and that τ_{cloud} is the average cloud optical depth, we can compute idealized effects of cloud changes on SDSR using the Beer–Lambert law:

$$E_{\text{surf}} = E_{\text{toa}} \exp(-\tau / \cos\phi), \tag{A1}$$

where τ and ϕ denote optical depth and solar zenith angle, respectively. The change in SSR per 1 % change in cloud cover can then be computed:

$$\begin{aligned}\Delta E_{\text{surf per 1 \%}} = 0.01 \times E_{\text{cloudy}} - E_{\text{clear sky}} = 0.01 \\ \times E_{\text{toa}} \exp(-\tau_{\text{cloud}} / \cos\phi \\ + \ln(E_{\text{clear sky}} / E_{\text{toa}})) + 0.99 \times E_{\text{clear sky}} \\ - E_{\text{clear sky}}.\end{aligned} \tag{A2}$$

Idealized computation for China. Assuming that ϕ is between 30 and 70°, that $E_{\text{clear sky}}$ is between 100 and 350 W m^{-2} and that $E_{\text{toa}} = 1362$ W m^{-2} in China, the theoretical effect of 1 % increase in cloud cover on all-sky SDSR is between -1 and -3.5 W m^{-2}, using the idealized computation described above.

Table A1. Global all-sky SDSR and cloud cover averaged over the years 1961–1966 (baseline values) as observed (GEBA for radiation, CRU for cloud cover) and as simulated in the ensemble mean of three ensemble members in the historical experiment by each of the models of this study. Data from both CRU and models are retrieved after co-location to all GEBA sites.

Model	SDSR (W m^{-2})	Cloud cover (%)
CESM2	186.3	63.9
NorESM2	186.8	55.6
CanESM5	189.5	56.2
GISS-E2-1-G	176.6	58.6
MRI-ESM2-0	193.8	56.2
CNRM-ES2-1	192.3	57.2
MIROC6	184.3	50.4
IPSL-CM6A-LR	185.7	54.5
CRU		**57.4**
GEBA	**171.6**	

Table A2. Details on the models used. IA: interactive aerosols. NIA: non-interactive aerosols.

Institution	Model	Resolution	Aerosol module	Complexity	Reference
NCAR	CESM2	1.25×0.9	MAM4	IA	Danabasoglu et al. (2020)
CCCma	CanESM5	2.81×2.81	CanAM4	IA	Swart et al. (2019)
CNRM-CERFACS	CNRM-ESM2-1	1.4×1.4	TACTIC_v2	IA	Séférian et al. (2019)
IPSL	IPSL-CM6A-LR	2.5×1.27	INCA fields	NIA	Boucher et al. (2020)
NCC	NorESM2-LM	2.5×1.875	OsloAero6	IA	Seland et al. (2020)
MRI	MRI-ESM2-0	1.125×1.125	MASINGAR mk-2r4c	IA	Yukimoto et al. (2019)
MIROC	MIROC6	1.4×1.4	SPRINTARS	IA	Tatebe et al. (2019)
NASA-GISS	GISS-E2-1-G	2.5×2.0	OMA fields	NIA	Kelley et al. (2020)

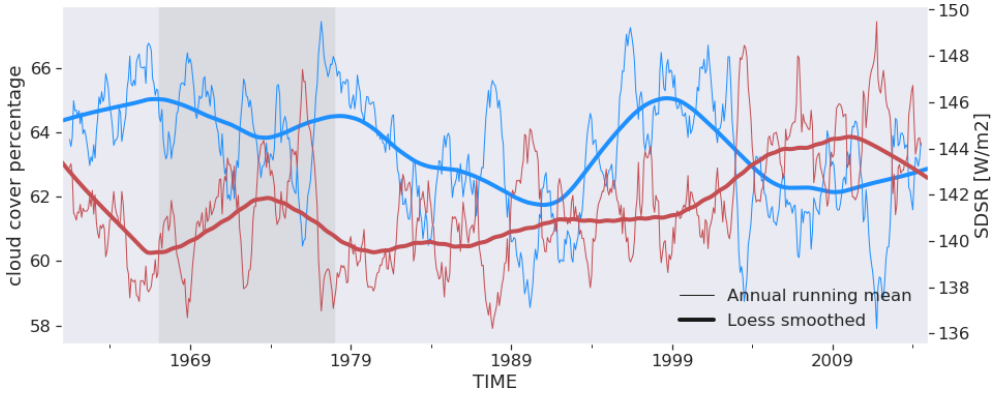

Figure A1. Time series of cloud cover (blue) and SDSR (red) between 1961 and 2014, co-located at GEBA sites in Europe. Thin lines show annual running means; bold lines show LOESS-smoothed variants. The shaded area delineates a period of interrupted dimming in Europe, between 1967 and 1978, which occurred simultaneously with a local minimum in the cloud cover trend.

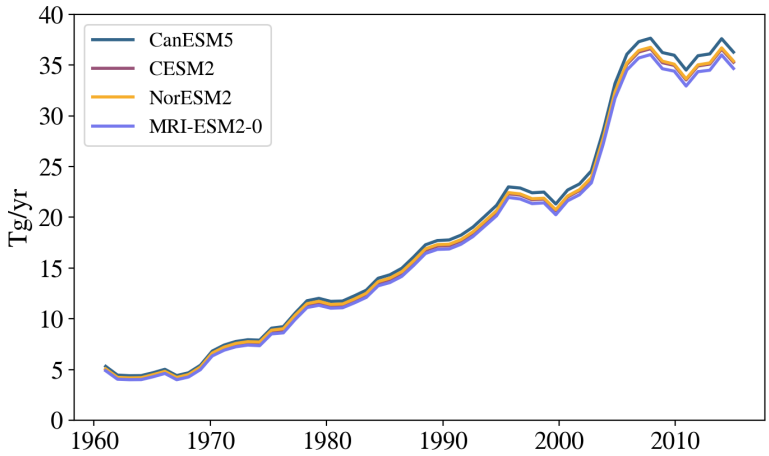

Figure A2. Emission of SO_2 in China, diagnosed by four of the models in this study. China is defined here as the area within latitudes [20–45° N] and longitudes [95–125° E].

Code and data availability. CMIP6 model outputs are freely available from the World Climate Research Programme (WCRP), 2011; https://esgf-node.llnl.gov/search/cmip6/ (WCRP, 2020). CRU TS v4.02 used for the cloud cover analysis is available from the University of East Anglia's website: https://crudata.uea.ac.uk/cru/data/hrg/ (University og East Anglia, 2020). The gap-filled GEBA data is available https://doi.org/10.5281/zenodo.4382033 (Storelvmo, 2020).

Author contributions. KOM wrote most of the article and did all analysis of CMIP6 data. MS and TS contributed to the design of the study and helped edit the text. DO, PN, JNSC, TT, NO, SEB, and GG contributed model data via the ESGF CMIP6 archive. IRJ and MW contributed with observational data, and IRJ wrote part of the Appendix. All the co-authors contributed to the analysis and gave feedback on the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study benefited greatly from the CMIP6 data infrastructure for handling and providing model data for analysis. Jan Griesfeller is thanked for data organization. High-performance computing and storage resources were provided by the Norwegian infrastructure for computational science (through projects NS2345K, NS9560K and NS9252K) and the Norwegian Meteorological Institute. Toshihiko Takemura was supported by the supercomputer system of the National Institute for Environmental Studies, Japan, and JSPS KAKENHI grant no. JP19H05669. Naga Oshima was supported by the Japan Society for the Promotion of Science (grant nos. JP18H03363, JP18H05292 and JP20K04070), the Environment Research and Technology Development Fund (JPMEERF20172003, JPMEERF20202003 and JPMEERF20205001) of the Environmental Restoration and Conservation Agency of Japan, the Arctic Challenge for Sustainability II (ArCS II), programme grant no. JPMXD1420318865, and a grant for the Global Environmental Research Coordination System from the Ministry of the Environment, Japan. GEBA is supported by the Federal Office of Meteorology and Climatology, MeteoSwiss, in the framework of GCOS Switzerland. Martin Wild is grateful for funding obtained from the Swiss National Science Foundation, grant no. 200020_188601.

Financial support. This research has been supported by the European Union's Horizon 2020 (FORCeS) (grant no. 821205) and the Research Council of Norway (KeyClim) (grant no. 295046).

Review statement. This paper was edited by Jui-Yuan Christine Chiu and reviewed by two anonymous referees.

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, Sci. Rep.-UK, 9, 953, https://doi.org/10.1038/s41598-018-37304-0, 2019.

Allen, R. J., Norris, J. R., and Wild, M.: Evaluation of multidecadal variability in CMIP5 surface solar radiation and inferred underestimation of aerosol direct effects over Europe, China, Japan, and India, J. Geophys. Res.-Atmos., 118, 6311–6336, https://doi.org/10.1002/jgrd.50426, 2013.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, Cambridge University Press, Cambridge, UK and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.016, 2013.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C. D., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Ãtha, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lavy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottla, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiablemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J. Adv. Model. Earth Sy., 12, e2019MS002010, https://doi.org/10.1029/2019MS002010, 2020.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., Ananthakrishnan, R., Miller, N., Denvil, S., Morgan, M., Pobre, Z., Bell, G. M., Doutriaux, C., Drach, R., Williams, D., Kershaw, P., Pascoe, S., Gonzalez, E., Fiore, S., and Schweitzer, R.: The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data, Future Gener. Comp. Sy., 36, 400–417, https://doi.org/10.1016/j.future.2013.07.002, 2014.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

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., Ole-

son, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L. v., 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 Sy., 12, e2019MS001916, https://doi.org/10.1029/2019MS001916, 2020.

Eddy, J. A., Gilliland, R. L., and Hoyt, D. V.: Changes in the solar constant and climatic effects, Nature, 300, 689–693, https://doi.org/10.1038/300689a0, 1982.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fan, C., Shannigrahi, S., DiBenedetto, S., Olschanowsky, C., Papadopoulos, C., and Newman, H.: Managing Scientific Data with Named Data Networking, in: Proceedings of the Fifth International Workshop on Network-Aware Data Management, Austin, TX, USA, 15–20 November 2015, 1–7, 2015.

Folini, D. and Wild, M.: The effect of aerosols and sea surface temperature on China's climate in the late twentieth century from ensembles of global climate simulations, J. Geophys. Res.-Atmos., 120, 2261–2279, https://doi.org/10.1002/2014JD022851, 2015.

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Santer, B. D., Stone, D., and Tebaldi, C.: The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6, Geosci. Model Dev., 9, 3685–3697, https://doi.org/10.5194/gmd-9-3685-2016, 2016.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Scientific Data, 7, 109, https://doi.org/10.1038/s41597-020-0453-3, 2020.

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, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Hoyt, D. V. and Schatten, K. H.: A discussion of plausible solar irradiance variations, 1700–1992, J. Geophys. Res.-Space, 98, 18895–18906, https://doi.org/10.1029/93JA01944, 1993.

Kaiser, D. P. and Qian, Y.: Decreasing trends in sunshine duration over China for 1954–1998: Indication of increased haze pollution?, Geophys. Res. Lett., 29, 1–4, https://doi.org/10.1029/2002GL016057, 2002.

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., Ackerman, A. S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B. I., Cruz, C. A., Genio, A. D. D., Elsaesser, G. S., Faluvegi, G., Kiang, N. Y., Kim, D., Lacis, A. A., Leboissetier, A., LeGrande, A. N., Lo, K. K., Marshall, J., Matthews, E. E., McDermid, S., Mezuman, K., Miller, R. L., Murray, L. T., Oinas, V., Orbe, C., GarcíaPando, C. P., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Shindell, D. T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yao, M.-S.: GISS-E2.1: Configurations and Climatology, J. Adv. Model. Earth Sy., 12, e2019MS002025, https://doi.org/10.1029/2019MS002025, 2020.

Koukouli, M. E., Theys, N., Ding, J., Zyrichidou, I., Mijling, B., Balis, D., and van der A, R. J.: Updated SO_2 emission estimates over China using OMI/Aura observations, Atmos. Meas. Tech., 11, 1817–1832, https://doi.org/10.5194/amt-11-1817-2018, 2018.

Liepert, B. G.: Observed reductions of surface solar radiation at sites in the United States and worldwide from 1961 to 1990, Geophys. Res. Lett., 29, 1–4, https://doi.org/10.1029/2002GL014910, 2002.

Lu, Z., Streets, D. G., Zhang, Q., Wang, S., Carmichael, G. R., Cheng, Y. F., Wei, C., Chin, M., Diehl, T., and Tan, Q.: Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000, Atmos. Chem. Phys., 10, 6311–6331, https://doi.org/10.5194/acp-10-6311-2010, 2010.

Michalsky, J., Dutton, E., Rubes, M., Nelson, D., Stoffel, T., Wesley, M., Splitt, M., and DeLuisi, J.: Optimal Measurement of Surface Shortwave Irradiance Using Current Instrumentation, J. Atmos. Oceanic Technol., 16, 55–69, https://doi.org/10.1175/1520-0426(1999)016<0055:OMOSSI>2.0.CO;2, 1999.

Nabat, P., Somot, S., Mallet, M., Sanchez-Lorenzo, A., and Wild, M.: Contribution of anthropogenic sulfate aerosols to the changing Euro-Mediterranean climate since 1980, Geophys. Res. Lett., 41, 5605–5611, https://doi.org/10.1002/2014GL060798, 2014.

Norris, J. R. and Wild, M.: Trends in aerosol radiative effects over China and Japan inferred from observed cloud cover, solar "dimming" and solar "brightening", J. Geophys. Res.-Atmos., 114, D00D15, https://doi.org/10.1029/2008JD011378, 2009.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geosci. Model Dev., 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

Ramanathan, V. and Vogelmann, A. M.: Greenhouse Effect, Atmospheric Solar Absorption and the Earth's Radiation Budget: From the Arrhenius-Langley Era to the 1990s, Ambio, 26, 38–46, 1997.

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: results from the Earth radiation budget experiment, Science, 243, 57–63, https://doi.org/10.1126/science.243.4887.57, 1989.

Romanou, A., Liepert, B., Schmidt, G. A., Rossow, W. B., Ruedy, R. A., and Zhang, Y.: 20th century changes in surface solar irradiance in simulations and observations, Geophys. Res. Lett., 34, L05713, https://doi.org/10.1029/2006GL028356, 2007.

Sanchez-Romero, A., Sanchez-Lorenzo, A., Calbó, J., González, J. A., and Azorin-Molina, C.: The signal of aerosol-induced changes in sunshine duration records: A review of the evidence, J. Geophys. Res.-Atmos., 119, 4657–4673, https://doi.org/10.1002/2013JD021393, 2014.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth Sys-

tem Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, J. Adv. Model. Earth Sy., 11, 4182–4227, https://doi.org/10.1029/2019MS001791, 2019.

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M.: The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-378, in review, 2020.

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J.-L., Nabat, P., Michou, M., Yukimoto, S., Cole, J., Paynter, D., Shiogama, H., O'Connor, F. M., Robertson, E., Wiltshire, A., Andrews, T., Hannay, C., Miller, R., Nazarenko, L., Kirkevåg, A., Olivié, D., Fiedler, S., Lewinschal, A., Mackallah, C., Dix, M., Pincus, R., and Forster, P. M.: Effective radiative forcing and adjustments in CMIP6 models, Atmos. Chem. Phys., 20, 9591–9618, https://doi.org/10.5194/acp-20-9591-2020, 2020.

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, Science, 327, 1219–1223, https://doi.org/10.1126/science.1182488, 2010.

Storelvmo, T.: Gap filled GEBA data [Data set], Zenodo, https://doi.org/10.5281/zenodo.4382033, 2020.

Storelvmo, T., Heede, U. K., Leirvik, T., Phillips, P. C. B., Arndt, P., and Wild, M.: Lethargic Response to Aerosol Emissions in Current Climate Models, Geophys. Res. Lett., 45, 9814–9823, https://doi.org/10.1029/2018GL078298, 2018.

Streets, D. G., Wu, Y., and Chin, M.: Two-decadal aerosol trends as a likely explanation of the global dimming/brightening transition, Geophys. Res. Lett., 33, L15806, https://doi.org/10.1029/2006GL026471, 2006.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Tang, W.-J., Yang, K., Qin, J., Cheng, C. C. K., and He, J.: Solar radiation trend across China in recent decades: a revisit with quality-controlled data, Atmos. Chem. Phys., 11, 393–406, https://doi.org/10.5194/acp-11-393-2011, 2011.

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, Geosci. Model Dev., 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019, 2019.

University og East Anglia: CRU TS v4.02 cloud cover data, available at: https://crudata.uea.ac.uk/cru/data/hrg/, last access: September 2020.

Wang, K. C., Dickinson, R. E., Wild, M., and Liang, S.: Atmospheric impacts on climatic variability of surface incident solar radiation, Atmos. Chem. Phys., 12, 9581–9592, https://doi.org/10.5194/acp-12-9581-2012, 2012.

Wang, Y. and Wild, M.: A new look at solar dimming and brightening in China: CHINA DIMMING AND BRIGHTENING REVISITED, Geophys. Res. Lett., 43, 11777–11785, https://doi.org/10.1002/2016GL071009, 2016.

Wang, Y. W. and Yang, Y. H.: China's dimming and brightening: evidence, causes and hydrological implications, Ann. Geophys., 32, 41–55, https://doi.org/10.5194/angeo-32-41-2014, 2014.

WCRP: CMIP6 project data, available at: https://esgf-node.llnl.gov/search/cmip6/, last access: December 2020.

Wild, M.: Global dimming and brightening: A review, J. Geophys. Res., 114, D00D16, https://doi.org/10.1029/2008JD011470, 2009.

Wild, M.: Decadal changes in radiative fluxes at land and ocean surfaces and their relevance for global warming, Wires Climate Change, 7, 91–107, https://doi.org/10.1002/wcc.372, 2016.

Wild, M. and Schmucki, E.: Assessment of global dimming and brightening in IPCC-AR4/CMIP3 models and ERA40, Clim Dynam, 37, 1671–1688, https://doi.org/10.1007/s00382-010-0939-3, 2011.

Wild, M., Gilgen, H., Roesch, A., Ohmura, A., Long, C. N., Dutton, E. G., Forgan, B., Kallis, A., Russak, V., and Tsvetkov, A.: From Dimming to Brightening: Decadal Changes in Solar Radiation at Earth's Surface, Science, 308, 847–850, https://doi.org/10.1126/science.1103215, 2005.

Wild, M., Ohmura, A., and Makowski, K.: Impact of global dimming and brightening on global warming, Geophys. Res. Lett., 34, L04702, https://doi.org/10.1029/2006GL028031, 2007.

Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., and König-Langlo, G.: The global energy balance from a surface perspective, Clim. Dynam., 40, 3107–3134, https://doi.org/10.1007/s00382-012-1569-8, 2013.

Wild, M., Ohmura, A., Schär, C., Müller, G., Folini, D., Schwarz, M., Hakuba, M. Z., and Sanchez-Lorenzo, A.: The Global Energy Balance Archive (GEBA) version 2017: a database for worldwide measured surface energy fluxes, Earth Syst. Sci. Data, 9, 601–613, https://doi.org/10.5194/essd-9-601-2017, 2017.

Wild, M., Hakuba, M. Z., Folini, D., Dörig-Ott, P., Schär, C., Kato, S., and Long, C. N.: The cloud-free global energy balance and inferred cloud radiative effects: an assessment based on direct observations and climate models, Clim. Dynam., 52, 4787–4812, https://doi.org/10.1007/s00382-018-4413-y, 2018.

Yang, S., Wang, X. L., and Wild, M.: Homogenization and Trend Analysis of the 1958–2016 In Situ Surface Solar Radiation Records in China, J. Climate, 31, 4529–4541, https://doi.org/10.1175/JCLI-D-17-0891.1, 2018.

Yang, S., Wang, X. L., and Wild, M.: Causes of Dimming and Brightening in China Inferred from Homogenized Daily Clear-Sky and All-Sky in situ Surface Solar Radiation Records (1958–2016), J. Climate, 32, 5901–5913, https://doi.org/10.1175/JCLI-D-18-0666.1, 2019.

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, J. Meteorol. Soc. Jpn., 97, 931–965, https://doi.org/10.2151/jmsj.2019-051, 2019.

Yunfeng, L., Daren, L., Xiuji, Z., Weiliang, L., and Qing, H.: Characteristics of the spatial distribution and yearly variation of aerosol optical depth over China in last 30 years, J. Geophys. Res.-Atmos., 106, 14501–14513, https://doi.org/10.1029/2001JD900030, 2001.

Paper II

II

Using ice cores to evaluate CMIP6 aerosol concentrations over the historical era

Kine Onsum Moseid, Michael Schulz, Anja Eichler, Margit Schwikowski, Joseph R. McConnell, Dirk Olivié, Allison S. Criscitiello, Karl J. Kreutz, Michel Legrand
Submitted to Journal of Geophysical Research: Atmospheres , 2021

Using ice cores to evaluate CMIP6 aerosol concentrations over the historical era

Kine Onsum Moseid[1,2], Michael Schulz[1,2], Anja Eichler[3,4], Margit Schwikowski[3,4,5], Joseph R. McConnell[6], Dirk Olivié[1], Alison S. Criscitiello[7], Karl J. Kreutz[8,9], Michel Legrand[10,11]

[1]Norwegian Meteorological Institute, Research Department, Oslo, Norway
[2]Department of Geoscience, University of Oslo, Oslo, Norway
[3]Laboratory of Environmental Chemistry, Paul Scherrer Institute, 5232 Villigen, Switzerland
[4]Oeschger Centre for Climate Change Research, University of Bern, 3012 Bern, Switzerland
[5]Department of Chemistry and Biochemistry, University of Bern, 3012 Bern, Switzerland
[6]Division of Hydrologic Sciences, Desert Research Institute, Reno, NV 89512
[7]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, Canada
[8]School of Earth and Climate Sciences, University of Maine, Orono, Maine, USA
[9]Climate Change Institute, University of Maine, Orono, Maine, USA
[10]Université Grenoble Alpes, CNRS, Institut des Géosciences de l'Environnement (IGE), Grenoble, France
[11]Laboratoire Interuniversitaire des Systèmes Atmosphériques, Université de Paris and Univ Paris Est Creteil, CNRS, LISA, F-75013, France

Key Points:

- The sulfate ice-core increase until 1970 and its subsequent decrease is well depicted by models
- The post-1950 increase of BC predicted by models is not confirmed by ice-core trends showing instead an early 20th century maximum
- Ice cores reveal possible error in CMIP6 emission inventories of BC in Europe

Corresponding author: Kine Onsum Moseid, `kristineom@met.no`

Abstract

The radiative effect of anthropogenic aerosols is one of the largest uncertainties in Earth's energy budget over the industrial period. This uncertainty is in part due to sparse observations of aerosol concentrations in the pre-satellite era. To address this lack of measurements, ice cores can be used, which contain the aerosol concentration record. To date, these observations have been under-utilised for comparison to aerosol concentrations found in state-of-the-art Earth system models. Here we compare long term trends in concentrations of sulfate and black carbon (BC) between 15 ice cores and 11 Earth system models over nine regions around the world during the period 1850-2000. We find that for sulfate concentration trends model results generally agree with ice core records, whereas BC concentration trends differ. Absolute concentrations of both investigated species are overestimated by the models, probably in part due to representation errors, but we assume that bias in relative trends are not altered by these. Ice cores in the European Alps and Greenland record a maximum BC concentration before 1950, while most Earth system models used in this study agree on a post-1950 maximum. We source this bias to an error in BC emission inventories in Europe. Emission perturbation experiments using NorESM2-LM support the observed finding that BC concentrations in Northern Greenland ice cores are recording European emissions. Errors in BC emission inventories have implications for all future and past studies where CMIP6 historical simulations are compared to observations relevant to aerosol forcing.

1 Introduction

Aerosols and aerosol precursors are both naturally and anthropogenically emitted into the atmosphere, and their total effect on climate is primarily by cooling the surface and thereby counteracting global warming (Storelvmo et al., 2016; P. Forster et al., 2021; Lohmann & Feichter, 2005). However, the historical aerosol forcing is highly uncertain (P. M. Forster et al., 2016; Schulz et al., 2006), both due to sparse aerosol observations in the pre-satellite era and our lack of understanding of aerosol microphysical processes (Lohmann, 2017). Constrained historical aerosol forcing estimates would allow ultimately for a better quantification of climate sensitivity (Bender, 2020; Bellouin et al., 2019), which is a key parameter in climate science (Sherwood et al., 2020).
Glaciers act as archives for deposited aerosols, and ice cores allow for the evaluation of aerosol concentration trends in the near and far past. Even though previous studies have compared ice core data to atmospheric models (Engardt et al., 2017; Bauer et al., 2013; Fagerli et al., 2007) they are an under-utilized source for Earth system model (ESM) evaluation.

ESMs are commonly used as a numerical tool for carrying out experiments to determine for instance aerosol forcing and climate sensitivity. The reconstruction of historical climate evolution by an ESM depends among other factors on the external forcing, and here of specific interest the evolution of aerosol emissions. Within the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016), a large number of ESMs participated in performing the historical experiments all using the so-called CMIP6 emission inventories as prescribed by Hoesly et al. (2018) and van Marle et al. (2017). Within these emission inventories we find the aerosol black carbon (BC) and the sulfate aerosol precursor sulfur dioxide which are both produce from fossil fuel burning. Together, they represent two crucial components for radiative forcing calculations, with major contributions to both the scattering and absorbing components of aerosol forcing.

A previous study has pointed to the potential errors in the CMIP6 emission inventory for the sulfate precursor sulphur dioxide (Moseid et al., 2020), possibly underestimating East Asian emissions. Emission inventories of aerosol Black Carbon (BC) as contained in CMIP6 (Hoesly et al., 2018) are associated with medium confidence according to the latest IPCC report (Naik et al., 2021). In addition, previous studies have suggested his-

torical BC emission can be substantially higher than depicted in current inventories (Bauer et al., 2013; Hodnebrog et al., 2014). Emission trends from the pre-satellite era (before the year 2000 for aerosol parameters) have an unquantified uncertainty and are often a result of scaling of more recent inventory years (Hoesly et al., 2018).

Here, we investigate the concentration of both sulfate and BC in ice as calculated in 11 ESMs participating in CMIP6, and compare them with ice core records from 15 cores in nine regions. We analyse the historical era (1850 to 2014), as we can take advantage of models using the recent best-guess CMIP6 emission inventory and results stored in the CMIP6 database. Our hypothesis is based on the idea that long term-trends in the concentration of an aerosol species in ice cores are a fingerprint of the aerosol's trend in general atmospheric burden and, thus, can be used to verify emission evolution over historical times, as compiled in the CMIP6 emission inventory. Assuming models represent the transport and deposition of an aerosol with comparable quality over the historical time span, the calculated concentrations trends should match the concentration evolution recorded in the ice cores with a relatively constant bias over time. An incorrect emission inventory in either sulfate or BC would reveal itself as a bias between observed and modelled concentration trend in ice for that specific aerosol. Since both sulfate and BC originate from similar source regions in the industrial era, at least on a large scale, the comparison between models and ice cores should show a highly consistent and correlated bias over time for both components.

Using the ensemble of CMIP6 historical model simulations offers the chance to investigate whether such bias is robust across different historical aerosol change reconstructions from a range of ESMs.

The following section will present the ice core and model data, and how the simulated concentrations in ice are calculated. Section 3 will present the results of our analysis, and a subsequent investigation of the potential causes for inter-model differences. Finally Section 4 discusses how the results match our hypothesis, and Section 5 concludes the study and what implications it may have.

2 Data and Methods

2.1 Ice core data

An overview of the location of the ice cores used in this study is shown in the map in Figure 1. We have selected ice cores where BC and sulfate concentrations were available with at least annual temporal resolution. All are from the Northern Hemisphere with the exception of one. Each ice core is presented individually (see Table 1) except those retrieved from Greenland. These eight ice cores are aggregated and averaged for a Northern and Southern Greenland region. To make sure there is no confusion with other atmospheric "concentration" metrics in use, we define the term "concentration" as being the concentration of an aerosol species (here sulfate and BC) measured in ice cores, when melting and analysing them. The methods for obtaining the aerosol concentrations from the ice cores used here are described in the publications listed in Table 1. Previously unpublished data and methods are described in more detail below.

2.1.1 Sulfate concentration

Most published sulfate concentration records in ice cores, including the ones we are using here (See Tab. 1), have been obtained using ion chromatography (IC) detecting the soluble fraction of sulphur in ice (Avak et al., 2019). However, the sulfate concentrations in the ice cores newly reported in this study were measured with Inductively Coupled Plasma Mass Spectrometry (ICP-MS). This applies to the Eclipse, McCall Glacier, Mt Oxford, and NGT_B19 ice cores. The ICP-MS method is based on mass spectrometry

Table 1. Overview of the ice cores used in this study along with respective references where BC and sulfate data can be found. Previously unpublished data is referenced as *This study*.

Site	Lat	Lon	BC	sulfate
Illimani	-16.62	-67.76	Osmont et al. (2019)	Kellerhals et al. (2010)
Eclipse	60.5	-139.5	*This study*	*This study*
McCall Glacier	69.3	-143.8	*This study*	*This study*
Mt Oxford	82.2	-73.2	*This study*	*This study*
Greenland				
Northern				
NGT_B19	78.0	-36.4	*This study*	*This study*
Tunu2013	78.0	-33.9	doi:10.18739/A2ZQ1G	Sigl et al. (2015)
NEEM_2011_S1	77.5	-51.1	Zennaro et al. (2014)	Sigl et al. (2013)
Humboldt	78.5	-56.8	McConnell (2010)	Sigl et al. (2013)
Southern				
Summit2010	72.6	-38.3	doi:10.18739/A2XV7T	doi:10.18739/A2XV7T
D4	71.4	-43.9	McConnell (2007)	McConnell (2007)
ACT11d	66.5	-46.3	*This study*	doi:10.18739/A2Z933
ACT2	66.0	-45.2	McConnell and Edwards (2008)	McConnell and Edwards (2008)
Col Du Dôme	45.8	6.9	*This study*	Preunkert et al. (2001) Legrand et al. (2013)
Colle Gnifetti	45.9	7.85	Sigl et al. (2018)	Engardt et al. (2017)
Mt Elbrus	42.4	42.4	Lim et al. (2017)	Preunkert et al. (2019)

and has the advantage of detecting all chemical forms of sulphur in the ice. A mix of analysis techniques can lead to discrepancies between measured concentrations of sulfate depending on the technique implemented (IC or ICP-MS). However, in historical times since 1850, the majority of the sulfur in ice cores is soluble sulfate, fully captured during analysis by both IC and ICP-MS. Therefore, the sulfate concentrations are very similar regardless of the technique used to measure them (McConnell et al., 2017). Yalcin and Wake (2001) found that the new ICP-MS-based sulphur measurements in the Eclipse core presented here are in close agreement with earlier IC based sulfate measurements from the same core. Sulfate measured by ICP-MS includes methanesulfonate (MSA), which has been shown to account for 3-5 % of the sulfate found in Greenland ice (Legrand et al., 1997). Sulfate can be subject of re-location at ice core sites influenced by meltwater percolation (see e.g. (Eichler et al., 2001)). To avoid such a postdepositional change of the original deposited signal, we chose ice core sites with negligible influence of melting.

2.1.2 Black Carbon

Concentrations of BC in all ice cores were determined with a Single Particle Soot Photometer (SP2, Droplet Measurement Technologies) coupled with a jet or ultrasonic nebulizer to aerosolize the molten ice core samples (Wendl et al., 2014; McConnell, 2007).

Although a number of annually resolved ice-core BC (e.g., Liu et al. (2020)) and sulfur (e.g., Sigl et al. (2014)) records spanning this period are available from Antarctica, here we chose to focus on ice-core record/model comparisons in the Northern Hemisphere and included only one comparison from a tropical site in Southern Hemisphere (Illimani).

2.2 Concentrations from ESMs

The Earth system is extremely complex and includes a multitude of interacting processes. As such, several ESMs have been developed with different ways to represent these interwoven processes. To compare these different models and the representations of physical processes within them, the Coupled Model Intercomparison Project Phase 6 (CMIP6) has been initiated as a collaborative effort across the ESM community (Eyring et al., 2016). CMIP6 consists of several model intercomparison projects (MIPs) that design experiments tailored to different aims and focus points; however, here we use only the historical experiment. Every ESM has to perform a set of basic experiments to participate in any of the MIPs of CMIP6. The historical experiment is one of them meant to simulate the climate from 1850 to 2014 and to allow for a comparison to observations and recent climate evolution. Its period was chosen to cover the times where the observational record is comprehensive enough that comparisons and evaluations between the real climate and simulations can be conducted. The historical experiment is forced and driven by a best guess of greenhouse gas concentrations and both natural and anthropogenic aerosol and aerosol precursor emissions.

Here we use the results from 11 of the ESMs that participated in CMIP6. They were chosen based on the availability of diagnostic outputs and variables needed to calculate the simulated concentrations of sulfate and BC at ice core sites in the historical experiment. The sum of dry and wet deposition of sulfate and BC along with total precipitation, colocated with the ice cores, was used to compute concentrations. The ESMs, their horizontal resolution, and corresponding references are listed in Table B2, and further details of the variables used to calculate the concentration of aerosols to compare with ice core data are found in supplementary Section B1.

2.3 Comparing ice core data to model data

Ice cores have been retrieved in topographically varying areas such as on high-alpine mountain glaciers in the European Alps, or in more smooth topographic areas, like on the top

of the ice cap of Greenland. In contrast, ESMs represent Earth's surface as a matrix of surface grid cells, at differing horizontal resolutions depending on the ESM (see Table B2). The topographic elevation within each grid cell in a model must represent both the peaks and the valleys within the area it covers. This means that the topography in each grid cell is a flat surface at the average altitude of the grid cell. Therefore, a point measurement such as an ice core taken at any location of a high mountain peak, will correspond to an ESM grid cell at a lower elevation, as shown in Table 2. Especially in mountain areas, we can expect the ESM to exhibit higher concentrations than measured at the ice core site due to the representation error. However, this bias is probably not changing much over historical time scales, and we assume that trends in models and in ice cores are expected to be correlated.

As we are interested in the long term trends of aerosol concentration and less so in a model's ability to represent local meteorological conditions, and since models show variability in deposition between close-by grid cells, we further extend the area from which model data are taken for the comparison of aerosol concentrations at the ice core site. In particular, we use the nearest neighbour method to find the grid cell closest to the ice core location, and then find the surrounding 3x3 grid cells and average them with equal weight.

ESMs do not output aerosol concentration directly, therefore it is calculated here based on the wet and dry deposition of the aerosol and total precipitation in the chosen area:

$$conc = \frac{\Sigma_1^9(wet + dry)}{\Sigma_1^9 prec_{ice+liq}} \tag{1}$$

where *conc* is the concentration of the aerosol component in question, and *wet* and *dry* refers to the sum of deposition of the aerosol in the 9 grid cells, divided by the total ice and liquid precipitation *prec* in the same 9 grid cells. Aerosol concentrations as calculated from the models and as measured in the ice cores are then averaged over ten year intervals, starting at the beginning of each decade. Altogether, this is done to compare the long-term trends in aerosol concentration rather than the inter-annual variability in the data, which would be primarily due to changes in meteorological conditions.

Table 2. Altitudes in meters at ice core sites and in model means with standard deviation shown in parenthesis. Model mean was averaged over 3x3 grid matrix surrounding ice core location and averaged together for the models used. For Northern and Southern Greenland we have used NGT_B19 and Summit2010, respectively.

Site	**Ice Core**	Model mean (std)
Illimani	**6300**	2303 (314)
Eclipse	**3017**	1031 (185)
McCall Glacier	**2400**	437 (127)
Mount Oxford	**2210**	452 (153)
Northern Greenland	**2270**	2318 (135)
Southern Greenland	**3258**	2189 (199)
Col Du Dôme	**4350**	719 (190)
Colle Gnifetti	**4452**	739 (174)
Mt Elbrus	**5115**	1001 (194)

While the concentration is straightforward to calculate, there is no way to determine the source regions of the aerosols at ice core sites in the CMIP6 simulations. Ad-

ditional simulations are therefore conducted with one of them, the ESM NorESM2-LM, as described in the following section.

2.4 Emission experiments with NorESM2-LM

To investigate which emission source areas affect aerosol concentration in the ice core locations, we have designed seven perturbation experiments and one control experiment. In a first set of experiments, the anthropogenic emissions of either BC or sulfur dioxide (aerosol precursor of sulfate) each are doubled in one of three regions at a time. The regions are defined according to the HTAP2 definitions as described in Galmarini et al. (2017) and represent Europe, Asia, and North America. This adds up to six experiments, and in a seventh experiment, the emissions from global wildfires were doubled, to investigate to what extent natural and anthropogenic biomass burning contribute to concentrations of BC. An overview of experiments is found in Table A1. To reduce inter-annual variability through forcing and feedback on the circulation and tracer transport in these perturbation experiments, the sea surface temperatures (SSTs) and sea-ice cover (SIC) were fixed to the SSTs and SIC fields extracted from the NorESM2-LM fully coupled historical experiment. We also performed one control experiment with unmodified emissions as in the historical simulation, using the same prescribed SSTs and SICs. In all of the experiments the aerosol emissions are based on the CMIP6 anthropogenic (Hoesly et al., 2018) and biomass-burning (van Marle et al., 2017) emission inventories. The experiment names and descriptions including the varying emission perturbations are presented in the supplementary Appendix A.

The aerosol concentrations at different ice core locations from these experiments are then used to calculate and estimate regional contributions. For this calculation we assume linearity, meaning that a doubling of emissions is expected to double the concentration. In order to assess the contribution of aerosols from each source region or wildfire, the aerosol concentration in the control simulation (histSST in Table A1) is simply subtracted from each of the individual experiments. The remainder is then the aerosol concentration believed to stem from the particular source region (or wildfire emissions) tested in the experiment.

3 Results

3.1 Sulfate and BC in ice cores and models

Figure 1 compares the simulated concentration of sulfate (A1-I1) and black carbon (BC, A2-I2) with the measured concentrations from seven ice cores and two composites from ice core in northern and southern Greenland. It displays the decadal average from the 11 models used in this study (blue solid line) and the model spread defined as the minimum and maximum decadal model average (shaded blue). The ice core concentrations (black solid line) are based on one ice core, except for the two areas Northern Greenland and Southern Greenland which are based on several ice cores (see map in Fig. 1 and Table 1).

The model means of the sulfate concentrations show a general increase until the mid- to late 1970s across Northern Hemisphere ice core areas, followed by a subsequent significant decrease in concentration. The model spread in decadal sulfate concentration is large across regions, especially in the decades before and after the 1970s. We find that the models in general show a sulfate concentration larger in magnitude than what is recorded by the ice cores. The models with the lowest concentrations are very close to the measured sulfate, in particular in Greenland and Colle Gnifetti. However, one can still identify a similar temporal evolution in ice core and models for sites such as Mt Oxford, Northern and Southern Greenland, Col Du Dôme, Colle Gnifetti, and, to some extent, Mt Elbrus. We also added a gray stippled line in Figure 1 showing five fold the decadal ice core

observation data in areas where the ice core data is outside the range of the models, which is every area except Illimani for sulfate. We refer to this graph as the five factor ice core data, and it helps visually to compare magnitude and trends between models and ice cores. For example, we see that both model means and ice core concentrations of sulfate show similar temporal trends with two maxima in the area surrounding McCall Glacier. However, there is a model-observation discrepancy as to when the two temporal maxima are largest in magnitude. The five factor lines also illustrate that the model mean concentration of sulfate in Col Du Dôme is approximately a factor of five higher than the ice core data, while in Northern and Southern Greenland the model mean concentration is higher by a factor of three to four.

The five factor ice data shows the ice core trends more clearly, and highlights that models in general are able to reproduce the evolution of sulfate concentrations, while they are unable to reproduce their magnitude. There may be several reasons the overestimation of the sulfate concentrations by the models. Recall from Section 2.3, all models are unable to resolve the real surface topography, meaning that the elevation in the ice core sites are not represented in the models (Table 2), explained as representation error above. Secondly, as we use a 3x3 grid box area surrounding the model grid cell closest to the ice core location (as explained in Section 2.3), grid boxes that are closer to aerosol emission sources than the actual ice core site are included in the model average, giving a larger absolute concentration value. This is especially relevant for ice cores close to sources such as Col Du Dôme and Colle Gnifetti in the European Alps. However, even if their magnitude differs, the sulfate comparison shows that the trends are well correlated, indicating consistency between the temporal evolution of sulfur emissions and sulfate concentrations in ice core archives.

In general, the absolute magnitude of BC concentrations is much lower than that of sulfate (see Figure 1 A2-I2). The temporal evolution of BC concentrations differs more prominently between ice cores and models than what was shown above for sulfate. Only at Mt Elbrus and, to some extent, Mt Oxford, comparable temporal trends appear between model and ice core data throughout the analysed time period. Modelled BC concentration at Illimani shows a sharp increase in the most recent period which is not measured in the ice core. We also find that while the model mean concentration of BC has maximum values post-1950 in the European Alps, the ice core data show a pre-1950 maximum of BC concentration in both sites (Col Du Dôme and Colle Gnifetti).

In fact the maximum in modelled BC concentration occurs after 1950 for all sites. However, at five out of nine areas, the maximum in the ice core data occurs before 1950. Also in both Greenland areas the comparison shows a clear discrepancy between the model data, with present-day ice core BC concentrations lower than pre-industrial values, and a distinct maximum in the early 1900s.
Another feature to note from the comparison is the inter-model range in BC concentrations in the European Alps, McCall Glacier and Mt Elbrus, which stays close to constant over time (Fig. 1 G2, H2, C2 and I2), as opposed to other areas like Illimani, Eclipse, Mt Oxford and both regions of Greenland where the inter-model range increases with time (Fig. 1 A2, B2, D2, E2, and F2). A constant model range indicates high model agreement in concentrations, while a diverging model range indicates that inter-model differences become more important with time. In the following we investigate two representative ice core sites, Colle Gnifetti and Northern Greenland, more closely, because they represent this difference in the evolution of the inter-model range of BC concentrations.

3.2 Inter-model differences at the European Alps and Northern Greenland

We find the largest model range at Colle Gnifetti (European Alps) for sulfate concentrations in the 1970s and -80s (Fig. 1 H1). In Figure 2A we can find the ESMs that are

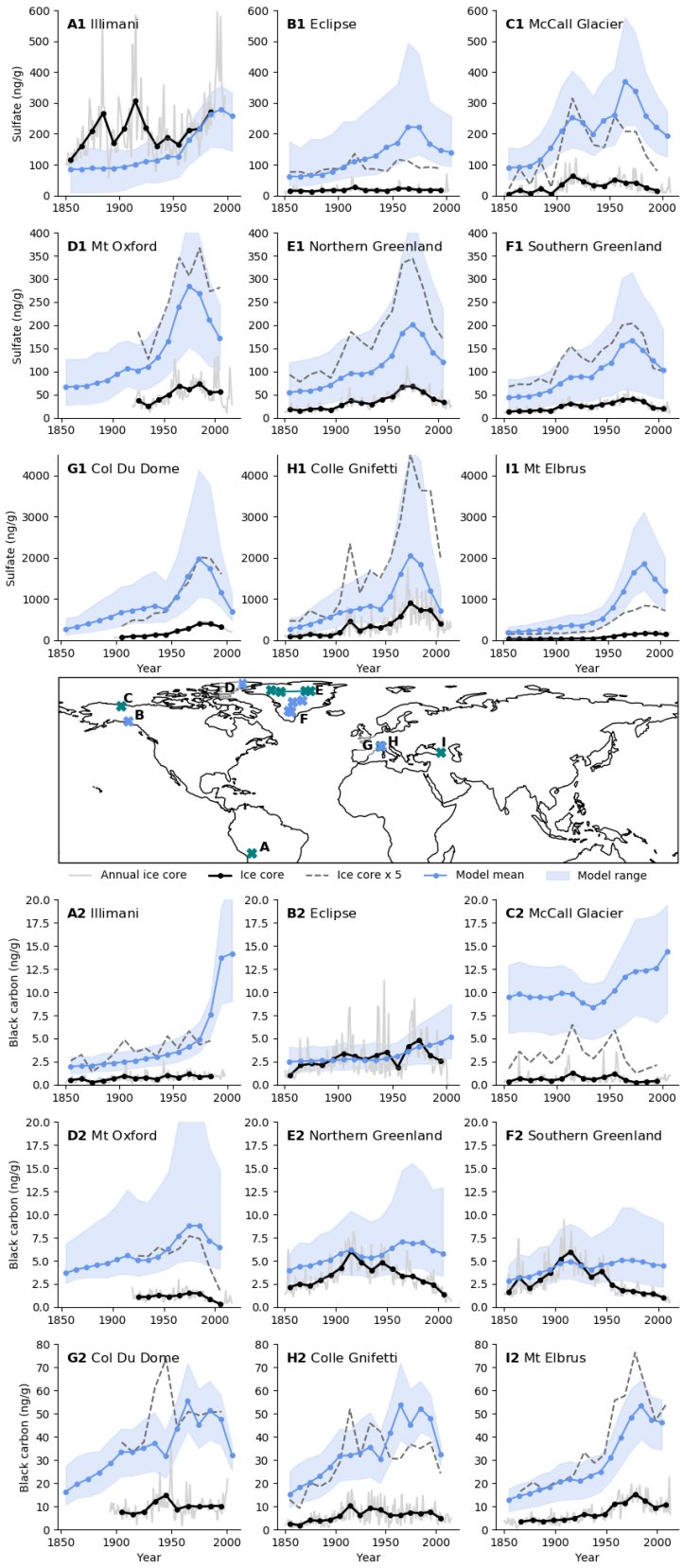

Figure 1. Decadal sulfate (A1-I1) and black carbon (A2-I2) concentrations [ng/g] in ice cores (black) and model mean of 11 models (blue), decadally averaged. The shading shows the maximum/minimum decadal average of the 11 models. The light gray solid line shows the annual ice core concentration. The stippled dark gray line shows the respective decadal ice core concentrations multiplied by a factor of 5.

responsible for this range. Models with maximum sulfate concentration in the high emission era are MIROC6, MPI-ESM-1-2-HAM and MRI-ESM2-0. These models have at the same time widely varying pre-industrial background sulfate concentrations. To further investigate inter-model trend differences we present the data as a percent change from their respective pre-industrial values of the period 1850-1865 (Fig. 2B). The ice core observations show a maximum increase of 800 % from pre-industrial values, which is in the middle of the model range of 400-1200 % change. The long-term relative trend in sulfate concentration at Colle Gnifetti is well represented in all models of this study.

Figures 2C and D show BC concentration at Colle Gnifetti in absolute values and percentage change from the pre-industrial era, respectively. Identical to Fig. 1 H2, in general the absolute model concentrations are higher than ice core values. Note also that there are evidently inter-model differences in pre-industrial concentration values. The percentage changes from pre-industrial values (Fig. 2D) clearly show a maximum BC concentration recorded in the ice core in the early 20th century, while all models agree on a late 20th century maximum. This was also apparent in the five factor ice core data shown in Figure 1 H2.

Figures 2E to H show absolute and percentage change in sulfate and BC concentrations according to the 11 models of this study together with the respective averages of four ice cores located in Northern Greenland (see Table 1). For the sulfate concentration, some models have comparable values to that of the ice cores, while others grossly overestimate the absolute concentrations. We note that the model with the highest concentration value is EC-Earth3-AerChem, which is the model with the lowest sulfate concentration at Colle Gnifetti (Fig. 2A). Likewise, similar opposite findings are found for MIROC6, which was the model with the highest sulfate concentrations at Colle Gnifetti, but has the lowest values in Northern Greenland. EC-Earth3-AerChem is the model of this study with the highest global lifetime for sulfate (6.7 days in the present day, see Fig C1), and MIROC6 is the model with the lowest global sulfate lifetime (2 days in the present day, see Fig C1), as computed from global burden and deposition. A long lifetime means that a particle is transported longer before being deposited. For these two models, global lifetimes explain why a model depositing in large magnitudes close to emission sources (i.e. Colle Gnifetti) does not deposit as much in a pristine remote area (i.e. Northern Greenland) and vice versa. However, for the rest of the models global lifetimes (see Fig C1) do not explain so straightforwardly the order in sulfate concentration differences between Europe and Northern Greenland as shown in Figures 2A and E. The graph showing percent change of sulfate concentration illustrates that the recorded changes from ice cores in Northern Greenland are in the middle of the model range. The relative sulfate concentration trends in Northern Greenland are well represented by all of the models in this study (Fig. 2F).

The absolute values of BC concentrations in Northern Greenland differ largely between models and ice core data. According to the ice core records there is a maximum of BC concentrations in the early 20th century, followed by a continuous negative trend after 1950 (Fig. 2G). The models do not represent this evolution, which is clearly shown in both Figures 2G and 2H. EC-Earth3-AerChem and GFDL-ESM4 have the largest BC concentrations in the end of the 20th century of about 15 ng/g (Fig. 2G), but when looking at the percentage change, multiple models have an overall positive trend since pre-industrial times. The models disagree not only with the ice core records, but also among each other regarding the long-term BC concentration trends in Northern Greenland. The differences in timing of BC concentration maxima are significant, and the source region of the BC arriving at the different ice core sites is of interest. An investigation using source region attribution within only one model has been thus added to this study.

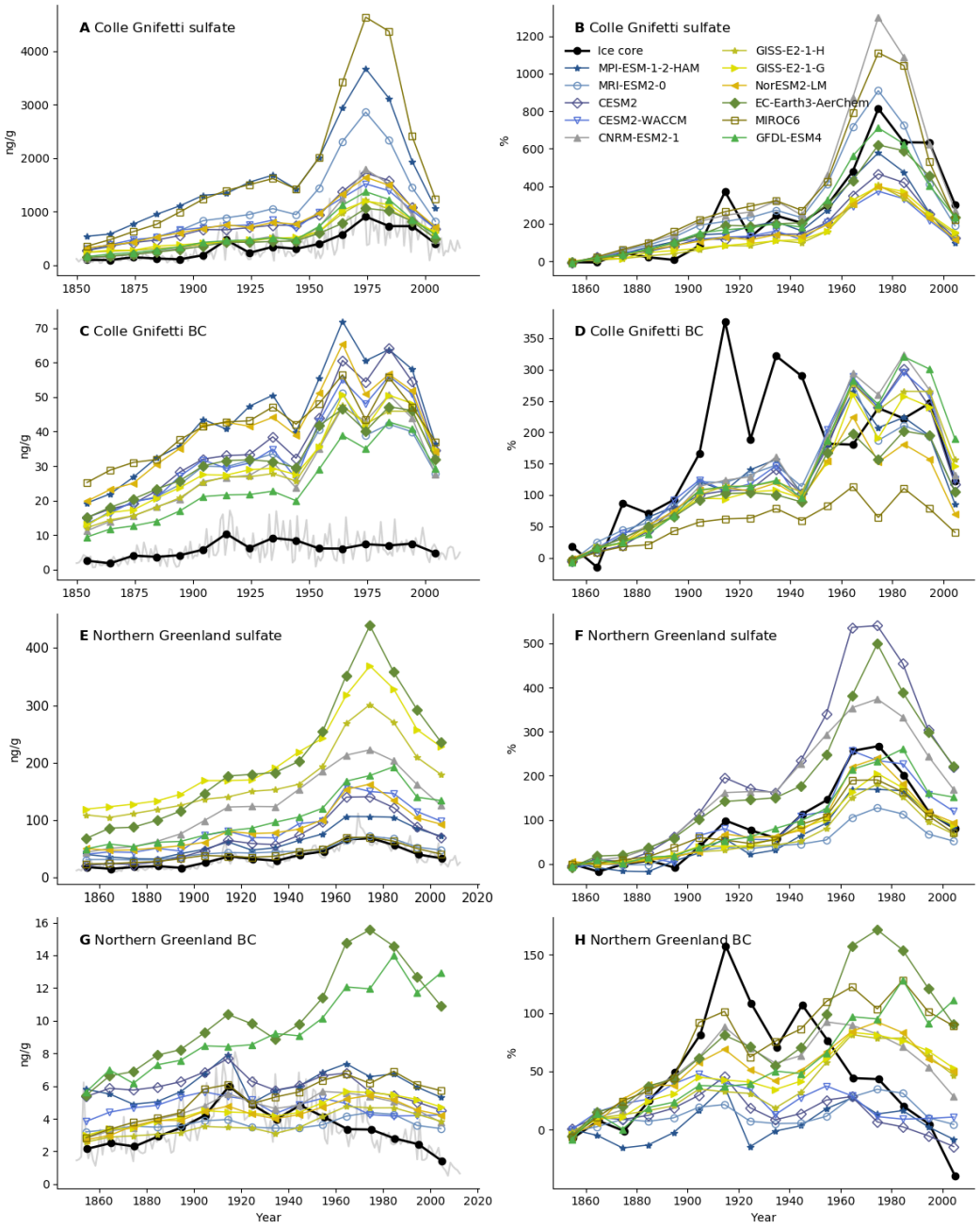

Figure 2. Concentrations of sulfate and BC in absolute values (left column) and percent change from pre-industrial (right column) for Colle Gnifetti (A-D) and Northern Greenland (E-H). Percent change is calculated using the first 15 years of each respective timeseries as baseline.

3.3 Emission region attribution using NorESM2-LM experiments

Concentration contributions from different source regions of both sulfate and BC at all ice core areas of this study have been calculated as explained in Section 2.4, and the result is shown in Figure 3 and Figure 4, respectively.

For most of our ice core areas the sum of sulfate contributions (bars in Fig. 3) does not add up to that of the reference simulation (dark green), which means that other sources than those we chose in our perturbation experiments contribute to the total sulfate concentration. Other sources comprise natural emissions of DMS and SO_2 from volcanic activity, as well as anthropogenic emissions from other regions of the world.
In fact, most of the sulfate concentration in Illimani can be attributed to other source areas (such as Southern America), while for Eclipse, McCall Glacier, and Mt Elbrus other sources account for about half of the simulated total sulfate concentration. North American sulfur dioxide emissions dominate among our selected perturbation sources in Eclipse, McCall Glacier and both Greenland areas, and one can note increasing Asian contributions in the end of our investigated period. Unsurprisingly, almost all sulfate found in Col Du Dôme and Colle Gnifetti can be attributed to European emissions according to NorESM2-LM. Biomass burning is not a large contributor to sulfate concentrations in any of the ice core sites, but acts as the definite main contributor to BC concentrations at Illimani (Fig. 4 A).

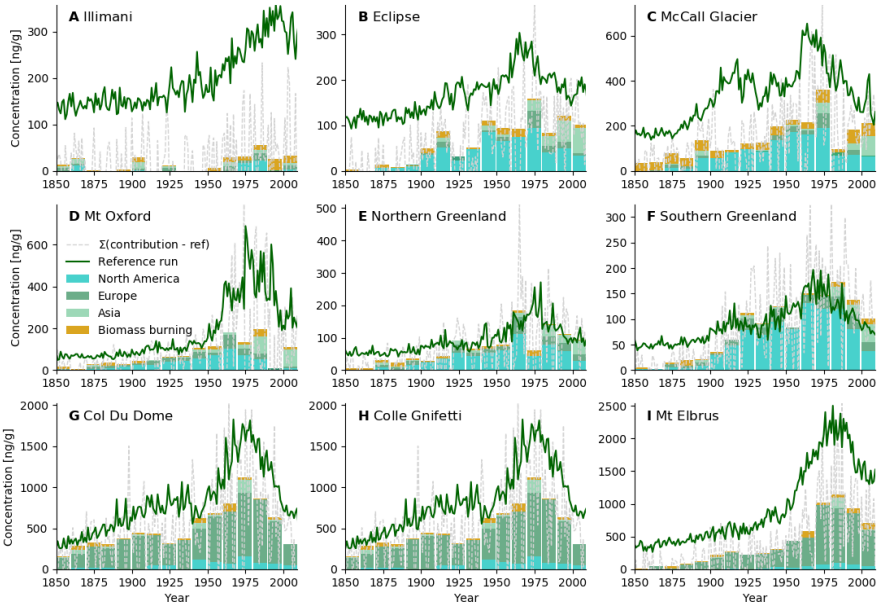

Figure 3. Decadal contributions of sulfate from three emission source regions and biomass burning according to NorESM2-LM. The dark green line represents the reference simulation, and the gray stippled line shows the sum of error for the contribution experiments.

Biomass burning is an important contributor to BC concentrations through the investigated period for all ice core areas except for the European Alps. The most diverse contribution sources of BC concentration is found in Mt Oxford and both Greenland areas. For example, in Southern Greenland in 1970-1980, all four contributors account for 20-

33 % of the total BC concentration in NorESM2-LM. Almost all of the simulated BC concentration at Col Du Dôme and Colle Gnifetti can be attributed to European anthropogenic emissions of BC, but more interestingly European emissions also dominate among the contributions in Northern Greenland. Between 1900 and 1990 European contributions account for 20-57 % of the total BC concentration there (Fig 4 E). Asian emissions of BC become more prevalent in the last decades at many ice core sites, but especially at Eclipse.

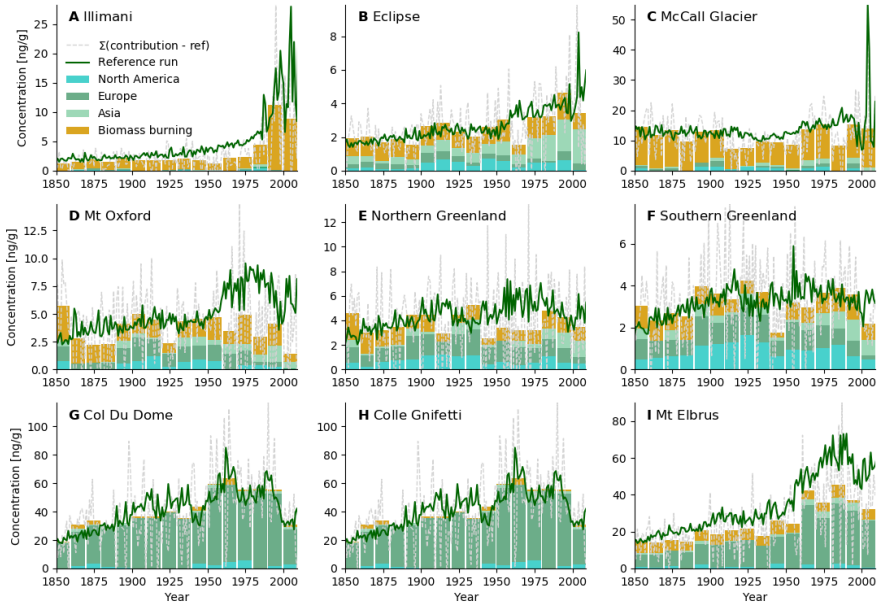

Figure 4. Decadal contributions of BC from three emission source regions and biomass burning. The dark green line represents the reference simulation, and the gray stippled line shows the sum of error for the contribution experiments.

Recall that NorESM2-LM contributions in theory are decompositions of the control simulation result, and as the previous Section showed, this control simulation does not accurately represent the BC concentration as recorded in ice cores in Northern Greenland and at Colle Gnifetti (Fig. 2D and 2H). Therefore, the contribution bars shown in Figure 4 E and H are a visualization of the components that add up to a biased evolution in BC concentration.

4 Discussion

4.1 Sulfate concentrations

The relative trend of sulfate concentrations from ice core records over the industrial era agree with that of the 11 Earth system models in this study in 7 out of 9 regions. We find inter-model agreement in sulfate concentration trends thus almost independent of ice core location, although there is disagreement between models regarding absolute magnitudes of concentration. This indicates that generally CMIP6 emission inventories represent historical trends in the sulfate precursor sulphur dioxide well. Anthropogenic sul-

phur dioxide emissions in Europe were at their maximum during the 1970s and -80s according to CMIP6 emission inventories (Figure A1 and Hoesly et al. (2018)), which is also when we find the largest model range.

By applying the NorESM2-LM models for source contributions, we found European SO_2 emissions as major sulfate sources for the European ice core sites, whereas North American sources dominate at Greenland sites. Fagerli et al. (2007) assessed contribution regions for sulfate deposition rates at Col Du Dôme and Colle Gnifetti using the EMEP model. They also concluded that the main contributor to sulfate found on Colle Gnifetti and Col Du Dome were European emission sources. Engardt et al. (2017) investigated sulphur concentration in Europe in the 20th century using two chemical transport models (EMEP MSC-W and MATCH) and ice core records from Colle Gnifetti. They found that both models represented non sea salt (nss) sulfate concentration trends well, comparable to Fagerli et al. (2007), Bauer et al. (2013), and this study. The maximum nss-sulfate concentration in EMEP MSC-W and MATCH at Colle Gnifetti was larger than the ice core recorded maximum nss-sulfate concentration by a factor of 3 and 9, respectively. We incorporated two NASA-GISS models in this study, and a previous generation of the atmospheric model (GISS-modelE) has been used to compare model to ice core records in Bauer et al. (2013). Three Greenland ice core sulfate records (Humboldt, D4, and ACT2, also used in the current study, see Table 1) from Bauer et al. (2013) agree well in trend with that using all the different setups of the GISS modelE using CMIP5 emissions (Lamarque et al., 2010). Together with the results of this study we suggest both CMIP5 and CMIP6 emission inventories of the aerosol precursor sulphur dioxide do capture overall emission trends in the Northern Hemisphere with a maximum in the 1970s and a decline since then. The ice core data available provide little constraint and verification for the East Asian sulfur emissions, since East Asian sources contribute only a small part to the sulfate recorded in the ice core archives (Figure 3).

There exist few previous multimodel studies evaluating sulfate concentrations using ice cores; all have found that although agreeing in trend, the models disagree in magnitude. The difference in concentration magnitude between models and ice cores is suggested to be in part a result of the lower elevation in grid cells in the model, and this rationale holds for all mountain sites apart from Greenland and possibly Illimani. Due to a closer proximity to emission sources, sulfate concentrations are higher at lower altitudes. Futhermore, the high-alpine glaciers are often in the free troposphere where aerosol concentrations in air are low (especially during cold season) (Engardt et al., 2017).

However, this representation error does not explain the bias at all ice core sites in the current study, as models represent high altitudes for Greenland sites, while still overestimating the sulfate concentration there. Even pre-industrial sulfate concentrations are overestimated in most models in Greenland. Note, however, that the model with the lowest concentration traces the evolution of the Greenland data quite well. The comparison at the Greenland sites suggests that the transport from anthropogenic sources in the Northern Hemisphere to these remote ice cap sites in Greenland is too efficient in most models.

The Illimani is the only site where ice core sulfate concentrations excees model estimates. The large variation in the observed record, the small increase after 1950, as well as the location on the western edge of South America suggest that other sources than fossil fuel burning are responsible for most of the sulfate found here. Diffuse volcanic and marine sulfur sources are incorporated in some models, but the comparison suggests that other sources, such as sulfur from e.g. explosive eruptions or from mining activities may play a role.

4.2 Black carbon concentrations

Trends in BC concentrations agree between decadal model mean and ice core records at Mt Elbrus, and to some extent Mt Oxford throughout the analysed period. The two ice core records from the European Alps show a pre-1950 maximum in BC concentration, while all models agree on a post-1950 maximum in this area. In both Greenland areas, however, the models disagree between each other in BC concentration trends in addition to a bias in the ice core data. As discussed above, the sulfate concentration trends are rather well represented by the simulations, and the models of this study perform well regarding formation, transport and deposition of sulfate. When models perform well in transport of one aerosol component they should perform well for another aerosol component, as transport within a model is a priori and not aerosol component dependent. However, emission amount, source regions and timing as well as processes leading to deposition are dependent on aerosol components. Thus, the apparently different temporal evolution of BC concentrations in ice core data and models can likely be due to discrepancies in either the CMIP6 BC emission inventory, or specific model deposition processes, or a combination of the two. As there is high inter-model agreement in the biased timing of the BC concentration maximum in the European Alps, we suggest errors in BC emission inventories is the source of this bias between models and observations. For Greenland this is slightly different. When multiple models are given the same input emission inventory, and we assume overall transport is correct, inter-model discrepancies in trends in BC concentrations point to inter-model deposition process differences. Such differences could reflect distinct parameterisations for a BC particle to become hydrophilic by coating of solubles. These will cause inter-model differences in BC lifetime, varying differently over time (See Fig. C2, Bond et al. (2013)). However none of the models, despite their differences in BC lifetime, represent the ice core recorded temporal evolution of BC concentration, and we suggest that the source of error in Northern and Southern Greenland is a combination of errors in emissions and differences in model deposition processes.

We investigated which emission source region contributed to the NorESM2-LM BC concentration in Northern Greenland and found that a large amount came from European sources. North American BC emissions have been thought to be a main contributor to BC in Greenland (Bauer et al., 2013) but on average in NorESM2-LM North America accounts only for 17 % of the total concentration in the northern part, and 28 % in the southern part of Greenland. Almost all of the BC at Colle Gnifetti and Col Du Dôme originates in Europe; this again points to wrong European emission data as possible reason for the bias between NorESM2-LM and the ice core data of these two sites. Since European emissions contribute to BC concentrations in Northern Greenland as well, at least part of the Northern Greenland bias for NorESM2-LM can be related to erroneous CMIP6 emission inventories of European anthropogenic BC.

Bauer et al. (2013) assessed the reconstruction of historical BC evolution between GISS-modelE and three Greenland ice cores, and found an early 20th century peak in the Greenland ice cores that is not present in the model data. This is what we found in Figure 2G and H. Bauer et al. (2013) suggest a missing source in emission inventories may be the cause of this bias, which supports our findings. While several ice core sites exhibit considerable representation errors for sulfate, Greenland and Eclipse are an exception. The bias of the model mean and ice core data between 1900 and 1950 is relatively small, although sulphate at the same time is underestimated by a factor of 3. This is another argument for a possible underestimation of BC emissions in this part of the century.

The differences in model range evolution in Figure 1 A2-I2 may be a result of inter-model differences in BC lifetimes. At some ice core sites the inter-model range is changing over time, while at others it stays rather constant. When investigating the sources for BC in nine ice core areas, according to NorESM2-LM (Fig. 4), we find that the areas with a close to constant model range (see McCall Glacier, Col Du Dome, Colle Gnifetti and Mt

Elbrus in Fig. 1) are relatively close to their emission sources. At sites where the emission sources are far away it seems that differences in BC life time enhance the inter-model range in periods with higher emissions.

At McCall Glacier, almost all BC concentration can be sourced to biomass burning in NorESM2-LM (e.g. boreal wild fires in Northern Canada and Siberia) and the model range in Figure 1 C2 stays near constant for this area. For sulfate our result suggest that the models have problems simulating the low concentrations at this Arctic high elevation site. In contrast, the nearby ice core site Eclipse has a diverging model range (Fig. 1 B2) while simultaneously as contributions from far-away sources such as Asia increase, according to NorESM2-LM. The absolute bias between models and data is small at this site, a location with larger exposure to inflow from the Pacific. The inter-model spread becomes larger as the far-away emission sources in Asia become more important in the last decades (Fig. 4). For the last decades the models simulate an increase in BC concentrations, consistent with an increase in Asian emissions. This is however not found in the ice core data, possibly because wildfire sources and their variability are dominating the ice core signal.

5 Conclusion

We have gathered sulfate and BC data from 15 ice cores in 9 regions and compared their concentrations to that calculated using 11 ESMs, which have participated in the recent CMIP6 exercise. The ice core data have been carefully compiled to provide a benchmark to test inter-decadal concentration evolution of BC and sulfate in a consistent manner. By investigating both components at the same time a more consistent picture of emissions and their evolution can be obtained.

Concentration trends as recorded in the ice cores agree with each other across large regions for sulfate, and this evolution is also captured in the modelled sulfate concentration trends. Both modelled and observed sulfate trend largely correlate to trends in global emissions of the sulfate precursor sulfur dioxide. We can conclude that the emission changes of sulfate precursors in the CMIP6 emission inventories are consistent with the observations presented here. East Asian emissions cannot be tested as rigorously because the location of the ice cores is not ideal to track them.

BC concentration trends vary across ice core sites, but observations from the European Alps and Greenland agree on a distinctive early 20th century maximum which is not present in the modelled BC concentration. In the European Alps there is high inter-model agreement on a bias in timing of the maximum BC concentration. Based on this we suggest CMIP6 emissions for BC in Europe are likely underestimated in the first half of the 20th century.
Errors in BC emission inventories have implications for all future and past studies where CMIP6 historical simulations are compared to observations relevant to aerosol forcing. This includes studies evaluating historical surface temperature, energy balance at the surface and top of atmosphere, ice nucleating particles, and historical cloud studies, to name a few.
European emissions contribute significantly to BC concentrations in 6 out of the 9 regions in this study according to NorESM2-LM, and erroneous input data likely contribute to the ice core-model bias in these regions. Emission region attribution studies with models other than NorESM2-LM would benefit the analysis of far-from-source ice core sites to further narrow the potential source of bias, together with a further investigation of BC deposition processes in ESMs.

Table A1. Emissions for CMIP6 are described in Hoesly et al. (2018) for anthropogenic emissions and in van Marle et al. (2017) for biomass burning emissions. Asia refers to South Asia + Eastern Asia + Central Asia (excluding Russia, Southeast Asia and Middle East). SST/SIC : prescribed sea surface temperature and sea-ice cover.

Experiment	Emission perturbation				Ocean model
	Species	Region	Sector	Size	
historical	–	–	–	–	Full ocean
histSST	–	–	–	–	SST/SIC prescribed from historical
histSST-so2x2nam	SO_2	North-Am.	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-so2x2eur	SO_2	Europe	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-so2x2asi	SO_2	Asia	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-bcx2nam	BC	North-Am.	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-bcx2eur	BC	Europe	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-bcx2asi	BC	Asia	Anthrop.	+100%	SST/SIC prescribed from historical
histSST-biox2	SO_2, BC	Global	Biom. burn.	+100%	SST/SIC prescribed from historical

Appendix A Contributions according to NorESM2-LM

Figure A1 shows global emission rates of SO_2 and BC in NorESM2-LM (black solid line), together with the contribution from anthropogenic activity in our chosen areas and from biomass burning globally. The sum of the contributions do not add up to the global emission rate, and this sheds light on how much of the global emission rate can be attributed to both natural and anthropogenic sources from the rest of the world, including shipping and volcanic contributions.

Appendix B inter-model differences in ice core areas

B1 Model data downloaded from ESGF

Table B2 shows an overview of the 11 models used in this study. For the historical simulations one ensemble member per model was downloaded, with the variant label r1i1p1f2 for CNRM-ESM2-1, r3i1p1f1 for MPI-ESM1-2-HAM, r1i1p3f1 for both GISS-models and r1i1p1f1 for the rest. The variables downloaded were pr, wetbc, drybc, wetso4, dryso4, emiso2, emibc, mmrso2, mmrbc, mmrso4, airmass, and areacella per model. The models EC-Earth3-AerChem, MPI-ESM1-2-HAM, CNRM-ESM2-1, and both GISS models did not have the variable "airmass" available, we then calculated airmass by using the the pressure at each vertical layer together with the `Python` package `geonum.atmosphere`.

Appendix C Lifetimes of BC and sulfate

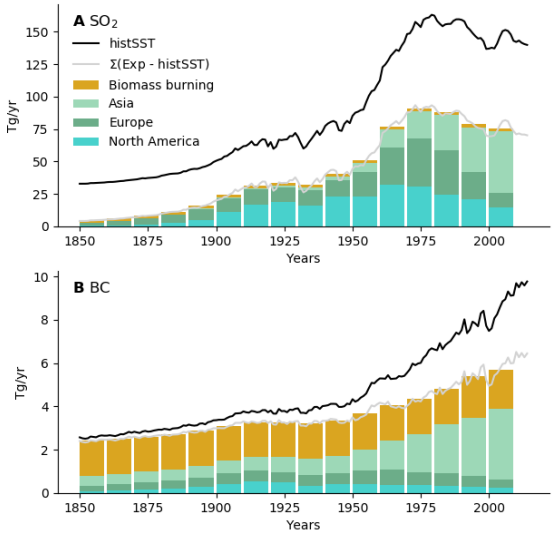

Figure A1. Emissions of SO_2 (top) and BC (bottom) as a function of time. The black line shows global emissions from the reference simulation (histSST in Table A1), the bars show decadal mean contribution per perturbation experiment together with the annual sum of error for the four contributions in grey.

Table B1. Altitudes in meters averaged over 3x3 grid matrix surrounding ice core location in the models used. Letters refer to locations as listed in Table 1. For Northern and Southern Greenland we have used NGT_B19 and Summit2010, respectively. ILI: Illimani, ECL: Eclipse, MCC: McCall Glacier, MTO: Mount Oxford, NGR: Northern Greenland, SGR: Southern Greeland, CDD: Col du Dome, CGN: Colle Gnifetti, MTE: Mt Elbrus.

Location	A	B	C	D	E	F	G	H	I
Abbreviation	ILI	ECL	MCC	MTO	NGR	SGR	CDD	CGN	MTE
Ice Core	**6300**	**3017**	**2400**	**2210**	**2270**	**3258**	**4350**	**4452**	**5115**
NorESM2-LM	1854	656	403	376	2087	1746	550	550	811
GISS-E2-1-H	2264	949	161	187	2178	2164	471	548	844
GISS-E2-1-G	2460	966	442	535	2170	2094	532	634	740
EC-Earth3-AerChem	2084	924	393	454	2206	1922	547	547	738
MPI-ESM-1-2-HAM	2068	939	339	218	2463	2237	674	674	1220
CESM2	2505	1238	584	562	2434	2338	922	922	1208
CESM2-WACCM	2508	1242	581	550	2426	2328	930	930	1211
GFDL-ESM4	3001	1267	589	586	2292	2316	1032	1032	1002
CNRM-ESM2-1	2282	1161	506	679	2424	2343	841	857	1021
MRI-ESM2-0	2007	933	387	439	2391	2382	658	658	1213
MIROC6	2300	1066	424	384	2428	2213	755	784	999

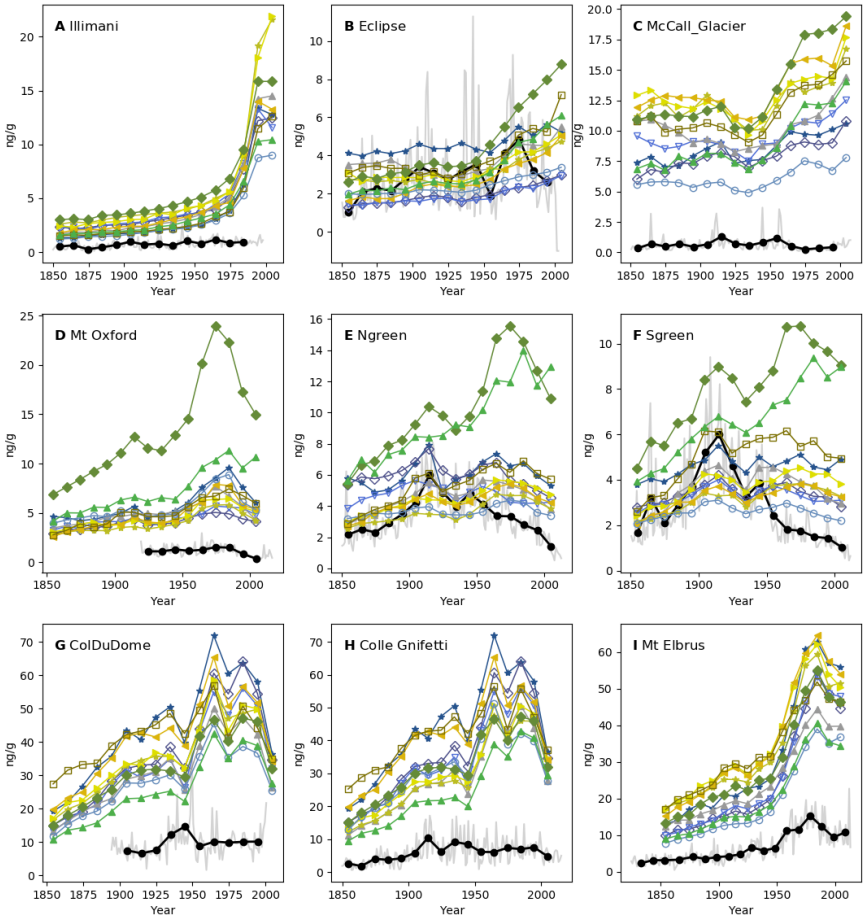

Figure B1. BC concentrations for the ice core sites shown in Figure 1 per model. The legend can be seen in Figure C2.

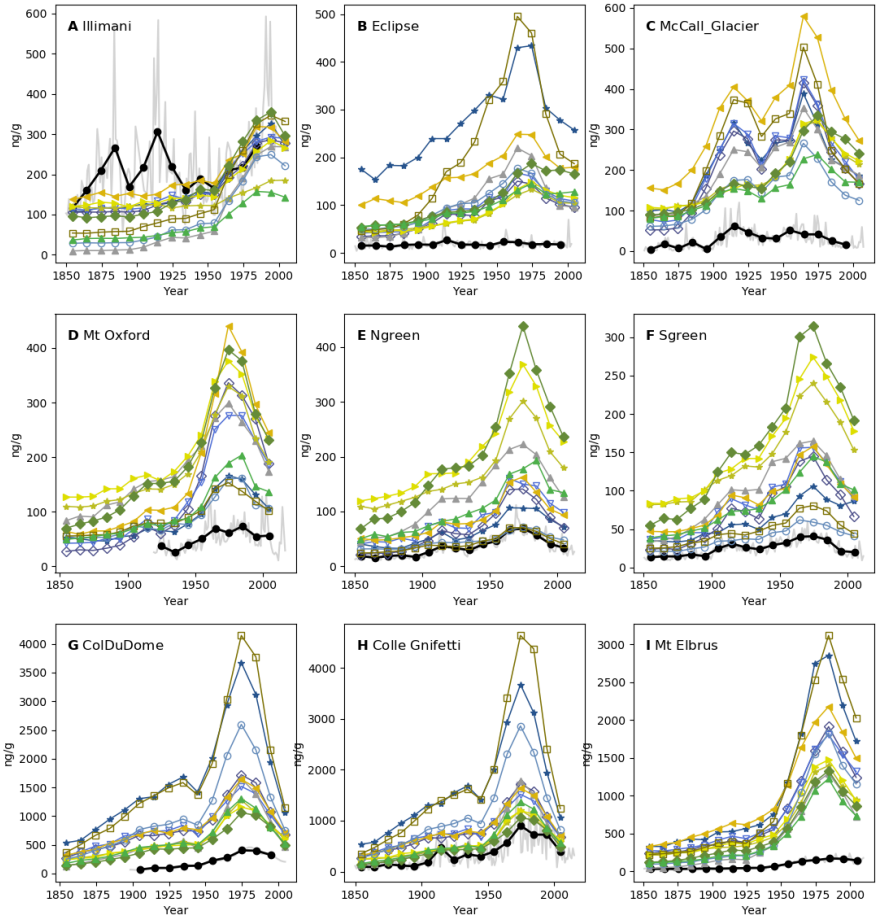

Figure B2. Sulfate concentrations for the ice core sites shown in Figure 1 per model. The legend can be seen in Figure C2.

Table B2. Earth system models used in this study

Institution	Model	Resolution	Reference
NCAR	CESM2	1.25x0.9	Danabasoglu et al. (2020)
NCAR	CESM2-WACCM	1.25x0.9	Danabasoglu et al. (2020)
CNRM-CERFACS	CNRM-ESM2-1	1.4x1.4	Séférian et al. (2019)
HAMMOZ-Consortium	MPI-ESM1-2-HAM	1.876x1.875	Tegen et al. (2019) Mauritsen et al. (2019)
EC-Earth-Consortium	EC-Earth3-AerChem	3.0x2.0	van Noije et al. (2021)
NOAA-GFDL	GFDL-ESM4	1.25x1.0	Dunne et al. (2020)
NCC	NorESM2-LM	2.5x1.875	(Seland et al., 2020)
MRI	MRI-ESM2-0	1.125x1.125	Yukimoto et al. (2019)
MIROC	MIROC6	1.4x1.4	Tatebe et al. (2019)
NASA-GISS	GISS-E2-1-G	2.5x2.0	Kelley et al. (2020)
NASA-GISS	GISS-E2-1-H	2.5x2.0	Kelley et al. (2020)

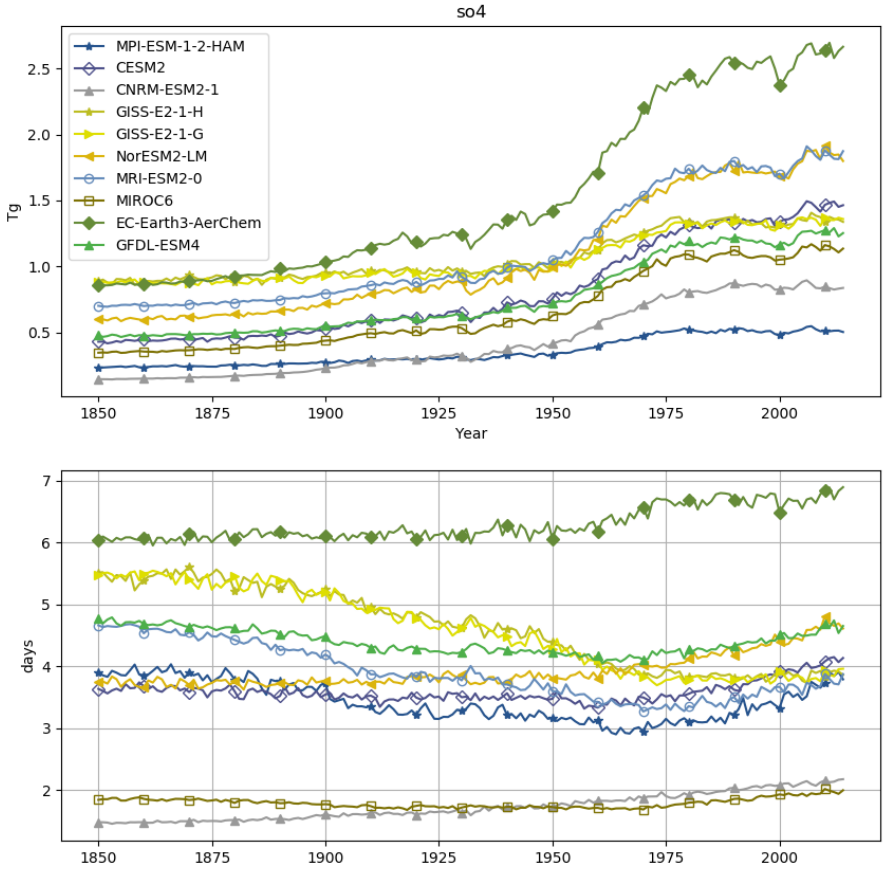

Figure C1. Global load and lifetime for sulfate. The model CESM2-WACCM has been removed from this figure as it is the only model containing volcanic emisssions of sulfur dioxide, which drastically affects loads and lifetime for sulfate in this model.

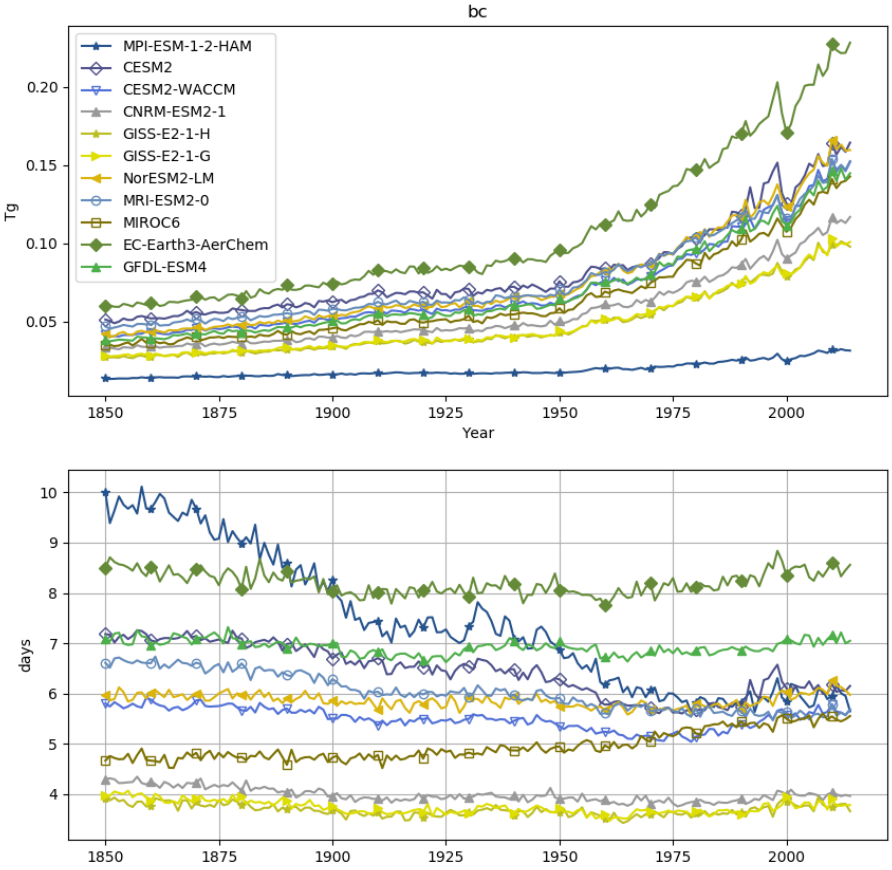

Figure C2. Global load and lifetime for BC

Acknowledgments

This study benefited greatly from the CMIP6 data infrastructure for handling and providing model data for analysis. Jan Griesfeller is thanked for data organization. High-performance computing and storage resources were provided by the Norwegian infrastructure for computational science (through projects NS2345K, NS9560K and NS9252K) and the Norwegian Meteorological Institute. A special thanks to Dr. Saehee Lim for recalculating the Mt Elbrus black carbon data into annual means for the purpose of this study.

The authors declare to have no competing interests.

Author contributions. K. O. Moseid carried out all data analysis developed the method, designed the emission experiments and wrote most of the paper. M. Schulz advised on method development and analysis, and helped write the paper. A. Eichler and M. Schwikowski provided ice core data, advised on analysis, and revised the paper. J. McConnell provided ice core data, advised on analysis, and revised the paper. D. Olivié helped design the emission experiments and performed the simulations, and revised the paper. A. S. Criscitiello, K. J. Kreutz, and M. Legrand provided previously unpublished ice core data and revised the paper.

Financial support. This research has been supported by the European Union's Horizon 2020 (FORCeS) (grant no. 821205) and the Research Council of Norway (KeyClim) (grant no. 295046).

Data Availability. CMIP6 model outputs are freely available from the World Climate Research Programme (WCRP),2011; https://esgf-node.llnl.gov/search/cmip6/ (WCRP, 2021). The previously published ice core data is available following references and DOI's presented in Table 1. The previously unpublished ice core data will be published for general access upon further review of this manuscript. The emission experiment data from NorESM2-LM will be published with a doi on zendo upon further review of this manuscript.

References

Avak, S. E., Trachsel, J. C., Edebeli, J., Brütsch, S., Bartels-Rausch, T., Schneebeli, M., ... Eichler, A. (2019). Melt-Induced Fractionation of Major Ions and Trace Elements in an Alpine Snowpack. *Journal of Geophysical Research: Earth Surface*, *124*(7), 1647–1657. Retrieved 2021-10-17, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019JF005026` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JF005026) doi: 10.1029/2019JF005026

Bauer, S. E., Bausch, A., Nazarenko, L., Tsigaridis, K., Xu, B., Edwards, R., ... McConnell, J. (2013, July). Historical and future black carbon deposition on the three ice caps: Ice core measurements and model simulations from 1850 to 2100. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7948–7961. Retrieved 2018-09-24, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50612` doi: 10.1002/jgrd.50612

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., ... Stevens, B. (2019). Bounding Global Aerosol Radiative Forcing of Climate Change. *Reviews of Geophysics*, *58*(1), e2019RG000660. Retrieved 2021-09-18, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000660` (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000660) doi: 10.1029/2019RG000660

Bender, F. A.-M. (2020). Aerosol Forcing: Still Uncertain, Still Relevant. *AGU Advances*, *1*(3), e2019AV000128. Retrieved 2021-09-12, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019AV000128` (_eprint:

https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019AV000128) doi: 10.1029/2019AV000128

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., ... Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, *118*(11), 5380–5552. Retrieved 2021-09-30, from https://onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50171 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50171) doi: 10.1002/jgrd.50171

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. Retrieved 2020-08-27, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001916 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001916) doi: 10.1029/2019MS001916

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., ... Zhao, M. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS002015. Retrieved 2021-09-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002015 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002015) doi: 10.1029/2019MS002015

Eichler, A., Schwikowski, M., & Gäggeler, H. W. (2001, January). Meltwater-induced relocation of chemical species in Alpine firn. *Tellus B: Chemical and Physical Meteorology*, *53*(2), 192–203. Retrieved 2021-10-17, from https://doi.org/10.3402/tellusb.v53i2.16575 (Publisher: Taylor & Francis _eprint: https://doi.org/10.3402/tellusb.v53i2.16575) doi: 10.3402/tellusb.v53i2.16575

Engardt, M., Simpson, D., Schwikowski, M., & Granat, L. (2017, January). Deposition of sulphur and nitrogen in Europe 1900–2050. Model calculations and comparison to historical observations. *Tellus B: Chemical and Physical Meteorology*, *69*(1), 1328945. Retrieved 2018-09-24, from https://doi.org/10.1080/16000889.2017.1328945 doi: 10.1080/16000889.2017.1328945

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016, May). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. Retrieved 2019-12-13, from https://www.geosci-model-dev.net/9/1937/2016/gmd-9-1937-2016.html doi: https://doi.org/10.5194/gmd-9-1937-2016

Fagerli, H., Legrand, M., Preunkert, S., Vestreng, V., Simpson, D., & Cerqueira, M. (2007, December). Modeling historical long-term trends of sulfate, ammonium, and elemental carbon over Europe: A comparison with ice core records in the Alps. *Journal of Geophysical Research: Atmospheres*, *112*(D23). Retrieved 2018-09-21, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008044 doi: 10.1029/2006JD008044

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., ... Zhang, X. (2021). The Earth's energy budget, climate feedbacks, and climate sensitivity. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., ... Schulz, M. (2016, October). Recommendations for diagnosing ef-

fective radiative forcing from climate models for CMIP6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12,460–12,475. Retrieved 2018-09-20, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025320 doi: 10.1002/2016JD025320

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., ... Dentener, F. (2017, January). Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats. *Atmospheric Chemistry and Physics*, *17*(2), 1543–1555. Retrieved 2021-09-21, from https://acp.copernicus.org/articles/17/1543/2017/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-17-1543-2017

Hodnebrog, O., Myhre, G., & Samset, B. H. (2014, September). How shorter black carbon lifetime alters its climate effect. *Nat Commun*, *5*(1), 5065. Retrieved 2021-09-23, from https://www.nature.com/articles/ncomms6065 (Bandiera_abtest: a Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Atmospheric science Subject_term_id: atmospheric-science) doi: 10.1038/ncomms6065

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., ... Zhang, Q. (2018, January). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*(1), 369–408. Retrieved 2019-04-25, from https://www.geosci-model-dev.net/11/369/2018/ doi: https://doi.org/10.5194/gmd-11-369-2018

Kellerhals, T., Brütsch, S., Sigl, M., Knüsel, S., Gäggeler, H. W., & Schwikowski, M. (2010). Ammonium concentration in ice cores: A new proxy for regional temperature reconstruction? *Journal of Geophysical Research: Atmospheres*, *115*(D16). Retrieved 2021-08-19, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012603 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD012603) doi: 10.1029/2009JD012603

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., ... Yao, M.-S. (2020). GISS-E2.1: Configurations and Climatology. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS002025. Retrieved 2020-08-27, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002025 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002025) doi: 10.1029/2019MS002025

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., ... Vuuren, D. P. v. (2010, August). Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, *10*(15), 7017–7039. Retrieved 2018-09-25, from https://www.atmos-chem-phys.net/10/7017/2010/acp-10-7017-2010.html doi: https://doi.org/10.5194/acp-10-7017-2010

Legrand, M., Hammer, C., De Angelis, M., Savarino, J., Delmas, R., Clausen, H., & Johnsen, S. J. (1997). Sulfur-containing species (methanesulfonate and SO4) over the last climatic cycle in the Greenland Ice Core Project (central Greenland) ice core. *Journal of Geophysical Research: Oceans*, *102*(C12), 26663–26679. Retrieved 2021-10-09, from https://onlinelibrary.wiley.com/doi/abs/10.1029/97JC01436 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/97JC01436) doi: 10.1029/97JC01436

Legrand, M., Preunkert, S., May, B., Guilhermet, J., Hoffman, H., & Wagenbach, D. (2013). Major 20th century changes of the content and chemical speciation of organic carbon archived in Alpine ice cores: Implications for the

long-term change of organic aerosol over Europe. *Journal of Geophysical Research: Atmospheres*, *118*(9), 3879–3890. Retrieved 2021-04-06, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50202 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50202) doi: https://doi.org/10.1002/jgrd.50202

Lim, S., Faïn, X., Ginot, P., Mikhalenko, V., Kutuzov, S., Paris, J.-D., ... Laj, P. (2017, March). Black carbon variability since preindustrial times in the eastern part of Europe reconstructed from Mt. Elbrus, Caucasus, ice cores. *Atmospheric Chemistry and Physics*, *17*(5), 3489–3505. Retrieved 2021-06-17, from https://acp.copernicus.org/articles/17/3489/2017/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-17-3489-2017

Liu, P., Kaplan, J. O., Mickley, L. J., Li, Y., Chellman, N. J., Arienzo, M. M., ... McConnell, J. R. (2020). Improved estimates of preindustrial biomass burning reduce the magnitude of aerosol climate forcing in the Southern Hemisphere. *Science Advances*, *7*(22), eabc1379. Retrieved 2021-10-26, from https://www.science.org/doi/10.1126/sciadv.abc1379 (Publisher: American Association for the Advancement of Science) doi: 10.1126/sciadv.abc1379

Lohmann, U. (2017, May). Why does knowledge of past aerosol forcing matter for future climate change? *Journal of Geophysical Research: Atmospheres*, *122*(9), 5021–5023. Retrieved 2018-09-24, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD026962 doi: 10.1002/2017JD026962

Lohmann, U., & Feichter, J. (2005, March). Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics*, *5*(3), 715–737. Retrieved 2021-09-22, from https://acp.copernicus.org/articles/5/715/2005/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-5-715-2005

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998–1038. Retrieved 2021-09-13, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001400 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018MS001400) doi: 10.1029/2018MS001400

McConnell, J. R. (2007, September). 20th century industrial black carbon emissions altered and climate forcing. *SCIENCE*, *317*.

McConnell, J. R. (2010, July). New Directions: Historical black carbon and other ice core aerosol records in the Arctic for GCM evaluation. *Atmospheric Environment*, *44*(21), 2665–2666. Retrieved 2021-08-19, from https://www.sciencedirect.com/science/article/pii/S1352231010002864 doi: 10.1016/j.atmosenv.2010.04.004

McConnell, J. R., Burke, A., Dunbar, N. W., Köhler, P., Thomas, J. L., Arienzo, M. M., ... Winckler, G. (2017, September). Synchronous volcanic eruptions and abrupt climate change ca17.7 ka plausibly linked by stratospheric ozone depletion. *PNAS*, *114*(38), 10035–10040. Retrieved 2021-09-12, from https://www.pnas.org/content/114/38/10035 (Publisher: National Academy of Sciences Section: Physical Sciences) doi: 10.1073/pnas.1705595114

McConnell, J. R., & Edwards, R. (2008, August). Coal burning leaves toxic heavy metal legacy in the Arctic. *PNAS*, *105*(34), 12140–12144. Retrieved 2018-12-05, from http://www.pnas.org/content/105/34/12140 doi: 10.1073/pnas.0803564105

Moseid, K. O., Schulz, M., Storelvmo, T., Julsrud, I. R., Olivié, D., Nabat, P., ... Gastineau, G. (2020, December). Bias in CMIP6 models as compared to observed regional dimming and brightening. *Atmospheric Chem-*

istry and Physics, *20*(24), 16023–16040. Retrieved 2021-04-09, from https://acp.copernicus.org/articles/20/16023/2020/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-20-16023-2020

Naik, V., Szopa, S., Adhikary, B., Artaxo Netto, P. E., Berntsen, T., Collins, W. D., ... Zanis, P. (2021). Short-lived climate forcers. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Osmont, D., Sigl, M., Eichler, A., Jenk, T. M., & Schwikowski, M. (2019, March). A Holocene black carbon ice-core record of biomass burning in the Amazon Basin from Illimani, Bolivia. *Climate of the Past*, *15*(2), 579–592. Retrieved 2021-09-12, from https://cp.copernicus.org/articles/15/579/2019/ (Publisher: Copernicus GmbH) doi: 10.5194/cp-15-579-2019

Preunkert, S., Legrand, M., Kutuzov, S., Ginot, P., Mikhalenko, V., & Friedrich, R. (2019, November). The Elbrus (Caucasus, Russia) ice core record – Part 1: reconstruction of past anthropogenic sulfur emissions in south-eastern Europe. *Atmospheric Chemistry and Physics*, *19*(22), 14119–14132. Retrieved 2021-08-20, from https://acp.copernicus.org/articles/19/14119/2019/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-19-14119-2019

Preunkert, S., Legrand, M., & Wagenbach, D. (2001, December). Sulfate trends in a Col du Dôme (French Alps) ice core: A record of anthropogenic sulfate levels in the European midtroposphere over the twentieth century. *J. Geophys. Res.*, *106*(D23), 31991–32004. Retrieved 2019-06-24, from http://doi.wiley.com/10.1029/2001JD000792 doi: 10.1029/2001JD000792

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., ... Takemura, T. (2006, November). Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmospheric Chemistry and Physics*, *6*(12), 5225–5246. Retrieved 2021-02-17, from https://acp.copernicus.org/articles/6/5225/2006/ (Publisher: Copernicus GmbH) doi: https://doi.org/10.5194/acp-6-5225-2006

Seland, O., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., ... Schulz, M. (2020, February). The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations. *Geoscientific Model Development Discussions*, 1–68. Retrieved 2020-08-27, from https://gmd.copernicus.org/preprints/gmd-2019-378/ (Publisher: Copernicus GmbH) doi: https://doi.org/10.5194/gmd-2019-378

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., ... Zelinka, M. D. (2020). An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. *Reviews of Geophysics*, *58*(4), e2019RG000678. Retrieved 2021-09-22, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000678 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000678) doi: 10.1029/2019RG000678

Sigl, M., Abram, N. J., Gabrieli, J., Jenk, T. M., Osmont, D., & Schwikowski, M. (2018, October). 19th century glacier retreat in the Alps preceded the emergence of industrial black carbon deposition on high-alpine glaciers. *The Cryosphere*, *12*(10), 3311–3331. Retrieved 2021-09-12, from https://tc.copernicus.org/articles/12/3311/2018/ (Publisher: Copernicus GmbH) doi: 10.5194/tc-12-3311-2018

Sigl, M., McConnell, J. R., Layman, L., Maselli, O., McGwire, K., Pasteris, D., ... Kipfstuhl, S. (2013). A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years. *Journal of Geophysical Research: Atmospheres*, *118*(3), 1151–1169. Retrieved 2021-08-19, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD018603 (_eprint:

https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2012JD018603) doi: 10.1029/2012JD018603

Sigl, M., McConnell, J. R., Toohey, M., Curran, M., Das, S. B., Edwards, R., ... Severi, M. (2014, August). Insights from Antarctica on volcanic forcing during the Common Era. *Nature Clim Change*, *4*(8), 693–697. Retrieved 2021-10-26, from https://www.nature.com/articles/nclimate2293 (Bandiera_abtest: a Cg_type: Nature Research Journals Number: 8 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Atmospheric chemistry;Palaeoclimate Subject_term_id: atmospheric-chemistry;palaeoclimate) doi: 10.1038/nclimate2293

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., ... Woodruff, T. E. (2015, July). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*(7562), 543–549. Retrieved 2018-12-10, from https://www.nature.com/articles/nature14565 doi: 10.1038/nature14565

Storelvmo, T., Leirvik, T., Lohmann, U., Phillips, P. C. B., & Wild, M. (2016, April). Disentangling greenhouse warming and aerosol cooling to reveal Earth's climate sensitivity. *Nature Geoscience*, *9*(4), 286–289. Retrieved 2018-09-26, from https://www.nature.com/articles/ngeo2670 doi: 10.1038/ngeo2670

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., ... Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182–4227. Retrieved 2020-08-27, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001791 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001791) doi: 10.1029/2019MS001791

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., ... Kimoto, M. (2019, July). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*(7), 2727–2765. Retrieved 2020-08-27, from https://gmd.copernicus.org/articles/12/2727/2019/ (Publisher: Copernicus GmbH) doi: https://doi.org/10.5194/gmd-12-2727-2019

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., ... Lohmann, U. (2019, April). The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 1: Aerosol evaluation. *Geoscientific Model Development*, *12*(4), 1643–1677. Retrieved 2021-10-08, from https://gmd.copernicus.org/articles/12/1643/2019/ (Publisher: Copernicus GmbH) doi: 10.5194/gmd-12-1643-2019

van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., ... van der Werf, G. R. (2017, September). Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015). *Geoscientific Model Development*, *10*(9), 3329–3357. Retrieved 2021-08-11, from https://gmd.copernicus.org/articles/10/3329/2017/ (Publisher: Copernicus GmbH) doi: 10.5194/gmd-10-3329-2017

van Noije, T., Bergman, T., Le Sager, P., O'Donnell, D., Makkonen, R., Gonçalves-Ageitos, M., ... Yang, S. (2021, September). EC-Earth3-AerChem: a global climate model with interactive aerosols and atmospheric chemistry participating in CMIP6. *Geoscientific Model Development*, *14*(9), 5637–5668. Retrieved 2021-09-13, from https://gmd.copernicus.org/articles/14/5637/2021/ (Publisher: Copernicus GmbH) doi: 10.5194/gmd-14-5637-2021

Wendl, I. A., Menking, J. A., Färber, R., Gysel, M., Kaspari, S. D., Laborde, M. J. G., & Schwikowski, M. (2014, August). Optimized method for black carbon analysis in ice and snow using the Single Particle Soot Photometer.

Atmospheric Measurement Techniques, *7*(8), 2667–2681. Retrieved 2021-09-16, from https://amt.copernicus.org/articles/7/2667/2014/ (Publisher: Copernicus GmbH) doi: 10.5194/amt-7-2667-2014

Yalcin, K., & Wake, C. P. (2001). Anthropogenic signals recorded in an ice core from Eclipse Icefield, Yukon Territory, Canada. *Geophysical Research Letters*, *28*(23), 4487–4490. Retrieved 2018-12-14, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013037 doi: 10.1029/2001GL013037

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., ... Ishii, M. (2019). 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. Ser. II*, *97*(5), 931–965. doi: 10.2151/jmsj.2019-051

Zennaro, P., Kehrwald, N., McConnell, J. R., Schüpbach, S., Maselli, O. J., Marlon, J., ... Barbante, C. (2014, October). Fire in ice: two millennia of boreal forest fire history from the Greenland NEEM ice core. *Climate of the Past*, *10*(5), 1905–1924. Retrieved 2021-08-30, from https://cp.copernicus.org/articles/10/1905/2014/ (Publisher: Copernicus GmbH) doi: 10.5194/cp-10-1905-2014

Paper III

Importance of BC lifetime for radiative effects in CMIP6 models

Kine Onsum Moseid, Michael Schulz, Trude Storelvmo
In prep, planned for Geophysical Research Letters

III