**Response to Reviewers**

We thank both reviewers for evaluating our manuscript. Below, we list our responses to each comment (in blue). We first note that all analyses have been updated based on current model availability. This includes the addition of more realizations and/or climate variables to the models previously used (e.g., we now have MIROC daily data and 2 additional CESM2-WACCM simulations, etc.). We have also added two additional models to the analysis: NorESM2-LM and UKESM1-0-LL. Our overall results and conclusions remain unchanged.

**Reviewer #1**

General:

Allen et al. introduce results of the AerChemMIP project on the impact of air quality measures on climate. This is a large exercise and certainly worth publishing. However, I think there are major shortcomings. The most apparent is the style. The paper is written as a report, stating what has been done and what is the outcome. While this is, of course, an essential part of a paper, it should contain much more. It is less written as a scientific paper that should motivate chosen assumptions, extract main new massages from results, and discuss uncertainties e.g. wrt. to the chosen assumptions. This is largely missing. For example the main message "Our findings suggest that future policies that aggressively target non-methane NTCF reductions will improve air quality, but will lead to additional surface warming" is shown in the end as being nothing new, but already covered by many other studies, as shown by the authors in line 345ff. So what is new? And this puts me actually in a difficult position, why should a paper be published which "just" confirms previous findings? I understand that IPCC deadlines have to be met, but more emphasis should be given to clearly describe what is new. More examples are given in the detailed comments below.

We have attempted to improve the writing, by placing more emphasis on motivation and assumptions. Although our main results support prior studies, given the sophistication of the models used here, as well as the relatively large number of models, we suggest that this work is the most comprehensive analysis on this topic to date. Overall, the structure of the manuscript is similar to the original submission. Reviewer #2 states: "…the manuscript is generally well written and well-structured and makes good use of the AerChemMIP simulations".

Major comments in addition to the writing style:
1) Structure: The method section is too short:

- More information on statistics should be given (see details below); Please explain why the multi-model trends are significant, although individual model trends are not. What trend model has been used? What exactly is tested?

More information on the statistics are included. Models with multiple realizations are first averaged to form the model mean. Individual model mean trends are calculated using least squares regression, and the corresponding trend significance is based on a two-tailed Student's t-test, where the null hypothesis of a zero regression slope is evaluated. Autocorrelation of the time series is also accounted for by using the effective sample size: $n \times (1 - r_1)/(1 + r_1)$, where $n$ is the number of years and $r_1$ is the lag-1 autocorrelation coefficient.

MMM trends and their significance are estimated in two ways, and both methods yield similar results. In the first approach, the overall multi-model mean time series is calculated as the mean over each model mean (i.e., each model has the same weight), and a similar procedure as above is used to determine the significance of the multi-model mean trend. However, we now use a weighted least squares regression (as suggested by Reviewer #2). Each value in the multi-model mean time series is weighted by $\frac{1}{\sigma_m^2}$, where $\sigma_m$ is the standard deviation across models.

We note that this does not change any of our results. For example, the global annual mean multi-model surface temperature trend changes from 0.062 without weighting to 0.060 K/decade with weighting, and both are significant at the 99% confidence level. Weighting the regression introduces negligible changes in our other climate and air quality trends.

In Figure 3, there was actually quite a bit of similarity in individual global mean model trends. All but one individual model surface temperature trend in Figure 3 was significant (MIROC6 the lone exception). Furthermore, all individual model precipitation, Hottest Day, PM2.5 an Ozone global mean trends were significant. Weaker results existed for the Wettest Day, and in particular, Consecutive Dry Days. It is therefore not surprising the multi-model mean trends were also significant. Using the weighted regression approach, we get similar conclusions.

To summarize, we now use a weighted least squares regression for the multi-model mean trends, and we get similar results as before.

In addition to estimating the magnitude and significance of the multi-model mean trend as just described, we also evaluate the multi-model mean trend and its significance relative to the individual model mean trends (e.g., Figure 5). Here, the MMM trend is estimated as the average of each model mean trend, and its uncertainty is estimated as plus/minus twice the standard error (i.e., the 95% confidence interval). This is calculated as: $2 \times \sigma / \sqrt{nm}$, where $\sigma$ is the standard deviation of model mean trends and *nm* is the number of models. If this confidence interval does not include zero, then the multi-model mean trend is significant at the 95% confidence level.

The corresponding 95% confidence intervals are now included in each of the global time series plots. As is the R^2 value of the multi-model mean trend.

We have also added the multi-model global mean trend (and others, including NH mid-latitudes, Tropics, SH mid-latitudes) and its uncertainty to the bar graphs in Figure 5.

- A motivation why exactly these climate/air quality/extreme indicators are chosen is missing;

As discussed in the Introduction, both PM2.5 and ozone are commonly used indicators of air quality. Both have been associated with adverse human health impacts. Surface temperature and precipitation are analyzed as these are arguably two of the most important climate variables. Changes in surface temperature are particularly relevant in the context of climate mitigation, as the goal of the Paris Agreement is to keep the increase in global mean surface temperature to well below 2°C above preindustrial values. Precipitation, and fresh water resources in general, are important to both human society and ecosystems. Perhaps more important than changes in the mean of a climate variable are changes in its extremes. Heat waves, for example, are a major cause of weather related fatalities. We focus here on the hottest and wettest day, as well as consecutive dry days, as these are frequently used extreme temperature and precipitation indices (e.g., Donat et al., 2013a,b). Furthermore, prior observational analyses have shown significant changes in all three quantities over the latter half of the 20$^{th}$ century. This information has been added.

- Part 3.1 is actually input to the study and should be moved from the results part to the method part.

Moved to the Methods section.

2) Statistics: I have strong concerns how the statistics are interpreted. If a difference is not statistical significant, there is no basis in discussing them. Please remove all parts, which interpret statistically insignificant differences.

These have been removed.

3) Discussion: How important are the choices made in the assumption section?

We assume the "assumption section" pertains to the future emission pathways used here. We have added more information on the assumptions pertaining to the two future emissions pathways analyzed here. Basically, our analysis assumes that NTCF policies can be enacted in the absence of GHG related climate policies (e.g., SSP1's air pollutant legislation and technological progress can be achieved in the SSP3 world). Furthermore, our results likely represent an upper bound, since our baseline/reference scenario lacks climate policy and has the highest levels of NTCFs. This has been clarified in the Future Scenarios Section.

In the Conclusion Section, we discuss the implications of the assumptions made in the weak and strong air quality control pathways used in this analysis. It is not possible to formally quantify these assumptions, as different NTCF mitigation simulations were not performed by AerChemMIP. We have added clarification in the conclusions:

Our simulations, however, do not account for CO2 reductions, implying the importance of simultaneous reductions in both CO2 and NTCFs. We note that it is difficult to reduce only the NTCF emissions while keeping CO2 emissions fixed (since there are co-emitted species, including SO2). If CO2 emissions are simultaneously reduced along with NTCFs, then the increase in global surface temperature and precipitation found here will be muted.

Detailed comments:

Abstract: "How future policies affect the abundance of NTCFs and their impact on climate and air quality remains uncertain." I am wondering whether this could be misunderstood in a way that for a given measure the impact remains uncertain. Most of the uncertainty comes from the uncertainty what measures will be taken, right?

Future climate and air quality are uncertain for two reasons. There is uncertainty due to the emissions pathway, and there is uncertainty in the corresponding climate response. Past IPCC reports have shown that both uncertainties are approximately of the same magnitude in the context of climate change. The latter uncertainty is due to uncertainty in climate sensitivity (e.g., 1.5-4.5 K per $2xCO_2$). As an aside, CMIP6 models tend to have a higher climate sensitivity than CMIP5 models, which has been related to clouds (e.g., Zelinka et al., 2020). Nonetheless, we have attempted to clarify this sentence, since the larger uncertainty for a given pathway is the climate response.

l13 "similar increases" what means similar here? Can a extreme weather index be similar to a temperature increase of 0.24 K? or is even 0.34 K similar to 1.1%. Please specify.

Re-worded.

l16 "ozone reductions.": I think it would be helpful to include half a sentence explaining the relation between aerosols and ozone.

We have added information in the Introduction. "…reductions in some precursor gases such as NOx and VOCs impact both ozone and aerosols (and perhaps $CH_4$). Reductions in NOx, for example, will promote cooling due to reduced tropospheric ozone, but the impact on $CH_4$ lifetime and aerosol formation will likely promote overall warming"

l20-21: I think the definition in Myhre et al 2013 is "We define 'near-term climate forcers'(NTCFs) as those compounds whose impact on climate occurs primarily within the first decade after their emission." It reads a little bit different from "that impact climate on relatively short time scales, typically within a few weeks to a decade after emission". Climate is defined on decadal timescales. To relate climate change to to weeks sounds weird. Concentration changes and RF can quickly react, but you started to discuss climatological changes in temperature and rain rates and those do not occur on weekly timescales. Please adapt the text.

We have adopted the reviewer's verbiage.

l28 should it be "-2.0 to−0.4" ?

Re-ordered.

l34 shouldn't methane be mentioned here as well, since it is a precursor for ozone? I think you are referring to table 8.6 in Myhre et al. 2013. Their tropospheric ozone area total ozone change and include effects from methane emissions.

Methane added here.

l34-37: Here you change from a concentration perspective (ozone) to an emission perspective (methane). Please clarify this, otherwise it seems to be inconsistent and double counting methane ozone effects. Especially the wording "Similarly," should be revised, since the view is exactly not similar.

Modified to concentration perspective.

l42-44 please clarify the sentence. How can a change in radiation, i.e. in W/m2, be balanced by evaporation in units m/s.

Changed evaporation to latent cooling.

l62 please clarify what you mean with "rapid". See also discussion above.

Clarified. Added "decadal".

l91 You mean the scenarios you are employing. ...

Added "Used here".

Section 2.1: I think it would be nice to have a motivation included. Currently, it reads like a report or namelist setting. Why is the reference without climate policy? etc. this should be motivated.

We have added motivation, assumptions, and clarity. Our analysis assumes that NTCF policies can be enacted in the absence of GHG related climate policies (e.g., SSP1's air pollutant legislation and technological progress can be achieved in the SSP3 world). Furthermore, our results likely represent an upper bound, since our baseline/reference scenario lacks climate policy and has the highest levels of NTCFs (i.e., to detect the largest signal, the reference is without climate policy).

Please also include a table showing the changes in relevant emissions, such as aerosol compounds and ozone precursors for some well-chosen times, e.g. 2015, 2035, 2055;or decadal? I think it is important to see the changes.

Figure 2 already shows changes in emissions by region, as well as over land.

l120 I find the abbreviation misleading. "lowAER" and "lowAERO3" are model group names. "low", however, is not referring to the models, but to the scenario, right? and at some point I though "AERO3" is the "AEROsol Group 3" and not aerosol-ozone. What about "Only-Aer" and "Aer-O3"; or "Aer+O3" ?

Changed abbreviations to "Aer" and "Aer+O3" l135 Please include what kind of linear model you are using y(t)=a+bt+err or y(t)=b(t-2035)+err ? Are you fitting one or two parameters? Often as a measure for the fitting quality the R^2 value or adjusted R^2 value is used. Why not here? I do not understand how the trend is tested. Are the individual model results fitted and then tested whether the mean trend is representing the range of models correctly? (At least the caption of Figure 7 might indicate something like this). How the statistics are treated is very important for the interpretation. Please include a thorough discussion here.

Our response to the concern over statistics is located above.

l159 Also here a motivation is missing. I understand that extreme values are important. But why is the max temperature chosen? Isn't that a statistically very difficult quantity, even among extreme value statistics? Why not using number of hot days, i.e.over 30◦C? This also concentrates on extremes, but includes a whole tail of a pdf (or estimated pdf).

The Hottest Day ("TXx") is chosen as it is a commonly used extreme temperature measure. See for example, Donat et al. (2013 a,b). As with other extreme temperature indices, significant increases in TXx were found (1951-2011) in two different data sets, GHCNDEX and HadGHCND. We also find significant TXx trends in the simulations analyzed here.

l162 please also add the respective time frame. Are you averaging over 10 or 30 years?

We are not sure what time frame the reviewer refers to. L162 states "Climate extremes are calculated at each grid box and then spatially averaged." There is no decadal averaging.

The hottest day (monthly maximum value of daily maximum temperature) and the wettest day (monthly maximum 1-day precipitation) are estimated for each month, and then averaged to obtain annual means. Consecutive dry days (CDD), defined as the maximum annual number of consecutive days with precipitation <1 mm/day, are estimated for each year.

l165ff: I would have expected this part in the scenario section. Please consider to move it there, since this is not a result from your paper, right? And then ignore my comment on the table (see above) ...

Moved to scenario section.

l167: Why is there a CO2 emission change at all, if you are considering NTCF changes only? Please explain. I don't think that this is a problem, but currently and certainly it confuses me.

Yes, there are small differences in CO2 emissions between the two scenarios. Methane reductions generate emissions abatement costs, which changes industrial outputs in all productin sectors and household consumption (Gidden et al., 2019). Energy consumption and CO2 emissions in all sectors are thus affected, which causes small differences between SSP3-7.0 and SSP3-LowNTCF.

However, AerChemMIP simulations use the same $CO_2$ emissions, based on SSP3-7.0 (as with methane). This has been clarified in the revision. We have also removed the SSP3-7.0-lowNTCF $CO_2$ (and $CH_4$) emissions from the plots, to avoid unnecessary confusion.

l192: Why are you discussing methane emission changes, if those are not relevant?

Deleted discussion on methane emission changes.

Figure 2: Trends are calculated as (2055-2015)/4 or with the regression method discussed in Section 2?

Emission trends are calculated using the same method as above. Trends are based on a least squares regression, with significance based on a two-tailed Student's t-test. We note that the emissions data is decadal after 2015, with monthly values for the year 2015, 2020, 2030, 2040, 2050, 2060, etc. We estimate the emissions in 2055 as the mean of the emissions in 2050 and 2060 at each grid box. We have added this information.

l 205: Please comment if the trends of the individual models are statistical significant. I miss a mathematical/statistical explanation in combination with a motivation why to test the multi-model mean and not, whether mean trend is significant wth respect to the variation in trends of the individual models.

Table 1 lists the trend (and whether it is significant at the 95% confidence level) for each model. All but one model yields a significant global mean increase in surface temperature. This has been clarified.

We initially did not evaluate the significance of the multi-model mean trend, relative to the individual model trends, for the global mean quantities. We did this for the regional trends (e.g., Figure 5). We do note that Figure 5 included land only averages. Nonetheless, we have now performed this analysis for the global mean quantities. The 95% confidence interval is now included in the global time series plots. We have also added the multi-model global mean trend and its uncertainty to the bar graphs in Figure 5 (and additional latitudinal bands).

l 206: For the regional trend an uncertainty range is given. Why not here?

Yes, we have now added this analysis for the global trends. The corresponding 95% confidence intervals are now included in each of the global time series plots. We have also added the multi-model global mean trend and its uncertainty to the bar graphs in Figure 5.

l213-l214: If a result is not statistical significant, there is no point in interpreting the result. Please delete the sentence.

We agree, and this statement has been modified. It is still interesting that similar (i.e., non-significant differences) global mean surface air temperature trends occur in Aer and Aer+O3 models. We acknowledge that this could be due to several factors, but one interpretation is weak surface cooling due to reductions in ozone.

l223: Keep in mind that the change was not statistical significant; so the results may not be inconsistent, but only noise. Please revise the discussion, based on what is inconsistent on a statistical significant basis.

Discussion edited and revised. As with surface temperature, non-significant differences between Aer and Aer+O3 ERF trends exist.

l 246: "Slightly larger (but not statistically significant) "; if not statistically significant, then they are not slightly larger! Please respect the statistics.

Edited. Similar increases occur in both Aer and Aer+O3 models.

l263-264: Please rephrase the sentence. I agree with the content, however, the formulation, starting with "however" suggests that there is either a shortcoming or something unexpected, etc. As the authors state this is by no means a surprise nor limitation of the aforementioned.

Deleted "However"

l262 I think somewhere it should made clear that a part of the warming is a reduced cooling from SO2 reduction and O3 reductions, right?

Yes, based on the information presented in the Introduction (e.g., radiative forcing), SO2 reductions will warm. But O3 reductions will cool. This has been clarified.

From the Introduction: Thus, reductions in some NTCFs, including non-absorbing aerosols, will warm the climate system, whereas reductions in other NTCFs, including absorbing aerosols, tropospheric ozone, and methane will cool the climate system.

l284 Is there some relation to the monsoon tipping points?

L284 states: "Furthermore, in agreement with prior studies, precipitation increases in several monsoon regions, including east Africa, south Asia, and east Asia." Thus, unlike the buildup of aerosols over the 20$^{th}$ century, future NTCF mitigation and continued increases in GHGs will likely accelerate the monsoons. Not exactly sure what the Reviewer wants us to change here.

Section 3.4: What about MAM/SON? Discussed are winter/summer differences. However, "Seasons" would imply more than that. I suggest to, at least, mention a general trend for MAM/SON and add the same figures in a supplement.

Added general trends in MAM/SON seasons. Added MAM/SON plots to supplement.

l339-342: This is important: see also above. If the difference is not statistical significant, there is no point in discussing or even highlighting it in the summary. Please remove this part!

We have rephrased. The lack of significant trend differences in Aer and Aer+O3 models is interesting. We acknowledge that this could be related to several factors. But one possible interpretation is weak surface cooling due to reductions in ozone. We feel as if our ability to compare Aer and Aer+O3 models is one of the novelties of this study. But again, we acknowledge this comes with caveats.

l359: It might be worth mentioning reduced warming, i.e. a net cooling. To avoid confusion about weather $CH_4$ itself has a cooling contribution.

Added "net cooling" here.

**Reviewer #2**

Allen et al. use model output from the AerChemMIP intercomparison project to evaluate 2015-2055 changes in climate variables associated with two future air quality control scenarios. By comparing a "weak" policy scenario to a "strong" policy scenario, they show increasing trends in temperature and precipitation over the period that are driven by near-term climate forcers (ozone and aerosols), suggesting a climate penalty associated with air quality improvements. The manuscript is generally well written and well structured and makes good use of the AerChemMIP simulations. It addresses an important question that is well suited to the scope of ACP. I do have a few concerns about the statistics and a few more minor comments and suggestions, discussed below.

GENERAL COMMENTS

1. The trends have been calculated using least squares regression. There is very little information on exactly how that was implemented, so from my reading it does not appear that this is a weighted least squares regression, or that the uncertainties have been accounted for in any other way. This is concerning because, looking at Fig. 3 for example, there is a large amount of variability in the individual models that are used to construct the multi-model means. I am not convinced by the robustness of some of the reported trends in the multi-model mean, or that they are truly statistically significant as stated. The multi-model mean trend calculations should be performed using a method that accounts for variability/uncertainty in the mean (e.g., weighted least squares, but there are other options) before the paper is publishable in ACP. In addition, some discussion of the method used and the influence of the variability/uncertainty is warranted.

More information on the statistics are included. Models with multiple realizations are first averaged to form the model mean. Individual model mean trends are calculated using least squares regression, and the corresponding trend significance is based on a two-tailed Student's t-test, where the null hypothesis of a zero regression slope is evaluated. Autocorrelation of the time series is also accounted for by using the effective sample size: $n\times(1-r_1)/(1+r_1)$, where $n$ is the number of years and $r_1$ is the lag-1 autocorrelation coefficient.

MMM trends and their significance are estimated in two ways, and both methods yield similar results. In the first approach, the overall multi-model mean time series is calculated as the mean over each model mean (i.e., each model has the same weight), and a similar procedure as above is used to determine the significance of the multi-model mean trend. However, we now use a weighted least squares regression (as suggested by the Reviewer). Each value in the multi-model mean time series is weighted by $\frac{1}{\sigma_m^2}$, where $\sigma_m$ is the standard deviation across models.

We note that this does not change any of our results. For example, the global annual mean multi-model surface temperature trend changes from 0.062 without weighting to 0.060 K/decade with weighting, and both are significant at the 99% confidence level. Weighting the regression introduces negligible changes in our other climate and air quality trends.

In Figure 3, there was actually quite a bit of similarity in individual global mean model trends. All but one individual model surface temperature trend in Figure 3 was significant (MIROC6 the lone exception). Furthermore, all individual model precipitation, Hottest Day, PM2.5 an Ozone global mean trends were significant. Weaker results existed for the Wettest Day, and in particular, Consecutive Dry Days. It is therefore not surprising the multi-model mean trends were also significant. Using the weighted regression approach, we get similar conclusions.

To summarize, we now use a weighted least squares regression for the multi-model mean trends, and we get similar results as before.

In addition to estimating the magnitude and significance of the multi-model mean trend as just described, we also evaluate the multi-model mean trend and its significance relative to the individual model mean trends (e.g., Figure 5). Here, the MMM trend is estimated as the average of each model mean trend, and its uncertainty is estimated as plus/minus twice the standard error (i.e., the 95% confidence interval). This is calculated as: $2\times\sigma/\sqrt{nm}$, where $\sigma$ is the standard deviation of model mean trends and $nm$ is the number of models. If this confidence interval does not include zero, then the multi-model mean trend is significant at the 95% confidence level.

The corresponding 95% confidence intervals are now included in each of the global time series plots. As is the R^2 value of the multi-model mean trend.

We have also added the multi-model global mean trend (and others, including NH mid-latitudes, Tropics, SH mid-latitudes) and its uncertainty to the bar graphs in Figure 5.

2. For the global trends in climate variables, it would help to contextualise the values associated with NTCFs by also providing the trends from the two individual scenarios(or at least from the one with weak NTCF control, as the other can be determined from the difference trends provided). Without this, it’s hard to tell how important the NTCF climate penalty is. I note that this is done in the figures for the regional trends, but not for the global trends. I would strongly

encourage the authors to add these in some form (for example, a figure in the SI equivalent to Fig. 3).

These have been added to Figure 5. The last set of bars (labeled "GL") now show the global mean trends for SSP3-7.0, SSP3-7.0-lowNTCF and their difference. Also included is the corresponding land (labeled "Ld") surface values (which were previously included). We have also added additional trends over various latitude bands.

3. The manuscript is very well structured and quite well written, but the heavy use of acronyms and technical identifiers (e.g., SSP3-7.0-lowNTCF, lowAERO3, etc.) makes it harder to read & follow than it needs to be. I would encourage the authors to simplify this wherever possible and then use a consistent, easy to interpret nomenclature throughout. For example, frequently the two scenarios are referred to as strong and weak air quality control, and these are much easier to interpret than SSP3-7.0 andSSP3-7.0-lowNTCF. I would suggest strong and weak air quality control could replaceSSP3-7.0 and SSP3-7.0-lowNTCF everywhere, in particular in figure legends and captions where the reader may not be referring back to the text. Similarly, NTCF mitigation is easier to interpret than SSP3-7.0-lowNTCFSSP3-7.0.

We have removed some of the acronyms. In particular, we now use strong and weak air quality control, as well as NTCF mitigation. We also use more straightforward acronyms for the two model subsets, Aer and Aer+O3, as suggested by Reviewer #1.

4. The manuscript cites a lot of "in prep" and "submitted" papers. In most cases, these are cited as part of long lists of other references, so they aren't really needed to make the points. If these are not at least in ACPD by the time of publication, they should not be included in the citation lists (except in cases where they are the only publications available to back-up the point).

All references have been updated.

SPECIFIC COMMENTS

L30: Does the net radiative effect here refer just to OC or to BC+OC? Please rephrase to clarify.

This statement has been clarified. This is the best estimate of net industrial-era climate forcing by all short-lived species from black-carbon-rich sources.

L59: Is this newer estimate of mortality for all air pollution or outdoor ambient only? Please rephrase to clarify.

This is for all air pollution. Clarified.

L90-108: This information would benefit from being summarised in a table listing the scenarios and some of the relevant information (e.g. air quality controls weak/strong,ozone and aerosols high/low, CH4 high/high, etc.) to make it easier for the reader to synthesise.

This information has been moved to this section ("Future Scenarios"), including Figures 1 and 2, which show the global evolution of emission species and the regional trends. Since only two scenarios are addressed in this manuscript, we only show results from SSP3-7.0 and SSP3-7.0-lowNTCF.

L120-122: I find this a bit confusing. What is the difference between CESM2 and CESM2-WACCM in this case? Is it the aerosol treatment? And if they are basically the same model, is it fair to include them as two separate data points in the multi-model means?

We have removed CESM2 from the analysis.

L141-144: So nitrate aerosol was not included in PM2.5 at all, even for the models that do include it? It would be nice to see how much uncertainty this adds, given nitrate can be an important component of aerosol loading in some regions. I'd suggest adding a version of the PM2.5 figures including nitrate to the supplement, and a brief discussion of the impacts of excluding nitrate either in the main text or in the supplement.

Only one model includes nitrate aerosol data (GFDL-ESM4). Globally (over land only), nitrate decreases by -0.0396 (-0.1165) μg/m^3. These trends are 17 and 20% of the magnitude of the corresponding PM2.5 trends. GFDL-ESM4 also archives ammonium. Globally (over land only), ammonium decreases by -0.0487 (-0.1168) μg/m^3. These trends are 21 and 20% of the magnitude of the corresponding PM2.5 trends. Thus, excluding nitrate and ammonium in GFDL-ESM4 leads to ~40% underestimation of the global PM2.5 trend.

CESM2-WACCM also archives ammonium. Here, however, the global (land) trends are much smaller at -0.00329 and -0.0081 μg/m^3, which leads to ~1% underestimation of the global PM2.5 trend.

This has been added to the revision, as have supplementary figures that show the spatial trend maps for nitrate and ammonium. We have also added a discussion and supplementary figures that compare archived versus estimated PM2.5 trends in 4 models (those 4 that included archived PM2.5).

L156: Is there a reference for these ground-based observations? Or is this the same GASSP observations mentioned above? If the latter, please state explicitly in the text.

This is referring to GASSP. Fixed.

L172-L180: This is confusing when paired with the figure. It is completely legitimate to not include the differences in CH4 pathways for this work, but anyone skimming quickly and focusing on the figures will miss that point. In my opinion, Figure 1 should only show what was used in this work, not scenarios that are not used here. I strongly encourage the authors to remove the SSP3-7.0-lowNTCF (right?) and difference lines from Figure 1. The comparison between the scenarios can be moved to the supplement if the authors feel it is important to include.

The SSP3-7.0-lowNTCF and difference $CH_4$ data have been deleted from Figure 1 (same for Figure 2). We have also removed SSP3-7.0-lowNTCF $CO_2$ from these figures. Both sets of AerChemMIP simulations use the same $CO_2$ and $CH_4$ data, based on SSP3-7.0.

L181-187: Similarly, I don't think this discussion belongs here. It is the first section of the results, yet it is mostly discussing what is not done in this work. I would suggest this could be removed entirely, or moved to the supplement or to the conclusions as part of a discussion of what future work should be done to build on what the authors have done here.

This discussion has been removed.

Sect. 3.2 and Figs 3-4: Generally speaking, is the changes in atmospheric composition (aerosols and ozone) that are driving the changes in climate. Thus it seems a bit odd to show and discuss the changes in climate variables BEFORE the changes in composition (ozone, PM2.5). I would suggest restructuring such that Fig.3 comes before Fig. 4, with the text order changed to match. (I note this is already the order used in the abstract and conclusions).

Re-ordered according to the reviewer's suggestion.

L204-205: Unless you rename & define the scenarios in the methods as discussed above, please clarify how "under NTCF mitigation" is defined here (I understand that itis the difference between the two scenarios, but that wasn't clear to me on first read).

Scenarios have been defined according to the reviewer's suggestion. SSP3-7.0 is referred to as weak air quality control and SSP3-7.0-lowNTCF is referred to as strong air quality control. Their difference (strong minus weak air quality control) is referred to as NTCF mitigation.

L211-218, L223-228 (and elsewhere): Much is made of the difference between the lowAER and lowAERO3 outcomes. Given that one of these only includes 3 models and the manuscript states explicitly that the difference is not significant, it is not justifiable to be interpreting this as a result. This appears to me to be over-interpretation of noise, and I would suggest this discussion be removed before publication in ACP.

We agree, and this statement has been modified. It is still interesting that similar (i.e., non-significant differences) global mean surface air temperature trends occur in Aer and Aer+O3 models. We acknowledge that this could be due to several factors, but one interpretation is weak surface cooling due to reductions in ozone.

L233: This land-only result appears to be insignificant for 75% of the models (including those that show increases) and so this statement should be removed or qualified.

Sentence has been deleted.

L238: CDD does not show a statistically significant increase in the overall MMM (or in the subset MMM or in any of the individual models bar one) – therefore should be removed from this sentence.

Deleted.

L255-256: Is the land-only warming pattern shown anywhere? Is the land-only warming weaker or stronger than the overall warming? If there is a difference, it would be useful to see an equivalent figure in the supplement. (And if there is not a difference,it's not clear why this is discussed separately.)

Table 1 shows that the land warming is stronger than that over both land and ocean. This has been clarified. Surface temperature trend patterns are included in the Supplement.

L264: "...forcing and response do not need to occur in the same regions." Can this be explained a bit more?

A sentence has been added.

L269-271: Do I understand Fig 6 bottom panel correctly that models don't agree about this feature? If so that would be worth stating in this discussion

Figure 6 bottom panels show the percentage of models that agree on the sign of the trend. Red colors indicate model agreement on a positive trend; blue colors indicate model agreement on a negative trend. White areas indicate lack of agreement on the sign of the trend. The caption has been clarified. About 70% of the models agree that the North Atlantic cools.

L307: For a discussion of seasonal patterns to make sense, consideration should be given to the different seasonalities of the two hemispheres. Figure 7 should either be separated or at least ordered/demarcated by hemisphere – I'd suggest NH extratropics, tropics, and SH tropics.

Figure 7 shows seasonal trends for each of our 12 world regions. Thus, this figure is already broken down into regional demarcations consistent with seasonality. Nonetheless, we have also added trends for several latitude bands, including those requested.

L370: Why is one model listed explicitly when all (including that model) are available from the same location? Also please spell out ESGF here and provide a link or doi.

Reference to GFDL deleted. ESGF spelled out, and a link is provided.

Figs 2, 5,7: regional legend labels on x-axis are impossible to read because they are so small. Perhaps give each region a number instead? Or include some other sort of key to make this clearer?

We have modified the x-axis on these figures. A key is now used.

Fig 3 caption (and elsewhere): Does “hottest day” refer to “surface temperature on hottest day”? Similarly for wettest day? Please clarify somewhere.

Extreme weather indices are defined in the Methodology section.

We also analyze climate extremes including the hottest day (monthly maximum value of daily maximum surface temperature), wettest day (monthly maximum 1-day surface precipitation) and consecutive dry days (CDD), defined as the maximum annual number of consecutive days with surface precipitation less than 1 mm/day. We focus on these three extreme indices since they are frequently used metrics for temperature and precipitation extremes. Prior observational analyses have shown significant increases (decreases) in the hottest and wettest day (CDD) over the latter half of the 20th century (Donat et al, 2013a,b). Climate extremes are based on daily data, and are calculated at each grid box and then spatially averaged.

Fig 3 caption: It seems the thin coloured lines show the trends for the individual models, but this has not been explicitly stated in the caption. Please update caption to clarify.

Thin (and non-black lines) show individual model mean trends. Line colors are denoted by the legend. We have also added this to the caption. Same for Figure 4.

Figs 3, 4, 5: why are different units used for the trends in the precip variables (mm/dayvs. %) in the global and regional trend figures? Same question for PM2.5 and O3. Can these be standardised to more easily compare?

Sure. We no longer use percent changes.

Fig 6d-f: these plots are not currently discussed in the text and therefore should per-haps move to the supplement (or be mentioned in the text)

A sentence pertaining to these panels has been added.

Table 1: I found this table hard to understand while trying to refer to it while reading the text. A few suggestions to improve the clarity. (1) Add lowAER and lowAERO3 identifiers above the list of relevant models in each sub-section so it’s easy to see which group is which. (2) If text and figures are re-ordered as suggested above, move PM2.5 and O3 columns to be left-most, followed by the climate variables. (3) Move the three “MMM total” lines either to a separate part of the table or (preferably) to a new table altogether as the numbers aren’t comparable to the lines above/below which makes it difficult to interpret (and already a lot to interpret in the table!). (4) For the lowAER models’ O3 response, replace 0.0 with n/a since these values are not included in the Overall MMM calculation (as is, looks like the overall will be an average of thelowAERO3 values and three zeros).

Made all suggested modifications to Table 1.

TECHNICAL COMMENTS

L139: “were are” –> “we are”

L357: “complex” –> “complexity”

Fig 2 caption: “astriks” –> “asterisks”

All have been fixed.

# Climate and air quality impacts due to mitigation of non-methane near-term climate forcers

Robert J. Allen[1], Steven Turnock[2], Pierre Nabat[3], David Neubauer[4], Ulrike Lohmann[4], Dirk Olivié[5], Naga Oshima[6], Martine Michou[3], Tongwen Wu[7], Jie Zhang[7], Toshihiko Takemura[8], Michael Schulz[5], Kostas Tsigaridis[9], Susanne E. Bauer[9], Louisa Emmons[10], Larry Horowitz[11], Vaishali Naik[11], Twan van Noije[12], Tommi Bergman[12,13], Jean-Francois Lamarque[14], Prodromos Zanis[15], Ina Tegen[16], Daniel M. Westervelt[17], Phillipe Le Sager[12], Peter Good[2], Sungbo Shim[18], Fiona O'Connor[2], Dimitris Akritidis[15], Aristeidis K. Georgoulias[15], Makoto Deushi[6], Lori T. Sentman[11], Jasmin G. John[11], Shinichiro Fujimori[19,20,21], and William J. Collins[22]

[1]Department of Earth and Planetary Sciences, University of California Riverside, Riverside, CA, 92521 USA
[2]Met Office Hadley Centre, Exeter, UK
[3]Centre National de Recherches Meteorologiques (CNRM), Universite de Toulouse, Meteo-France, CNRS, Toulouse, France
[4]Institute of Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[5]Norwegian Meteorological Institute, Oslo, Norway
[6]Meteorological Research Institute, Japan Meteorological Agency
[7]Beijing Climate Center, China Meteorological Administration, Beijing, China
[8]Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan
[9]Center for Climate Systems Research, Columbia University, NASA Goddard Institute for Space Studies, USA
[10]Atmospheric Chemistry Observations and Modelling Lab, National Center for Atmospheric Research, Boulder, CO, USA
[11]DOC/NOAA/OAR/Geophysical Fluid Dynamics Laboratory. Biogeochemistry, Atmospheric Chemistry, and Ecology 10 Division, Princeton, USA
[12]Royal Netherlands Meteorological Institute, De Bilt, Netherlands
[13]Finnish Meteorological Institute, Helsinki, Finland
[14]NCAR/UCAR, Boulder, CO, USA
[15]Department of Meteorology and Climatology, School of Geology, Aristotle University of Thessaloniki, Thessaloniki, Greece
[16]Leibniz Institute for Tropospheric Research, Leipzig, Germany
[17]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA
[18]National Institute of Meteorological Sciences, Seogwipo-si, Jeju-do, Korea
[19]Department of Environmental Engineering, Kyoto University, C1-3 361, Kyotodaigaku Katsura, Nishikyoku, Kyoto city, Japan
[20]Center for Social and Environmental Systems Research, National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba, Ibaraki, 305-8506, Japan
[21]International Institute for Applied System Analysis (IIASA), Schlossplatz 1, A-2361, Laxenburg, Austria
[22]Department of Meteorology, University of Reading, Reading, RG6 6BB, UK

**Correspondence:** Robert J. Allen (rjallen@ucr.edu)

**Abstract.** Over the next few decades, policies that optimally address both climate change and air quality are essential. Although targeting near-term climate forcers (NTCFs), defined here as aerosols, tropospheric ozone and precursor gases, should improve air quality, NTCF reductions will also impact climate. **The climate impacts of future policies that address the abundance of NTCFs and air quality remain uncertain.** Here, we quantify the 2015-2055 climate and air quality effects of non-methane NTCFs using state-of-the-art chemistry-climate model simulations conducted for the Aerosol and Chemistry Model Intercomparison Project (AerChemMIP). Simulations are driven by two future scenarios featuring similar increases in greenhouse gases (GHGs) but with “weak" (SSP3-7.0) versus “strong" (SSP3-7.0-lowNTCF) levels of air quality control measures. Unsurprisingly, we find significant improvements in air quality under NTCF mitigation (strong versus weak air quality controls). Surface fine particulate matter ($PM_{2.5}$) and ozone ($O_3$) decrease by $-2.2$ $\mu$g m$^{-3}$ and $-4.6$ ppb, respectively, over global land surfaces, with larger reductions in some regions including south and southeast Asia. Non-methane NTCF mitigation, however, leads to additional climate change due to the removal of aerosol which causes a net warming effect, including global mean surface temperature and precipitation increases of 0.25K and 0.03 mm day$^{-1}$, respectively. **Similarly, increases in extreme weather indices, including the hottest and wettest day, also occur.** Regionally, the largest warming and wetting occurs over Asia, including central and north Asia (0.66K and 0.03 mm day$^{-1}$), south Asia (0.47K and 0.17 mm day$^{-1}$) and east Asia (0.46K and 0.15 mm day$^{-1}$). Relatively large warming and wetting of the Arctic also occur at 0.59K and 0.04 mm day$^{-1}$, respectively. Similar surface warming occurs in model simulations with aerosol-only mitigation, implying weak cooling due to ozone reductions. Our findings suggest that future policies that aggressively target non-methane NTCF reductions will improve air quality, but will lead to additional surface warming, particularly in Asia and the Arctic. Policies that address other NTCFs including methane, as well as carbon dioxide emissions, must also be adopted to meet climate mitigation goals.

## 1 Introduction

Near-term climate forcers (NTCFs) are **those chemical species whose impact on climate occurs primarily within the first decade after their emission** (Myhre et al., 2013). This set of compounds includes ozone, aerosols, and their precursor gases, as well methane ($CH_4$) which is also a well-mixed greenhouse gas (GHG). Other well-mixed GHGs, including carbon dioxide ($CO_2$) and nitrous oxide ($N_2O$), possess much longer atmospheric lifetimes and impact climate on decadal to centennial time scales.

NTCFs have important impacts on the climate system and human health, as they perturb the radiative balance of Earth and contribute to air pollution. The total aerosol radiative effect, estimated as an effective radiative forcing (ERF), is $-0.9$ W m$^{-2}$ with a 90% confidence range of $-1.9$ to $-0.1$ W m$^{-2}$ (Boucher et al., 2013). A more recent review revised the 90% confidence range to more negative values (**$-$2.0 to $-$0.4 W m$^{-2}$**) (Bellouin et al., 2020). Moreover, not all aerosols have a negative forcing, as black carbon (BC) from anthropogenic fossil and biofuel emissions possesses a radiative forcing of +0.40 (0.05 to 0.80) W m$^{-2}$. BC, however, is often associated with co-emission of organic matter. **The best estimate of net industrial-era climate forcing by all short-lived species from black-carbon-rich sources, including open burning emissions, is slightly negative but with relatively large uncertainty bounds of $-$1.45 to +1.29 W m$^{-2}$ (Bond et al., 2013).** Thus, changes in BC emissions that are different from changes in non-absorbing aerosols will lead to differing ERF changes. **Tropospheric ozone, which is formed in the atmosphere through chemical reactions between nitrogen oxides ($NO_x$), carbon monoxide (CO), and volatile organic compounds (VOCs) including methane in the presence of sunlight, also exhibits a positive forcing of +0.40$\pm$0.2 W m$^{-2}$ (Myhre et al., 2013). The radiative forcing of changes in methane concentrations is estimated**

**at 0.48±0.05 W m$^{-2}$** (Myhre et al., 2013). We note that these estimates are currently being updated as part of the Coupled Model Intercomparison Project version 6 (CMIP6) (Pincus et al., 2016; Eyring et al., 2016). Thus, reductions in some NTCFs, including non-absorbing aerosols, will warm the climate system, whereas reductions in other NTCFs, including absorbing aerosols, tropospheric ozone, and methane will cool the climate system. **Things become more complex from an emissions perspective, as reductions in some precursor gases such as $NO_x$ and VOCs impact ozone, methane and aerosols (Myhre et al., 2013). Reductions in $NO_x$, for example, will promote cooling due to reduced tropospheric ozone, but the impact on $CH_4$ lifetime and aerosol formation will likely promote overall warming (Fiore et al., 2015).**

NTCFs also perturb the hydrological cycle. Energetic constraints and modeling studies show that anthropogenic aerosols lead to reduced global mean precipitation (Ramanathan et al., 2001; Wilcox et al., 2013; Samset et al., 2016). Aerosol induced reductions in surface solar radiation will be partially balanced by reductions in **latent cooling**, leading to corresponding rainfall reductions. In the case of absorbing aerosols−particularly in the boundary layer−atmospheric heating stabilizes the atmosphere and reduces convection, also leading to an overall decrease in precipitation (Ming et al., 2010; Ban-Weiss et al., 2012; Stjern et al., 2017; Allen et al., 2019a; Johnson et al., 2019). The buildup of aerosols during the 20th century has likely masked the expected increase in global mean precipitation due to GHG-induced warming (Liepert et al., 2004; Wu et al., 2013; Salzmann, 2016; Richardson et al., 2018). Furthermore, the hemispheric contrast in aerosol forcing has likely shifted the tropical rainbelt southward, which is associated with a weakening of the west African Monsoon and the occurrence of the Sahel drought of the mid-1980s (Rotstayn and Lohmann, 2002; Biasutti and Giannini, 2006; Allen and Sherwood, 2011; Ackerley et al., 2011; Chang et al., 2011; Biasutti, 2013; Hwang et al., 2013; Dong et al., 2014; Allen et al., 2015; Undorf et al., 2018). The observed precipitation decrease during recent decades over most of the areas affected by the South and East Asia monsoon can also be explained by the dominance of aerosol radiative effects suppressing precipitation over the expected precipitation enhancement due to increased GHGs (Wang et al., 2013; Song et al., 2014; Li et al., 2015; Xie et al., 2016; Krishnan et al., 2016; Guo et al., 2016; Lau and Kim, 2017; Zhang et al., 2017; Lin et al., 2018; Liu et al., 2018).

NTCFs are also a source of air pollution, including surface ozone ($O_3$) and fine particulate matter less than 2.5 $\mu$m in diameter ($PM_{2.5}$). Air pollution has negative impacts on human health, including exacerbation of cardiovascular and respiratory diseases, and cancer. Recent estimates show air pollution is the 4th-highest ranking risk factor for death globally, responsible for $\sim$7 million premature deaths per year, with 4.2 million of these annual deaths attributable to ambient air pollution (WHO, 2016; Cohen et al., 2017; Butt et al., 2017). **A more recent study suggests the global total excess mortality rate due to all air pollution is 8.79 million per year (95% confidence interval of 7.11-10.41 million per year), leading to a global mean loss of life expectancy of 2.9 years (Lelieveld et al., 2019).**

Future reductions in emissions of NTCFs are necessary for improved air quality, but will likely yield relatively rapid (i.e., decadal) climate responses due to their short atmospheric lifetimes (relative to GHGs). Samset et al. (2018) show that complete removal of present-day anthropogenic aerosol emissions induces a global mean surface heating of 0.5-1.1K and a precipitation increase of 2-4.6%. Similar large, near-term increases in global warming and precipitation are predicted by other studies that assume a rapid removal of anthropogenic aerosols (Brasseur and Roeckner, 2005; Andreae et al., 2005; Ramanathan and Feng, 2008; Raes and Seinfeld, 2009; Kloster et al., 2010; Arneth et al., 2009; Matthews and Zickfeld, 2012; Rotstayn et al., 2013;

Wu et al., 2013; Westervelt et al., 2015; Salzmann, 2016; Hienola et al., 2018; Richardson et al., 2018; Lelieveld et al., 2019). Furthermore, future aerosol reductions will likely shift the tropical rainbelt northward and may strengthen precipitation in several monsoon regions, including West Africa, South Asia, and East Asia (Levy et al., 2013; Allen, 2015; Rotstayn et al., 2015; Allen and Ajoku, 2016; Westervelt et al., 2017; Zhao et al., 2018; Westervelt et al., 2018; Scannell et al., 2019; Zanis et al., 2020). In contrast to the above studies, however, Shindell and Smith (2019) show that the time required to transform power generation, industry and transportation leads to largely offsetting climate impacts of $CO_2$ and sulfur dioxide (a precursor of sulfate aerosol), implying no conflict between climate and air-quality objectives. Their simulations use a simple emissions-based climate model, Finite Amplitude Impulse Response (FAIR) (Smith et al., 2018), and it is not known if this result also applies to fully coupled chemistry-climate models.

Despite the rich literature, the impact of NTCF mitigation on climate and air quality remains uncertain. Part of this uncertainty stems from the idealized nature of many of the prior studies (e.g., instantaneous removal of all aerosols), simplified treatment of aerosols and chemically reactive gases, as well as a lack of a sufficiently large number of models performing identical simulations with which to quantify model diversity and robust responses. **The Aerosol and Chemistry Model Intercomparison Project (AerChemMIP) (Collins et al., 2017), part of CMIP6 (Eyring et al., 2016), quantifies the climate and air quality impacts of aerosols and chemically reactive gases. Here, we use AerChemMIP and the Scenario Model Intercomparison Project (ScenarioMIP, O'Neill et al., 2016) to quantify the climate and air quality impacts due to non-methane NTCF mitigation (aerosols and ozone only) through analysis of two future emission scenarios−one with weak (SSP3-7.0) and one with strong (SSP3-7.0-lowNTCF) levels of air quality control measures. NTCF mitigation is defined here as the difference between these two scenarios, SSP3-7.0-lowNTCF − SSP3-7.0.** Models include an interactive representation of tropospheric aerosols and atmospheric chemistry, allowing for the quantification of chemistry-climate interactions. We show that non-methane NTCF reductions improve air quality, but also lead to additional climate change including surface warming. Policies that address other NTCFs including $CH_4$, as well as $CO_2$ emissions, must also be undertaken. Methods are presented in Section 2 and results are discussed in Section 3. Conclusions appear in Section 4.

## 2 Methods

### 2.1 AerChemMIP Models

**Nine coupled ocean-atmosphere climate models performed the SSP3-7.0 and SSP3-7.0-lowNTCF simulations, including CNRM-ESM2-1 (Séférian et al., 2019; Michou et al., 2019), MIROC6 (Takemura et al., 2005, 2009; Tatebe et al., 2019), MPI-ESM1-2-HAM (Mauritsen et al., 2019; Neubauer et al., 2019; Tegen et al., 2019), NorESM2-LM (Seland et al., 2020), BCC-ESM1 (Wu et al., 2019, 2020), GFDL-ESM4 (John et al., 2018; Horowitz et al., 2018; Dunne et al., submitted; Horowitz et al., submitted), CESM2-WACCM (Emmons et al., 2020; Gettelman et al., 2019; Tilmes et al., 2019), UKESM1-0-LL (Sellar et al., 2019) and MRI-ESM2-0 (Yukimoto et al., 2019).** However, the first four models (CNRM-ESM2-1, MIROC6, MPI-ESM1-2-HAM, NorESM2-LM) lack interactive tropospheric chemistry schemes and therefore include identical ozone evolution in both SSP3-7.0 and SSP3-7.0-lowNTCF simulations (as recommended by AerChemMIP).

**As NTCF mitigation only includes the effects of aerosols in these four models, we refer to these models as "Aer". The remaining five models, including BCC-ESM1, GFDL-ESM4, CESM2-WACCM, UKESM1-0-LL and MRI-ESM2-0, include interactive atmospheric chemistry and aerosols, and therefore both aerosol and ozone reductions are included. These models are referred to as "Aer+O3".**

In addition to coupled simulations, models also performed analogous fixed-SST experiments to quantify the effective radiative forcing (ERF). The ERF is calculated from the top-of-the-atmosphere (TOA) flux differences between atmosphere-only simulations with identical SSTs but differing composition (Forster et al., 2016; Pincus et al., 2016). The above scenarios (SSP3-7.0 and SSP3-7.0-lowNTCF) are repeated with prescribed SSTs. These SSTs (and sea ice) are taken from the monthly mean evolving values from one ensemble member of the coupled SSP3-7.0 ScenarioMIP run (Collins et al., 2017). MPI-ESM1-2-HAM used daily mean SST and sea ice. The differences in radiative fluxes between the weak and strong air quality control scenarios yield the TOA transient ERF due to NTCF mitigation.

### 2.2 Model Data and Methodology

All models performed at least one realization each of SSP3-7.0 and SSP3-7.0-lowNTCF. **CNRM-ESM2-1, MIROC6, UKESM1-0-LL, NorESM2-LM, CESM2-WACCM and BCC-ESM1 performed three realizations of each experiment.** For these models, the model mean response (average over the three realizations) is shown. **The multi-model mean (MMM) is obtained by averaging each model's mean response (i.e., each model has the same weight).** Only one realization exists for the corresponding fixed-SST experiments. Unless otherwise mentioned, all analyses are based on annual means. All data is spatially interpolated to a 2.5°x2.5° grid using bilinear interpolation.

**Model trends are calculated using least-squares regression, and the corresponding trend significance is based on a two-tailed Student's $t$-test, where the null hypothesis of a zero regression slope is evaluated. Multi-model mean trends and their significance are calculated using two different methods. In the first method, MMM trends are calculated from the multi-model mean time series using a weighted least-squares regression, where each value in the multi-model mean time series is weighted by $1/\sigma_m^2$, where $\sigma_m$ is the standard deviation across models. We note that the MMM trends and significance are very similar with and without weighting the regression. Autocorrelation of the time series is also accounted for by using the effective sample size, defined as $n(1-r_1)/(1+r_1)$, where $n$ is the number of years and $r_1$ is the lag-1 autocorrelation coefficient.**

**We also quantify the significance of the multi-model mean trend relative to each individual model mean trend. Here, the MMM trend is calculated as the average of the individual model mean trends and its uncertainty is calculated as plus/minus twice the standard error (i.e., the 95% confidence interval), which is $2\sigma/\sqrt{n}_m$, where $\sigma$ is the standard deviation of the trends and $n_m$ is the number of models. If this confidence interval does not include zero, then the multi-model mean trend is significant at the 95% confidence level. Both methods yield similar conclusions as to the magnitude and significance of the MMM trends.**

Climate variables analyzed include monthly mean surface temperature (Ts) and precipitation (Precip). **Surface temperature and precipitation are analyzed as these are arguably two of the most important climate variables. Changes in surface**

**temperature are particularly relevant in the context of climate mitigation, as the goal of the Paris Agreement is to keep the increase in global mean surface temperature to well below 2°C above preindustrial values (IPCC, 2018). Precipitation, and fresh water resources in general, are important to both human society and ecosystems.**

**As discussed in the Introduction, both $PM_{2.5}$ and ozone are commonly used indicators of air quality, and both have been associated with adverse human health impacts (WHO, 2016; Cohen et al., 2017; Butt et al., 2017). Air quality is therefore quantified from surface $PM_{2.5}$ and surface $O_3$.** These monthly mean fields are obtained from the model level closest to the surface. Unfortunately, few models archived sub-monthly aerosol or ozone data, so **we are** unable to analyze changes in daily or sub-daily maximum $PM_{2.5}$ or $O_3$ pollution. **Furthermore, only four models directly archive $PM_{2.5}$ (with differing methodologies), and not all models include the same aerosol species (e.g., nitrate aerosol; Supplement).** Thus, we approximate $PM_{2.5}$ in all models using the following equation (Fiore et al., 2012; Silva et al., 2017): $PM_{2.5}$ = BC + OA + $SO_4$ + 0.1xDU + 0.25xSS, where BC is black carbon, OA is organic aerosol, $SO_4$ is sulfate aerosol, DU is dust and SS is sea salt. This formula assumes 100% of the BC, OA and $SO_4$ is fine mode, whereas 25% of the sea salt and 10% of the dust is fine mode. The SS and DU factors are likely dependent on the model and its size distribution. In the case of CNRM-ESM2-1, sensitivity tests were used to estimate a much smaller SS factor of 0.01. This smaller factor addresses the large SS size range of up to 20 $\mu$m in this model (P. Nabat 2019, personal communication, November 27th). **Although this approach likely introduces some uncertainties (see Section 3.1), it provides first and foremost an estimate of $PM_{2.5}$ for all models, as well as a consistent estimate for all models.**

CMIP6 model evaluation of air quality metrics, including surface $O_3$ and $PM_{2.5}$ (as approximated here), is quantified in a companion paper (Turnock et al., 2020). To summarize, CMIP6 models generally underestimate $PM_{2.5}$ over most regions relative to ground based observations from the Global Aerosol Synthesis and Science Project (GASSP) (Reddington et al., 2017). This in part is due to the absence of nitrate aerosol, and may also be related to misrepresentation of secondary organic aerosol. A similar $PM_{2.5}$ underestimation occurs over Europe and North America relative to the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2) aerosol reanalysis product (Buchard et al., 2017; Randles et al., 2017). **In contrast, CMIP6 models overestimate $PM_{2.5}$ relative to MERRA2 over south and east Asia, contrary to the evaluation using GASSP observations**. Compared to surface $O_3$ measurements from Tropospheric Ozone Assessment Report (TOAR) (Schultz et al., 2017), CMIP6 models consistently overestimate surface ozone during both summer and winter across most regions, potentially due to the coarse resolution of global models simulating excess $O_3$ production.

**Perhaps more important than changes in the mean of a climate variable are changes in its extremes. Heat waves, for example, are a major cause of weather-related fatalities. Thus, we also analyze climate extremes including the hottest day (monthly maximum value of daily maximum surface temperature), wettest day (monthly maximum 1-day surface precipitation) and consecutive dry days (CDD), defined as the maximum annual number of consecutive days with surface precipitation $< 1$ mm day$^{-1}$. We focus on these three extreme indices since they are frequently used metrics for temperature and precipitation extremes. Prior observational analyses have shown significant increases in the hottest and wettest day, and decreases in CDD over the latter half of the 20th century (Donat et al., 2013a, b). Climate extremes are based on daily data, and are calculated at each grid box and then spatially averaged.**

## 2.3 Future Scenarios: SSP3-7.0 and SSP3-7.0-lowNTCF

**As part of ScenarioMIP, a set of future emissions pathways have been developed for CMIP6 (Eyring et al., 2016). These scenarios, referred to as Shared Socio-economic Pathways (SSPs) (O'Neill et al., 2014; van Vuuren et al., 2014; Gidden et al., 2019), link socioeconomic and technological innovation to provide future trajectories of emissions, including different levels of controls on air quality pollutants. The medium strength of pollution control corresponds to current legislation (CLE) until 2030 and progresses three-quarters of the way towards maximum technically feasible reduction (MTFR) thereafter. Strong pollution control exceeds CLE and progresses ultimately towards MTFR. Weak pollution control assumes delays to the implementation of CLE and makes less progress towards MTFR than the medium scenario (Rao et al., 2017). The rate of progress is different for high, medium and low-income countries. By encompassing a wide range of possible futures, these scenarios provide a large sample space of potential emissions through the 21st century.**

**To detect the largest signal, AerChemMIP uses the SSP3-7.0 "Regional Rivalry" without climate policy ($\sim$7.0 W m$^{-2}$ at 2100) (Fujimori et al., 2017) as the reference scenario, which has the highest levels of NTCFs and "weak" levels of air quality control measures (O'Neill et al., 2014; Rao et al., 2017). The perturbation scenario SSP3-7.0-lowNTCF uses the same socio-economic scenario, but with "strong" levels of air quality control measures (Gidden et al., 2019). Basically, the emissions drivers (population, GDP, energy and land-use) are based on SSP3, but the emissions factors of air pollutants that are related to NTCFs are associated with a Sustainability pathway represented by SSP1 in conjunction with the stringent climate policy equivalent of stabilizing the radiative forcing to around 2.6 W m$^{-2}$. Assumptions include the following: SSP3-7.0-lowNTCF can reduce $CH_4$ as if SSP1's stringent climate mitigation policy is implemented in the SSP3 world; SSP1's air pollutant legislation and technological progress can be achieved in the SSP3 world; other species (e.g., CFCs, HFCs and $SF_6$) are identical to the SSP3 baseline. Although methane reductions are included in the lowNTCF scenario, they are not included in the lowNTCF experiment. Differences between these two scenarios are designed to evaluate a SSP3 world in which NTCF-related policies are enacted in the absence of other GHG-related climate policies. We note that this is a bit of an idealization, as it is difficult to reduce only the NTCF emissions while keeping $CO_2$ emissions fixed (since there are co-emitted species, including $SO_2$). Moreover, our results (e.g., the magnitude of the surface temperature increase) likely represent an upper bound as our baseline scenario lacks climate policy and contains the highest levels of NTCFs**.

Differences in climate, effective radiative forcing, chemical composition and air quality between the two scenarios will be solely due to the alternative air quality control measures. These experiments cover the time frame from 2015 to 2055, as this is when reductions in aerosol and ozone precursor emissions are expected to be significant, particularly in some world regions. Here, we define NTCF mitigation as the difference between the strong (low NTCF) and weak (high NTCF) air quality control scenarios (i.e., SSP3-7.0-lowNTCF minus SSP3-7.0). Although methane reductions are included in the strong air quality control scenario, AerChemMIP protocol specifies unchanged levels of WMGHGs, including methane, between the strong and the weak air quality control simulations (Collins et al., 2017). Thus, our results quantify non-methane NTCF mitigation (aerosols and ozone only).

Figure 1 shows the 2015-2055 global mean time series of $CO_2$, aerosol species and gaseous precursor emissions for SSP3-7.0 (weak air quality control) and SSP3-7.0-lowNTCF (strong air quality control). **Emissions shown here comes directly from the CMIP6 forcing datasets, which were downloaded from the input datasets for Model Intercomparison Project (input4MIPS) served by the Earth System Grid Federation. We note that the emissions data is decadal after 2015, with monthly values for the year 2015, 2020, 2030, 2040, 2050, 2060, etc. We estimate the emissions in 2055 as the mean of the emissions in 2050 and 2060 at each grid box. Only weak air quality control $CO_2$ and $CH_4$ emissions are shown, as AerChemMIP simulations include the same change in $CO_2$ and $CH_4$ emissions based on the weak air quality control scenario. By 2055, $CO_2$ and $CH_4$ increase by 65% and 50% (relative to 2015), respectively.** In contrast to $CO_2$ and $CH_4$, however, very different non-methane NTCF evolution occurs. Under weak air quality control, global emissions of all aerosols and gaseous precursors (except $SO_2$) increase by 5-15% by 2055. In contrast, strong air quality control yields strong emission reductions in all species, ranging from ~30% for VOCs to 55% for $SO_2$. Thus, NTCF mitigation (SSP3-7.0-lowNTCF−SSP3-7.0) yields emission reductions of all aerosols and gaseous precursors by ~40-55%.

The corresponding 2015-2055 regional emission trends (relative to 2015) are shown in Figure 2. **As with climate and air quality trends, emission trends are estimated using least-squares regression. Consistent with the global mean time series of emissions (Fig. 1), $CO_2$ emissions increase under weak air quality control (and in both sets of AerChemMIP simulations), with larger increases in south and north Africa, south Asia, and southeast Asia. Similarly, $CH_4$ emissions increase in all world regions under weak air quality control (and in both sets of AerChemMIP simulations), with larger increases in south and north Africa, and south Asia.** Most world regions also show increases in BC, $SO_2$ and organic carbon (OC) under weak air quality control, but strong decreases under strong air quality control. NTCF mitigation (strong minus weak air quality control) shows large (~20% decade$^{-1}$) BC decreases in central America, central and north Asia, east Asia and southeast Asia. Most world regions exhibit a 10-20% decade$^{-1}$ reduction in $SO_2$ emissions under NTCF mitigation, with a large decrease in south Asia at −28% decade$^{-1}$. Similarly, OC and CO emissions decrease by ~10-20% decade$^{-1}$. Relatively large OC reductions also occur in east Asia, south Asia and southeast Asia. $NO_x$ and VOC emissions also decrease in all world regions under NTCF mitigation (although this is only a decrease relative to non-mitigated emissions for $NO_x$ in south Asia and for VOC in east Asia).

# 3 Results

## 3.1 Global Climate and Air Quality Trends

Figure 3 shows the 2015-2055 global annual mean time series for air quality under NTCF mitigation. By design, NTCF mitigation leads to significant decreases in air pollution, in terms of both surface $PM_{2.5}$ and $O_3$. All models yield significant global mean decreases in both quantities, with an overall MMM decrease of $-0.23\ \mu g\ m^{-3}$ decade$^{-1}$ for $PM_{2.5}$ and −1.19 ppb decade$^{-1}$ for $O_3$ (Table 1). Over the 2015-2055 year time period, these rates of change correspond to global mean decreases of

$-0.92$ $\mu$g m$^{-3}$ and $-4.76$ ppb, respectively. Larger $PM_{2.5}$ decreases occur over land only at $-2.20$ $\mu$g m$^{-3}$, whereas similar $O_3$ decreases occur over land only at $-4.55$ ppb. **Similar $PM_{2.5}$ trends occur in Aer+O3 and Aer models over land only ($-$0.59 versus $-$0.44 $\mu$g m$^{-3}$ decade$^{-1}$, respectively), as well as over both land and ocean ($-$0.26 versus $-$0.16 $\mu$g m$^{-3}$ decade$^{-1}$, respectively.)** Note that the MMM over all models for $O_3$ does not include Aer models, as they yield negligible change in surface ozone (by design).

**As mentioned in the Methods section, as only four models directly archive $PM_{2.5}$ (with differing methodologies), and not all models include the same aerosol species (Supplement), we approximate $PM_{2.5}$. Comparing estimated $PM_{2.5}$ trends to those from the actual $PM_{2.5}$ as calculated and archived by four models (GFDL-ESM4, NorESM2-LM, MRI-ESM2-0 and MPI-ESM1-2-HAM) yields reasonably good results. The global annual multi-model mean trend in estimated (actual) $PM_{2.5}$ for this four model subset is $-$0.24 ($-$0.28) $\mu$g m$^{-3}$ decade$^{-1}$ (Supplementary Figure 1). Over land only, the corresponding trends are $-$0.56 ($-$0.65) $\mu$g m$^{-3}$ decade$^{-1}$. Thus, the estimated global mean and land-only $PM_{2.5}$ trends are about 85% as large as those based on archived $PM_{2.5}$ (underestimation by a similar amount exists in all four models, with the largest underestimation in GFDL-ESM4). Larger differences exist in some world regions, particularly south Asia, where the estimated (actual) $PM_{2.5}$ is $-$4.08$\pm$0.70 ($-$4.71$\pm$1.36) $\mu$g m$^{-3}$ decade$^{-1}$ (Supplementary Figure 2). Some of this underestimation is due to the aforementioned lack of nitrate and ammonium aerosol in our estimated $PM_{2.5}$. However, other factors also contribute, as the estimated $PM_{2.5}$ trends in all four models underestimate the actual $PM_{2.5}$ trends, but not all of these models include nitrate and ammonium species.**

**GFDL-ESM4 is the lone model that archived nitrate aerosol data. Globally (over land only), nitrate decreases by $-$0.04 ($-$0.12) $\mu$g m$^{-3}$ decade$^{-1}$, with maximum decrease over east Asia and in particular south Asia (Supplementary Figure 3). These trends are 17 and 20% (13 and 15%) of the magnitude of the estimated (actual) global and land-only $PM_{2.5}$ trend. GFDL-ESM4 also archives ammonium, and similar changes occur (Supplementary Figure 3). Globally (over land only), ammonium decreases by $-$0.05 ($-$0.12) $\mu$g m$^{-3}$ decade$^{-1}$, with maximum decreases over both south Asia and east Asia. These trends are 21 and 20% (16 and 15%) of the magnitude of the estimated (actual) global and land-only $PM_{2.5}$ trend. Thus, excluding nitrate and ammonium in GFDL-ESM4 leads to $\sim$30-40% underestimation of the global and land-only PM2.5 trend. The relatively large decreases in nitrate and ammonium in south Asia helps to explain the relatively large difference in estimated and actual $PM_{2.5}$ trend in this region (Supplementary Figure 2). In addition to GFDL-ESM4, CESM2-WACCM also archives ammonium (Supplementary Figure 3). Here, however, the global and land-only ammonium trends are an order of magnitude smaller than those in GFDL-ESM4, which leads to $\sim$1% underestimation of the corresponding (estimated) PM2.5 trends.**

The 2015-2055 global annual mean time series for climate variables under NTCF mitigation are shown in Figure 4. All but one model (MIROC6) shows significant global annual mean surface warming in response to NTCF mitigation (**Table 1 lists the trends for each model**). Averaged over all models, global mean surface warming is 0.06 K decade$^{-1}$, or 0.25 K over the 2015-2055 time period (Table 1). We note that this warming will continue past 2055, as these transient simulations have not reached radiative equilibrium. Similar conclusions exist over land only, where the multi-model mean (MMM) warming is even larger at 0.36 K over the entire time period (Table 1). Enhanced land warming is consistent with the land-sea warming contrast (Sutton et al., 2007; Joshi et al., 2008), which may also act to increase aerosol burden itself (Allen et al., 2016, 2019b), implying a climate change penalty to air quality. **Interestingly, models that include both aerosol and ozone reductions (Aer+O3) yield similar surface warming relative to the models that include aerosol reductions (Aer) alone (0.07 versus 0.06 K decade$^{-1}$, respectively). Although this could be due to several factors (e.g., small sample size, internal climate variability, different model parameterizations, feedbacks, etc.) it suggests weak surface cooling due to reductions in ozone. Such an interpretation is consistent with the negative forcing from aerosol increases dominating the positive forcing due to ozone increases over the historical period (Naik et al., 2013). Simulations with a single model, running both coupled and uncoupled chemistry experiments, would help isolate this effect.**

Warming in response to NTCF mitigation is consistent with the corresponding increase in ERF. All but two models (BCC-ESM1, GFDL-ESM4) yield a significant increase in ERF, with a MMM of 0.44 W m$^{-2}$ over the entire time period (Table 1). Over land only, this increases to 0.59 W m$^{-2}$. Although not significant, Aer+O3 models yield a weaker trend in global mean ERF than Aer models, at 0.07 versus 0.17 W m$^{-2}$ decade$^{-1}$. This is consistent with ozone reductions driving a decrease in ERF in Aer+O3 models, offsetting part of the ERF increase due to aerosol reductions (Turnock et al., 2019).

All models also yield a significant increase in global annual mean precipitation (Table 1), with an overall MMM of 0.008 mm day$^{-1}$ decade$^{-1}$. Aer+O3 and Aer models yield similar increases in global mean precipitation at 0.009 and 0.005 mm day$^{-1}$ decade$^{-1}$, respectively. Somewhat less robust results occur over land only. Although all models yield an increase in precipitation over land, it is only significant in four models.

Similar, but less robust responses also occur in climate extremes, particularly those based on precipitation. Globally significant increases in the surface temperature of the hottest day occur in all but one model (MIROC6 is the exception). The multi-model mean also yields a significant trend at 0.06 K decade$^{-1}$. The wettest day significantly increases in about half of the models and in the overall MMM at 0.053 mm day$^{-1}$ decade$^{-1}$. A mixed signal exists for CDD, with four models yielding a positive trend and three models yielding a negative trend. The overall MMM yields 0.08 days per year decade$^{-1}$, but lacks significance.

Thus, from a global mean perspective, NTCF mitigation leads to significant improvements in air quality based on both $PM_{2.5}$ and $O_3$, but also significant climate change in most metrics. This includes increases in surface temperature and precipitation, as well as corresponding increases in most climate extremes, particularly the hottest day and to lesser extent the wettest day. Except for surface temperature and the hottest day, less robust results generally occur over land only. CDD yields a mixed signal, with lack of significance in the multi-model mean.

### 3.2 Regional Climate and Air Quality Trends

Figure 5 shows the regional climate and air pollution trends for weak and strong air quality control and the effect of NTCF mitigation. We include both Aer and Aer+O3 models in this analysis to maximize the signal to noise ratio (except for ozone changes). The aforementioned response differences between these two model subsets are generally not significant. Consistent with increased aerosol and precursor gas emissions (Figures 1-2), air quality metrics generally show significant increases under weak air quality control, particularly $O_3$ where all 12 world regions exhibit an increase. In contrast, strong air quality control yields decreases in both $PM_{2.5}$ and $O_3$ for nearly all world regions. The overall effect of NTCF mitigation is thus a robust decrease in air pollution, in terms of both $PM_{2.5}$ and $O_3$, over all 12 world regions, as well as the Arctic, Northern Hemisphere (NH) midlatitudes, and Tropics. Over all land surfaces, the $PM_{2.5}$ decrease is $-0.55\pm0.08$ $\mu gm^{-3}$ decade$^{-1}$. Regionally, decreases in $PM_{2.5}$ range from $-0.05\pm0.01$ $\mu gm^{-3}$ decade$^{-1}$ over Canada to $-3.8\pm0.69$ $\mu gm^{-3}$ decade$^{-1}$ in south Asia. Relatively large $PM_{2.5}$ decreases also occur over east Asia, southeast Asia and north Africa at $-2.1\pm0.27$ $\mu gm^{-3}$ decade$^{-1}$, $-0.78\pm0.16$ $\mu gm^{-3}$ decade$^{-1}$, and $-0.82\pm0.20$ $\mu gm^{-3}$ decade$^{-1}$, respectively. The relatively large $PM_{2.5}$ decreases over east Asia, southeast Asia, and south Asia are generally consistent with the relatively large reductions in aerosol species, including BC, $SO_4$ and OC (Figure 2).

Similar results exist for $O_3$, with a robust decrease over land of $-1.11\pm0.22$ ppb decade$^{-1}$. Regionally, $O_3$ decreases range from $-2.41\pm0.33$ ppb decade$^{-1}$ over central America and $-1.97\pm0.20$ ppb decade$^{-1}$ over southeast Asia to $-0.86\pm0.11$ ppb decade$^{-1}$ over Australia. Relatively large $O_3$ decreases also occur over south Asia ($-1.55\pm0.93$ ppb decade$^{-1}$), as well as north Africa ($-1.7\pm0.25$ ppb decade$^{-1}$). Notably, a weak $O_3$ decrease occurs in east Asia ($-0.45\pm0.51$ ppb decade$^{-1}$), which may be related to relatively weak VOC reductions (Figure 2). In addition to significant reduction in the Arctic, the other latitudinal bands also exhibit significant reductions in $O_3$.

Over all 12 world regions, significant surface warming occurs in both the weak and strong air quality control scenarios, due to continued increases in $CO_2$ (and $CH_4$). **More importantly, NTCF mitigation−due to reduced cooling from reductions in non-absorbing aerosol (e.g., sulfate)−also yields significant warming, with a significant increase in land-only surface temperature of 0.09±0.02 K decade$^{-1}$.** Significant warming also occurs in all but one world region (Australia is the lone exception) due to NTCF mitigation, ranging from $0.05\pm0.02$ K decade$^{-1}$ over southeast Asia to $0.16\pm0.05$ K decade$^{-1}$ over central and north Asia. Relatively large warming also occurs over east Asia ($0.11\pm0.05$ K decade$^{-1}$) and south Asia ($0.12\pm0.04$ K decade$^{-1}$; see also Supplementary Figure 4). Furthermore, large warming of the Arctic (60-90N) occurs ($0.15\pm0.09$ K decade$^{-1}$), particularly in the East Siberian and Beaufort Seas, north of western Canada/Alaska and around the Canadian Arctic Archipelago (Figure 6). This result is consistent with recent studies showing high Arctic sensitivity to aerosol reductions (Acosta Navarro et al., 2016; Lewinschal et al., 2019; Westervelt et al., 2020). **Other latitudinal bands also significantly warm, including the NH midlatitudes (30-60N), Tropics (30S-30N), and Southern Hemisphere (SH) midlatitudes (60S-30S) at 0.10±0.03, 0.05±0.01, 0.03±0.02 K decade$^{-1}$, respectively.**

**Warming is consistent with the increase in ERF, with most world regions yielding significant positive ERF trends. Little correspondence exists between regions that warm the most and their ERF trend. This is not necessarily surprising, as forcing and response do not need to occur in the same regions, due to climate feedbacks, remote teleconnection and other processes.** For example, central and north Asia and the Arctic warm the most, but there is not a particularly large increase in their regional ERF. Similarly, southeast Asia warms the least, but this region features a relatively large ERF increase.

Significant increases in the hottest day also occur, with larger increases under strong relative to weak air quality controls. NTCF mitigation yields significant increases in the hottest day for all but two world regions (Australia and south America are the exceptions). A significant increase in the hottest day also occurs over all land regions ($0.09\pm0.03$ K decade$^{-1}$), with five of six models yielding a significant increase (Table 1; MIROC6 is the exception). **Thus, NTCF mitigation unmasks the**

**warming due to GHG increases resulting in robust increases in both surface air temperature and the hottest day over nearly all world regions. We note that the lone area with cooling is the north Atlantic (around Iceland and southwest of Svalbard; Fig. 6), which may be associated with a weakening of the Atlantic Meridional Overturning Circulation (AMOC) (Delworth and Dixon, 2006; Cai et al., 2006; Menary et al., 2013). Figures 6d-f show that this cooling is a robust feature, with ∼70% of the models yielding cooling here.**

**Over all land surfaces, a significant precipitation increase also occurs in both scenarios at 0.012±0.005 and 0.022±0.006 mm day$^{-1}$ decade$^{-1}$ under weak and strong air quality control, respectively. Thus, NTCF mitigation−by unmasking GHG-induced warming−also yields a significant increase in land precipitation at 0.011±0.003 mm day$^{-1}$ decade$^{-1}$.** The effect of NTCF mitigation on precipitation over individual world regions, however, has mixed significance and ranges from 0.003±0.035 mm day$^{-1}$ decade$^{-1}$ over Australia to 0.044±0.022 mm day$^{-1}$ decade$^{-1}$ over south Asia. Note that some world regions exhibit decreases in precipitation under both weak and strong air quality control (e.g., central America), such that NTCF mitigation yields a weaker decrease (as opposed to an absolute increase). In addition to south Asia, a significant precipitation increase also occurs over central and north Asia (0.008±0.005 mm day$^{-1}$ decade$^{-1}$), east Asia (0.038±0.014 mm day$^{-1}$ decade$^{-1}$) and the Arctic (0.010±0.005 mm day$^{-1}$ decade$^{-1}$). Although southeast Asia also exhibits a relatively large increase in precipitation, it is not significant (0.019±0.041 mm day$^{-1}$ decade$^{-1}$). Both south and north Africa yield precipitation increases, but the bulk of the African precipitation increase occurs over East Africa (Supplementary Figure 5). **From a latitudinal perspective, in addition to the Arctic, the NH midlatitudes, Tropics and SH midlatitudes all experience a significant increase in precipitation at 0.010±0.005, 0.011±0.001, and 0.004±0.001 mm day$^{-1}$ decade$^{-1}$, respectively.** Thus, NTCF mitigation generally increases precipitation in most world regions (although, in some regions, this is a smaller decrease) but the signal is less robust than that for surface temperature. Furthermore, in agreement with prior studies (Levy et al., 2013; Westervelt et al., 2017; Zhao et al., 2018; Westervelt et al., 2018; Scannell et al., 2019), precipitation increases in several monsoon regions, including east Africa, south Asia, and east Asia.

Precipitation extremes, including the wettest day and in particular CDD, also exhibit regional uncertainty under NTCF mitigation, with most regions lacking a robust response. Similar to the significant increases in mean precipitation, significant increases in the wettest day also occur in central and north Asia, east Asia, south Asia, and the Arctic. The NH midlatitudes and Tropics (but not the SH midlatitudes) also experience a robust increase in the wettest day. NTCF mitigation also yields robust CDD increases in south America and south Africa, and robust CDD decreases in Canada and the Arctic. Outside of the Arctic, no other latitudinal bands yield a robust CDD response under NTCF mitigation.

### 3.3 Seasonal Climate and Air Quality Trends

Figure 7 shows the regional surface temperature, precipitation and air quality responses during June-July-August (JJA) and December-January-February (DJF). Seasonal air pollution, including both $O_3$ and $PM_{2.5}$, exhibit robust decreases in nearly all world regions under NTCF mitigation. Over land regions, slightly larger $O_3$ decreases occur during JJA relative to DJF, at −1.41±0.16 ppb decade$^{-1}$ and −0.86±0.34 ppb decade$^{-1}$, respectively. This seasonal contrast is more pronounced over the NH midlatitudes, where the JJA (DJF) decrease is −1.87±0.17 (−0.72±0.52) ppb decade$^{-1}$. In contrast, slightly larger

$PM_{2.5}$ decreases occur during DJF relative to JJA, at $-0.67\pm0.12$ $\mu$g m$^{-3}$ decade$^{-1}$ and $-0.48\pm0.08$ $\mu$g m$^{-3}$ decade$^{-1}$, respectively. As with the annual mean, the largest JJA and DJF $O_3$ reductions occur over central America and southeast Asia (and north Africa during DJF). The lone regional increase in $O_3$ occurs during DJF in east Asia at $0.95\pm0.52$ ppb decade$^{-1}$. The largest JJA decreases in $PM_{2.5}$ occur in east Asia ($-1.64\pm0.31$ $\mu$g m$^{-3}$ decade$^{-1}$) and south Asia ($-1.67\pm0.32$ $\mu$g m$^{-3}$ decade$^{-1}$) . These regions also exhibit large DJF decreases in $PM_{2.5}$, particularly south Asia at $-5.55\pm1.2$ $\mu$g m$^{-3}$ decade$^{-1}$.

NTCF mitigation yields similar warming in both seasons (see also Supplementary Figures 6-7). Over all land surfaces, JJA warming is $0.09\pm0.02$ K decade$^{-1}$; DJF warming is $0.09\pm0.04$ K decade$^{-1}$. Consistent with the annual mean warming, relatively large JJA warming also occurs in central and north Asia ($0.16\pm0.06$ K decade$^{-1}$), south Asia ($0.10\pm0.05$ K decade$^{-1}$) and east Asia ($0.10\pm0.06$ K decade$^{-1}$), as well as Canada ($0.14\pm0.05$ K decade$^{-1}$). DJF warming is largest in similar regions as JJA, including central and north Asia and south Asia ($0.20\pm0.12$ and $0.13\pm0.04$ K decade$^{-1}$) and east Asia ($0.13\pm0.12$ K decade$^{-1}$). Arctic warming is most pronounced during DJF, where the rate of warming is about double that during JJA ($0.23\pm0.16$ versus $0.12\pm0.05$ K decade$^{-1}$). Similar JJA and DJF warming occurs for the NH midlatitudes (0.11 versus $0.10\pm0.03$ K decade$^{-1}$), Tropics (0.05 versus $0.04\pm0.02$ K decade$^{-1}$) and SH midlatitudes ($0.03\pm0.02$ K decade$^{-1}$ for both seasons). As with the annual mean warming, central and north Asia, east Asia, and south Asia generally warm the most during JJA and DJF, with large Arctic warming during DJF.

Regional seasonal precipitation responses continue to exhibit relatively large uncertainty, as most world regions lack a robust response (see also Supplementary Figures 8-9). Central and north Asia, east Asia and south Asia yield robust JJA increases in precipitation under NTCF mitigation at $0.015\pm0.008$, $0.053\pm0.034$, and $0.089\pm0.047$ mm day$^{-1}$ decade$^{-1}$, respectively. The increase in south and east Asia precipitation is consistent with aerosol reductions driving enhanced monsoonal flow. Interestingly, there is also a significant increase in south Asian precipitation during DJF. Canada and north Africa in particular also exhibit robust increases in DJF precipitation. As with the annual mean, most of the increase in DJF precipitation over Africa occurs in east Africa (Supplementary Figure 9).

**Similar results generally exist for the other seasons, March-April-May (MAM) and September-October-November (SON) (Supplementary Figure 10). The largest decrease in $O_3$ occurs in central America, south Asia and southeast Asia, as well as north Africa. The largest $PM_{2.5}$ decreases occur in east Asia, south Asia, and southeast Asia. Over all land surfaces, MAM and SON surface warming are both $0.09\pm0.02$ K decade$^{-1}$. Maximum MAM (SON) regional warming occurs in central and north Asia (Arctic) at $0.18\pm0.05$ ($0.17\pm0.09$) K decade$^{-1}$. Relatively large MAM warming also occurs in east Asia ($0.13\pm0.07$ K decade$^{-1}$) and south Asia ($0.11\pm0.06$ K decade$^{-1}$); relatively large SON warming occurs in Canada ($0.15\pm0.06$ K decade$^{-1}$) and Europe ($0.13\pm0.05$ K decade$^{-1}$). Precipitation responses are again less robust, although east Asia experiences robust increases in both seasons ($0.03\pm0.02$ in MAM and $0.06\pm0.03$ mm day$^{-1}$ decade$^{-1}$ in SON). Relatively large SON precipitation increases also occur for central America and south Asia.**

## 4 Conclusions

Under the experimental protocols of ScenarioMIP (O'Neill et al., 2016) and AerChemMIP (Collins et al., 2017), we have analyzed future chemistry-climate simulations to assess the impact of non-methane NTCF mitigation of climate and air quality from 2015-2055. Simulations show robust decreases in air pollution in nearly all world regions. Over global land, surface $PM_{2.5}$ and $O_3$ decrease by $-2.2$ $\mu$g m$^{-3}$ and $-4.6$ ppb, respectively, with larger reductions in some world regions including south and southeast Asia. **However, NTCF mitigation unmasks the warming due to GHG increases, resulting in additional global warming and precipitation increases of 0.25K and 0.03 mm day$^{-1}$, respectively. Similarly, increases in extreme weather indices also occur, including the hottest and wettest day. All but one world region (minus Australia) yields robust warming in response to NTCF mitigation, with the largest warming (and wetting) occurring over Asia, including central and north Asia, east Asia and south Asia. Relatively large warming also occurs over the Arctic at 0.59K, more than double the global mean warming. Interestingly, models that include both aerosol and ozone reductions (Aer+O3) yield similar warming (and wetting) relative to models that include aerosol reductions alone (Aer). This suggests a weak cooling effect due to ozone reductions, or other possible effects from interactive chemistry and aerosol that need to be further explored. For example, aerosol formation may be reduced due to changes in oxidants (from $O_3$ reductions), which would lead to more surface warming in Aer+O3. Simulations with a single model, running both coupled and uncoupled chemistry experiments, would help isolate this effect.**

Our results are consistent with several studies that have shown aerosol reductions will unmask GHG warming, resulting in large, near-term increases in global surface temperature and precipitation (Brasseur and Roeckner, 2005; Andreae et al., 2005; Ramanathan and Feng, 2008; Raes and Seinfeld, 2009; Kloster et al., 2010; Arneth et al., 2009; Matthews and Zickfeld, 2012; Rotstayn et al., 2013; Wu et al., 2013; Westervelt et al., 2015; Salzmann, 2016; Hienola et al., 2018; Richardson et al., 2018; Samset et al., 2018; Lelieveld et al., 2019). Shindell and Smith (2019), however, show that the time required to transform power generation, industry and transportation leads to largely offsetting climate impacts of $CO_2$ and sulfur dioxide, implying no conflict between climate and air-quality objectives. There, a 1.5°C mitigation pathway is used, with gradual phasing out of fossil fuel combustion, which leads to relatively small change in the near-future warming. Furthermore, Shindell and Smith (2019) include methane mitigation, which compensates the relatively small near-term future warming from $SO_2$ reductions.

Our simulations, however, do not account for $CO_2$ or $CH_4$ reductions, implying the importance of simultaneous reductions in both WMGHGs and NTCFs. We note that it is difficult to reduce only the NTCF emissions while keeping $CO_2$ emissions fixed (since there are co-emitted species, including $SO_2$). **If WMGHG emissions are simultaneously reduced along with non-methane NTCFs, then the increase in global surface temperature and precipitation found here will be muted (and perhaps, offset). Moreover, our results (e.g., the magnitude of the surface temperature increase) likely represent an upper bound as our baseline scenario lacks climate policy and contains the highest levels of NTCFs**. The lowNTCF scenario, however, can be used to provide forcing and response sensitivities under current climate, which could be used by intermediate **complexity** models for testing out more scenarios which include complex NTCF-$CO_2$ reduction scenarios. Furthermore, the AerChemMIP SSP3-7.0-lowNTCF simulations used in this study do not account for reductions of methane, which is another

NTCF, the reduction of which would promote **net cooling (i.e., reduced warming)**. **As the strong air quality control pathway includes reductions of methane, additional AerChemMIP simulations are being conducted that include the effects of all NTCFs, including aerosols, ozone precursor gases and methane.** It is likely that inclusion of methane reductions will offset some of the warming reported here, and also impact tropospheric $O_3$ and air quality. Although not addressed in this study, we also note the potential role of hydrofluorocarbon (HFC) mitigation through the Kigali Amendment, particularly for the late 21st century. Efficient implementation of the Kigali Amendment and national regulations is estimated to lead to relatively small cooling ($<0.07°C$) by 2050, but this increases to cooling of 0.2-0.4°C by 2100 (WMO, 2018). Nonetheless, cleaning the air while keeping global warming below the 1.5-2°C Paris Agreement climate target will likely require simultaneous cuts in both NTCFs and carbon dioxide.

*Data availability.* **CMIP6 model data can be freely downloaded from the Earth System Grid Federation (ESGF) server at https://esgf-node.llnl.gov/projects/cmip6/. The emissions data used here can also be downloaded from the ESGF via the input datasets for Model Intercomparison Projects (input4MIPs) at https://esgf-node.llnl.gov/search/input4mips/.**

*Author contributions.* R.J.A. performed the analysis and wrote the paper. D.N., U.L. and I.T. performed MPI-ESM1-2-HAM simulations. P.N. and M.M. performed CNRM-ESM2-1 simulations. T.W. and J.Z. performed BCC-ESM1 simulations. N.O. and M.D. performed MRI-ESM2-0 simulations. T.T. performed MIROC6 simulations. L.E. and J-F.L. performed CESM2 and CESM2-WACCM simulations. L.H., V.N. L.S., and J. J. performed GFDL-ESM4 simulations. W.J.C., J-F.L. and M.S. originally conceived the AerChemMIP project, including the low NTCF simulations analyzed here. All authors contributed to editing the manuscript.

*Competing interests.* The authors declare that they have no competing financial interests.

*Acknowledgements.* T. Takemura was supported by the supercomputer system of the National Institute for Environmental Studies, Japan, and JSPS KAKENHI Grant Number JP19H05669. D. Neubauer acknowledges funding from the European Union's Horizon 2020 research and innovation programme project FORCeS under grant agreement No 821205. D. Neubauer and I. Tegen acknowledge a grant for computing resources from the Deutsches Klimarechenzentrum (DKRZ) under project ID 1051. The CESM project is supported primarily by the National Science Foundation. This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. S. Shim was supported by the Korea Meteorological Administration Research and Development Program "Development and Assessment of IPCC AR6 Climate Change Scenario", grant agreement number 1365003000. D.O. performed NorESM2-LM simulations. M. Deushi and N. Oshima were supported by the Japan Society for the Promotion of Science (grant numbers: JP18H03363, JP18H05292, and JP20K04070) and the Environment Research and Technology Development Fund (2-1703, 2-2003, and 5-2001) of the Environmental Restoration and Conservation Agency, Japan.

## References


Ackerley, D., Booth, B. B. B., Knight, S. H. E., Highwood, E. J., Frame, D. J., Allen, M. R., and Rowell, D. P.: Sensitivity of twentieth-century Sahel rainfall to sulfate aerosol and CO2 forcing, Journal of Climate, 24, 4999–5014, doi: 10.1175/JCLI–D–11–00 019.1, 2011.

Acosta Navarro, J. C., Varma, V., Riipinen, I., Seland, Ø., Kirkevåg, A., Struthers, H., Iversen, T., Hansson, H. C., and Ekman, A. M. L.: Amplification of Arctic warming by past air pollution reductions in Europe, Nature Geoscience, 9, 277–281, doi: 10.1038/ngeo2673, 2016.

Allen, R. J.: A 21st century northward tropical precipitation shift caused by future anthropogenic aerosol reductions, Journal of Geophysical Research: Atmospheres, 120, 9087–9102, doi: 10.1002/2015JD023 623, 2015.

Allen, R. J. and Ajoku, O.: Future aerosol reduction and widening of the nothern tropical belt, J. Geophys. Res., pp. 1–22, doi: 10.1002/2016JD024 803, 2016.

Allen, R. J. and Sherwood, S. C.: The impact of natural versus anthropogenic aerosols on atmospheric circulation in the Community Atmosphere Model, Climate Dyn., 36, 1959–1978, doi: 10.1007/s00 382–010–0898–8, 2011.

Allen, R. J., Evan, A. T., and Booth, B. B. B.: Interhemispheric Aerosol Radiative Forcing and Tropical Precipitation Shifts during the Late Twentieth Century, Journal of Climate, 28, 8219–8246, doi: 10.1175/JCLI–D–15–0148.1, 2015.

Allen, R. J., Landuyt, W., and Rumbold, S. T.: An increase in aerosol burden and radiative effects in a warmer world, Nature Climate Change, 6, 269–274, doi: 10.1038/nclimate2827, 2016.

Allen, R. J., Amiri-Farahani, A., Lamarque, J.-F., Smith, C., Shindell, D., Hassan, T., and Chung, C. E.: Observationally-constrained aerosol-cloud semi-direct effects, npj Climate and Atmospheric Science, 2, 16, doi: 10.1038/s41 612–019–0073–9, 2019a.

Allen, R. J., Hassan, T., Randles, C. A., and Su, H.: Enhanced land-sea warming contrast elevates aerosol pollution in a warmer world, Nature Climate Change, 9, 300–305, doi: 10.1038/s41 558–019–0401–4, 2019b.

Andreae, M. O., Jones, C. D., and Cox, P. M.: Strong present-day aerosol cooling implies a hot future, Nature, 435, 1187–1190, doi: 10.1038/nature03 671, 2005.

Arneth, A., Unger, N., Kulmala, M., and Andreae, M. O.: Clean the Air, Heat the Planet?, Science, 326, 672–673, doi: 10.1126/science.1181 568, 2009.

Ban-Weiss, G. A., Cao, L., Bala, G., and Caldeira, K.: Dependence of climate forcing and response on the altitude of black carbon aerosols, Climate Dyn., 38, 897–911, doi: 10.1007/s00 382–011–1052–y, 2012.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, e2019RG000 660, doi: 10.1029/2019RG000 660, 2020.

Biasutti, M.: Forced Sahel rainfall trends in the CMIP5 archive, Journal of Geophysical Research: Atmospheres, 118, 1613–1623, doi: 10.1002/jgrd.50 206, 2013.

Biasutti, M. and Giannini, A.: Robust Sahel drying in response to late 20th century forcings, Geophysical Research Letters, 33, doi: 10.1029/2006GL026 067, 2006.

Bond, T. C., Doherty, S. J., Hahey, D. W., and et al.: Bounding the role of black carbon in the climate system: A scientific assessment, J. Geophys. Res., 118, 5380–5552, doi:10.1002/jgrd.50 171, 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. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Tech. rep., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Brasseur, G. P. and Roeckner, E.: Impact of improved air quality on the future evolution of climate, Geophysical Research Letters, 32, doi: 10.1029/2005GL023 902, 2005.

Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., and Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies, Journal of Climate, 30, 6851–6872, doi: 10.1175/JCLI–D–16–0613.1, 2017.

Butt, E. W., Turnock, S. T., Rigby, R., Reddington, C. L., Yoshioka, M., Johnson, J. S., Regayre, L. A., Pringle, K. J., Mann, G. W., and Spracklen, D. V.: Global and regional trends in particulate air pollution and attributable health burden over the past 50 years, Environmental Research Letters, 12, 104 017, doi: 10.1088/1748–9326/aa87be, 2017.

Cai, W., Bi, D., Church, J., Cowan, T., Dix, M., and Rotstayn, L.: Pan-oceanic response to increasing anthropogenic aerosols: Impacts on the Southern Hemisphere oceanic circulation, Geophysical Research Letters, 33, doi: 10.1029/2006GL027 513, 2006.

Chang, C.-Y., Chiang, J. C. H., Wehner, M. F., Friedman, A. R., and Ruedy, R.: Sulfate Aerosol Control of Tropical Atlantic Climate over the Twentieth Century, Journal of Climate, 24, 2540–2555, doi: 10.1175/2010JCLI4065.1, 2011.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L., and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, The Lancet, 389, 1907–1918, doi: 10.1016/S0140–6736(17)30 505–6, 2017.

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, Geoscientific Model Development, 10, 585–607, doi: 10.5194/gmd–10–585–2017, 2017.

Delworth, T. L. and Dixon, K. W.: Have anthropogenic aerosols delayed a greenhouse gas-induced weakening of the North Atlantic thermohaline circulation?, Geophysical Research Letters, 33, doi: 10.1029/2005GL024 980, 2006.

Donat, M., Alexander, L., Yang, H., Durre, I., Vose, R., and Caesar, J.: Global Land-Based Datasets for Monitoring Climatic Extremes, Bulletin of the American Meteorological Society, 94, 997–1006, doi: 10.1175/BAMS–D–12–00 109.1, 2013a.

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J. H., Willett, K. M., Aguilar, E., Brunet, M., Caesar, J., Hewitson, B., Jack, C., Klein Tank, A. M. G., Kruger, A. C., Marengo, J., Peterson, T. C., Renom, M., Oria Rojas, C., Rusticucci, M., Salinger, J., Elrayah, A. S., Sekele, S. S., Srivastava, A. K., Trewin, B., Villarroel, C., Vincent, L. A., Zhai, P., Zhang, X., and Kitching, S.: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, J. Geophys. Res.-Atmos., 118, 2098–2118, doi: 10.1002/jgrd.50 150, 2013b.

Dong, B., Sutton, R. T., Highwood, E., and Wilcox, L.: The Impacts of European and Asian Anthropogenic Sulfur Dioxide Emissions on Sahel Rainfall, Journal of Climate, 27, 7000–7017, doi: 10.1175/JCLI–D–13–00 769.1, 2014.

Dunne, J. P., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., C.A.Stock, Zadeh, N., Balaji, V., Blanton, C., Dunne, K., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P., Griffies, S., Guo, H., Hallberg, R., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P., Nikonov, S., Paynter, D., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B., Robinson, T., Schwarzkopf, M., Sentman, L., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): Model description and simulation characteristics, Journal of Advances in Modeling Earth Systems, p. 2019MS002008, submitted.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., Marsh, D., Mills, M. J., Tilmes, S., Bardeen, C., Buchholz, R. R., Conley, A., Gettelman, A., Garcia, R., Simpson, I., Blake, D. R., Meinardi, S., and Pétron, G.: The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 882, doi: 10.1029/2019MS001 882, 2020.

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, doi: 10.5194/gmd–9–1937–2016, 2016.

Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsøren, S., Eyring, V., Folberth, G. A., Ginoux, P., Horowitz, L. W., Josse, B., Lamarque, J.-F., MacKenzie, I. A., Nagashima, T., O'Connor, F. M., Righi, M., Rumbold, S. T., Shindell, D. T., Skeie, R. B., Sudo, K., Szopa, S., Takemura, T., and Zeng, G.: Global air quality and climate, Chem. Soc. Rev., 41, 6663–6683, doi: 10.1039/C2CS35 095E, 2012.

Fiore, A. M., Naik, V., and Leibensperger, E. M.: Air Quality and Climate Connections, Journal of the Air & Waste Management Association, 65, 645–685, doi: 10.1080/10962 247.2015.1040 526, 2015.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., Andrews, T., Pincus, R., and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, Journal of Geophysical Research: Atmospheres, 121, 12,460–12,475, doi: 10.1002/2016JD025 320, 2016.

Fujimori, S., Hasegawa, T., Masui, T., Takahashi, K., Herran, D. S., Dai, H., Hijioka, Y., and Kainuma, M.: SSP3: AIM implementation of Shared Socioeconomic Pathways, Global Environmental Change, 42, 268–283, doi: 10.1016/j.gloenvcha.2016.06.009, 2017.

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., Tilmes, S., Vitt, F., Bardeen, C. G., McInerny, J., Liu, H.-L., Solomon, S. C., Polvani, L. M., Emmons, L. K., Lamarque, J.-F., Richter, J. H., Glanville, A. S., Bacmeister, J. T., Phillips, A. S., Neale, R. B., Simpson, I. R., DuVivier, A. K., Hodzic, A., and Randel, W. J.: The Whole Atmosphere Community Climate Model Version 6 (WACCM6), Journal of Geophysical Research: Atmospheres, 124, 12 380–12 403, doi: 10.1029/2019JD030 943, 2019.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, Geoscientific Model Development, 12, 1443–1475, doi: 10.5194/gmd–12–1443–2019, 2019.

Guo, L., Turner, A. G., and Highwood, E. J.: Local and Remote Impacts of Aerosol Species on Indian Summer Monsoon Rainfall in a GCM, Journal of Climate, 29, 6937–6955, doi: 10.1175/JCLI–D–15–0728.1, 2016.

Hienola, A., Partanen, A.-I., Pietikäinen, J.-P., O'Donnell, D., Korhonen, H., Matthews, H. D., and Laaksonen, A.: The impact of aerosol emissions on the 1.5 °C pathways, Environmental Research Letters, 13, 044 011, doi: 10.1088/1748–9326/aab1b2, 2018.

Horowitz, L. W., Naik, V., Sentman, L. T., Paulot, F., Blanton, C., McHugh, C., Radhakrishnan, A., Rand, K., Ginoux, P., and Paynter, D. J.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 AerChemMIP, Version YYYYMMDD[1].Earth System Grid Federation., p. https://doi.org/10.22033/ESGF/CMIP6.1404, 2018.

Horowitz, L. W., Naik, V., Paulot, F., Ginoux, P. A., Dunne, J. P., Mao, J., Schnell, J., Chen, X., He, J., Lin, M., Lin, P., Malyshev, S., Paynter, D., Shevliakova, E., and Zhao, M.: The GFDL Global Atmospheric Chemistry-Climate Model AM4.1: Model Description and Simulation Characteristics, Journal of Advances in Modeling Earth Systems, submitted.

Hwang, Y.-T., Frierson, D. M. W., and Kang, S. M.: Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century, Geophysical Research Letters, 40, 2845–2850, doi: 10.1002/grl.50 502, 2013.

IPCC: Summary for Policymakers. In: Global Warming of 1.5 °C. An IPCC Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, Tech. rep., [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)], World Meteorological Organization, Geneva, Switzerland, 32 pp., 2018.

John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Vahlenkamp, H., Zadeh, N. T., Gauthier, P., Ginoux, P., Harrison, M., Horowitz, L., Malyshev, S., Naik, V., Paynter, D. J., Ploshay, J., Silvers, L., Stock, C., Winton, M., Zeng, Y., and Dunne, J. P.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 ScenarioMIP., Version YYYYMMDD[1].Earth System Grid Federation., p. https://doi.org/10.22033/ESGF/CMIP6.1414, 2018.

Johnson, B. T., Haywood, J. M., and Hawcroft, M. K.: Are Changes in Atmospheric Circulation Important for Black Carbon Aerosol Impacts on Clouds, Precipitation, and Radiation?, Journal of Geophysical Research: Atmospheres, 124, 7930–7950, doi: 10.1029/2019JD030 568, 2019.

Joshi, M. M., Gregory, J. M., Webb, M. J., Sexton, D. M. H., and Johns, T. C.: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change, Climate Dyn., 30, 455–465, doi:10.1007/s00 382–007–0306–1, 2008.

Kloster, S., Dentener, F., Feichter, J., Raes, F., Lohmann, U., Roeckner, E., and Fischer-Bruns, I.: A GCM study of future climate response to aerosol pollution reductions, Climate Dynamics, 34, 1177–1194, doi: 10.1007/s00 382–009–0573–0, 2010.

Krishnan, R., Sabin, T. P., Vellore, R., Mujumdar, M., Sanjay, J., Goswami, B. N., Hourdin, F., Dufresne, J.-L., and Terray, P.: Deciphering the desiccation trend of the South Asian monsoon hydroclimate in a warming world, Climate Dynamics, 47, 1007–1027, doi: 10.1007/s00 382–015–2886–5, 2016.

Lau, W. K.-M. and Kim, K.-M.: Competing influences of greenhouse warming and aerosols on Asian summer monsoon circulation and rainfall, Asia-Pacific Journal of Atmospheric Sciences, 53, 181–194, doi: 10.1007/s13 143–017–0033–4, 2017.

Lelieveld, J., Klingmüller, K., Pozzer, A., Burnett, R. T., Haines, A., and Ramanathan, V.: Effects of fossil fuel and total anthropogenic emission removal on public health and climate, Proceedings of the National Academy of Sciences, 116, 7192–7197, doi: 10.1073/pnas.1819989 116, 2019.

Levy, H., Horowitz, L. W., Schwarzkopf, M. D., Ming, Y., Golaz, J.-C., Naik, V., and Ramaswamy, V.: The roles of aerosol direct and indirect effects in past and future climate change, J. Geophys. Res, 118, 4521–4532, doi: 10.1002/jgrd.50 192, 2013.

Lewinschal, A., Ekman, A. M. L., Hansson, H.-C., Sand, M., Berntsen, T. K., and Langner, J.: Local and remote temperature response of regional $SO_2$ emissions, Atmospheric Chemistry and Physics, 19, 2385–2403, doi: 10.5194/acp–19–2385–2019, 2019.

Li, X., Ting, M., Li, C., and Henderson, N.: Mechanisms of Asian Summer Monsoon Changes in Response to Anthropogenic Forcing in CMIP5 Models, Journal of Climate, 28, 4107–4125, doi: 10.1175/JCLI–D–14–00 559.1, 2015.

Liepert, B. G., Feichter, J., Lohmann, U., and Roeckner, E.: Can aerosols spin down the water cycle in a warmer and moister world?, Geophys. Res. Lett., 31, L06 207, doi: 10.1029/2003GL019 060, 2004.

Lin, L., Xu, Y., Wang, Z., Diao, C., Dong, W., and Xie, S.-P.: Changes in Extreme Rainfall Over India and China Attributed to Regional Aerosol-Cloud Interaction During the Late 20th Century Rapid Industrialization, Geophysical Research Letters, 45, 7857–7865, doi: 10.1029/2018GL078 308, 2018.

Liu, L., Shawki, D., Voulgarakis, A., Kasoar, M., Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S. G., Boucher, O., Faluvegi, G., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shindell, D., and Takemura, T.: A PDRMIP Multimodel Study on the Impacts of Regional Aerosol Forcings on Global and Regional Precipitation, Journal of Climate, 31, 4429–4447, doi: 10.1175/JCLI–D–17–0439.1, 2018.

Matthews, H. D. and Zickfeld, K.: Climate response to zeroed emissions of greenhouse gases and aerosols, Nature Climate Change, 2, 338–341, doi: 10.1038/nclimate1424, 2012.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, doi: 10.1029/2018MS001 400, 2019.

Menary, M. B., Roberts, C. D., Palmer, M. D., Halloran, P. R., Jackson, L., Wood, R. A., Müller, W. A., Matei, D., and Lee, S.-K.: Mechanisms of aerosol-forced AMOC variability in a state of the art climate model, Journal of Geophysical Research: Oceans, 118, 2087–2096, doi: 10.1002/jgrc.20 178, 2013.

Michou, M., Nabat, P., Saint-Martin, D., Bock, J., Decharme, B., Mallet, M., Roehrig, R., Séférian, R., Sénési, S., and Voldoire, A.: Present-day and historical aerosol and ozone characteristics in CNRM CMIP6 simulations, Journal of Advances in Modeling Earth Systems, p. doi: 10.1029/2019MS001816, 2019.

Ming, Y., Ramaswamy, V., and Persad, G.: Two opposing effects of absorbing aerosols on global-mean precipitation, Geophys. Res. Lett., 37, L13 701, doi: 10.1029/2010GL042 895, 2010.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Tech. rep., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Naik, V., Horowitz, L. W., Fiore, A. M., Ginoux, P., Mao, J., Aghedo, A. M., and Levy II, H.: Impact of preindustrial to present-day changes in short-lived pollutant emissions on atmospheric composition and climate forcing, Journal of Geophysical Research: Atmospheres, 118, 8086–8110, doi: 10.1002/jgrd.50 608, 2013.

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stier, P., Partridge, D. G., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., and Lohmann, U.: The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 2: Cloud evaluation, aerosol radiative forcing, and climate sensitivity, Geoscientific Model Development, 12, 3609–3639, doi: 10.5194/gmd–12–3609–2019, 2019.

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and van Vuuren, D. P.: A new scenario framework for climate change research: the concept of shared socioeconomic pathways, Climatic Change, 122, 387–400, doi: 10.1007/s10 584–013–0905–2, 2014.

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

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, doi: 10.5194/gmd–9–3447–2016, 2016.

Raes, F. and Seinfeld, J. H.: New Directions: Climate change and air pollution abatement: A bumpy road, Atmospheric Environment, 43, 5132 – 5133, doi: 10.1016/j.atmosenv.2009.06.001, 2009.

Ramanathan, V. and Feng, Y.: On avoiding dangerous anthropogenic interference with the climate system: Formidable challenges ahead, Proceedings of the National Academy of Sciences, 105, 14 245–14 250, doi: 10.1073/pnas.0803838 105, 2008.

Ramanathan, V., Crutzen, P. J., Lelieveld, J., Mitra, A. P., and et al.: Indian Ocean Experiment: An integrated analysis of the climate forcing and effects of the great Indo-Asian haze, J. Geophys. Res., 106, 28,371–28,398, 2001.

Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., and Flynn, C. J.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation, Journal of Climate, 30, 6823–6850, doi: 10.1175/JCLI–D–16–0609.1, 2017.

Rao, S., Klimont, Z., Smith, S. J., Dingenen, R. V., Dentener, F., Bouwman, L., Riahi, K., Amann, M., Bodirsky, B. L., van Vuuren, D. P., Reis, L. A., Calvin, K., Drouet, L., Fricko, O., Fujimori, S., Gernaat, D., Havlik, P., Harmsen, M., Hasegawa, T., Heyes, C., Hilaire, J., Luderer, G., Masui, T., Stehfest, E., Strefler, J., van der Sluis, S., and Tavoni, M.: Future air pollution in the Shared Socio-economic Pathways, Global Environmental Change, 42, 346–358, doi: 10.1016/j.gloenvcha.2016.05.012, 2017.

Reddington, C. L., Carslaw, K. S., Stier, P., Schutgens, N., Coe, H., Liu, D., Allan, J., Browse, J., Pringle, K. J., Lee, L. A., Yoshioka, M., Johnson, J. S., Regayre, L. A., Spracklen, D. V., Mann, G. W., Clarke, A., Hermann, M., Henning, S., Wex, H., Kristensen, T. B., Leaitch, W. R., Pöschl, U., Rose, D., Andreae, M. O., Schmale, J., Kondo, Y., Oshima, N., Schwarz, J. P., Nenes, A., Anderson, B., Roberts, G. C., Snider, J. R., Leck, C., Quinn, P. K., Chi, X., Ding, A., Jimenez, J. L., and Zhang, Q.: The Global Aerosol Synthesis and Science Project (GASSP): Measurements and Modeling to Reduce Uncertainty, Bulletin of the American Meteorological Society, 98, 1857–1877, doi: 10.1175/BAMS–D–15–00 317.1, 2017.

Richardson, T. B., Forster, P. M., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Hodnebrog, Ø., Kasoar, M., Kirkevåg, A., Lamarque, J.-F., Myhre, G., Olivié, D., Samset, B. H., Shawki, D., Shindell, D., Takemura, T., and Voulgarakis, A.: Drivers of precipitation change: An energetic understanding, Journal of Climate, 31, 9641–9657, doi: 10.1175/JCLI–D–17–0240.1, 2018.

Rotstayn, L. D. and Lohmann, U.: Tropical rainfall trends and the indirect aerosol effect, J. Climate, 15, 2103–2116, 2002.

Rotstayn, L. D., Collier, M. A., Chrastansky, A., Jeffrey, S. J., and Luo, J. J.: Projected effects of declining aerosol in RCP4.5: unmasking global warming?, Atmos. Chem. Phys., 13, 10 883–10 905, 2013.

Rotstayn, L. D., Collier, M. A., and Luo, J.-J.: Effects of declining aerosols on projections of zonally averaged tropical precipitation, Environmental Research Letters, 10, 044 018, doi: 10.1088/1748–9326/10/4/044 018, 2015.

Salzmann, M.: Global warming without global mean precipitation increase?, Science Advances, 2, doi: 10.1126/sciadv.1501 572, 2016.

Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø., Andrews, T., Faluvegi, G., Fläschner, D., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shindell, D., Shine, K. P., Takemura, T., and Voulgarakis, A.: Fast and slow precipitation responses to individual climate forcers: A PDRMIP multimodel study, Geophysical Research Letters, 43, 2782–2791, 10.1002/2016GL068 064, 2016.

Samset, B. H., Sand, M., Smith, C. J., Bauer, S. E., Forster, P. M., Fuglestvedt, J. S., Osprey, S., and Schleussner, C.-F.: Climate Impacts From a Removal of Anthropogenic Aerosol Emissions, Geophysical Research Letters, 45, 1020–1029, doi: 10.1002/2017GL076 079, 2018.

Scannell, C., Booth, B. B. B., Dunstone, N. J., Rowell, D. P., Bernie, D. J., Kasoar, M., Voulgarakis, A., Wilcox, L. J., Acosta Navarro, J. C., Seland, Ø., and Paynter, D. J.: The Influence of Remote Aerosol Forcing from Industrialized Economies on the Future Evolution of East and West African Rainfall, Journal of Climate, 32, 8335–8354, doi: 10.1175/JCLI–D–18–0716.1, 2019.

Schultz, M., Schröder, S., Lyapina, O., Cooper, O., Galbally, I., Petropavlovskikh, I., von Schneidemesser, E., Tanimoto, H., and et al.: Tropospheric Ozone Assessment Report: Database and Metrics Data of Global Surface Ozone Observations, Elem. Sci. Anth., 5, p. 58, doi: 10.1525/elementa.244, 2017.

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, doi: 10.1029/2019MS001 791, 2019.

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M.: The Norwegian Earth System Model, NorESM2 – Evaluation of the CMIP6 DECK and historical simulations, Geoscientific Model Development Discussions, 2020, 1–68, doi: 10.5194/gmd–2019–378, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, doi: 10.1029/2019MS001 739, 2019.

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, Cambridge University Press, 2012.

Shindell, D. and Smith, C. J.: Climate and air-quality benefits of a realistic phase-out of fossil fuels, Nature, 573, 408–411, doi: 10.1038/s41 586–019–1554–z, 2019.

Silva, R. A., West, J. J., Lamarque, J.-F., Shindell, D. T., Collins, W. J., Faluvegi, G., Folberth, G. A., Horowitz, L. W., Nagashima, T., Naik, V., Rumbold, S. T., Sudo, K., Takemura, T., Bergmann, D., Cameron-Smith, P., Doherty, R. M., Josse, B., MacKenzie, I. A., Stevenson, D. S., and Zeng, G.: Future global mortality from changes in air pollution attributable to climate change, Nature Climate Change, 7, 647–651, doi: 10.1038/nclimate3354, 2017.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, Geoscientific Model Development, 11, 2273–2297, doi: 10.5194/gmd–11–2273–2018, 2018.

Song, F., Zhou, T., and Qian, Y.: Responses of East Asian summer monsoon to natural and anthropogenic forcings in the 17 latest CMIP5 models, Geophysical Research Letters, 41, 596–603, doi: 10.1002/2013GL058 705, 2014.

Stjern, C. W., Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø., Andrews, T., Boucher, O., Faluvegi, G., Iversen, T., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shawki, D., Shindell, D., Smith, C. J., Takemura, T., and Voulgarakis, A.: Rapid Adjustments Cause Weak Surface Temperature Response to Increased Black Carbon Concentrations, Journal of Geophysical Research: Atmospheres, 122, 11,462–11,481, doi: 10.1002/2017JD027 326, 2017.

Sutton, R. T., Dong, B., and Gregory, J. M.: Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations, Geophys. Res. Lett., 34, L02 701, doi:10.1029/2006GL028 164, 2007.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, Journal of Geophysical Research: Atmospheres, 110, doi: 10.1029/2004JD005 029, 2005.

Takemura, T., Egashira, M., Matsuzawa, K., Ichijo, H., O'ishi, R., and Abe-Ouchi, A.: A simulation of the global distribution and radiative forcing of soil dust aerosols at the Last Glacial Maximum, Atmospheric Chemistry and Physics, 9, 3061–3073, doi: 10.5194/acp–9–3061–2009, 2009.

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, doi: 10.5194/gmd–12–2727–2019, 2019.

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., Rast, S., Kokkola, H., Schultz, M., Schroeder, S., Daskalakis, N., Barthel, S., Heinold, B., and Lohmann, U.: The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 1: Aerosol evaluation, Geoscientific Model Development, 12, 1643–1677, doi: 10.5194/gmd–12–1643–2019, 2019.

Tilmes, S., Hodzic, A., Emmons, L. K., Mills, M. J., Gettelman, A., Kinnison, D. E., Park, M., Lamarque, J.-F., Vitt, F., Shrivastava, M., Campuzano-Jost, P., Jimenez, J. L., and Liu, X.: Climate Forcing and Trends of Organic Aerosols in the Community Earth System Model (CESM2), Journal of Advances in Modeling Earth Systems, 11, 4323–4351, doi: 10.1029/2019MS001 827, 2019.

Turnock, S. T., Smith, S., and O'Connor, F. M.: The impact of climate mitigation measures on near term climate forcers, Environmental Research Letters, 14, 104 013, doi: 10.1088/1748–9326/ab4222, 2019.

Turnock, S. T., Allen, R. J., Andrews, M., Bauer, S. E., Emmons, L., Good, P., Horowitz, L., Michou, M., Nabat, P., Naik, V., Neubauer, D., O'Connor, F. M., Olivié, D., Schulz, M., Sellar, A., Takemura, T., Tilmes, S., Tsigaridis, K., Wu, T., and Zhang, J.: Historical and future changes in air pollutants from CMIP6 models, Atmospheric Chemistry and Physics Discussions, 2020, 1–40, doi: 10.5194/acp–2019–1211, 2020.

Undorf, S., Polson, D., Bollasina, M. A., Ming, Y., Schurer, A., and Hegerl, G. C.: Detectable Impact of Local and Remote Anthropogenic Aerosols on the 20th Century Changes of West African and South Asian Monsoon Precipitation, Journal of Geophysical Research: Atmospheres, 123, 4871–4889, doi: 10.1029/2017JD027 711, 2018.

van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R., and Winkler, H.: A new scenario framework for Climate Change Research: scenario matrix architecture, Climatic Change, 122, 373–386, doi: 10.1007/s10 584–013–0906–1, 2014.

Wang, T., Wang, H. J., Otterå, O. H., Gao, Y. Q., Suo, L. L., Furevik, T., and Yu, L.: Anthropogenic agent implicated as a prime driver of shift in precipitation in eastern China in the late 1970s, Atmospheric Chemistry and Physics, 13, 12 433–12 450, doi: 10.5194/acp–13–12 433–2013, 2013.

Westervelt, D. M., Horowitz, L. W., Naik, V., Golaz, J.-C., and Mauzerall, D. L.: Radiative forcing and climate response to projected 21st century aerosol decreases, Atmospheric Chemistry and Physics, 15, 12 681–12 703, doi: 10.5194/acp–15–12 681–2015, 2015.

Westervelt, D. M., Conley, A. J., Fiore, A. M., Lamarque, J.-F., Shindell, D., Previdi, M., Faluvegi, G., Correa, G., and Horowitz, L. W.: Multimodel precipitation responses to removal of U.S. sulfur dioxide emissions, Journal of Geophysical Research: Atmospheres, 122, 5024–5038, doi: 10.1002/2017JD026 756, 2017.

Westervelt, D. M., Conley, A. J., Fiore, A. M., Lamarque, J.-F., Shindell, D. T., Previdi, M., Mascioli, N. R., Faluvegi, G., Correa, G., and Horowitz, L. W.: Connecting regional aerosol emissions reductions to local and remote precipitation responses, Atmospheric Chemistry and Physics, 18, 12 461–12 475, doi: 10.5194/acp–18–12 461–2018, 2018.

Westervelt, D. M., Mascioli, N. R., Fiore, A. M., Conley, A. J., Lamarque, J.-F., Shindell, D. T., Faluvegi, G., Previdi, M., Correa, G., and Horowitz, L. W.: Local and remote mean and extreme temperature response to regional aerosol emissions reductions, Atmospheric Chemistry and Physics, 20, 3009–3027, doi: 10.5194/acp–20–3009–2020, 2020.

WHO: Ambient air pollution: A global assessment of exposture and burden of disease, Tech. Rep. ISBN: 9789241511353, World Health Organization, 2016.

Wilcox, L. J., Highwood, E. J., and Dunstone, N. J.: The influence of anthropogenic aerosol on multi-decadal variations of historical global climate, Environ. Res. Lett., 8, doi:10.1088/1748–9326/8/2/024 033, 2013.

WMO: Scientific Assessment of Ozone Depletion: 2018, Tech. Rep. Global Ozone Research and Monitoring Project–Report No. 58, 588 pp., Geneva, Switzerland., World Meteorological Organization, 2018.

Wu, P., Christidis, N., and Stott, P.: Anthropogenic impact on Earth's hydrological cycle, Nature Climate Change, 3, 807 EP –, doi: 10.1038/nclimate1932, 2013.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y., and Liu, X.: The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6, Geoscientific Model Development, 12, 1573–1600, doi: 10.5194/gmd–12–1573–2019, 2019.

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

Xie, X., Wang, H., Liu, X., Li, J., Wang, Z., and Liu, Y.: Distinct effects of anthropogenic aerosols on the East Asian summer monsoon between multidecadal strong and weak monsoon stages, Journal of Geophysical Research: Atmospheres, 121, 7026–7040, doi: 10.1002/2015JD024 228, 2016.

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

Zanis, P., Akritidis, D., Georgoulias, A. K., Allen, R. J., Bauer, S. E., Boucher, O., Cole, J., Johnson, B., Deushi, M., Michou, M., Mulcahy, J., Nabat, P., Olivie, D., Oshima, N., Sima, A., Schulz, M., and Takemura, T.: Fast responses on pre-industrial climate from present-day aerosols in a CMIP6 multi-model study, Atmospheric Chemistry and Physics Discussions, 2020, 1–32, doi: 10.5194/acp–2019–1201, 2020.

Zhang, L., Wu, P., and Zhou, T.: Aerosol forcing of extreme summer drought over North China, Environmental Research Letters, 12, 034 020, doi: 10.1088/1748–9326/aa5fb3, 2017.

Zhao, A. D., Stevenson, D. S., and Bollasina, M. A.: The role of anthropogenic aerosols in future precipitation extremes over the Asian Monsoon Region, Climate Dynamics, 52, 6257–6278, doi: 10.1007/s00 382–018–4514–7, 2018.

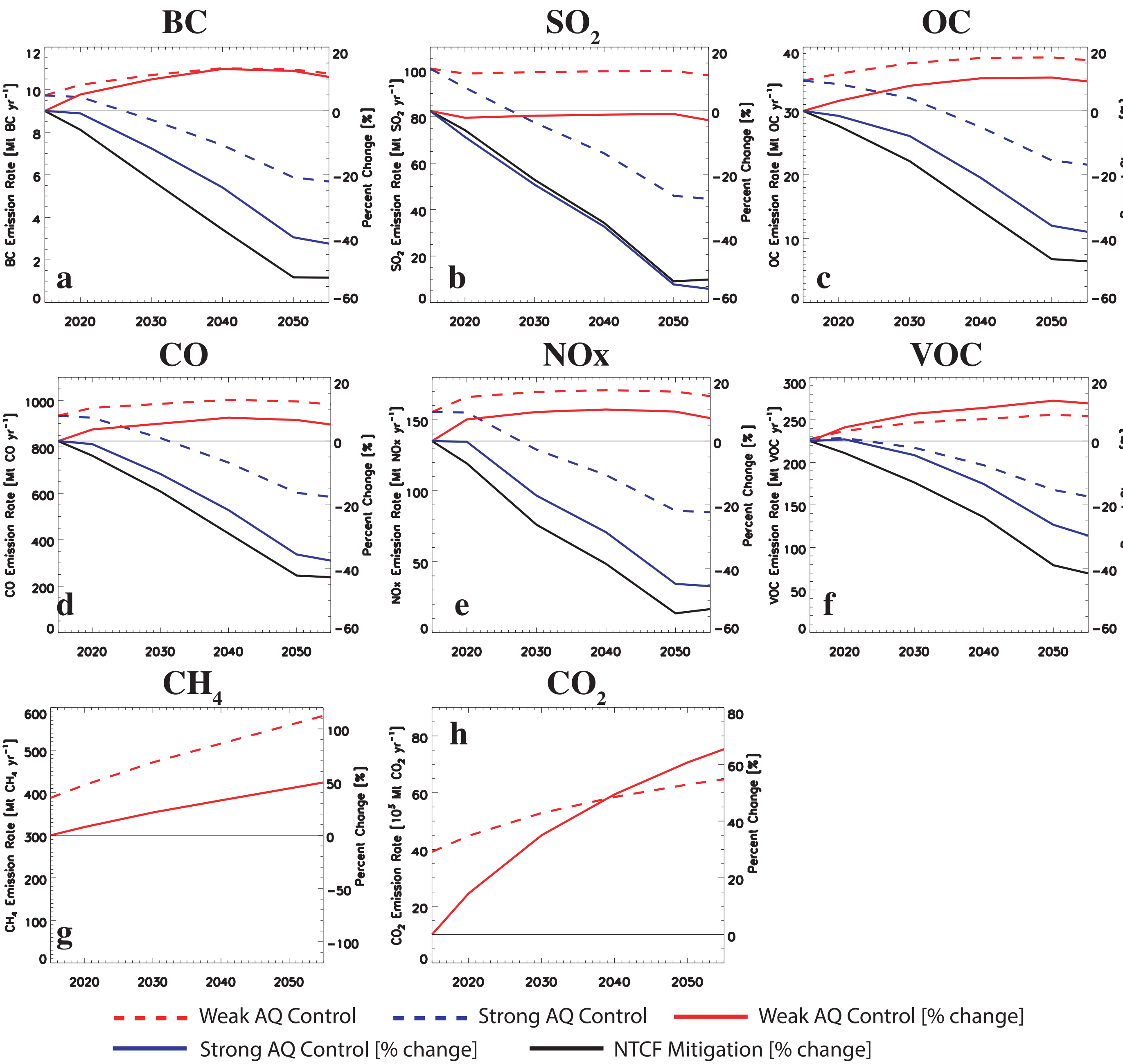


**Figure 1. 2015-2055 global mean $CO_2$, NTCF and precursor gas emissions.** Panels show (a) black carbon (BC); (b) sulfur dioxide ($SO_2$); (c) organic carbon (OC); (d) carbon monoxide (CO); (e) nitrogen oxides ($NO_x$); (f) volatile organic compounds (VOC); (g) methane ($CH_4$); and (h) carbon dioxide ($CO_2$) emissions for weak (red) and strong (blue) air quality control. Also included is the percent change (relative to 2015) for weak (red solid) and strong (blue solid) air quality control, and NTCF mitigation (black solid). Emission units for species X are Mt X $yr^{-1}$. Percent change units are %. **Only weak air quality control $CO_2$ and $CH_4$ emissions are shown, as AerChemMIP simulations include the same change in $CO_2$ and $CH_4$ emissions based on the weak air quality control scenario. Emissions data comes directly from the CMIP6 forcing datasets, which were downloaded from the input datasets for Model Intercomparison Project (input4MIPS).**

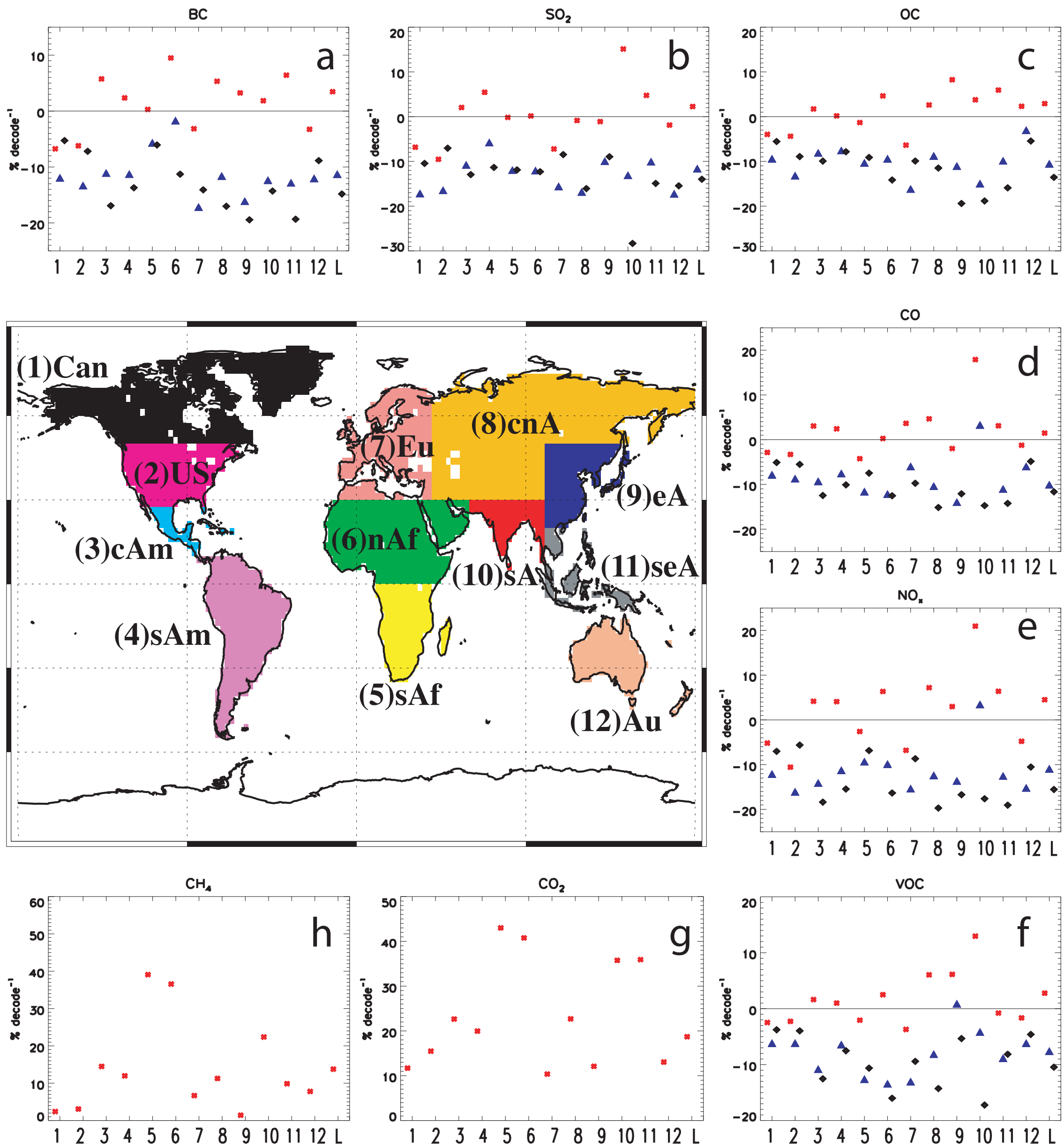


**Figure 2. 2015-2055 regional mean $CO_2$, NTCF and precursor gas emission trends.** Regional 2015-2055 emission trends for (a) black carbon (BC); (b) sulfur dioxide ($SO_2$); (c) organic carbon (OC); (d) carbon monoxide (CO); (e) nitrogen oxides ($NO_x$); (f) volatile organic compounds (VOC); (g) methane ($CH_4$); and (h) carbon dioxide ($CO_2$) for weak (red **asterisks**) and strong air quality control (SSP3-7.0-lowNTCF; blue triangles) and NTCF mitigation (SSP3-7.0-lowNTCF−SSP3-7.0; black diamonds). Center map shows the corresponding color coded world regions, based on Seneviratne et al. (2012). **The following abbreviations are used: Canada = 1 (Can; black), United States = 2 (US; magenta), central America = 3 (cAm; sky blue), south America = 4 (sAm; purple), south Africa = 5 (sAf; yellow), north Africa = 6 (nAf; green), Europe = 7 (Eu; pink), central and north Asia = 8 (cnA; orange), east Asia = 9 (eA; navy), south Asia = 10 (sA; red), southeast Asia = 11 (seA; gray), and Australia = 12 (Au; beige). The average over these 12 land regions is abbreviated as "L". Trend units are % decade$^{-1}$ (relative to 2015). Only weak air quality control $CO_2$ and $CH_4$ emission trends are shown, as AerChemMIP simulations include the same change in $CO_2$ and $CH_4$ emissions based on the weak air quality control scenario. Emissions data comes directly from the CMIP6 forcing datasets, which were downloaded from the input datasets for Model Intercomparison Project (input4MIPS).**

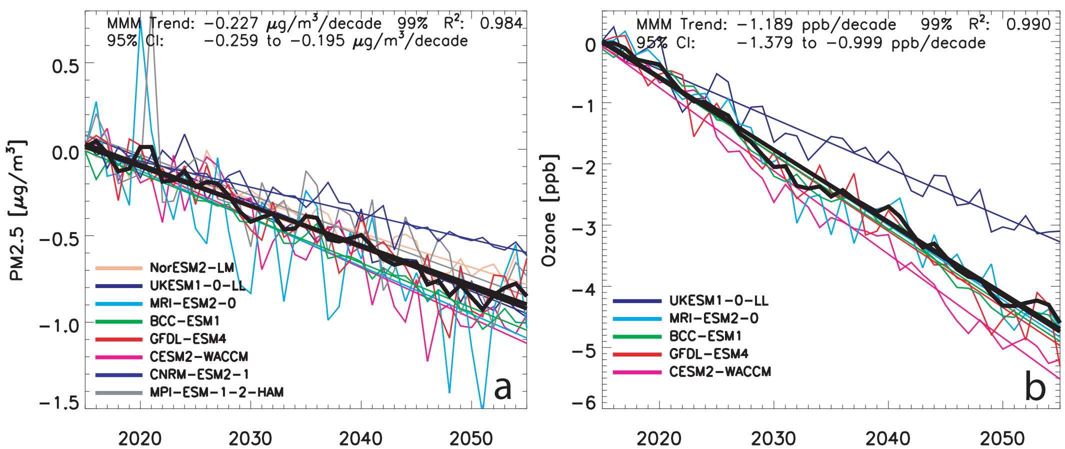


**Figure 3. 2015-2055 time series of global annual mean air pollution due to NTCF mitigation.** Panels show (a) surface particulate matter ($PM_{2.5}$) [$\mu$g m$^{-3}$] and (b) surface ozone [ppb] for NTCF mitigation. **The multi-model mean time series, and the corresponding trend estimated using a weighted least-squares regression, are included as thick black lines. The multi-model mean (MMM) trend, its significance and $R^2$ value, are also included, as is the 95% confidence interval (CI). Individual model mean trends are also included as defined by the legend.**

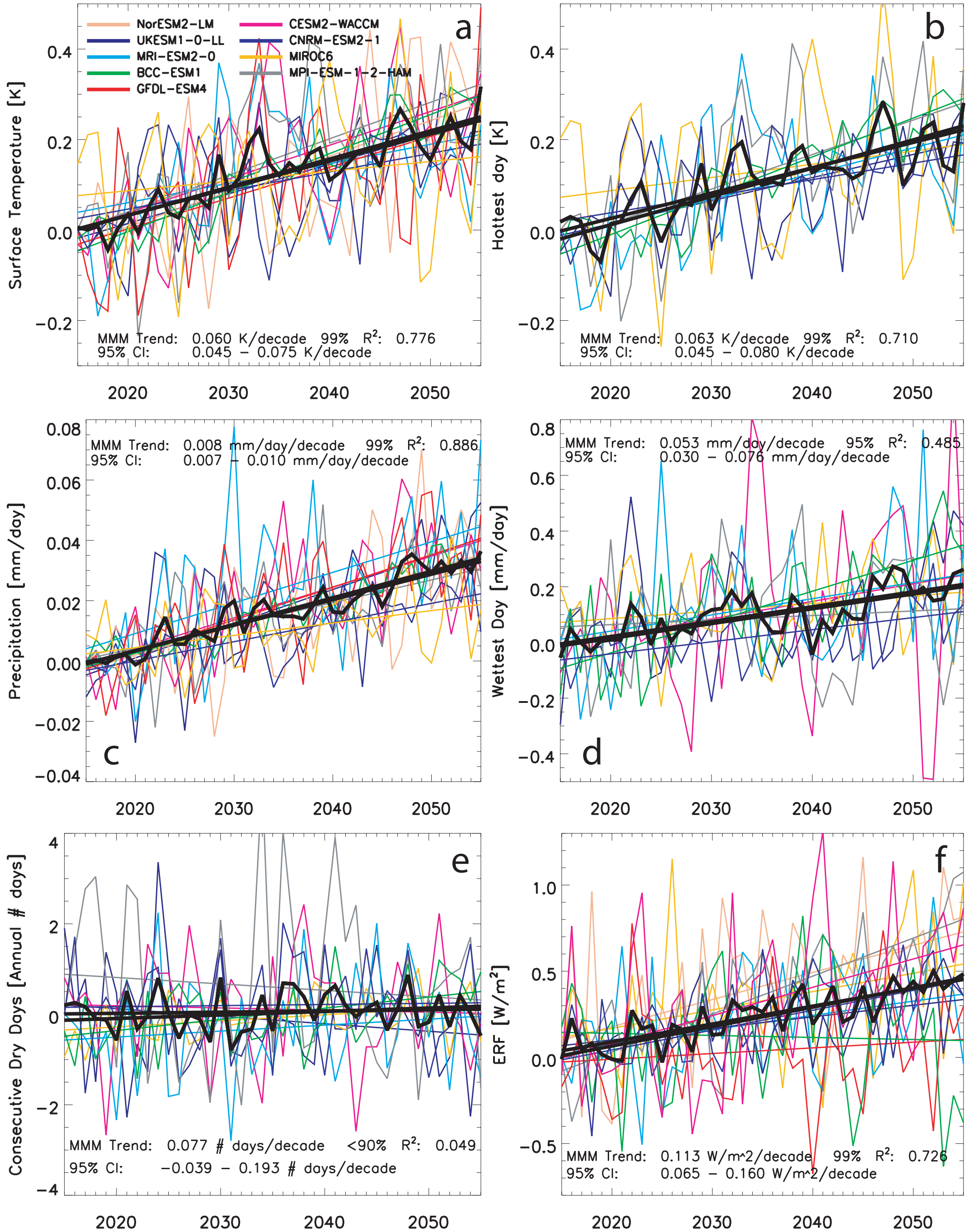


**Figure 4. 2015-2055 time series of global annual mean climate variables due to NTCF mitigation.** Panels show (a) surface temperature [K]; (b) hottest day [K]; (c) precipitation [mm day$^{-1}$]; (d) wettest day [mm day$^{-1}$]; (e) consecutive dry days [annual number of days]; and (f) effective radiative forcing (ERF) [W m$^{-2}$] for NTCF mitigation. **The multi-model mean time series, and the corresponding trend estimated using a weighted least-squares regression, are included as thick black lines. The multi-model mean (MMM) trend, its significance and $R^2$ value, are also included, as is the 95% confidence interval (CI). Individual model mean trends are also included as defined by the legend.**

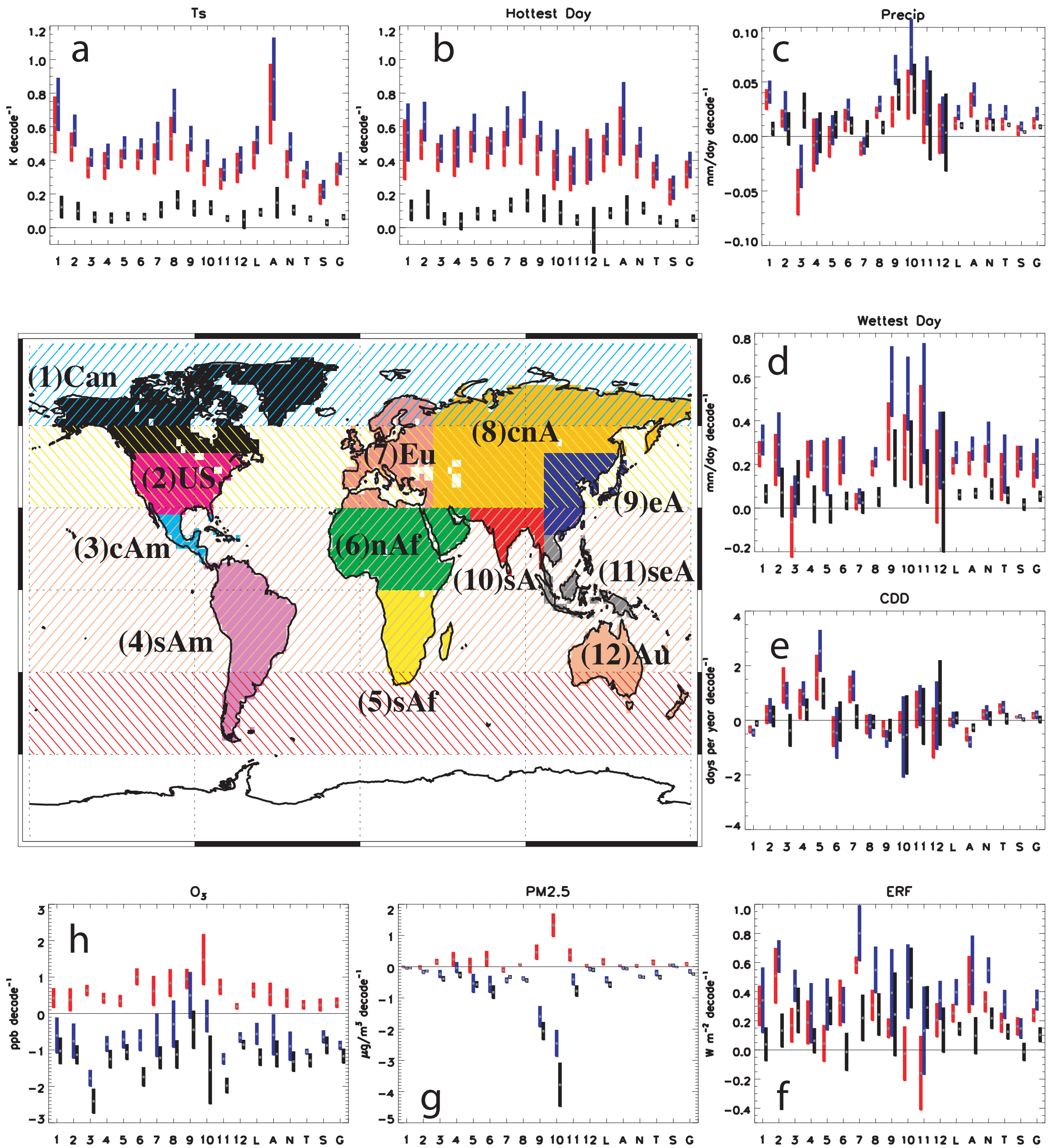


**Figure 5. Regional climate and air pollution responses to NTCF mitigation.** Bar plots show regional 2015-2055 trends in (a) surface temperature ($T_s$); (b) hottest day; (c) precipitation (Precip); (d) wettest day; (e) consecutive dry days (CDD); (f) effective radiative forcing (ERF); (g) surface particulate matter ($PM_{2.5}$) and (h) ozone ($O_3$) for weak (red) and strong (blue) air quality control, and NTCF mitigation (black). **Bar center (gray horizontal line) shows the multimodel mean trend, estimated as the average of each model's mean trend. Bar length represents the 95% confidence interval, estimated as $2\sigma/\sqrt{n}$, where $\sigma$ is the standard deviation of the individual model mean trends and $n$ is the number of models. Center map shows the corresponding color coded world regions for each bar plot (as in Fig. 2). The average over these 12 land regions is abbreviated as "L". Also included is the Arctic ("A"; 60-90N; light blue hatched region); NH mid-latitudes ("N"; 30-60N; yellow hatched region); Tropics ("T"; 30S-30N; beige hatched region); SH mid-latitudes ("S"; 60-30S; red hatched region); and the global mean ("G"). Trend units are K decade$^{-1}$ for $T_s$ and hottest day; mm day$^{-1}$ decade$^{-1}$ for Precip and wettest day; $\mu$g m$^{-3}$ decade$^{-1}$ for $PM_{2.5}$; ppb decade$^{-1}$ for $O_3$; days per year decade$^{-1}$ for CDD; and W m$^{-2}$ decade$^{-1}$ for ERF.**

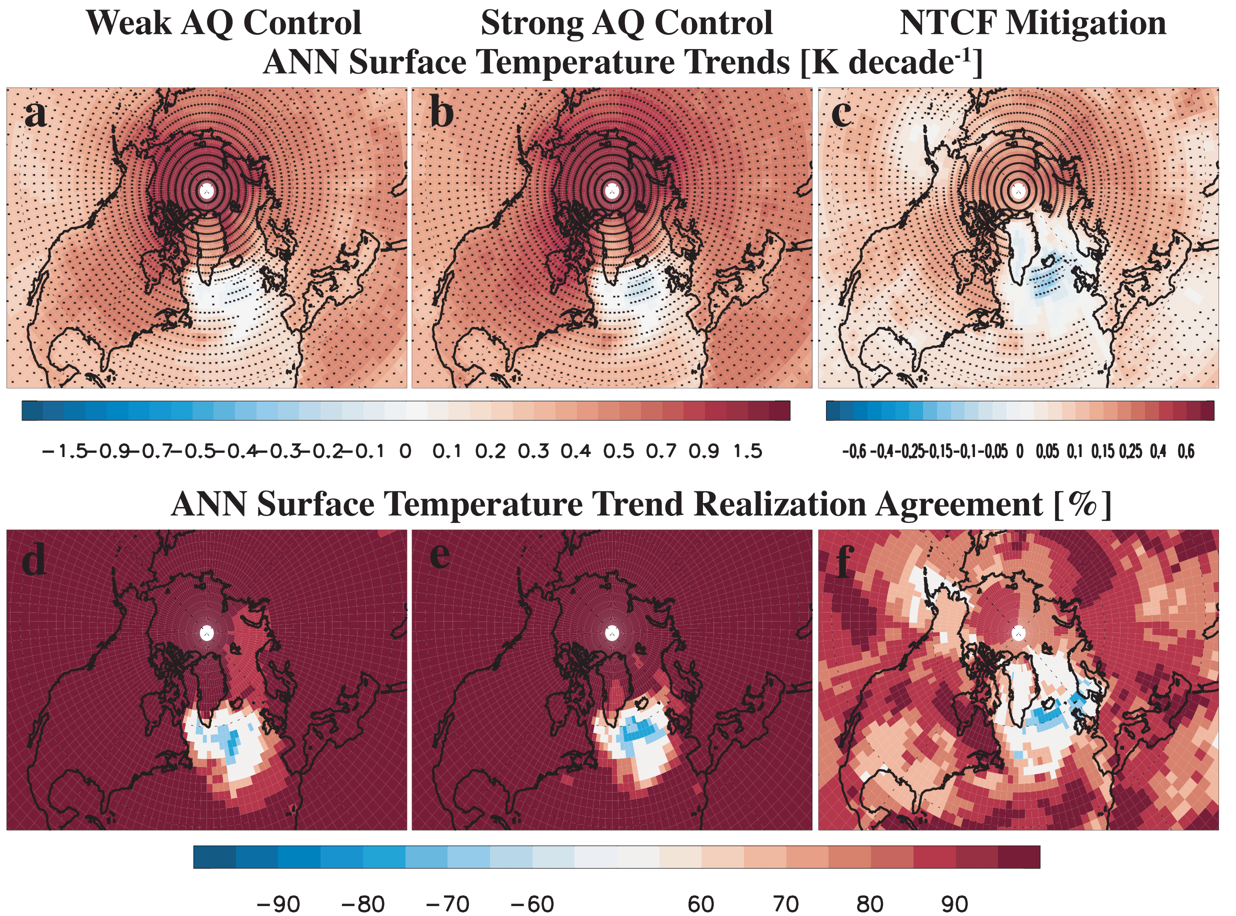


**Figure 6. 2015-2055 annual mean surface temperature trends and model trend realization agreement over the Arctic**. Surface temperature (a-c) trends [K decade$^{-1}$] and (d-f) model trend realization agreement [%] for (left panels) weak air quality control; (middle panels) strong air quality control and (right panels) NTCF mitigation. Stippling denotes trend significance at the 95% confidence level based on a standard $t$-test. Trend realization agreement represents the percentage of models that agree on the sign of the trend. **Red colors indicate model agreement on a positive trend; blue colors indicate model agreement on a negative trend. White areas indicate lack of agreement on the sign of the trend.**

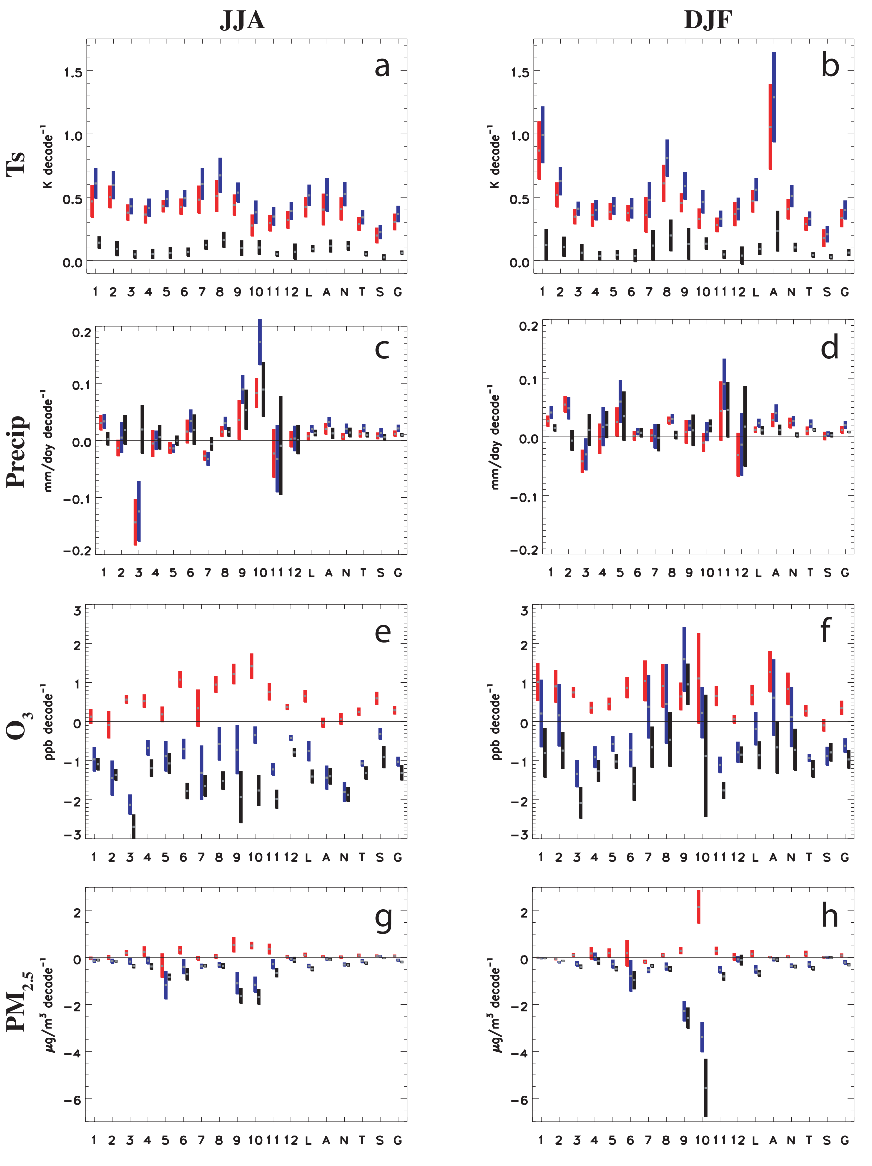


**Figure 7. Regional climate and air pollution seasonal responses to NTCF mitigation.** Bar plots show regional 2015-2055 June-July-August (JJA; left panels) and December-January-February (DJF; right panels) trends in (a-b) surface temperature ($T_s$); (c-d) precipitation (Precip); (e-f) surface ozone ($O_3$); and (g-h) surface particulate matter ($PM_{2.5}$) for weak (red) and strong (blue) air quality control, and NTCF mitigation (black). Bar center (gray horizontal line) shows the multimodel mean trend, estimated as the average of each model's mean trend. Bar length represents the 95% confidence interval, estimated as $2\sigma/\sqrt{n}$, where $\sigma$ is the standard deviation of the individual model mean trends and $n$ is the number of models. World regions are identical to those in Figure 5. Trend units are K decade$^{-1}$ for $T_s$; mm day$^{-1}$ decade$^{-1}$ for Precip; $\mu$g m$^{-3}$ decade$^{-1}$ for $PM_{2.5}$; and ppb decade$^{-1}$ for $O_3$.

**Table 1. Air pollution and climate responses to NTCF mitigation**. Annual mean 2015-2055 trends in surface particulate matter ($PM_{2.5}$), ozone ($O_3$), surface temperature ($T_s$), precipitation (Precip), hottest day, wettest day, consecutive dry days (CDD) and the effective radiative forcing (ERF) for NTCF mitigation. First set of numbers is the global mean trend; second set of numbers is the land-only trend. Trends significant at the 95% confidence level are denoted by bold font based on a $t$-test. Trend units are K decade$^{-1}$ for $T_s$ and hottest day; mm day$^{-1}$ decade$^{-1}$ for Precip and wettest day; $\mu$g m$^{-3}$ decade$^{-1}$ for $PM_{2.5}$; ppb decade$^{-1}$ for $O_3$; days per year decade$^{-1}$ for CDD; and W m$^{-2}$ decade$^{-1}$ for ERF. The first five models include both aerosol and ozone changes (Aer+O3 models); bottom four models include only aerosol changes (Aer models). MMM is the multi-model mean and the last row ("MMM Total") shows the total change over the entire 2015-2055 time period based on all models.

| | | | | Aer+O3 Models | | | | |
|---|---|---|---|---|---|---|---|---|
| | $PM_{2.5}$ | $O_3$ | $T_s$ | Precip | Hottest Day | Wettest day | CDD | ERF |
| UKESM1-0-LL | **−0.26/−0.67** | **−0.81/−0.81** | **0.07/0.09** | **0.011/0.017** | **0.05/0.05** | **0.055/0.100** | −0.17/−0.36 | **0.07**/0.02 |
| BCC-ESM1 | **−0.26/−0.51** | **−1.22/−1.13** | **0.09/0.14** | **0.009/0.010** | **0.09/0.14** | **0.111/0.095** | **0.25**/0.40 | −0.01/0.07 |
| GFDL-ESM4 | **−0.24/−0.57** | **−1.25/−1.26** | **0.07/0.08** | **0.011**/0.004 | n/a | n/a | n/a | 0.03/**0.13** |
| CESM2-WACCM | **−0.29/−0.78** | **−1.36/−1.39** | **0.08/0.10** | **0.010**/0.007 | n/a | 0.63/0.033 | −0.05/−0.06 | **0.16/0.23** |
| MRI-ESM2-0 | **−0.28/−0.58** | **−1.22/−1.42** | **0.04/0.07** | **0.010**/0.007 | **0.06/0.09** | **0.058**/0.041 | 0.13/0.26 | **0.08/0.16** |
| MMM | **−0.26/−0.59** | **−1.19/−1.11** | **0.07/0.10** | **0.009**/0.012 | **0.07/0.11** | **0.064/0.067** | 0.12/0.07 | **0.07/0.12** |
| | | | | Aer Models | | | | |
| | $PM_{2.5}$ | $O_3$ | $T_s$ | Precip | Hottest Day | Wettest day | CDD | ERF |
| CNRM-ESM2-1 | **−0.15/−0.41** | n/a | **0.04/0.05** | **0.006/0.014** | **0.04/0.05** | **0.043/0.068** | 0.03/0.01 | **0.12/0.16** |
| MIROC6 | n/a | n/a | 0.02/0.04 | **0.004**/0.010 | 0.03/0.06 | 0.027/0.024 | **0.13/0.41** | **0.11/0.21** |
| MPI-ESM1-2-HAM | **−0.23/−0.58** | n/a | **0.08/0.13** | **0.008**/0.010 | **0.08/0.12** | 0.016/0.041 | −0.14/−0.05 | **0.22/0.17** |
| NorESM2-LM | **−0.20/−0.48** | n/a | **0.08/0.11** | **0.010/0.012** | n/a | n/a | n/a | **0.16/0.16** |
| MMM | **−0.16/−0.44** | n/a | **0.06/0.09** | **0.005/0.009** | **0.04/0.10** | **0.039/0.061** | **0.07**/0.14 | **0.17/0.19** |
| | | | | All Models | | | | |
| | $PM_{2.5}$ | $O_3$ | $T_s$ | Precip | Hottest Day | Wettest day | CDD | ERF |
| MMM | **−0.23/−0.55** | **−1.19/−1.11** | **0.06/0.09** | **0.008/0.011** | **0.06/0.09** | **0.053/0.054** | 0.08/0.09 | **0.11/0.15** |
| MMM Total | −0.92/−2.20 | −4.76/−4.55 | 0.25/0.36 | 0.032/0.045 | 0.26/0.36 | 0.212/0.221 | 0.32/0.37 | 0.44/0.59 |