Author response

We thank the reviewers for their critical assessment of our work. In the following we address their concerns point by point. Please note that these responses have been updated and may differ from the Author Comment as was published as a reply to reviewers on ACPD on 31st of August 2020. The point by point responses to the reviewers include quotes from the new manuscript and work as a list of changes done to the manuscript as well as responses to the reviewer comments. The manuscript have undergone large literary changes throughout, in regards to the non-literary changes we have

- Added three more models to the analysis.
- Added three ensemble members per model for the *historical* experiment.
- Added a discussion surrounding the observed SDSR in China (the trend reversal) as Section 4.1.
- Added an explanation of both the direct and indirect aerosol effect, and included the indirect effect in the analysis of clear sky versus all sky radiation in Section 3.4.
- Changed the methodology of timeseries shown from a running mean to interval means, to make it easier for the reader divide the model results.
- Added Section 3.1 and Figure 1 on model variability.
- Changed Figure 3 of old manuscript into Table 2 in new manuscript for clarity.

Reviewer 1

General comments

Reviewer Point P 1.1 — Moseid et al. compare surface downwelling shortwave radiation from CMIP6 models and from ground stations. They show the discrepancy between modeled and observed SDSR is partly caused by erroneous aerosol and aerosol precursor emission inventories, thus providing important information for the evaluation of ESM. While the research topic is essential, the methodology can be improved to clarify the impacts of clouds and cloud-aerosol interaction. Instead of using all-sky SDSR, I would suggest the authors compare the sunny-day SDSR from CMIP6 and from ground stations throughout the whole text.

Reply: We agree that the manuscript should include a description of the impact of clouds and cloud-aerosol interactions. A new part was added in line 26-38 in the revised manuscript:
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). Insolation at the top of the atmosphere changes on millennial timescales when the Earth's orbital parameters change, but the solar 11-year cycle nor solar historical time variations have created multidecadal important trends in surface radiation (Eddy et al., 1982). 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 effect. The direct effect is the scatter or absorption directly caused by a dry aerosol, and the indirect effect is how aerosols change properties in clouds. These properties includes both a change in cloud lifetime and most importantly a change in cloud albedo, making the cloud appear brighter (Boucher et al., 2013).

And Section 3.4 "Clear sky and cloud cover in China" has been improved throughout, including changing Figure 3 into a table that is easier to read and analyze, and change the focus from only the aerosol direct effect to the inclusion of the indirect effect as well.
Unfortunately neither GEBA nor CMIP6 models provide sunny day SDSR. Previous studies such as Allen et al. (2013) have used the GEBA data set to create a clear sky proxy for a selection of stations to compare with the clear sky flux variable of CMIP models. However, this is beyond the scope of our study.

Reviewer Point P 1.2 — To be more accurate, I would also suggest the authors compare the SDSR conditions on the atmospheric relative humidity, which is associated with the scattering from water vapor.
Note that the clear-sky SDSR from climate models is usually used for calculating cloud radiative forcing and is not the same as sunny-day SDSR.

Reply: We are looking at longtime fluctuations in SDSR. Water vapor has not changed in a sufficient magnitude in recent decades to have an effect on decadal fluctuations in SDSR (Wang and Yang (2014), Yang et al. (2019), Wild (2009), Hoyt and Schatten (1993), Ramanathan and Vogelmann (1997), Solomon et al. (2010)). This was added in the new version of the manuscript in line 26-38 as cited above. We therefore assume in this study that the SDSR effects of water vapor scattering is negligible.

Minor comments

Reviewer Point P 1.3 — The title: I would not use the "1961-2014" in the title. It provides little information.

Reply: Removed.

Reviewer Point P 1.4 — The title: compare to -> compare with.

Reply: We changed the title to *Bias in CMIP6 models as compared to observed regional dimming and brightening*

Reviewer Point P 1.5 — The title: maybe the authors should include "aerosol", which is the theme of the paper

Reply: We changed the title to *Bias in CMIP6 models as compared to observed regional dimming and brightening*. Although we do agree that aerosols are relevant in this paper, we feel this title explains in general the findings of this paper in simple terms.

Reviewer Point P 1.6 — Figure 3: Please double check the cloud fraction and the calculation of anomaly. If the trend is reversed, it explains everything.

Reply: This is double checked and the presented Figure was correct. In the new version of the manuscript this Figure has been made into a table. (Table 2 in the revised version)

Reviewer 2

General comments

Reviewer Point P 2.1 — It would improve the paper if more background information in the introduction section was provided on the key drivers of SDSR i.e. clouds and greenhouse gases can also influence SDSR in addition to aerosol effects.

Reply: We thank the reviewer for the comment and agree more background information should be provided regarding SDSR. We have added a more detailed description of what can influence SDSR in line 26-38 in the introduction, as cited in our reply to reviewer point P1.1.

Reviewer Point P 2.2 — Throughout the paper there are numerous mentions to the fact that aerosols play a key role in the dimming signal of SDSR observed and simulated across all regions. However, the same cannot be said for the observed brightening signal in more recent years. A key question seems to be why are aerosols a key driver in the dimming but not brightening?
If the emission inventories and aerosols were in error throughout the whole period of study then surely the models would not be able to simulate the temporal evolution of both phenomenon across all regions?

Reply: We respond to this point in two parts - first the role of aerosols in brightening:
We would like to point the reviewer to the studies by Allen et al. (2013), Chiacchio et al. (2015) and Wild (2012), which show that indeed aerosols are a key driver to the observed brightening in recent years. The reduction of anthropogenic aerosol emission leads to brightening. We would also like to thank the reviewer for mentioning this point that we did not explicitly make in the original manuscript. The following text has now been added in line 46-49:
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.

Then the final question on emission inventories: Correct, this is why we are proposing errors in emission inventories as a possible reason for discrepancies in the regions where the models are not able to simulate the temporal evolution of dimming (and brightening).

Reviewer Point P 2.3 — The paper states that the CMIP6 models are able to represent the observed SDSR signal over Europe relatively well. However, I think there are a few interesting discrepancies which should be discussed further. Prior to 1980 the observations do not show much of a dimming signal (in fact the observed anomaly is slightly positive at times) but the CMIP6 models do show a consistent dimming signal. Is there a specific reason for the absence of a dimming in the observations, when we know there were large concentrations of aerosols over Europe at this time? Contrary to what was mentioned in point 2 above Europe is the only region where there is a simulated brightening signal in both the model and observations, implying that models are able to reproduce brightening signal over certain regions. It would be good to know if there a reason for this over Europe and does it occur over other regions like for example North America.

Reply: The referee is right that Figure 1b in the old manuscript does show some interesting discrepancies in the beginning of the time period in the study that was not mentioned in the original manuscript. The observational data set used in this study starts in 1961, and the anomaly shown in the figure is made as a difference from the the mean value of SDSR from 1961-1966. Since the European dimming started before 1960 (Wild, 2012) the "true" European SDSR anomaly might not be achieved using this data set, as is also seen by the weak European dimming in Storelvmo et al. (2018) using the same data set. In the new version of the manuscript we have added in line 197-200:
The dimming in Europe is believed to have started before the start time (1961) of the observational data set used here (Wild, 2009), which partly explains why the dimming in Figure 2 (b) 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 to changes in cloud cover is found in Section A1 the Appendix.
The observed and simulated brightening in Europe are quite comparable and we therefore propose that the emission inventories of aerosols in Europe are estimated well.
North America has not been shown in Figure 1, but is included here in the reply as supplementary Figure S1. See Leibensperger et al. (2012b) and Leibensperger et al. (2012a) for a closer look at the climatic effects in North America due to anthropogenic aerosol emissions. We chose Europe and Asia as areas of focus to give the readers a clean impression of one example region where the models perform well and one example region where they do not perform well.

Reviewer Point P 2.4 — For the analysis over China the paper suggests that the error between the models and observations of SDSR are due to the errors in emission inventories that translate into errors in the calculation of atmospheric burden of aerosols. (1)Are we certain that the errors in the emission inventories are that large to account for the discrepancy in model and observed SDSR? Is there an estimate of the uncertainty for the CMIP6 emission inventory and how does CMIP6 compare to other global and regional emission inventories? (2) Can these differences explain some of the inconsistencies of models with observations? I am not convinced that the observed trend reversal in SDSR over China in 1990 can be explained by errors in the emission inventories alone. (3) Are we anticipating a slowing down of SO2 emissions in China from the 1990s onwards? As far as I understand, anthropogenic emissions of aerosols and their precursors (particularly SO2) have largely been increasing over China up until 2010 when air pollutant control measures were then implemented to reduce emissions. Therefore, if aerosols were driving the temporal change in SDSR over China a dimming signal should have been observed up until this point, but it isn't. This is present in the observed and simulated change in SDSR over India but not China. (4) How do this discrepancy match with the conclusions of the paper and what else could be driving the SDSR trend over China throughout this period? I think this needs to be explored further in the paper as the assumed underlying trend in emissions (and therefore aerosols) and SDSR do not seem to match over China and from what I can tell it cannot be reconciled by errors in the emission inventories alone.

Reply: To make it easier for the reader we have marked up numbers to the questions in the reviewer point. (1) We are not certain that the errors in the emission inventories are large enough to account for the discrepancy in model and observed SDSR alone, but we suggest that this error plays an important role. Unfortunately there is no estimate of uncertainty in the CMIP6 emission inventories, but this is planned to be included in the next generation of CMIP emission inventories (see Hoesly et al. (2018)). Due to the lacking estimation of uncertainty we do not have evidence to say that errors in emission inventories are too small to cause a discrepancy between model and observation.
(2) The CMIP6 emission inventory is made using CEDS that makes datasets based on EDGAR, as is described in more detail in Hoesly et al. (2018). There are probably some differences between the CMIP6 data set and other regional emission data sets, but this study does not look further into finding such differences. We propose at least some of the discrepancy between model and observed SDSR is caused by errors in emission inventories, but we do not have enough evidence to claim that all discrepancies are emission driven. We recognize that the original manuscript may have given the wrong impression to the reader that errors in emission inventories *alone* cause all discrepancies, and this has now been addressed and made clearer throughout the text.
(3) During the review process we found more information regarding the observed trend reversal in the GEBA data in China. According to the CMIP6 data set of sulfate emission we do not expect a slow down of emitted SO_2 from 1990, but rather from around 2005. The observed trend reversal in SDSR does therefore not fit with CMIP6 emitted sulfate. However, previous studies have found that the trend reversal in SDSR is to a considerable extent caused by the fact that the measurement devices at the Chinese radiation network stations were replaced with new ones between 1991 and 1993, which caused a spurious upward jump in the records (Wang and Wild (2016), Wang and Yang (2014), Yang et al. (2019)). With this new information we have added a Section 4.1 "The trend reversal in China" under Section 4 "Discussion" where we compare our results to that of Yang et al. (2018) where the "jump" has been removed by homogenization. The main point from this section can be summarized as in line 359-360:
Models do not accurately represent the strength of dimming, or the evolutionary pattern of SDSR observed in China with or without the early 1990s brightening (the "jump").
(4) The conclusions of the paper propose that errors in anthropogenic aerosol emission inventories play a role in the discrepancy seen between simulated and observed SDSR. Even if the trend reversal in the observed SDSR in China was to be an artifact, the models would still largely underestimate the magnitude of dimming. With regards to the trend reversal, the assumed underlying trend of increasing sulfate emission until 2010 as proposed by the reviewer (and CMIP6) is being questioned in this manuscript, as even though Wang and Wild (2016) suggests most of the "jump" is an artifact, they still estimate that 20% is real. We thank the reviewer for this comment and hope the new Section 4.1 is satisfactory.

Reviewer Point P 2.5 — Only limited discussion within the paper is provided on clouds and aerosol-cloud interactions, which needs to be improved throughout the paper. Within section 3.3 a link is made between cloud cover change and SDSR but how much of an influence do clouds have on all-sky SDSR? How reliable are the observed and simulated cloud cover changes and can some uncertainty bounds be placed on them? Is a regional cloud cover change of 1-2% significant in terms of SDSR? In figure 3 the temporal change in observed cloud cover is similar to that in observed SDSR so even if clouds can't explain the magnitude and all of the observed change then surely they must be exerting some influence on SDSR? Is it possible to compare a clear-sky derived observed SDSR to that from model simulations to eliminate any influence of clouds on the signal?

Reply: We thank the reviewer for this comment, and would refer to our reply to reviewer 1 point 1 (P1.1) where we cite lines from our new manuscript regarding clouds' role in all sky SDSR. In the appendix of the new version of the manuscript (Section A3) we have added an idealized estimation for how much a theoretical 1 % cloud cover increase in China would affect SDSR. Line 435-436:
... 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.
Previous studies have found that the link between cloud cover and SDSR trends depends on what region you are looking into. In Europe cloud cover has most of an effect on SDSR on the shorter time scales, and the dimming and following brightening observed in Europe is dominantly caused by changes in anthropogenic aerosol emission and thereby the aerosol absorption and scattering (Norris and Wild, 2007). In China cloud cover made a negligible contribution to all sky SDSR trend in GEBA until 1989. After 1980 the heavily discussed trend reversal is observed in China, and Norris and Wild (2009) suggests half of the observed brightening between 1990 and 2002 is caused by a reduction in cloud cover. Please note that this paper was published before the proposal of the trend reversal being an artifact of a change in instrumentation (Wang and Wild, 2016). This complicates the story and is the reason Norris and Wild (2009) is not discussed in the new version of our manuscript.
In North America cloud cover is found to have played an important role in the observed brightening (Long et al., 2018). Other studies have made clear-sky derived observed SDSR (Norris and Wild (2007), Norris and Wild (2009)) when assessing the cloud signal for Europe and China (mentioned above in this reply), but this goes beyond the scope of our study.

Reviewer Point P 2.6 — The previous comparison of CMIP5 models to observed SDSR by Storelvmo et al., (2018) is mentioned throughout this study, with similar results presented here for CMIP6 models. A key question is therefore why has there been no improvement in simulating observed SDSR between CMIP5 and CMIP6 models? This is despite some changes to individual aerosol schemes within models and also different historical aerosol precursor emission datasets being used. Some discussion is needed on what is continually missing from the models and what are the model developments to focus on to improve the future simulation of SDSR.

Reply: To answer this question we must first find out whether the source of the error is within the model's codes or within the emission inventories, or a combination. Storelvmo et al. (2018) argues that the discrepancy between observed and modelled SDSR may be attributed to errors in the treatment of processes that translate aerosol emissions into clear-sky and all-sky radiative forcings. Here, we show that simulated SDSR develops similarly in time, but opposite in sign, to simulated atmospheric burden of SO2. By doing this we narrow down the potential source of error by suggesting that the atmospheric burden in the models are at fault, and that the processes translating burden into clear-sky and all-sky radiative forcings are behaving as expected. The final answer of what is at fault is still not found, but we suggest to have found a piece of the puzzle in the emission inventories.
It is important to note that Storelvmo et al. (2018) included all CMIP5 models, and we "only" include eight models.
We thank the reviewer for this comment and have updated the end of the conclusion in the new manuscript. Lines 415-420 is added:
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 an important part in earth system models' ability to simulate the past for the right reasons.
Since the SDSR measurements are not only sensitive to aerosol effects, 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 sulphur, organic carbon and nitrate in ice cores, as well as daily temperature range records.

Reviewer Point P 2.7 — Further details are required, either in Table 1 or a new table, on each of the CMIP6 models used in this study. In particular, it would be useful to know horizontal resolution and some information on the individual chemistry and aerosol schemes in each model. This could provide useful information to the reader of the potential causes of discrepancies between models. In addition, it would be good to have a record somewhere of the actual output used from the ESGF (e.g. temporal period, variant ID, CMIP table ID etc). Furthermore, if there is data now available for additional CMIP6 models then it would be useful to include it, as long as it further informs the current study.

Reply: We thank the reviewer for this comment and we have added Table A2 and Section A2 in the Appendix of the paper listing information such as variant ID, variables, references to model documentation, horizontal resolution and aerosol scheme. More data has been published since the first submission of this paper, and we have therefore decided to include more models in this study. Three models have been added (GISS-E2-1-G, IPSL-CM6A-LR, MRI-ESM2-0) to the analysis as more data was released.

Reviewer Point P 2.8 — The methods section (2.3) appears to lack important details of what model data is being used (see point 7) and how the gridded model data has actually been compared to the observations which are at point locations. In calculating the regional means at observation locations, do the number of sites used change over time period and does this have any impact on the results? Furthermore, in the results section the clear-sky SDSR is discussed but is not mentioned in the methods section. I also think that it is important to use multiple ensemble for meaning purposes when using coupled experiments members from models so that the internal variability in each model can be shown (this would give a range of variability important to show on some of the Figures for certain variables).

Reply: We thank the reviewer for this comment and agree that the methods section was indeed lacking both clarity and details. In Section 2.3 "Methods" in the new manuscript we have added line 142-144:
All model output and CRU results have been co-located to GEBA station locations using the nearest neighbour 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 neighbour interpolation, as every station has been weighted equally.
We tackle the question regarding number of sites used in time in both Section 2.1 " Observations" with line 80-83 added in the new manuscript:
This allows for all 1487 stations to have data on each time step, so that all regions have a complete record and the same amount of stations throughout the entire time period in question.
And in Section 2.3 "Methods" in line 139-140:
The number of stations per region remains constant throughout the time period.
In the new version of the manuscript we have added three ensemble members per model for the historical simulation. Both inter-annual variability and inter-ensemble variability is shown in Figure 1 of the new manuscript, that is presented in the new section in results called Section 3.1 " Model variability". We have changed Figure 1 in the old manuscript into Figure 2 of the new manuscript, where we present ensemble means per model, and show shading for the standard deviations of the total 24 ensemble members.

Reviewer Point P 2.9 — A General comment on the figures is that they could be improved to make them easier to read by using better colours (I found the light green very bright), tick marks on the axis and line types that are easier to distinguish between different model experiments. Also, if it is possible to include a measure of observational and model uncertainty on any of the figures then this would improve the comparisons. When values from figures are continually referred to in the text it would help the reader if there is reference table containing some of the key numbers included (like the supplementary table).

Reply: We thank the reviewer and have chosen a different color chart for the figures, more tick marks, and different line types to better differ the graphs. Variability and uncertainty is shown in the new Figure 1 and Figure 2 as explained in the previous reply (P2.8). We are currently not referring to specific values until Section 3.4 "Clear sky SDSR and cloud cover in China" where we have changed the previous Figure 3 into a table to make the point more clear and the discussion easier to follow.

Minor comments

Reviewer Point P 2.10 — Page 1, line 9 – Reword this sentence as mentioning SO2 emissions, which are not aerosols, and then other aerosols relevant to SDSR. Be more precise in this statement.

Reply: Changed to line 11: *The emissions of SO_2 used in the models show no pattern that could explain the observed SDSR evolution over China.* as we are mostly looking at sulfate throughout the paper.

Reviewer Point P 2.11 — Page 1, line 13 – Can you say how much error is associated with aerosols and emission inventories that might contribute to error in SDSR?

Reply: Unfortunately the emission inventory data set for CMIP6 does not have estimates of uncertainty, which is why we chose the word "partly" in line 13 as we have no evidence telling us how much of the discrepancy can be attributed to emission estimates.

Reviewer Point P 2.12 — Page 2, Line 30 – Is this statement true across all regions? What about for Europe?

Reply: This statement is only true globally based on previous studies. Added the word global in line 49: *Previous studies show that historical simulations from ESMs do not reproduce the global transient development of SDSR as observed (Storelvmo et al. (2018), Wild (2009)).*

Reviewer Point P 2.13 — Page 2, line 35 – For the introduction it would be good to include a bit more detail on what the GEBA observations on their own show before introducing any comparisons to models.

Reply: We thank the reviewer for pointing this out, as the introduction to global dimming mentioned several citations that all used GEBA to identify dimming (and regional brightening), which was not explicitly mentioned. This has now been clarified in the text in line 55-56: *In this study we use gap-filled data based the GEBA dataset. The GEBA dataset is the observational dataset as used in the citations in the previous paragraph, together with several recent CMIP6 historical model experiments from eight ESMs to investigate the aerosol effect in the time period 1961-2014, globally and regionally.*

Reviewer Point P 2.14 — Page 2, line 46 – here the study says that two observational datasets are used but only one has been mentioned in the previous paragraph. Please include details of what is the second dataset used in this study.

Reply: The second observational data set has been added in line 59-60: *We also use observational cloud cover data to briefly assess the role of cloud cover in the historical development of SDSR.*

Reviewer Point P 2.15 — Page 2, line 47 – please reword sentence "An explanation of the methods used to obtain and analyse the data complete Section 2."

Reply: Changed to line 69: *The methods used to obtain and analyse the data finalize Section 2.*

Reviewer Point P 2.16 — Page 3, line 57 – it would be good to include the error in the observations on all figures to show the uncertainty in the observations.

Reply: Unfortunately sources of error in observation differs from station to station and we only have a general estimation of error from the instruments used. In addition to the instrumental error presented in line 76 we have chosen to include a light grey line with the highly variable yearly observational data in the background of Figure 2 in the new version of the manuscript.

Reviewer Point P 2.17 — Page 3, line 60 – Please clarify if this temporal gap filling technique allows for all 1487 stations to have a complete record of observations over the entire 1961-2014 and how this technique impacts the observations. If the number of stations used changes over the entire time period then it could be important for the analysis.

Reply: This has been clarified and is cited in reply to P2.8.

Reviewer Point P 2.18 — Page 3, line 74 – insert 'is' between "these the"

Reply: Fixed.

Reviewer Point P 2.19 — Page 4, line 93 – replace 'stales' with "stalls"

Reply: Fixed.

Reviewer Point P 2.20 — Page 4, line 94-95 – "So these experiments will show to what extent the removal of cloud cover change from global warming has an effect on SDSR." – I am sure that this is the case as there will be still be variability in the cloud fields simulated by climate models in these experiments. In addition, as the aerosol fields are changing in these experiments, they will also impact the simulated clouds in the models. Therefore, to make this statement further evidence would be required from each model that the cloud fields are being properly constrained to isolate their impacts on SDSR.

Reply: We are not stating that all cloud cover change is removed, only the cloud cover change that is induced by global warming - as global warming essentially is removed in these experiments. Cloud properties will change with the aerosol fields in the models, so this experiment has not removed all cloud changes - only the cloud cover changes induced by global warming. Changed the wording to line 119-121: *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.*

Reviewer Point P 2.21 — Page 4, line 107 – It would be good to show on a figure the spatial distribution of the GEBA observations within each defined region.

Reply: Storelvmo et al. (2018)s Figure 1 is an excellent figure showing the spatial distribution of the stations used in both this and her study in addition to the trends of the stations in colours. I have added a reference to that figure in line 136.

Reviewer Point P 2.22 — Page 4, line 110-112 – Please clarify exactly how anomalies have been calculated. Are anomalies calculated for each individual observation site within a region first before then calculating a regional mean value?

Reply: Clarified. New line 146-150: *When a result is shown as an anomaly, as opposed to an absolute value, the general formula has been to subtract the mean of the first five years of the investigated time period (1961-2014) from the timeseries in question. To clarify - first an average value per year per region is calculated, and then a new mean is created from the first five years of this timeseries. This 5-year-mean is then subtracted from each year in the timeseries for the region in question and presented as an anomaly.*

Reviewer Point P 2.23 — Page 4, line 112 - Supplementary table number is not shown

Reply: Fixed.

Reviewer Point P 2.24 — Page 4, line 113 – Provide more information on exactly what model data has been obtained from the ESGF (perhaps in a separate table) e.g. CMIP table ID, variant label etc. (see general comment 8)

Reply: We thank the reviewer for this request and a table has been added as cited in the reply to P2.7

Reviewer Point P 2.25 — Page 4, line 115 – I think it would be more prudent to use more ensemble members for coupled experiments and with this an idea of the internal variability for each model could be obtained for variables such as cloud cover and SDSR.

Reply: Three ensemble member have been used in the historical experiment, see reply to P2.8 for citation.

Reviewer Point P 2.26 — Page 4, line 116 – It is not clear if the 10-year running mean is used for the model data, observation data or both?

Reply: Running means have been exchanged for 6-year-intervals means in most figures in the new manuscript. the only exception if Figure 4 which shows SO_4 burdens form models as a 10-year running mean, while the observation is shown as yearly data. This is clarified in the figure caption.

Reviewer Point P 2.27 — Page 5, line 121 – it is hard to see from Figure 1 a) as to whether the global SDSR representation in the models is similar to the observations at all. There is clearly a difference in magnitude but there does not appear to be a strong dimming signal in many of the models. Is this just the scale on the figure or is there not much change in the model at all? Can the Figure be improved in any way to make this easier to see?

Reply: Figure 1a) corresponds to Figure 2a) in the new manuscript. The models generally do not represent the *global* change in SDSR as observed. We have included gray shading for the ensemble standard deviations and changed the method from a running mean to 6-year-interval-means to show clearly the weak signal in the models in Figure 2 of the new manuscript.

Reviewer Point P 2.28 — Page 5, line 122 – Change "these discrepancy originate" to "this discrepancy originates"

Reply: Fixed.

Reviewer Point P 2.29 — Page 5, line 125 – More discussion on European model observational differences (see general comments point 3)

Reply: This discussion has been added and is cited in reply to P2.3.

Reviewer Point P 2.30 — Page 5, line 135 – I think that this is only true for certain models as others seem to have opposite temporal changes compared to observations e.g. NorESM2.

Reply: We agree and this line has been removed.

Reviewer Point P 2.31 — Page 5, line 138 – It is hard to say without tick marks on the figures as to whether the end points in models are similar to the observations. For example, is a -10 Wm-2 anomaly in 2014 from GEBA considered to be similar to a -6 Wm-2 from NorESM2?

Reply: We agree that this statement was questionable in the old manuscript. By adding more models to the analysis the remark of similar end points between model and observations became blatantly wrong and we have removed all statements regarding this.

Reviewer Point P 2.32 — Page 5, line 140 – please explain what "temporal forcing evolution" means in this context.

Reply: This line has been removed due to the added discussion of the trend reversal in China in observations.

Reviewer Point P 2.33 — Page 6, line 156-157 – does this imply that the greenhouse gases impact on SDSR over China throughout this period?

Reply: When adding more models to the study this implication became untrue, and the statement has been removed in the new manuscript.

Reviewer Point P 2.34 — Page 6, line 157-158 – I am not sure this is true for all models. The temporal evolution of SDSR from CanESM5 seems quite different in the historical and piClim-histall but perhaps not so much in MIROC6.

Reply: With new models added to the study the entire RFMIP paragraph has been updated to line 239-244: *Recall that the experiments of RFMIP utilize pre-industial SST's, meaning essentially there is no global warming in these experiments. In the RFMIP experiments shown in Figure 3(c) 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 (stipled lines/piClim-histaer). This implies that greenhouse gases without their global warming effect do not affect all sky SDSR in a significant way over China throughout the period.*

Reviewer Point P 2.35 — Page 6, line 167 – Aerosols have a key role in dimming but not it appears brightening – why not? (see general comment 2)

Reply: We thank the reviewer for this comment and refer til our reply in P2.2

Reviewer Point P 2.36 — Page 6, lines 168-169 – similar to point above in that there are differences between these simulations which don't appear to be the temporal driver of SDSR but perhaps can influence it? It would be good to show the actual differences between models in these simulations and what influence other factors (like clouds and greenhouse gases) can have on SDSR.

Reply: Clouds and greenhouse gases can influence SDSR, but are, as mentioned in the introduction, not a dominant driver of multidecadal SDSR changes. It is therefore expected to see small differences between these simulations. The overall picture of models showing dimming with anthropogenic aerosol emissions, and no dimming without it remains whether or not you include greenhouse gases or SST changes. This has been clarified in line 246-250: *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 containing pre-industrial aerosols only. This points to anthropogenic aerosol emissions playing a key role in global 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 in most models.*

Reviewer Point P 2.37 — Page 6, line 173 – how has all-sky SDSR been decomposed into clear-sky?

Reply: This is a diagnostic that is output from the models. The general idea is that clear-sky SDSR from models represents the amount of sunlight reaching the surface if all shortwave effects from clouds were removed. Clear-sky SDSR is not to be confused with sunny day SDSR which is from actual cloud free days.

Reviewer Point P 2.38 — Page 6, line 179-180 – Can the clear-sky and all-sky changes be shown on the same figure to compare differences?

Reply: This figure has been replaced by Table 2 in the new manuscript showing changes in cloud cover, all sky SDSR and clear sky SDSR for three different time periods.

Reviewer Point P 2.39 — Page 6, line 182-189 –How have the changes in model cloud cover been calculated? This needs to be in the methods section. Also line 183-184 states that cloud cover changes mask the clear-sky SDSR signal. This implies that the clear-sky decrease would have been even larger without changes to clouds indicating that clouds do have an important influence on SDSR in models. I think this needs to be explained more - see general comment section 5 for more details.

Reply: Cloud cover is a standard output from climate models and has not been calculated by the authors, and the source of the data has been added in Section A2 the Appendix of the new manuscript. The effect of clouds in SDSR has been added and is cited in our reply to P2.5.

Reviewer Point P 2.40 — Page 7, line 193 – "session" should be "section"

Reply: Fixed.

Reviewer Point P 2.41 — Page 7, line 194 – "In this session we found the clear-sky SDSR to be stronger than all-sky SDSR, indicating the simulated dimming is primarily caused by aerosol-radiation interactions." But also that clouds have an influence on SDSR too.

Reply: This sentence has been removed in the renewal of the manuscript. But in general clouds have an influence on SDSR which is clarified and cited in our reply to P2.5. but our findings are that aerosols effects are the dominating cause of dimming (that includes the aerosol indirect effect which is a change in cloud radiative properties).

Reviewer Point P 2.42 — Page 7, line 205 – "SO2 burden" is mentioned but should this not be SO4 burden.

Reply: Fixed.

Reviewer Point P 2.43 — Page 7, line 205-206 – Given that all models have the same SO2 emissions, do we know why the changes in SO4 burden are so different between NorESM2 and CESM2? Could this indicate some of the potential problems in translating emissions into atmospheric burden or aerosols, which lead to errors in SDSR?

Reply: Burdens are a result of emission, aerosol formation, transport and deposition. The emissions in both models are the same but the other processes dependent on many different parameterisations within each model. The atmospheric circulation in CESM2 and NoreSM2 differs, among other things, so for example a sulfate particle may be brought higher up in the atmosphere in NorESM2 - giving sulfate a longer lifetime and thereby making NorESM2 have a higher sulfate burden. In addition to this these burdens are calculated using co-location to point locations as described in the methods section, and this is where transportation plays a role.

Reviewer Point P 2.44 — Page 7, line 210 – can a more scientific term be used than "real story".

Reply: Definitely. Added line 314: *Assuming GEBA data provide a reasonable representation of the historical development of SDSR and implicitly sulphur 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.*

Reviewer Point P 2.45 — Page 7, line 210-211 – This sentence makes the assumption that aerosols are the sole driving force in SDSR and that it is only the emissions and removal processes that could be in error. Other potential causes could be mentioned like the model translation of emissions to burden which leads to the larger differences in simulated SO4 burdens between models. Also see major comments above.

Reply: As cited in the previous answer we have added "aerosol formation, transport" in this sentence. The model burden is translated from emission, transportation and removal processes. We are assuming the model translation from emission to burden is behaving as expected in the two regions of special interest, and an explanation for this is found in the Dicussion in line 372-383:
Here, we can see an anti-correlation between simulated SO_4 *burdens from Figure 4 (a) and (b), and simulated SDSR from Figure 2 (b) 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 2(b)), along with simulated* SO_4 *burden anti-correlating relatively well with observed SDSR in Europe (Fig 4 (a)). This means that the model code translating burdens into SDSR in Europe can simulate changes in SDSR as a consequence of changed 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; 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 correct 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.* If there is an error in burden than the error is sourced in either emissions, transportation or removal processes.

Reviewer Point P 2.46 — Page 7, line 212 – "the precursor of SO2", should this not be SO4?

Reply: Fixed.

Reviewer Point P 2.47 — Page 7, line 215-218 – Should we be expecting a trend reversal in SO2 emissions over China between 1980 and 1990? At this point in time emissions would have been increasing over China and emissions have only begun to reduce recently (since 2010). See general comment point 4

Reply: A section discussing the trend reversal was added and is cited in our reply to P2.4.

Reviewer Point P 2.48 — Page 8, line 235 – Is it possible to include the clear-sky proxy from GEBA here and compare to that from models on Figure 3 to show how well models simulate the aerosol radiation interactions?

Reply: Unfortunately that is not easily done and is beyond the scope of this study. There is currently an NSF project working on creating clear sky proxies at ETH Zurich lead by Dr Martin Wild.

Reviewer Point P 2.49 — Page 8, line 238 – change "shown in Figure displayed" to "(Fig. 2) show"

Reply: Fixed.

Reviewer Point P 2.50 — Page 8, line 242 – But the magnitude of the dimming was not sufficient to reproduce that observed (same as Allen?) and implies emissions are not high enough historically?

Reply: Correct.

Reviewer Point P 2.51 — Page 8, line 247 - change “burden of SO2” to “burden of SO4”.

Reply: Fixed.

Reviewer Point P 2.52 — Page 8, Lines 246-249 - The study only shows change in SDSR is opposite to SO4 burden over Europe and not the case over China so can we really say that the process of translating burden to forcings are ok? What about over other regions? Might not just be due to errors in atmospheric burdens, but other factors combining?

Reply: A new section answering this question has been added in the Dicussion and has been cited in our reply to P2.45.

Reviewer Point P 2.53 — Page 8, Line 250 – “The models of this study ...” changed to “The models used in this study ...”

Reply: Fixed.

Reviewer Point P 2.54 — Page 9, line 254-255 – Should we expect a reversal of emissions across China over this period?

Reply: A section discussing the trend reversal was added and is cited in our reply to P2.4.

Reviewer Point P 2.55 — Page 9, line 256 – Is this referring to Figure 3 in Hoesly et al., (2018)? Make clearer.

Reply: Fixed.

Reviewer Point P 2.56 — Page 9, line 258 – should we expect BC and OC to influences SDSR much? Need to mention these aerosol components earlier in the manuscript if going to mention now as no introduction to them at all. All discussion previously has been made about SO4 so why suddenly bring them in now?

Reply: We agree that this is confusing. This sentence has been removed from the manuscript.

Reviewer Point P 2.57 — Page 9, line 259-261 – Do these studies give an uncertainty in emission inventories and can this be used to see if it can account for the differences between model and observed SDSR.

Reply: Unfortunately none of these studies presents number for uncertainty in emission inventories, but Aas et al. (2019) show annual average trend in sulfate in aerosols from 2000-2015 and found that the standard deviation was larger than the actual trend for East Asia, and none of the locations used in that study was located in China Aas et al. (2019)[Tab. 1].

Reviewer Point P 2.58 — Page 9, line 270 – change “CMIP6 experiment models” to “experiments, CMIP6 models”

Reply: Fixed.

Reviewer Point P 2.59 — Page 9, line 273 – mention that the dimming is underestimated by the models.

Reply: Fixed.

Reviewer Point P 2.60 — Page 9, line 276-279 – Would we not have anticipated the SO4 burden to have increased across China over this period as SO2 emissions are anticipated to have also increased? Are the errors in SO4 burden and SO2 emissions really that large to account for the observed discrepancy in SDSR? More work to back up this statement and other factors should be included in conclusions. Uncertainty in emission inventories probably do contribute to this but the trend changes in SDSR and anticipated emission changes don’t match for China, so this cannot be the sole reason and needs to be expanded on. see general comment point 4.

Reply: As we do not know the estimation uncertainty for emission we do not have evidence to rule out that the emission inventories can have large errors. The observed trend reversal in China have a new discussion section which is cited in our reply to P2.4.

Reviewer Point P 2.61 — Page 10, line 285-287 – how would these future investigations improve our understanding of SDSR temporal evolution?

Reply: A comparison between different observational datasets such as GEBA and ice cores will give a unique insight in aerosol presence before the satellite era, especially if the emission inventories are wrong. Satellite products can be used to compare to CRU cloud cover data and give a full picture. Model experiments with different aerosol emission as input can disentangle the role of different aerosols on SDSR assuming their translation from emission to SDSR is correct. Have added in the manuscript "Since the SDSR measurements are not only sensitive to aerosol effects, further studies could include... "

Reviewer Point P 2.62 — Fig. 2b – why is CanESMS so different in Hist-Nat and does show that other drivers influence the SDSR trend?

Reply: We do not know why CanESM5 differs from the others in it’s hist-nat experiment, but this single experiment is unfortunately not enough evidence to say that other drivers influence the SDSR trend - except for in CanESM5 only.

Reviewer Point P 2.63 — Fig. 3b – Can the uncertainty in cloud cover from observations and models be shown?

Reply: The section explaining the CRU dataset has been improved in line 84-89:
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^o latitude/longitude resolution grid covering continental areas. Information on interpolation methods and procedures used to create the gridded data set are 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.
As cloud cover is a secondary observed variable we have added line 289 in the Results section regarding clear sky and cloud cover data in China:
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.
Uncertainty in cloud cover from models is hard to quantify with only three ensemble members, but Table A1 shows the different baseline values for cloud cover in each model, which can be seen with a spread of 50%-64%.

Reviewer Point P 2.64 — Fig. 4 – CESM2 seems to show a small change, can you confirm that this model has interact aerosols included? If not then why such a small change compared to others?

Reply: Aerosols interact with the climate in CESM2. We have added a Table in the appendix (Table A2) showing which models that have interactive aerosols and which that don't.

References

Aas, Wenche, Augustin Mortier, Van Bowersox, Ribu Cherian, Greg Faluvegi, Hilde Fagerli, Jenny Hand, Zbigniew Klimont, Corinne Galy-Lacaux, Christopher M. B. Lehmann, Cathrine Lund Myhre, Gunnar Myhre, Dirk Olivié, Keiichi Sato, Johannes Quaas, P. S. P. Rao, Michael Schulz, Drew Shindell, Ragnhild B. Skeie, Ariel Stein, Toshihiko Takemura, Svetlana Tsyro, Robert Vet, and Xiaobin Xu (2019), "Global and regional trends of atmospheric sulfur." *Scientific Reports*, 9, 953, URL `https://www.nature.com/articles/s41598-018-37304-0`.

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, 6311–6336, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50426`.

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens, and X.Y. Zhang (2013), *Clouds and Aerosols*, book section 7, 571–658. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, URL `www.climatechange2013.org`.

Chiacchio, Marc, Fabien Solmon, Filippo Giorgi, Paul Stackhouse, and Martin Wild (2015), "Evaluation of the radiation budget with a regional climate model over Europe and inspection of dimming and brightening." *Journal of Geophysical Research: Atmospheres*, 120, 1951–1971, URL `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014JD022497`.

Eddy, John A., Ronald L. Gilliland, and Douglas V. Hoyt (1982), "Changes in the solar constant and climatic effects." *Nature*, 300, 689–693, URL `https://www.nature.com/articles/300689a0`. Number: 5894 Publisher: Nature Publishing Group.

Harris, Ian, Timothy J. Osborn, Phil Jones, and David Lister (2020), "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset." *Scientific Data*, 7, 109, URL `https://www.nature.com/articles/s41597-020-0453-3`. Number: 1 Publisher: Nature Publishing Group.

Hoesly, Rachel M., Steven J. Smith, Leyang Feng, Zbigniew Klimont, Greet Janssens-Maenhout, Tyler Pitkanen, Jonathan J. Seibert, Linh Vu, Robert J. Andres, Ryan M. Bolt, Tami C. Bond, Laura Dawidowski, Nazar Kholod, June-ichi Kurokawa, Meng Li, Liang Liu, Zifeng Lu, Maria Cecilia P. Moura, Patrick R. O'Rourke, and Qiang Zhang (2018), "Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS)." *Geoscientific Model Development*, 11, 369–408, URL `https://www.geosci-model-dev.net/11/369/2018/`.

Hoyt, Douglas V. and Kenneth H. Schatten (1993), "A discussion of plausible solar irradiance variations, 1700-1992." *Journal of Geophysical Research: Space Physics*, 98, 18895–18906, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/93JA01944`. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/93JA01944.

Kaiser, Dale P. and Yun Qian (2002), "Decreasing trends in sunshine duration over China for 1954–1998: Indication of increased haze pollution?" *Geophysical Research Letters*, 29, 38–1–38–4, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GL016057`.

Leibensperger, E. M., L. J. Mickley, D. J. Jacob, W.-T. Chen, J. H. Seinfeld, A. Nenes, P. J. Adams, D. G. Streets, N. Kumar, and D. Rind (2012a), "Climatic effects of 1950–2050 changes in US anthropogenic aerosols – Part 1: Aerosol trends and radiative forcing." *Atmospheric Chemistry and Physics*, 12, 3333–3348, URL `https://www.atmos-chem-phys.net/12/3333/2012/`. Publisher: Copernicus GmbH.

Leibensperger, E. M., L. J. Mickley, D. J. Jacob, W.-T. Chen, J. H. Seinfeld, A. Nenes, P. J. Adams, D. G. Streets, N. Kumar, and D. Rind (2012b), "Climatic effects of 1950–2050 changes in US anthropogenic aerosols – Part 2: Climate response." *Atmospheric Chemistry and Physics*, 12, 3349–3362, URL `https://www.atmos-chem-phys.net/12/3349/2012/acp-12-3349-2012.html`. Publisher: Copernicus GmbH.

Long, C. N., E. G. Dutton, J. A. Augustine, W. Wiscombe, M. Wild, S. A. McFarlane, and C. J. Flynn (2018), "Significant decadal brightening of downwelling shortwave in the continental United States." *Journal of Geophysical Research: Atmospheres*, URL `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008JD011263%4010.1002/%28ISSN%29216` Publisher: John Wiley & Sons, Ltd.

Nabat, Pierre, Samuel Somot, Marc Mallet, Arturo Sanchez-Lorenzo, and Martin Wild (2014), "Contribution of anthropogenic sulfate aerosols to the changing Euro-Mediterranean climate since 1980." *Geophysical Research Letters*, 41, 5605–5611, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL060798`.

Norris, Joel R. and Martin Wild (2007), "Trends in aerosol radiative effects over Europe inferred from observed cloud cover, solar "dimming," and solar "brightening"." *Journal of Geophysical Research: Atmospheres*, 112, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007794`. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD007794.

Norris, Joel R. and Martin Wild (2009), "Trends in aerosol radiative effects over China and Japan inferred from observed cloud cover, solar "dimming," and solar "brightening"." *Journal of Geophysical Research: Atmospheres*, 114, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011378`. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD011378.

Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. Hartmann (1989), "Cloud-radiative forcing and climate: results from the Earth radiation budget experiment." *Science*, 243, 57–63.

Ramanathan, V. and Andrew M. Vogelmann (1997), "Greenhouse Effect, Atmospheric Solar Absorption and the Earth's Radiation Budget: From the Arrhenius-Langley Era to the 1990s." *Ambio*, 26, 38–46, URL `https://www.jstor.org/stable/4314548`. Publisher: [Springer, Royal Swedish Academy of Sciences].

Romanou, A., B. Liepert, G. A. Schmidt, W. B. Rossow, R. A. Ruedy, and Y. Zhang (2007), "20th century changes in surface solar irradiance in simulations and observations." *Geophysical Research Letters*, 34, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL028356`. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006GL028356.

Solomon, Susan, Karen H. Rosenlof, Robert W. Portmann, John S. Daniel, Sean M. Davis, Todd J. Sanford, and Gian-Kasper Plattner (2010), "Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming." *Science*, 327, 1219–1223, URL `https://science.sciencemag.org/content/327/5970/1219`. Publisher: American Association for the Advancement of Science Section: Research Article.

Storelvmo, Trude, Ulla K. Heede, Thomas Leirvik, Peter C. B. Phillips, Philipp Arndt, and Martin Wild (2018), "Lethargic Response to Aerosol Emissions in Current Climate Models." *Geophysical Research Letters*, 0, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078298`.

Wang, Y. W. and Y. H. Yang (2014), "China's dimming and brightening: evidence, causes and hydrological implications." *Ann. Geophys.*, 32, 41–55, URL `https://www.ann-geophys.net/32/41/2014/`.

Wang, Yawen and Martin Wild (2016), "A new look at solar dimming and brightening in China: CHINA DIMMING AND BRIGHTENING REVISITED." *Geophys. Res. Lett.*, 43, 11,777–11,785, URL `http://doi.wiley.com/10.1002/2016GL071009`.

Wild, Martin (2009), "Global dimming and brightening: A review." *Journal of Geophysical Research*, 114, URL `http://doi.wiley.com/10.1029/2008JD011470`.

Wild, Martin (2012), “Enlightening Global Dimming and Brightening.” *Bull. Amer. Meteor. Soc.*, 93, 27–37, URL `https://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00074.1`.

Wild, Martin (2016), “Decadal changes in radiative fluxes at land and ocean surfaces and their relevance for global warming.” *Wiley Interdisciplinary Reviews: Climate Change*, 7, 91–107, URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.372`.

Wild, Martin, Hans Gilgen, Andreas Roesch, Atsumu Ohmura, Charles N. Long, Ellsworth G. Dutton, Bruce Forgan, Ain Kallis, Viivi Russak, and Anatoly Tsvetkov (2005), “From Dimming to Brightening: Decadal Changes in Solar Radiation at Earth’s Surface.” *Science*, 308, 847–850, URL `https://science.sciencemag.org/content/308/5723/847`. Publisher: American Association for the Advancement of Science Section: Report.

Wild, Martin, Maria Z. Hakuba, Doris Folini, Patricia Dörig-Ott, Christoph Schär, Seiji Kato, and Charles N. Long (2018), “The cloud-free global energy balance and inferred cloud radiative effects: an assessment based on direct observations and climate models.” *Clim Dyn*, URL `https://doi.org/10.1007/s00382-018-4413-y`.

Yang, Su, Xiaolan L. Wang, and Martin Wild (2018), “Homogenization and Trend Analysis of the 1958–2016 In Situ Surface Solar Radiation Records in China.” *J. Climate*, 31, 4529–4541, URL `http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0891.1`.

Yang, Su, Xiaolan L. Wang, and Martin Wild (2019), “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, URL `http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0666.1`.

Yunfeng, Luo, Lu Daren, Zhou Xiuji, Li Weiliang, and He Qing (2001), “Characteristics of the spatial distribution and yearly variation of aerosol optical depth over China in last 30 years.” *Journal of Geophysical Research: Atmospheres*, 106, 14501–14513, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD900030`.

1 Supplementary Material

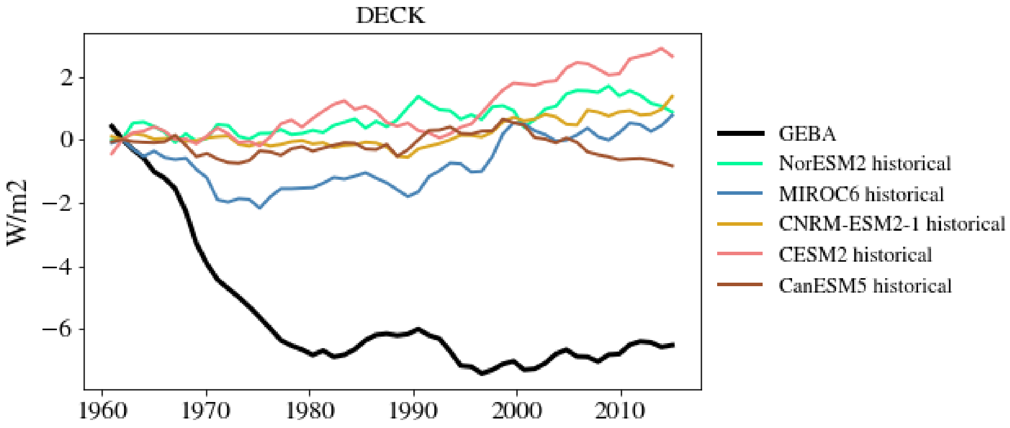

Figure S1: SDSR anomaly North America, model results are co-located to GEBA stations following the methodology as described in Moseid et al 2020 in prep

Bias in CMIP6 models as compared to observed regional dimming and brightening~~trends (1961-2014)~~

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)

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, ~~and previous~~ 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 ~~five~~ 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. ~~The models do not reproduce the observed trend reversal in SDSR in China in the late 1980s, which is attributed to a change in the emission~~ We suggest that this may be attributed to missing emissions of sulfur dioxide in ~~this region~~China, sulfur dioxide being a precursor to sulfate, which is a highly reflective aerosol and responsible for more reflective clouds. The emissions of ~~SO_2 show no sign of a trend reversal~~ sulfur dioxide used in the models do not show a temporal pattern that could explain ~~the~~ observed SDSR evolution over China~~, and neither do other aerosols relevant to SDSR~~. The results from various aerosol emission perturbation experiments from DAMIP, RFMIP and AerChemMIP ~~suggest that its likely, that aerosol effects are responsible for the dimmingsignal, although not its full amplitude. Simulated~~ 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 largely, suggesting that cloud cover changes ~~in the different models are not correlated with observed changes over China~~are not a dominant cause to 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 interpreting 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 ~~used~~ evaluated based on their ability to reproduce the climate evolution of the past 165 years, and ~~sparse~~ the sparsity of aerosol-related observations in the pre-satellite era ~~play~~ 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 ESMs future climate projections. ~~Surface~~

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). Insolation at the top of the atmosphere changes on millennial timescales when the Earth's orbital parameters change, but the solar 11-year cycle nor solar historical time variations have created multidecadal important trends in surface radiation (Eddy et al., 1982). 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 effect. The direct effect is the scatter or absorption directly caused by a dry aerosol, and the indirect effect is how aerosols change properties in clouds. These properties 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~~, and the~~. The Global Energy Balance Archive (GEBA) dataset contains measurements of SDSR as far back as in 1922 (Wild et al., 2017)~~.~~ ~~As such, it~~, and as such represents a unique and valuable ~~data set~~ 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, ~~as~~ which is expected if the cause of dimming ~~was in fact regionally varying~~ 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, ~~called~~ 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 ~~as observed~~ (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.

~~Here we use~~ In this study we use gap-filled data based on the GEBA dataset, together with several ~~very~~ recent CMIP6 historical model experiments from ~~five ESMs~~ 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 ~~historical~~ 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~~for the right reason~~.

The paper is structured as follows: In Section 2 we begin by presenting the two observational datasets used, followed by a detailed description of the experiments simulated by the ~~five~~ eight models chosen to be part of this study. ~~An explanation of the~~ The methods used to obtain and analyse the data ~~complete~~ finalize Section 2. The results are presented in Section 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 Section 5.

2 Data and Methods

2.1 Observations

The Global Energy Balance Archive (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 data set has been complemented by a machine learning technique (random forests (Breiman, 2001)) as explained in Storelvmo et al. (2018) to cover ~~temporal gaps~~ 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 amount of stations throughout the entire time period in question.

Monthly mean cloud cover data is ~~taken from the Climate Research Unit Time Series~~ provided by the Climatic Research Unit (University of East Anglia) and NCAS, and we are using version 4.02 of this dataset (CRU)~~, which~~. CRU covers the period 1901-2017 ~~(Harris et al., 2014). CRU~~ (Harris et al., 2020) and consists of a climatology made from measurements at meteorological stations around the globe, interpolated to a 0.5^o latitude/longitude resolution grid covering continental areas. Information on interpolation methods and procedures used to create the gridded data set are 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

~~Five ESMs~~ Eight climate models (NorESM2, CanESM5, MIROC6, CESM2~~and~~, 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 ~~on~~ in our diagnostic variables. An overview of models and experiments ~~covering the proposed CMIP6 reference and perturbation studies~~ 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 ~~this is the one experiment for which~~ 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. ~~The experiment tracing~~ One experiment traces exclusively the impact of ~~exclusively the~~ 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 ~~perturbation~~ perturbations due to the evolution of the natural forcing, e.g. from stratospheric aerosols from volcanoes 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 ~~earth system~~ climate model attributes responses over the historical period to the different climate ~~forcers~~forcings.

While DAMIP provides a good framework for one of the main questions in CMIP6, namely how the Earth system responds to forcing, ~~the RFMIPintercomparison~~ RFMIP, the Radiative Forcing Model Intercomparison Project, focuses on understanding the forcing itself. ~~The Radiative Forcing Model Intercomparison Project (RFMIP)~~ 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 ~~fixed~~ prescribed to pre-industrial values. One experiment includes both anthropogenic and natural aerosol emissions (~~piClim-histall~~*piClim-histall*) while the other only includes anthropogenic ~~emissions (piClim-histaer~~aerosol emissions (*piClim-histaer*). When sea surface temperatures are kept to pre-industrial values, the global surface temperature development ~~stales, and one can say we have~~ stalls, and the simulation will keep to first order a pre-industrial climate. Sea surface temperatures ~~can also~~ changes would have an effect on cloud cover, which in turn can affect SDSR. ~~So these experiments will show to what extent the removal of cloud cover change from global warming has an effect on SDSR . In addition, these RFMIP experiments are therefore useful to investigate how, or if, aerosol effects are dependent on~~ 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 ~~effect~~ specific effect of aerosols and other near-term climate forcers (NTCF) ~~can have~~ on climate. NTCFs include methane, tropospheric ozone, aerosols and their precursors (Collins et al., 2017). Three experiments have been selected from AerChemMIP, ~~two of~~ *histSST*, with all forcing agents included, and two perturbations which have pre-industrial aerosols emissions: (*hist-piAer*) and (~~hist-piAer) and~~ *hist-piNTCF*). The hist-piNTCF experiment has in addition pre-industrial ~~NTCFs (hist-piNTCF), respectively, while the last experiment has prescribed~~ NTCF levels for ozone and methane. A difference in these two simulations would only appear if ozone or methane 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 ~~(histSST), with all forcing agents included~~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 ~~see wether~~ differentiate between historical changes in aerosol and tropospheric ozone, or ~~wether~~ 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.

2.3 Methods

The GEBA stations have been divided into regions based on the country and continent~~each GEBA station is registered to~~. The number of stations in a region is presented together with the first results in the caption of Figure 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 Figure 1 in Storelvmo et al. (2018).

All model output and CRU results have been co-located to GEBA station locations using the nearest neighbour 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 neighbour 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. ~~Every station has been weighted equally.~~ 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 five years of the investigated time period (1961-2014)~~from the timeseries in question. These~~, from the timeseries in question. To clarify - first an average value per year per region is calculated, and then a new mean is created from the first five years of this timeseries. This 5-year-mean is then subtracted from each year in the timeseries 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 timeseries over 54 years(1961-2014) in nine equal intervals and average these intervals together. When the atmospheric burdens of SO_4 is shown together with observed SDSR from GEBA the timeseries 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 ~~supplementary Table ??~~the Appendix in Table A1.

The model data has been retrieved from The Earth System Grid Federation (ESGF) (Cinquini et al., 2014). ESGF is a data ~~mangement~~ 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 ~~per experiment~~from the rest of experiments shown, as not every experiment had ~~the option of providing~~ requested more than one simulation. ~~Since we are working with values that are highly variable a centered running mean of 10 years has been used as a smoothing technique.~~ Table A2 in the Appendix shows the resolutions, aerosol schemes, and aerosol complexity of the models in this study, and Section A3 explains the variables and variant labels downloaded.

3 Results

3.1 Model variability

Figure 1 shows 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 Figure 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 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 is showing the most similar SDSR evolution compared to the observed data according to Figure 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 ~~five~~ eight models is presented together with GEBA data in Figure 2. Panel (a) of this figure corresponds to the results shown in Figure 1. Each model graph in Figure 2 represents the ensemble mean of the model in question averaged over 6 years, based on three ensemble members. GEBA data is shown in black, also as 6-year averages, but with the yearly time series shown in grey in the background. Model simulations show ~~similar patterns~~ small changes of global SDSR ~~to observations, but are remarkably different in magnitude~~compared to observations (fig 2a). Global SDSR is observed to decrease over the 1487 stations until late 1980's 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 W/m^2 between models and observations. To further identify from where ~~these discrepancy originate~~this discrepancy originates, we consider ~~the~~ some geographical regions separately. Asia and Europe are relevant regions in regards to anthropogenic aerosol emissions (as explained in Section 1) and thereby also relevant to global dimming and brightening. The historical SDSR evolution in Europe and Asia are presented in Figure 2 (b) and (c), respectively. European SDSR is relatively well represented by the model simulations~~, while the~~. The yearly GEBA timeseries has values within the shaded area, that is showing 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 the 1961 (Wild, 2009), which partly explains why the initial European dimming in Figure 2 (b) 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 to changes in cloud cover is found in Section A1 in the Appendix.

There is generally a large discrepancy between model simulations and observations of SDSR in Asia, as seen in Figure 2(c). The ground stations in Asia show a noticeable trend reversal in SDSR ~~around the early~~ in the transition from 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 ~~evolutions~~ evolution for each of these countries are shown in Figure 2 (d), (e), and (f), respectively.

Figure 2 (d) 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 ~~very~~ similar to that of China~~, and Japan is~~ . Japan is downwind of the Asian continent and thus believed to be ~~heavily~~ influenced by aerosol emissions from China~~from 1980 and onwards~~. Model simulations do not capture the magnitude of dimming in Japan~~but similar SDSR temporal tendencies can be identified in both observations and simulations~~, or the apparent brightening in the 1990s. The timing of minimum SDSR occurs differently in between models, which was also seen in Figure 2 (a).

Observations from China (Figure 2 (f)) show a trend reversal in SDSR similar to the one identified in Figure 2 (c) for Asia as a whole~~. In general the historical model simulations have similar end points as the observations in China and~~ , with the minimum value found in 1989. We note that China consists of 119 GEBA stations while Asia as a whole ~~. However, their temporal evolution does not show the observed trend reversal around the late 1980s~~ 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, ~~but rather a continuous negative trend throughout the period. This in turn suggests that the temporal forcing evolution of the last half century in the ESMs is not consistent with observations for Asia~~meaning none of them show a similar trend reversal to the one from the observational data set. This trend reversal is a source of discussion within the field, and a thorough assessment, relevant to the conclusions from this study, is found in Section 4.1.

3.3 Dimming and brightening over China in various CMIP6 experiments

~~The CMIP6 framework consists of many simulations that can help investigate dimming and brightening (as explained in Section 2.2).~~ 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 ~~Tabel~~ Table 1. Figure 5 (a) shows ~~the historical simulations~~ perturbed historical simulations as performed in DAMIP together with observations of SDSR~~as previously seen in Figure 2 (f). Figure 5 (b), (c), and (d)shows the SDSR from experiments in DAMIP, RFMIP and AerChemMIP, respectively. Out of the three experiments in DAMIP, only one of them contains the evolution of~~ . 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*~~, and this experiment clearly diverges from the other DAMIP experiments over time~~). 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 start- and end-points comparable to the observations~~with continuous dimming throughout the period, but also still without the trend reversal seen in the observed temporal evolution of SDSR. SDSR in the experiments *hist-nat* and *hist-GHG* do not show signs of dimming or brightening over the investigated period ~~.~~

~~In the RFMIP experiments, where both piClim-histaer and piClim-histall contain anthropogenic aerosol emissions, all simulations show a continuous dimming throughout the period, but like in the historical simulations there is no apparent trend reversal in the late 1980s. Two of the models (NorESM2 and CanESM5) exhibit a more negative SDSR when letting evolving aerosols impact the radiation alone, without GHGs. By comparing the historical with the piClim-histall experiments, one can also note that the choice of the coupling and sea surface temperatures do not seem to affect SDSR largely~~in China, which confirms that water vapor or stratospheric aerosols are not dominant for simulated dimming.

Out of the three experiments from AerChemMIP only ~~histSST~~ *histSST* has prescribed sea surface temperatures, and contains anthropogenic aerosol emissions. This is ~~clear from how histSST~~ consistent with the time evolution of SDSR in *histSST* as the simulations diverge from the other ~~simulation~~ simulations as time progresses ~~shown in Figure 5 (d).~~ (Fig 5b). Keeping in mind that *histSST* also has anthropogenic 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-ESM2-0 presents the strongest dimming in both DAMIP and AerChemMIP. The simulations with pre-industrial aerosols (~~hist-piAer~~*hist-piAer*) and pre-industial near term climate forcers, including aerosols and ozone (~~hist-piNTCF~~*hist-piNTCF*) show very small or negligible changes in the SDSR over the time period considered.

Recall that the experiments of RFMIP utilize pre-industial SST's, meaning essentially there is no global warming in these experiments. In the RFMIP experiments shown in Figure 5(c) 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 (stipled lines/ *piClim-histaer*). This implies that greenhouse gases without their global warming effect do not affect all sky SDSR in a significant way over China throughout the period.

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 ~~containing pre-industrial aerosols only~~using aerosols maintained at constant pre industrial levels. This points to anthropogenic aerosol emissions playing a key role in ~~global~~ 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 trend reversal or distinct flattening of the dimming is identified in any of the simulations in which dimming is identified, and therefore none of the model simulations show a temporal evolution of SDSR close to the one seen in observations over China.

~~All-sky~~ All sky SDSR changes can be further decomposed ~~into a clear-sky~~ by the models into a clear sky contribution as well as a contribution from changes in cloud cover ~~and~~/or other cloud properties. In the next section we present the decomposed contributions to ~~all-sky~~ all sky SDSR in China to further understand the discrepancy seen in Figure 5.

3.4 Clear sky SDSR and cloud cover in China

~~Clear-sky SDSRover China for the historical CMIP6 simulation is shown together with all-sky SDSR over China from GEBA in Figure **??** (a). If the simulated dimming is primarily caused by aerosol-radiation interactions, the dimming is stronger in the clear-sky SDSR for all models compared to the all-sky SDSR . This is exactly what we see in Figure **??** (a). All modelsand observation show a change in behaviour in the late 1990s until 2010, where models show a steepening of their dimming trend while the observations go from a brightening trend to a SDSR stabilisation. This can be related to the cloud cover change presented in Figure **??** (b), where all models except for MIROC6 show a decrease in cloud cover over the same period. A decrease in cloud cover would entail a brightening ,~~ 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 data sets of this study. Between the years 1961 and ~~will therefore act as a mask for the steep decrease in clear sky SDSR. The simulated cloud cover changes are presented together with~~ 1989 GEBA shows a strong negative change in all sky SDSR in Figure 2 (f). 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 Figure 2(f)), 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 do 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 mainly changes the radiative properties of clouds by making them appear brighter, not by changing cloud cover. Therefore, a weak change in cloud cover followed by a strong change in all sky and clear sky SDSR point to both direct and indirect aerosol effect being the primary cause of SDSR change.

In the second time period GEBA shows a positive change (which will be further discussed in Section 4.1), and CRU shows a cloud cover ~~observations from CRU in Figure **??** (b).The transient change in~~ change of +3.0 %. Intuitively, an increase in cloud cover ~~presented by CRU are, if anything, opposite of what they would have to be to explain the observed All-sky SDSR~~ 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 both cloud cover changes, 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 effect 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 period. 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 acts as a brightening mask for clear sky dimming nor is 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 is found in Section A2 in the Appendix, indicating that such an increase could result in a dimming of 1-3 Wm^{-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.

~~The pronounced trend reversal in observed all sky SDSR in the late 1980s in China is neither identified in all sky SDSR, clear sky SDSR, nor cloud cover in any of the model simulations.~~ In section 3.3 ~~,~~ we showed that ~~a~~ dimming was only apparent in simulations that included anthropogenic aerosol emissions. In this ~~session~~ section we found the clear-sky SDSR to be close in value or even stronger than all-sky SDSR, indicating the simulated dimming is primarily caused by ~~aerosol-radiation interactions~~aerosol effects. Table 2 underlines previous findings, dimming in models are 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 ~~actual~~ *presence* of an aerosol is ~~of course what scatters~~ the cause for scattered shortwave radiation, and the ~~emissions~~ emission of its precursor is only an indirect indicator of ~~this presence.~~ 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 ~~the simulated change~~ here also the simulated anomalies in burden of SO_4 in the various models over Europe, a location where dimming and brightening ~~was~~ is fairly well represented in simulations, and over China, where dimming and brightening ~~was~~ is poorly represented in simulations (Figure 4 (a) and (b) respectively). ~~As expected if~~ The sulfate burdens are co-located to GEBA station locations in the respective regions. As expected, sulfate aerosols have ~~in fact played~~ an important role in European dimming and brightening, ~~the simulated burden~~ as the simulated burdens of SO_4 ~~shows~~ show a strikingly similar pattern (but with opposite sign) as the observed SDSR over Europe for all models. The maximum ~~burden is~~ 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. ~~NorESM2~~ MRI-ESM2-0 is the model with the largest changes, and ~~CESM2~~ GISS-E2-1-G is the model with the smallest changes in SO~~$_2$~~$_4$ burden. The same is observed over China, where ~~NorESM2 has double the~~ MRI-ESM2-0 has an SO_4 burden at the end of the time period ~~than the next model~~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 ~~to~~ the simulated SO_4 burden timeseries 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 Figure 4 (b). All the simulated historical SO_4 burdens increase until 2010, showing no signs of ~~the trend reversal identified in the GEBA data.~~ either a trend reversal or a flattening of aerosol induced dimming. Assuming GEBA data ~~shows the real story~~provide a reasonable representation of the historical development of SDSR and implicitly sulfur burdens in China, the problem in SO_4 burden must come from either the emissions~~or in~~ , aerosol formation, transport or the removal processes of SO_4.

~~Figure A2 shows emitted sulfur dioxide over China, the precursor of SO_2, for four of the models in this study. Emission sources are not expected to be at the same locations as GEBA stations, so the results shown Figure A2 is for a defined area as stated in the figure caption. Recall we are looking for signs of the observed trend reversal present in GEBA in Figure 4 (b) around the late 1980s. Figure A2 displays no trend reversal in SO_2 emissions between 1980 and 1990. The simulated burden~~ It appears, however, that the simulated burdens of SO_4 co-located to GEBA stations in China ~~presents a similar behaviour as the~~ follow quite closely the time series of emitted SO_2 ~~over China. Therefore~~ in the climate models over China (shown in Appendix Figure 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 correlate with burden which again anti correlates with SDSR changes, the temporal development of SDSR seen in GEBA cannot be ~~expected~~ explained from the current emission inventories, given sulfate aerosols play an important ~~part~~ role in SDSR in China.

~~Aerosol emissions in China and Asia as a whole has increased greatly over the last century. This includes more than just sulfur dioxide. Especially black carbon (BC) and organic carbon (OC) has had a strong increase in emissions in China. In the next section we will consider these aerosols and their potential influence on SDSR, together with a general discussion on our results.~~

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 ~~of the past~~ is an important reason for the large spread in climate projections for the future. ~~GEBA~~ Here, we investigate the effect of aerosols in GEBA which provides valuable observations of historical shortwave radiation at the surface~~that are of great value for model evaluation~~.

We have shown that a subset of models participating in CMIP6 do 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 ~~have~~ has not improved from CMIP5 to CMIP6. We find that ~~while most models have similar change in SDSR as observationsin the most recent years,~~ most models show an underestimation of changes in SDSR 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 magnitude of modelled clear-sky SDSR during the dimming period compared to a clear-sky proxy based on GEBA data.

~~China stands out as a region of interest as the observed SDSR shows a trend reversal in the mid 1980s that is not reproduced in the *historical* simulation by any of the models of this study.The RFMIP experiments shown in Figure 5displayed that~~ The simulated SDSR on decadal timescales over China does not differ significantly when comparing the RFMIP experiments (Figure 5) to the historical experiment. RFMIP experiments have pre industrial sea surface temperatures~~did not noticeably affect SDSR on decadal timescales over~~ , 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 to cloud cover, not aerosol effects. Table 2 showed inconclusive connections between modelled 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 a trend reversal in the late 1980s. This trend reversal is the topic of discussion in the next section.

4.1 The trend reversal in China

Several studies have tried to explain the trend reversal 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 reversal (Tang et al., 2011) , while recent studies propose the trend reversal is to a considerable extent an artifact of a nation wide change in SDSR measurements (Wang and Wild (2016) Yang et al. (2018)). The change in SDSR measurements include a replacement of SDSR instrumentation in addition to an update in the classification of SDSR stations, and Wang and Wild (2016) conclude that the upward trend ("jump" between 1990-1999) 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) Figure 10). The newly homogenized data exhibit a significant dimming trend (-6.13 +- 0.47 W/m^2/decade) between 1958-1990, a flattening of the curve in 1991-2005, followed by a brightening trend (6.13 +- 1.77 W/m^2/decade) between the years 2005-2016. We can use Figure 2(f) to compare our model data to these homogenized data, and see that even without a larger "jump" in the data there are still large discrepancies between model and observation, both in the form and magnitude of the brightening period after 1990. All models show dimming in the flattening period of the new homogenized data. All models apart from CanESM5 show an averaged negative trend between the 6-year-means of 2003-2008 and 2009-2014, where the homogenized data show a brightening. Models do not accurately represent the strength of dimming, or the time evolution of SDSR observed in China with or without the early 1990s brightening.

4.2 Aerosol effect on dimming

Out of all the experiments presented in Table 1 and Figure 5, only those containing anthropogenic aerosol emissions showed dimming~~in China~~. 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) argues that the discrepancy seen between observed and modelled CMIP5 model mean global SDSR can be attained to errors in the treatment of processes that translate aerosol emissions into clear-sky and all-sky radiative forcings. Here, we ~~show that simulated SDSR develops similarly in time, but opposite in sign, to simulated atmospheric burden of SO_2. By doing this we narrow down the potential source of error by suggesting that the atmospheric burden in the models are at fault, and that the processes translating burden into clear-sky and all-sky radiative forcings are behaving as expected.Atmospheric burdens are a result of emissions, gas-to-particle conversion, and wet-removal. The models of this study do a fairly good job in representing SDSR~~ can see an anti-correlation between simulated SO_4 burdens from Figure 4 (a) and (b), and simulated SDSR from Figure 2 (b) 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 2(b)), along with simulated SO_4 burden anti-correlating relatively well with observed SDSR in Europe (Fig A2 (a)). This means that the model code translating burdens into SDSR in Europe can simulate changes in SDSR as a consequence of changed in aerosol emissions. If models translate burden into SDSR correctly in Europe, ~~so we assume both emissions and subsequent processes are well represented here. The temporal development of SDSR is represented poorly in Asia, and specifically in China . Following the above logic this discrepancy could be rooted in errors in emissions or removal processes. The modeled emissions of SO_2 over China showed no trace of the trend reversal in observed SDSR between 1980 and 1990. Assuming sulfate burden is responsible for the observed trend reversal, we argue that errors in emissions inventories in China could be part of the problem.~~ 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; 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 correct 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) (Figure 3~~corresponding to Hoesly et al. (2018).This figure also shows emission inventories of black carbon and organic carbon in China, and a closer look shows that neither of these aerosol emissions show tendencies matching a trend reversal in observed SDSR between 1980 and 1990.~~).

Hoesly et al. (2018) have pointed to the need to study in the future 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 ~~was~~ were largest in Asia, even ~~though~~ if their study only went back to 1990.

The modeled emissions of SO_2 as shown in Figure A2 over China showed no trace of either the trend reversal("jump") in our observed SDSR timeseries nor patterns similar to the proposed new homogenized data 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

~~An earlier study has~~ 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 ~~compare~~ 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 ~~experiment~~ models reproduce the transient development of SDSR well in Europe, but poorly in Asia. ~~Observations in Asia exhibit a trend reversal in SDSR in the late 1980s that is primarily driven by SDSR changes in China.~~ The multiple historical and ~~historical~~ associated perturbation experiments performed under CMIP6 reveal ~~, that , in China,~~ that only those simulations containing anthropogenic aerosol emissions show dimming~~. None of the simulations exhibit the observed trend reversal over China in the late 1980s (brightening)~~, 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~~ Sulfur dioxide emissions over China show ~~no~~ neither sign of the observed trend reversal ~~in SDSR and neither does black carbon nor organic carbon emissions~~ nor of the brightening-followed-flattening in SDSR. 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 Figure 2(f). 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 ~~an important part~~ important in earth system models~~ability to simulate the past for~~

~~the right reasons~~.

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 ~~sulphur~~sulfur, organic carbon and nitrate in ice cores, as well as daily temperature range records.

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

Experiment	NorESM2	CanESM5	MIROC6	~~MIP~~ CESM2	CNRM-ESM2-1	~~Forcing agents~~GISS-E2-1-G	IPSL-CM6A-LR
~~CMIP6~~ historical	x	x	x	x	x	~~All~~x	x
~~DAMIP~~ hist-aer	x	x	x			~~Anthr. Aer~~ x	x
hist-GHG	x	x	x	x		~~Anthr. GHG~~x	x
hist-nat	x	x		x		~~Volc. and solar~~x	x
~~RFMIP~~ piClim-histaer	x	x	x			~~Anthr. Aer~~x	
piClim-histall	x	x	x			~~All~~x	x
~~AerChemMIP~~ hist-piAer	x		x			~~Volc.,solar,GHG~~	
hist-piNTCF	x		x		x	~~Volc.,solar,GHG~~	
histSST	x		x		x	~~All~~	

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, colocated and extracted at Chinese GEBA station locations. Changes in cloud cover are from CRU gridded data and represent means colocated 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		

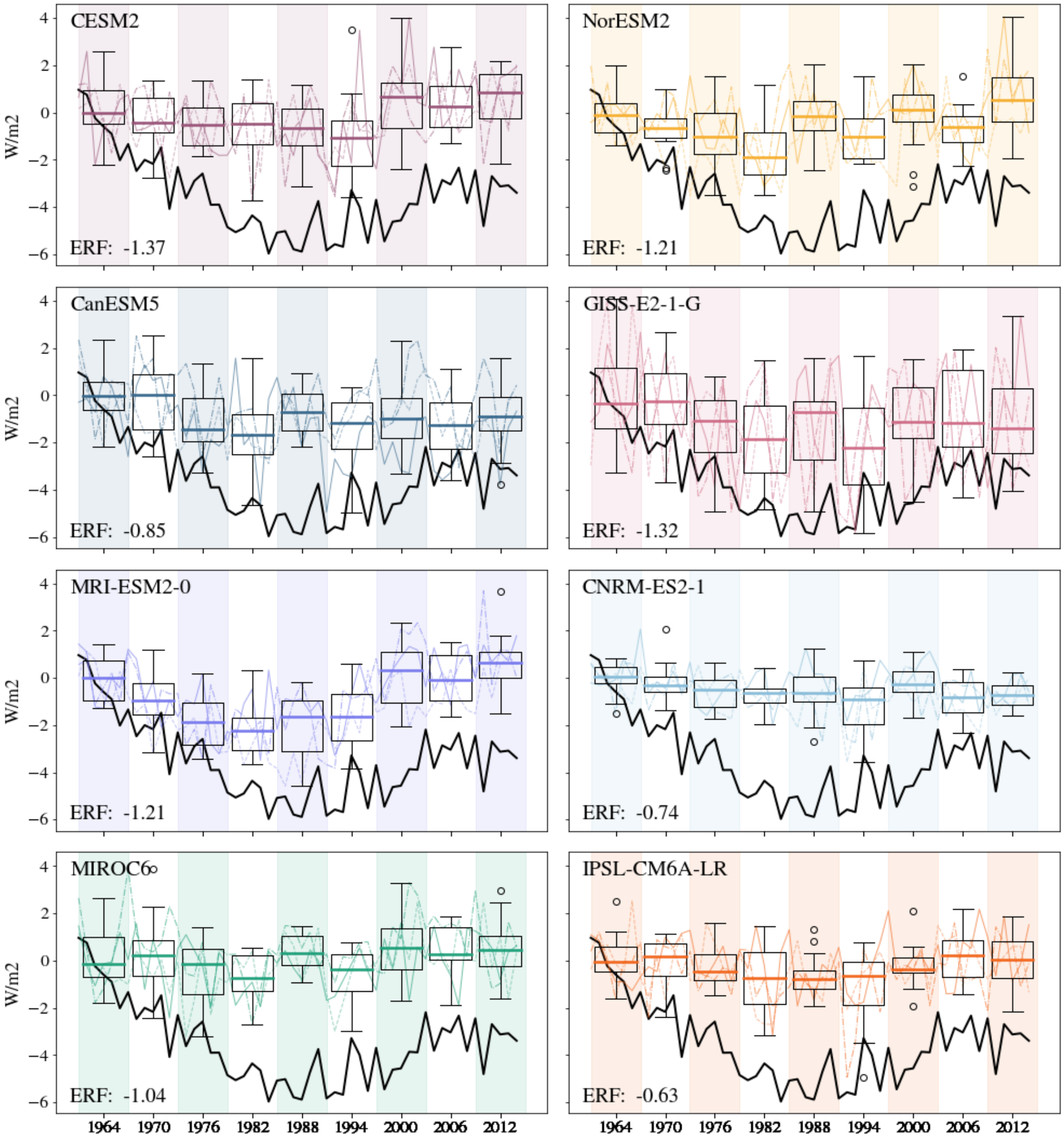

Figure 1. ~~Surface~~ Global surface downwelling shortwave radiation (SDSR) anomaly at the surface for GEBA (black) and ~~five~~ three ensemble members for the historical simulation of the eight models in this study. The boxes are made for 6-year intervals (shaded in background) based on 6 yearly means 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.

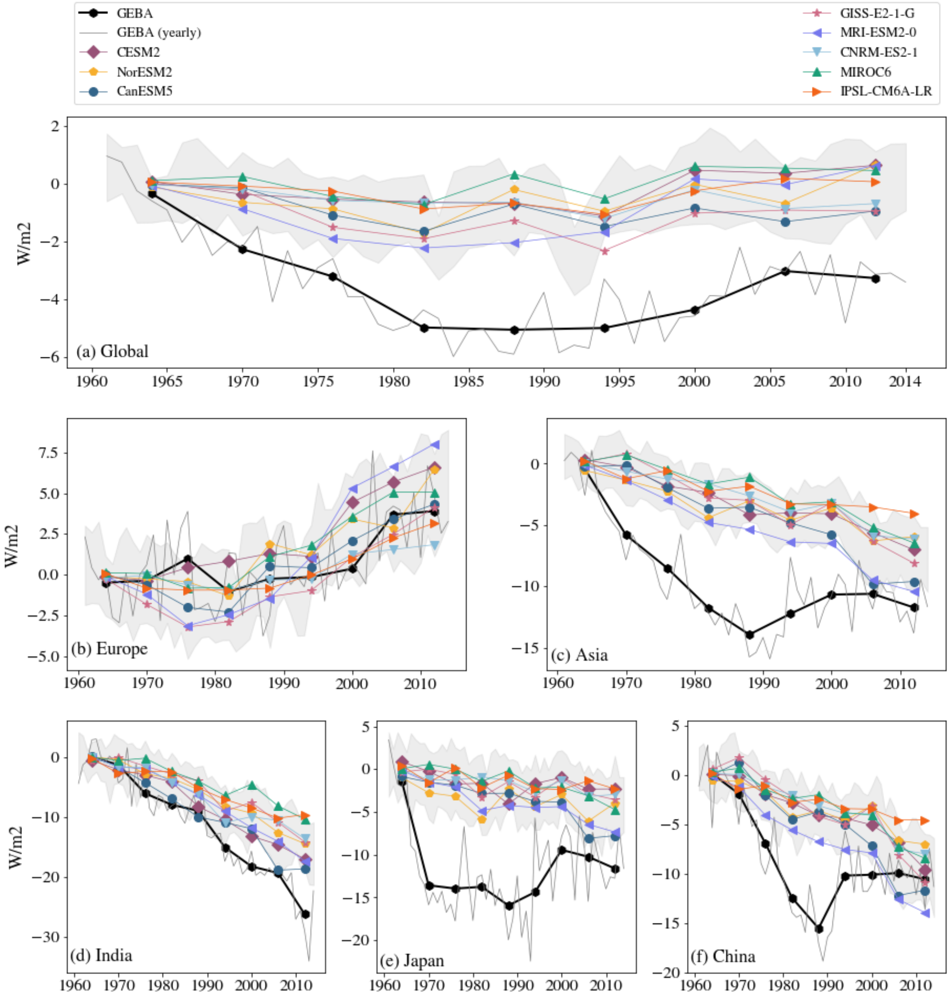

Figure 2. Six-year-averages of surface downwelling shortwave radiation (SDSR) anomaly at the surface for GEBA and eight earth system models. Results are co-located at (a) all GEBA stations (1487), (b) European (503), (c) Asian (311), (d), Indian (15), (e) Japanese (100), and (f) Chinese (119) stations.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

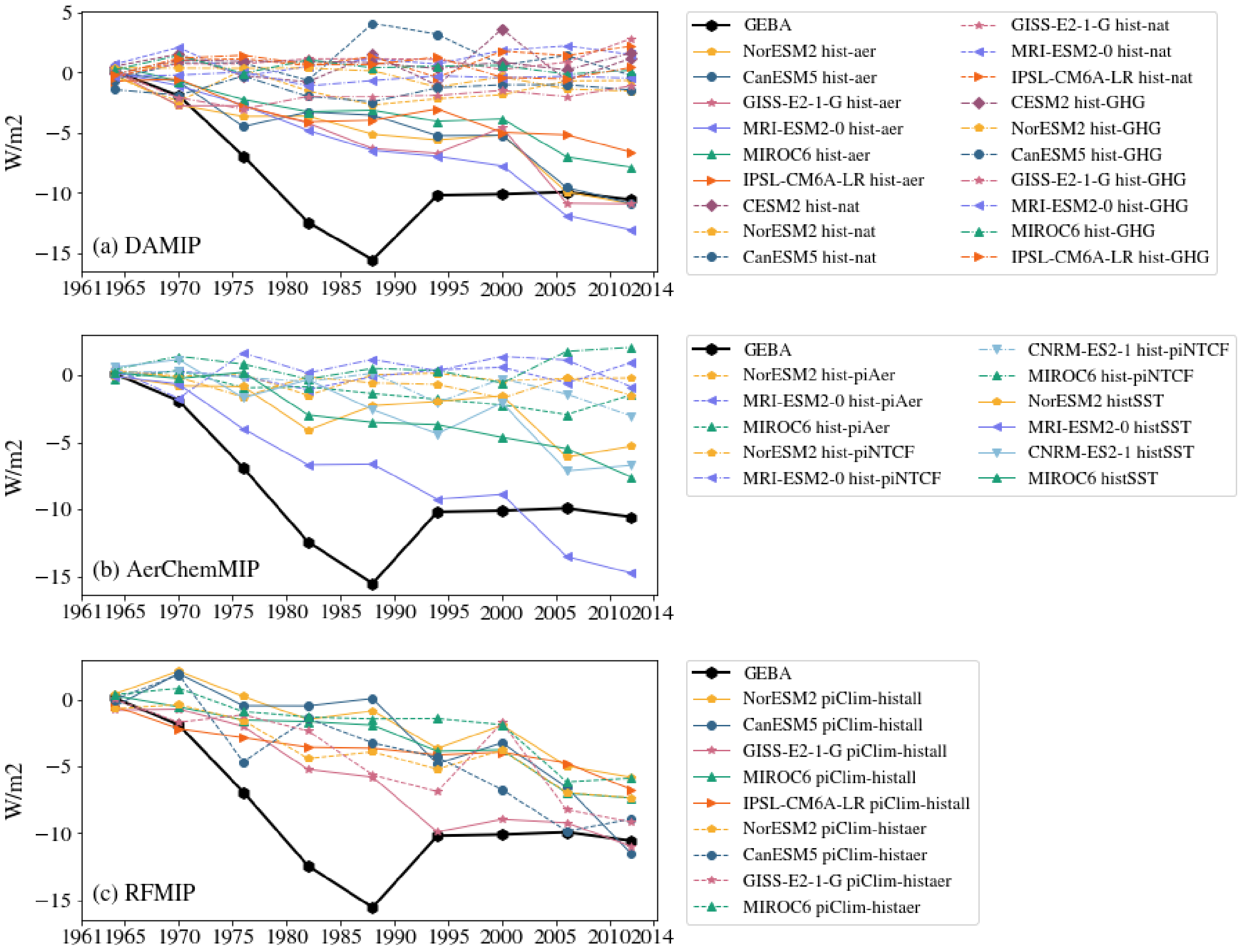

Figure 3. SDSR anomaly in China for all the CMIP6 simulations as listed in Table 1. All model results are co-located at GEBA station locations registered to China ~~. SDSR and cloud analysis at Chinese GEBA (~~119 stations~~in historical experiments: a)~~~~Clear sky SDSR anomaly together with all sky GEBA SDSR anomaly~~. The entire 54-year period has beed divided into intervals of 6 years and ~~(b) cloud cover anomaly~~ then averaged together ~~with corresponding CRU~~ to make nine data points as shown by the markers.

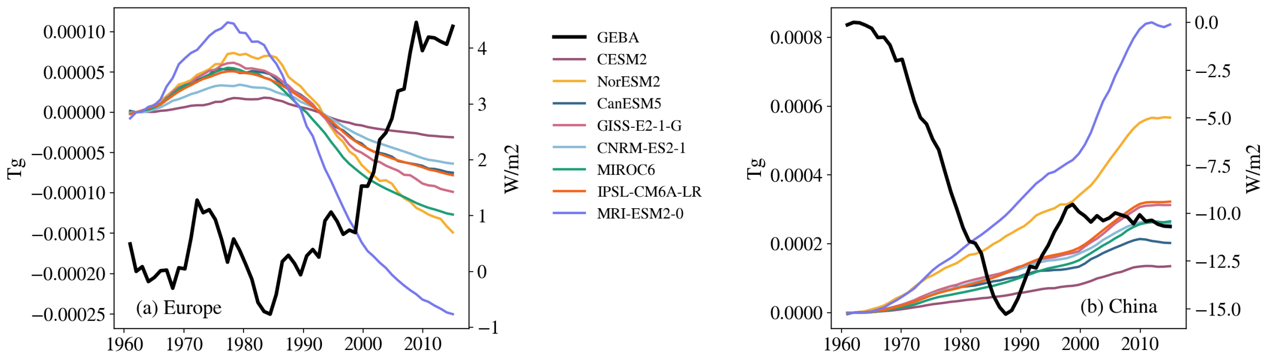

Figure 4. Anomaly of simulated atmospheric load of sulfate per model together with observed all sky SDSR anomaly in (a) Europe and (b) China. The GEBA data are shown as yearly anomalies, while the atmospheric loads have been smoothed using a 10-year running mean technique as explained in Section 2.3.

~~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°N–45°N, and longitudes 95°E–125°E.~~

Appendix A: ~~tables~~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 to the maximum in SDSR. The peak is not reproduced in the historical runs of earth system models studied herein (see Figure 2(b). 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 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:

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

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:

$$\Delta E_{\text{surf per 1\%}} = 0.01 \times E_{\text{cloudy}} - E_{\text{clear sky}}$$

$$= 0.01 \times E_{\text{toa}} \exp(-\tau_{\text{cloud}} / \cos\phi + \ln(E_{\text{clear sky}} / E_{\text{toa}})) + 0.99 \times E_{\text{clear sky}} - E_{\text{clear sky}}$$

Idealized computation for China: Assuming that ϕ is between 30° and 70°, that $E_{\text{clear sky}}$ is between 100 W/m^2 and 350 W/m^2 and that E_{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.

A3 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 was downloaded, with the variant labels r[1,2,3]i1p1f[1,2] for the variables rsds, rsds and clt. In addition the variable loadso4 and areacella was 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.

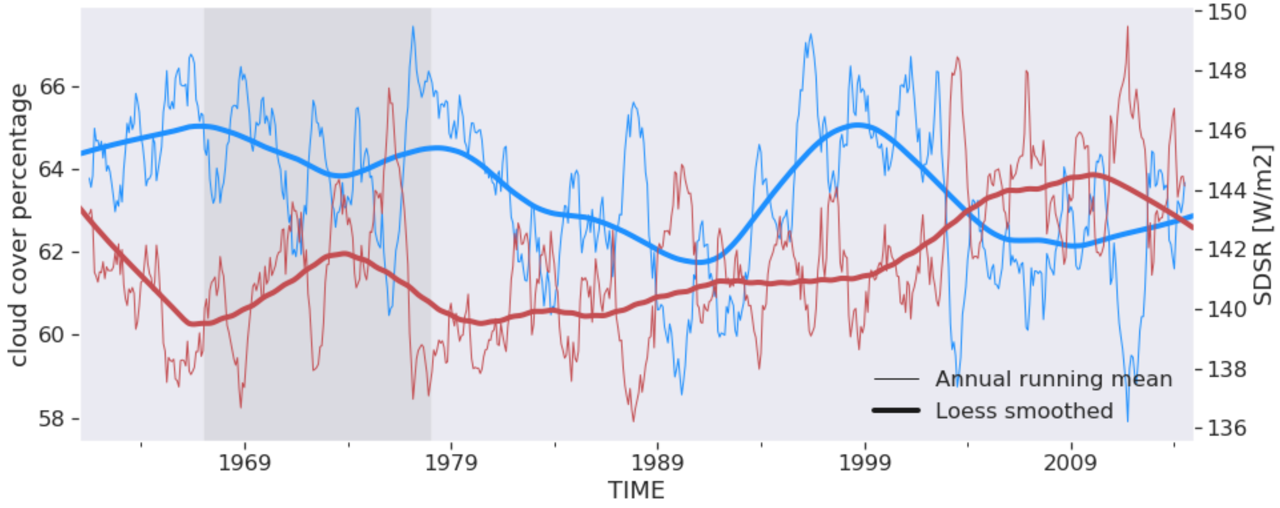

Figure A1. Timeseries of cloud cover (blue) and SDSR (red) between 1961 and 2014, co-located at GEBA sites in Europe. Thin lines show annual running means; bold lines show LOESS-smoothed variants. The shaded area delineates a period of interrupted dimming in Europe, between 1967 and 1978, which occurred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

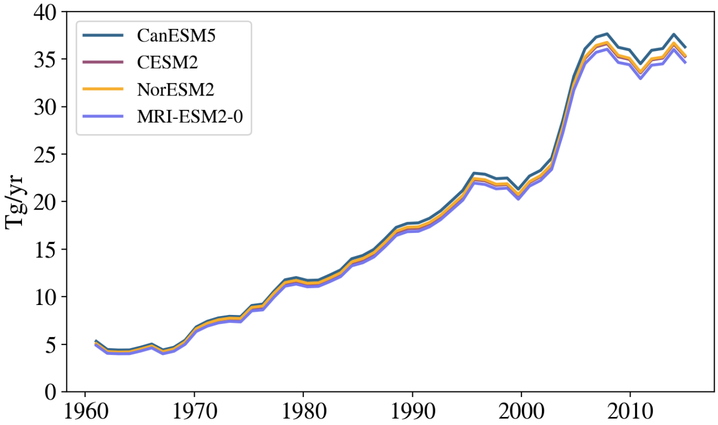

Figure A2. Emission of SO_2 in China, diagnosed by four of the models in this study. China is defined here as the area within latitudes [20°N–45°N], and longitudes [95°E–125°E].

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 ~~at~~ to all GEBA sites.

Model	~~Observation~~ SDSR [W/m^2]	Cloud Cover [%]
CESM2	186.3	63.9
NorESM2	186.8	55.6
CanESM5	~~MIROC6~~ 189.5	~~CESM2~~ 56.2
GISS-E2-1-G	~~CNRM-EMS2-1~~176.6	58.6
~~SDSR~~ MRI-ESM2-0	~~176.7~~ 193.8	~~187.6~~56.2
CNRM-ES2-1	~~189.7~~192.3	~~184.3~~57.2
MIROC6	~~186.3~~ 184.3	~~192.4~~ 50.4
~~Cloud Cover~~ IPSL-CM6A-LR	~~58.5~~ 185.7	~~55.3~~ 54.5
CRU	~~56.0~~	~~50.3~~ **57.4**
GEBA	~~63.8~~ **171.6**	~~57.3~~

Table A2. Details on the models used. IA: interactive aerosols NIA: not interactive aerosols.

Institution	Model	Resolution	Aerosol module	Complexity	Reference
NCAR	CESM2	1.25x0.9	MAM4	IA	Danabasoglu et al. (2020)
CCCma	CanESM5	2.81x2.81	CanAM4	IA	Swart et al. (2019)
CNRM-CERFACS	CNRM-ESM2-1	1.4x1.4	TACTIC_v2	IA	Séférian et al. (2019)
IPSL	IPSL-CM6A-LR	2.5x1.27	INCA fields	NIA	Boucher et al. (2020)
NCC	NorESM2-LM	2.5x1.875	OsloAero6	IA	Seland et al. (2020)
MRI	MRI-ESM2-0	1.125x1.125	MASINGAR mk-2r4c	IA	Yukimoto et al. (2019)
MIROC	MIROC6	1.4x1.4	SPRINTARS	IA	Tatebe et al. (2019)
NASA-GISS	GISS-E2-1-G	2.5x2.0	OMA fields	NIA	Kelley et al. (2020)

Author contributions. KM wrote most of the article and did all analysis of CMIP6 data. MS and TS contributed to design of the study and helped editing the text. DO, PN, JC and TT, NO, SB, GG contributed model data via the ESGF CMIP6 archive. IJ and MW contributed with observational data, and IJ wrote part of the appendix. All co-authors contributed to the analysis and gave feedback to the manuscript.

Competing interests. No competing interests

Acknowledgements. This study benefited greatly from the CMIP6 data infrastructure for handling and providing model data for analysis. KM, MS and TS acknowledge funding from the European Union's Horizon 2020 project FORCeS under grant agreement No 821205. We acknowledge support from the Research Council of Norway funded project KeyClim (295046). Jan Griesfeller is thanked for data organisation. 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. TT was supported by the supercomputer system of the National Institute for Environmental Studies, Japan, and JSPS KAKENHI Grant Number JP19H05669. NO was supported by the Japan Society for the Promotion of Science (grant numbers: 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), Program Grant Number 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 Switzelrand. MW acknowledges fundings obtained from the Swiss National Science foundation grand

References

Aas, W., Mortier, A., Bowersox, V., Cherian, R., Faluvegi, G., Fagerli, H., Hand, J., Klimont, Z., Galy-Lacaux, C., Lehmann, C. M. B., Myhre, C. L., Myhre, G., Olivié, D., Sato, K., Quaas, J., Rao, P. S. P., Schulz, M., Shindell, D., Skeie, R. B., Stein, A., Takemura, T., Tsyro, S., Vet, R., and Xu, X.: Global and regional trends of atmospheric sulfur, Scientific Reports, 9, 953, https://doi.org/10.1038/s41598-018-37304-0, https://www.nature.com/articles/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, Journal of Geophysical Research: Atmospheres, 118, 6311–6336, https://doi.org/10.1002/jgrd.50426, https://agupubs.onlinelibrary.wiley.com/doi/abs/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, book section 7, p. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.016, www.climatechange2013.org, 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., Éthé, 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., Lévy, 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., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, 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, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 010, https://doi.org/10.1029/2019MS002010, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002010, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002010, 2020.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 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 Generation Computer Systems, 36, 400–417, https://doi.org/10.1016/j.future.2013.07.002, http://www.sciencedirect.com/science/article/pii/S0167739X13001477, 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, https://www.geosci-model-dev.net/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., Oleson, 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), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 916, https://doi.org/10.1029/2019MS001916, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001916, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/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, https://www.nature.com/articles/300689a0, number: 5894 Publisher: Nature Publishing Group, 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, Geoscientific Model Development, 9, 1937–1958, https://doi.org/https://doi.org/10.5194/gmd-9-1937-2016, https://www.geosci-model-dev.net/9/1937/2016/gmd-9-1937-2016.html, 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, NDM ’15, pp. 1:1–1:7, ACM, New York, NY, USA, https://doi.org/10.1145/2832099.2832100, http://doi.acm.org/10.1145/2832099.2832100, event-place: Austin, Texas, 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, Journal of Geophysical Research: Atmospheres, 120, 2261–2279, https://doi.org/10.1002/2014JD022851, https://agupubs.onlinelibrary.wiley.com/doi/abs/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, Geoscientific Model Development, 9, 3685–3697, https://doi.org/https://doi.org/10.5194/gmd-9-3685-2016, https://www.geosci-model-dev.net/9/3685/2016/, 2016.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, https://doi.org/10.1002/joc.3711, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.3711, 2014.

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, https://www.nature.com/articles/s41597-020-0453-3, number: 1 Publisher: Nature Publishing Group, 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), Geoscientific Model Development, 11, 369–408, https://doi.org/https://doi.org/10.5194/gmd-11-369-2018, https://www.geosci-model-dev.net/11/369/2018/, 2018.

Hoyt, D. V. and Schatten, K. H.: A discussion of plausible solar irradiance variations, 1700-1992, Journal of Geophysical Research: Space Physics, 98, 18 895–18 906, https://doi.org/10.1029/93JA01944, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/93JA01944, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/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?, Geophysical Research Letters, 29, 38–1–38–4, https://doi.org/10.1029/2002GL016057, https://agupubs.onlinelibrary.wiley.com/doi/abs/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., García-Pando, 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, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 025, https://doi.org/10.1029/2019MS002025, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002025, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002025, 2020.

Liepert, B. G.: Observed reductions of surface solar radiation at sites in the United States and worldwide from 1961 to 1990, Geophysical Research Letters, 29, 61–1–61–4, https://doi.org/10.1029/2002GL014910, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GL014910, 2002.

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, https://journals.ametsoc.org/doi/full/10.1175/1520-0426%281999%29016%3C0055%3AOMOSSI%3E2.0.CO%3B2, 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, Geophysical Research Letters, 41, 5605–5611, https://doi.org/10.1002/2014GL060798, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL060798, 2014.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, https://doi.org/https://doi.org/10.5194/gmd-9-3447-2016, https://www.geosci-model-dev.net/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, https://www.jstor.org/stable/4314548, publisher: [Springer, Royal Swedish Academy of Sciences], 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, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL028356, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL028356, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/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, Journal of Geophysical Research: Atmospheres, 119, 4657–4673, https://doi.org/10.1002/2013JD021393, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021393, 2014.

Seland, , 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, Geoscientific Model Development Discussions, pp. 1–68, https://doi.org/https://doi.org/10.5194/gmd-2019-378, https://gmd.copernicus.org/preprints/gmd-2019-378/, publisher: Copernicus GmbH, 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., Pincus, R., and Forster, P. M.: Effective radiative forcing and adjustments in CMIP6 models, preprint, Radiation/Atmospheric Modelling/Troposphere/Physics (physical properties and processes), https://doi.org/10.5194/acp-2019-1212, https://acp.copernicus.org/preprints/acp-2019-1212/acp-2019-1212.pdf, 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, https://science.sciencemag.org/content/327/5970/1219, publisher: American Association for the Advancement of Science Section: Research Article, 2010.

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, Geophysical Research Letters, 0, https://doi.org/10.1029/2018GL078298, https://agupubs.onlinelibrary.wiley.com/doi/abs/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, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026471, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026471, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/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), Geoscientific Model Development, 12, 4823–4873, https://doi.org/https://doi.org/10.5194/gmd-12-4823-2019, https://gmd.copernicus.org/articles/12/4823/2019/, publisher: Copernicus GmbH, 2019.

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 System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, Journal of Advances in Modeling Earth Systems, 11, 4182–4227, https://doi.org/10.1029/2019MS001791, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001791, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001791, 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, Atmospheric Chemistry and Physics, 11, 393–406, https://doi.org/https://doi.org/10.5194/acp-11-393-2011, https://acp.copernicus.org/articles/11/393/2011/, publisher: Copernicus GmbH, 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, Geoscientific Model Development, 12, 2727–2765, https://doi.org/https://doi.org/10.5194/gmd-12-2727-2019, https://gmd.copernicus.org/articles/12/2727/2019/, publisher: Copernicus GmbH, 2019.

Wang, K. C., Dickinson, R. E., Wild, M., and Liang, S.: Atmospheric Impacts on Climatic Variability of Surface Incident Solar Radiation, https://doi.org/10.5194/acp-12-9581-2012, https://repositories.lib.utexas.edu/handle/2152/43322, accepted: 2016-10-28T19:53:01Z, 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, 11,777–11,785, https://doi.org/10.1002/2016GL071009, http://doi.wiley.com/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, https://www.ann-geophys.net/32/41/2014/, 2014.

Wild, M.: Global dimming and brightening: A review, Journal of Geophysical Research, 114, https://doi.org/10.1029/2008JD011470, http://doi.wiley.com/10.1029/2008JD011470, 2009.

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

Wild, M. and Schmucki, E.: Assessment of global dimming and brightening in IPCC-AR4/CMIP3 models and ERA40, Clim Dyn, 37, 1671–1688, https://doi.org/10.1007/s00382-010-0939-3, https://link.springer.com/article/10.1007/s00382-010-0939-3, company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 7 Publisher: Springer-Verlag, 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, https://science.sciencemag.org/content/308/5723/847, publisher: American Association for the Advancement of Science Section: Report, 2005.

Wild, M., Ohmura, A., and Makowski, K.: Impact of global dimming and brightening on global warming, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL028031, https://agupubs.onlinelibrary.wiley.com/doi/abs/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 Dyn, 40, 3107–3134, https://doi.org/10.1007/s00382-012-1569-8, 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 System Science Data, 9, 601–613, https://doi.org/https://doi.org/10.5194/essd-9-601-2017, https://www.earth-syst-sci-data.net/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 Dyn, https://doi.org/10.1007/s00382-018-4413-y, 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, http://journals.ametsoc.org/doi/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, http://journals.ametsoc.org/doi/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, Journal of the Meteorological Society of Japan. Ser. II, 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, Journal of Geophysical Research: Atmospheres, 106, 14 501–14 513, https://doi.org/10.1029/2001JD900030, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD900030, 2001.