Supplementary Information for

Clean air policies are key for successfully mitigating Arctic warming

Knut von Salzen, Cynthia H. Whaley, Susan C. Anenberg, Rita Van Dingenen, Zbigniew Klimont, Mark G. Flanner, Rashed Mahmood, Stephen R. Arnold, Stephen Beagley, Rong-You Chien, Jesper H. Christensen, Sabine Eckhardt, Annica M. L. Ekman, Nikolaos Evangeliou, Greg Faluvegi, Joshua S. Fu, Michael Gauss, Wanmin Gong, Jens L. Hjorth, Ulas Im, Srinath Krishnan, Kaarle Kupiainen, Thomas Kühn, Joakim Langner, Kathy S. Law, Louis Marelle, Dirk Olivié, Tatsuo Onishi, Naga Oshima, Ville-Veikko Paunu, Yiran Peng, David Plummer, Luca Pozzoli, Shilpa Rao, Jean-Christophe Raut, Maria Sand, Julia Schmale, Michael Sigmond, Manu A. Thomas, Kostas Tsigaridis, Svetlana Tsyro, Steven T. Turnock, Minqi Wang, Barbara Winter

*Corresponding author. Email: knut.vonsalzen@ec.gc.ca

Content of this file:

Supplementary Notes 1-6:
>   Supplementary Note 1: Emissions
>   Supplementary Note 2: Climate and air quality models
>   Supplementary Note 3: Model Evaluation
>   Supplementary Note 4: Projections of future climate in the climate models
>   Supplementary Note 5: Projections of future air pollutants in the models
>   Supplementary Note 6: Health impacts

Supplementary Figures 1-11:
>   Supplementary Figure 1: Arctic Council countries annual mean greenhouse gas and
>   particulate matter component emissions
>   Supplementary Figure 2: Arctic Council Asian Observer countries annual mean
>   greenhouse gas and particulate matter component emissions
>   Supplementary Figure 3: Global and Arctic O_3 and $PM_{2.5}$ model percent biases and the
>   multi-model mean (mmm) biases
>   Supplementary Figure 4: Measured and modelled annual mean time series of surface-
>   level O_3 at 11 Arctic ground stations
>   Supplementary Figure 5: Measured and modelled annual mean time series of surface-
>   level BC at 11 Arctic ground stations
>   Supplementary Figure 6: Measured and modelled annual mean time series of surface-
>   level SO_4 at 13 Arctic locations
>   Supplementary Figure 7: Monthly mean AOD comparison between AERONET
>   observations and model simulations
>   Supplementary Figure 8: The 1995 to 2015 AOD trend
>   Supplementary Figure 9: Projected global and Arctic mean temperature changes in 2030
>   Supplementary Figure 10: Impacts of changes in air pollution on future global climate
>   Supplementary Figure 11: Projected global and Arctic mean temperature changes in 2050

Supplementary Acknowledgements

Supplementary References

Supplementary Note 1: Emissions

Supplementary Figures 1 and 2 present total emissions of greenhouse gases and selected (important from the perspective of radiative forcing) air pollutants for Arctic Council member countries (Canada and United States, Kingdom of Denmark, Finland, Iceland, Norway, the Russian Federation, and Sweden) and Arctic Council Asian Observers (Japan, People's Republic of China, Republic of India, Republic of Korea, Republic of Singapore). Emissions are shown for the key CMIP6 scenarios, including historical estimates from the CEDS dataset[1,2], and the data and scenario set used in this work developed with the GAINS[3,4,5] model and referred to as AMAP (Arctic Monitoring and Assessment Programme) or ECLIPSE V6b; for a description of scenarios see the main text.

While for CO_2, the GAINS model historical data and projections (relying on the International Energy Agency statistics and the World Energy Outlook scenarios[6]) are consistent with CEDS and two CMIP6 SSP scenarios (SSP2-4.5 for CLE, MFR, and CFM; and SSP1-1.9 for MFR_SDS). There are noticeable differences for methane (Supplementary Figure 1A and Supplementary Figure 2A). For Arctic Council countries, a strong decline in CH_4 in the 1990-2000 decade is associated with the collapse of the Soviet Union and both CEDS and GAINS show this, however, GAINS estimates are higher due higher losses from venting. The growth towards 2015 is driven primarily by increasing production of gas, including unconventional gas sources in North America. For the Arctic Council Asian Observer countries, CEDS and GAINS align well until about 2000 but then CEDS emissions grow much stronger than GAINS, primarily due to difference in estimates for coal mining in China. It appears that CEDS assumes constant emission factors for China post 2005 and is similar to earlier estimates of EDGAR inventory, while GAINS is similar to recent EDGAR (not shown). Inverse modelling studies[7,8] indicate lower coal mining emissions in China than CEDS and even slightly lower than GAINS. At the global level, GAINS model CH_4 trajectory shows good alignment with inverse modelling study for the post 2000 period[9,10]. Overall mitigation potential, relative to baseline, is similar for SSP and GAINS scenarios, with SSP1-2.6 and MFR_SDS (sustainable development) aligning closer also in absolute terms.

Globally, CEDS emissions of aerosol species (sulfur, BC, and OC) are consistently higher than GAINS emissions[11] (Fig. 1), both in totals but also on a sectoral basis, with major differences over southeast Asia. On the other hand, the GAINS emission decline more rapidly from 2015 to 2030, compared to the SSP emissions. Emissions in the SSP1-2.6 scenario and the MFR scenario are similar in 2050.

Sulfur emissions trajectories for the Arctic Council are similar for the past (CEDS and GAINS) and follow comparable trends in the future (SSP vs GAINS, Supplementary Figure 1B). For the Asian Observers (Supplementary Figure 2B), however, there are some notable differences that appear from about 2005, with much stronger and continued increase in CEDS while GAINS growth is slower and a decline from about 2007 is estimated, resulting in a much lower value for 2015 that is a starting point for projections. Key reasons are linked to assumptions about the effectiveness of policies in China where GAINS relies on more recent studies for China[10] - not available at the time of CEDS development - showing a strong decline of sulfur emissions. This has important implications for the projections, while having somehow

3

similar trends, show different absolute emission levels especially for 2030 and the SSP2-4.5 (comparable in terms of energy consumption to CLE); by 2050 the deep mitigation cases are rather similar. Notably, recent updates of the CEDS database that also include extension towards 2017, confirm the historical trends estimated for China in the GAINS model[12].

For carbonaceous particles (BC and OC), the CEDS, GAINS, and SSPs are very similar for the Arctic Council (Supplementary Figure 1B). Only for BC, historical GAINS numbers are somewhat higher than CEDS owing to differences in emissions from oil and gas flaring in Russia and generally larger estimates for agricultural residue burning using bottom up data rather than remote sensing products. For Asian Observers, however, the differences are much larger (Supplementary Figure 2B). Historical values deviate strongly from 2000, with CEDS showing very strong growth while GAINS is stable with a decline in later years. The key reasons include higher CEDS BC emissions from industrial use of coal and coke production in China, open burning of municipal waste in India (affecting mostly OC) as well as moderate change in emissions from diesel vehicles. GAINS has assumed lower emission factors for coal use in industry, transformation in coke sector in China towards larger integrated plants, decline in BC emission factor from diesel trucks and cars, and own estimates of municipal solid waste generation and rates of open burning (the latter lower than the study CEDS relied on[13]). Two recently published studies evaluating BC emissions in the last decade in China confirm the trends estimated with GAINS[14,15] and the recent update of the CEDS emission dataset shows very similar trajectories to GAINS as well[12]. The trends and overall mitigation potential in respective SSP and GAINS scenarios for BC and OC is similar, however, in absolute terms GAINS attains much lower BC and OC emissions for the Arctic Observer countries by 2050 - among major reason for additional potential in GAINS is assumption about successful transformation in the waste sector virtually eliminating open burning of trash.

Supplementary Note 2: Climate and air quality models

Participating in this study were 8 Chemical Transport Models (CTMs), 2 Chemistry Climate Models (CCMs), 3 global climate models (GCMs), and 5 Earth System Models (ESMs) for a total of 18 models. A list of these models is provided (Supplementary Data 1), along with some details about meteorology, emissions and primary references. For individual model descriptions, we refer the reader to the 2021 AMAP Short-Lived Climate Forcers (SLCF) assessment report and Whaley et al.[16].

All models used forest fire emissions from the Coupled Model Intercomparison Project Phase 6 (CMIP6), except for DEHM, which used GFAS; EMEP MSC-W, which used FINN; GEM-MACH, which used CFFEPS; and GEOS-Chem, OsloCTM, and WRF-Chem, which all used GFEDv4[2]. All models used AMAP emissions for the agricultural waste burning, except for GEM-MACH, which used Canadian and U.S. inventories for North America, and AMAP emissions everywhere else; and UKESM1, which used CMIP6 for the historical time period, though AMAP emissions[5] for future time periods.

The 18 models have highly varying horizontal and vertical resolution, as well as vertical extent. For example, 8 models (DEHM, EMEP MSC-W, FLEXPART, GEM-MACH, GEOS-CHEM, MATCH, Oslo-CTM, WRF-Chem) have an air quality focus, usually with relatively

high horizontal resolution, a lower atmospheric lid (top modelled level), focused on the troposphere, and with many trace gas and aerosol species, and chemical reactions. Two models (CMAM and MATCH-SALSA are chemistry climate models (CCMs), meaning they are focused on climate, with a lower horizontal resolution, higher atmospheric lid (including the stratosphere), and a considerable amount of atmospheric chemistry - particularly in the stratosphere. The remaining 3 atmospheric models (CanAM5-PAM, CIESM-MAM7, and ECHAM-SALSA) have a tropospheric climate focus. They have fewer chemical reactions, thus, can be run for long periods of time, and simulate important atmospheric processes for climate. Five of the participating models (CESM, GISS-E2.1, MRI-ESM2, NorESM, and UKESM1) are ESMs, simulating the atmosphere in a similar way to the climate models mentioned above, but also including interactive land surface and ocean. In addition to atmospheric processes, the ESMs can simulate changes to ocean temperatures and salinity, sea ice thickness and extent, and land-air exchange in an interactive way, whereas these are prescribed in the 13 atmospheric models mentioned above.

The ESM future simulations were done with 3 to 5 ensemble members to provide a range of future climate conditions given the internal variability.

Supplementary Note 3: Model Evaluation

Models used to provide results in the main manuscript were evaluated by comparing their SLCF output to ground, aircraft, ship, and satellite-based measurements. Most of that evaluation is available in the 2021 AMAP SLCF Assessment Report and further documented by Whaley et al.[16].

Generally, the multi-model mean provided the best results for most of the SLCF species compared to any one particular model (e.g. Supplementary Figure 3), and it is the multi-model mean that the climate and health impacts in the main paper are based on. This is optimal use of these model datasets.

Evaluation of simulated methane, black carbon, sulfate, ozone, and fine particulate matter

The annual, multi-model mean biases for each SLCF species in the historical model simulations (2008-9 and 2014-15 evaluation periods) are described below, including their implications for climate and health impacts:

- Annual mean, simulated surface-level Arctic (>60°N) and global methane is biased slightly high (+1.3% in the Arctic) compared to measurements and biased slightly low in the free troposphere (e.g., -0.7% at 600 hPa), which means that confidence in the warming from methane shown in the main paper is high.

- Annual mean, surface-level Arctic black carbon is biased low compared to measurements (-17%), but there is up to a factor of 2 uncertainty in BC measurements from different techniques. The warming from black carbon is due to its concentrations – not just at the surface – but throughout the atmospheric column. There is good agreement between models and observations over black carbon source regions, but simulated concentrations are underestimated at higher altitudes in the troposphere and generally increase with

altitude and proximity to the Arctic according to comparisons with aircraft measurements (e.g., -80% at 6 km in altitude for the Arctic region). This means that the warming due to black carbon shown in the main paper may be underestimated.

- Annual mean Arctic sulfate was biased low compared to measurements (-21%) but that bias is close to the 20% uncertainty on sulfate measurements. Vertical profiles of sulfate and radiative forcings were not evaluated in this project so modelled temperature impacts are uncertain.

- Globally, and in the Arctic, modelled surface-level O_3 is biased high (e.g., +6% in the Arctic). This means that the detrimental health impacts of ozone may be overestimated.

- Higher in the atmosphere (in the free troposphere) ozone is an effective greenhouse gas, and there it is slightly underestimated by models (-10 to -20% compared to satellite measurements, which have ~15% uncertainty range). This means that the modelled warming impact from O_3 may be slightly underestimated.

- Annual mean $PM_{2.5}$ was biased low globally (-26%), which means that the health impacts due to $PM_{2.5}$ exposure may be underestimated.

- Annual mean aerosol optical depth – giving an indication of the aerosol load in the atmosphere – was biased low (-23.2% in Arctic and -25% globally) suggesting that the overall impacts of aerosols on climate and human health may be underestimated.

- The differences and uncertainties in both the satellite-based observations and model simulations of cloud properties were high over the Arctic, especially over the regions covered by multiyear sea ice and during the polar winter months. This implies that the poor representation of the processes relevant to the interactions among aerosol-precursor gases, aerosols and clouds could be a major source of uncertainties in the climate impacts in this assessment.

There were seasonal differences in the model biases, which complicates the health and climate impact uncertainties. For example, the seasonal cycle of ozone has health implications, since ozone only mainly exceeds health thresholds in the summertime (photochemical smog) and $PM_{2.5}$ exceeds health thresholds mainly in the wintertime (haze). At mid-latitudes, where human populations are much higher, the models simulated the seasonal cycles well – thus the modelled health impacts should be fairly accurate as far as the seasonal cycles are concerned. However, the differences in simulated seasonal cycles of ozone, black carbon, and sulfate are particularly large in the Arctic. The coarse resolution models cannot simulate the local Arctic smog well, and due to large uncertainties in fire parameterizations, the summertime aerosol concentrations are often overestimated by models.

There were regional differences in model biases as well. For example, surface-level ozone had a much smaller positive bias in North America and Europe than in Asia, and modelled $PM_{2.5}$ was biased high in North America, but low in Asia. These, too, further complicate the interpretation of the health and climate impacts.

Supplementary Figure 3 shows the annual mean model biases for each model and the

multi-model mean (mmm) for O_3 and $PM_{2.5}$, globally (at all measurement locations), and in the Arctic (only measurement locations with latitude $> 60°N$). Positive and negative model biases tend to cancel out in the multi-model mean, except for global O_3 comparisons, where almost all models are biased high (mainly due to the high bias at mid-latitudes where most of the measurements occur).

Modelled and measured trends of Arctic O_3, CH_4, BC, and SO_4

The simulated historical trends for the time period 1990(5) to 2015 were compared to measured trends for O_3, BC, and SO_4. Their annual mean results are shown in Supplementary Figures 4 to 6, respectively. In addition, modelled and measured trend values for different seasons are included in the AMAP 2021 SLCF assessment report for O_3, BC, and SO_4.

Arctic O_3 trends are negligible at most Arctic locations from both models and measurements (Supplementary Figure 4), though when analyzed seasonally, there are some small (<1%) significant trends that the models do not capture. The relatively steady annual mean Arctic O_3 in observations and simulations are likely due to competing changes in O_3 precursors (e.g., decreasing or increasing NOx depending on region, increasing CH_4 globally, but decreasing CO).

Only three models simulated methane over the longer time period, and they capture the increasing Arctic CH_4 trend (not shown) when compared to observations at Alert, Canada.

Black carbon (Supplementary Figure 5) and sulfate (Supplementary Figure 6) concentrations show similar changes over the long time series. At most Arctic locations (e.g., shown clearly at Alert, NV), the BC and SO_4 trends are decreasing in the 1990s, then level off in the 2000s, consistent with slowing BC emission reductions in Europe, North America, and globally). The models capture these changes over time quite well. The similarity in model biases between black carbon and sulfate for individual models indicates a common source of error (e.g., aerosol wet deposition[17]).

Modelled and measured trends of aerosol optical depth

Aerosol optical depth (AOD), a proxy for aerosol concentrations in the atmosphere, was used to investigate recent trends in atmospheric concentrations of aerosols in observations and model simulations. 5 models simulated AOD over the period 1995-2015 and their trends were assessed. Supplementary Figure 7 compares AOD from 22 AErosol RObotic NETwork (AERONET[18]) stations selected based on their temporal coverage during 1995-2015. The stations for which the monthly mean values were available for at least 60% of this evaluation period were used. The results show that for most of the station locations in the US, Canada and Europe the observed aerosol concentrations are declining during this period, which is consistent with the reductions in emissions. These trends were well captured by all models although some models tend to overestimate the observed seasonal variations especially during summer. Similarly, the models were also able to capture the increasing AOD trend in Kanpur, India, suggesting that the models simulated well the aerosol concentrations in regions becoming more important for anthropogenic pollution in recent decades. However, over highly clean locations

such as at Mauna Loa the AOD is overestimated by all models. This is likely related to the difficulty in modelling very small background aerosol concentrations using coarse model grids.

The spatial features of simulated AOD trends are evaluated in Supplementary Figure 8 where they are compared to the Advanced Very High Resolution Radiometer (AVHRR) based estimation[19,20]. The modeled and AVHRR AOD data sets were remapped to a uniform 1° × 1° spatial grid before performing the trend analysis. Supplementary Figure 8 shows that, similar to the trends discussed above, the dominant features include decreasing AOD trends along eastern coasts of North America, the Atlantic coast of Europe and in the Mediterranean region, which is consistent with decreasing aerosol emissions. Over the Arabian Sea, Bay of Bengal, southwest coasts of Africa and Eastern coast of China the AVHRR AOD showed an increasing trend. The models are also able to capture these significantly negative trends along Eastern coast of the US, western coast of Europe and in the Mediterranean region and also the increasing trends in Arabian sea, Bay of Bengal and southwest coasts of Africa. However, the models differ in magnitudes and sign at some locations. For further discussions on uncertainties in observed AOD please see the AMAP SLCF Assessment Report.

Supplementary Note 4: Projections of future climate in the climate models

Five fully-coupled Earth system models with interactive ocean and sea ice components were used to determine the full climate impacts of the future AMAP emission scenarios in Fig. 2, including temperature, precipitation, and sea ice changes: NorESM-happi (3 ensemble members), CESM2 (4 ensemble members), MRI-ESM2 (5 ensemble members), GISS-E2.1 (3 ensemble members), and UKESM1 (3 ensemble members). The models were initialized from CMIP6 historical simulations, with the simulations for 4 models (CESM2, MRI-ESM2, GISS-E2.1, and UKESM1) branched off from the year 2000 and NorESM-happi branched off from 2011. Between branch off time and the year 2020, all simulations follow the CLE emission scenario, while the CLE, MFR, and CFM scenarios were used after that. Fig. 2 shows the multi-model median temperatures for 2046 - 2055 for these models, after averaging results from the ensemble members for each model, with 5-95% confidence intervals (± 1.64 σ), relative to 1995 – 2014. To calculate global and Arctic temperatures relative to 1880 - 1920, we added temperature offsets that were obtained from the GISS Surface Temperature Analysis, GISTEMP v4[21,22].

In addition to the simulations for AMAP scenarios, we use results from archived simulations with CMIP6 global ESMs for the SSP scenarios in Fig. 2. All results for the first available ensemble member were used to calculate multi-model median temperatures and their confidence ranges, without screening, from between 29 (SSP3-7.0) and 35 (SSP2-4.5 and SSP5-8.5) CMIP6 models (Supplementary Data 2).

Supplementary Note 5: Projections of future air pollutants in the models

Future air quality will be primarily driven by changes in emissions but will be impacted by changes in climate, which go beyond changes in air pollutant sources[23,24]. The combined 3D air quality and climate multi-model ensemble used here provides a harmonized approach for the analysis of air quality changes and climate-related processes, caused by changes in key fossil

fuel air pollutants.

We find that model biases for historical conditions can be substantial, largely consistent with a separate analysis of CMIP6 models. These biases are likely to affect future projections of air quality. In particular, $PM_{2.5}$ concentrations tend to be underestimated in many models (see sections above). Although there are differences in key model features, air quality and climate models share common biases, which provides evidence that parameterizations of air pollutant processes are often at the root of model biases rather than differences in meteorological processes and model resolution, for instance. Specifically, seasonally varying concentration biases in many of the models are likely related to the representation of wildfires and mineral dust aerosol. Similarities in model biases and the persistence of known model biases over many years of model development by the modelling community highlight the need for future improvements of parameterizations of air pollutant processes and emissions from the terrestrial and oceanic biosphere.

Changes in biospheric emissions may lead to important climate and air quality feedbacks[24], which will need to be addressed for future assessments of SLCF emission impacts on human health. The models used here specify wildfire emissions according to the SSP2-4.5 scenario, for all the simulated AMAP scenarios. Some models simulate changes in emissions from the ocean, sea salt, or mineral dust. None of the models simulates changing emissions from permafrost thaw. Given these limitations, an analysis of the different approaches for naturally emitted air pollutants and their impacts on future climate and human health is beyond the scope of this study.

Regardless of different air quality and climate model biases, simulated multi-model mean air pollutant concentrations from anthropogenic sources agree reasonably well with observations, often within the accuracy of the observational data sets. Additionally, air pollution trends agree well across the global models that simulated future changes in air quality, despite considerable differences in simulated chemical processes and emissions from natural sources (climate models: CanAM5-PAM, CESM2.1, GISS-E2.1, MRI-ESM2, UKESM1; air quality model: EMEP MSC-W; Supplementary Data 1). For instance, for Asian Observers, the difference in mean $PM_{2.5}$ concentration in 2050 between the MFR and CLE scenarios amounts to about 45% of the anthropogenic $PM_{2.5}$ concentrations in 2015 in this region (or about 6 μg m-3), with a model range from -53.4% to -43.4%.

For the analysis of simulated $PM_{2.5}$ concentrations in this assessment, we only consider three key aerosol chemical species, which are strongly affected by fossil fuel emissions and are simulated in all models: BC, OC, and sulfate. Although ammonium and nitrate are predicted by some of the models, these aerosol species are not explicitly included in the analysis of $PM_{2.5}$ concentrations, given the limited amount of data from the multi-model ensemble and model validation options. Instead, we approximate $PM_{2.5}$ concentrations by assuming that sulfate is fully neutralized by ammonium, the usual approach in models without interactive ammonium and nitrate processes.

Multi-model mean population-weighted anthropogenic $PM_{2.5}$ concentrations in 2015 (Fig. 4) are based on simulations with the following global models: CanAM5-PAM, CESM2.1,

CIESM-MAM7, ECHAM6-SALSA, EMEP MSC-W, GEOS-Chem, GISS-E2.1, MRI-ESM2, Oslo CTM, and UKESM1. Future multi-model mean concentrations in years 2030 and 2050 are based on a subset of these models that were available (CanAM5-PAM, CESM2.1, EMEP MSC-W, GISS-E2.1, MRI-ESM2, and UKESM1).

Projected future changes in speciated population-weighted anthropogenic $PM_{2.5}$ concentrations in emulator simulations (Fig. 4) are driven by changes in emissions of sulfur, OC, and BC from upstream oil and gas production sources, combined fossil and bio fuel sources related to energy consumption, and marine shipping. The emulator projections do not account for changes in natural emission sources, particularly including wildfires. Furthermore, we do not consider changes in sea salt and mineral dust $PM_{2.5}$ species, or changes in oxidant concentrations. However, anthropogenic $PM_{2.5}$ concentrations in 2015, i.e. the base year for the emulator simulations, are derived from combined emulator and downscaled 3D multi-model $PM_{2.5}$ concentrations, which are available for sulfate, OC, and BC in 2015. Therefore, natural sources of these species are implicitly included in the emulator baseline concentrations.

We compared the simulated $PM_{2.5}$ concentrations to observation-based estimates of $PM_{2.5}$ concentrations in different regions (Fig. 4), based on satellite retrievals[25]. This estimate does not include contributions of sea salt and mineral dust to total $PM_{2.5}$. Although uncertainties are substantial, this estimate agrees reasonably well with in-situ observations, which also tend to produce higher concentrations than the multi-model mean.

Simulations of ozone were conducted by a subset of the available air quality and climate models. Multi-model mean results for 2015 (Fig. 5) are based on 3-hourly concentrations from simulations with CESM2.1, EMEP MSC-W, MRI-ESM2, and UKESM1, which are all global models. There are no emulator results available for ozone concentrations and impacts of emission changes on regional ozone concentrations were analyzed using the parameterization by Turnock et al.[26] by AMAP.

Supplementary Note 6: Health impacts

$PM_{2.5}$- and ozone-attributable mortality are calculated using the TM5-FASST tool[27]. Methods generally follow those used for the Global Burden of Disease (GBD) 2017 Study. For $PM_{2.5}$, we include mortality from ischemic heart disease, stroke, chronic obstructive pulmonary disease, lung cancer, lower respiratory infection, and diabetes mellitus type 2. For ozone, we include only chronic obstructive pulmonary disease.

TM5-FASST provides pollution-attributable mortality estimates using the common population-attributable fraction approach, described in Equation S1:

Equation S1: $AM_{a,h} = Pop * M_{c,h} * (RR_{a,h}-1)/RR_{a,h}$

Where AM is pollution-attributable mortality for each age group a and health endpoint h, Pop is population, M is the baseline mortality rate for each country c and health endpoint h, RR is relative risk estimates for each age group a and health endpoint h from the GBD 2017 Study. Equation X is applied in each grid cell and then summed across grid cells to national boundaries.

For PM$_{2.5}$, RR estimates for each health endpoint (and, for the cardiovascular health endpoints, each 5-year age band) are computed from fitted functions representing the median exposure-response curve (and 95% confidence intervals) from GBD 2017[28] (A. Cohen 2019, personal communication) giving RR estimates for each annual PM$_{2.5}$ concentration value. For ozone, RR is calculated using Equation S2:

Equation S2: RR = exp (beta * dO$_3$)

Where beta is the concentration-response factor calculated from RR = 1.06 (95% CI 1.02, 1.10) per 10 ppb change in the 6-month average of the 1-hour daily maximum ozone concentration, and dO$_3$ is the change in the maximum 6-month average of the 1-hour daily maximum ozone concentration simulated for these emission scenarios. We use theoretical minimum risk exposure levels below which no health impacts are calculated of 29.1 ppbV for ozone, consistent with the GBD 2019 Study. For PM$_{2.5}$, the minimum risk threshold depends on the health endpoint and age group, and ranges between 3.7 and 4.3 µg/m³ for the median exposure-response function.

Country-specific baseline mortality rates for 2015 are from the GBD 2017 Study and rates for future years are projected by applying ratios of future estimated rates from the GBD 2016 foresight study (https://vizhub.healthdata.org/gbd-foresight/) to 2015 rates. We assume the year 2050 base mortality rates to be equal to the 2040 values, the last available year in the GBD2016 foresight data. Present and projected gridded population estimates are from Jones and O'Neill[29], while country-specific age structures are from the UN 2017 revision World Population Prospects.

We estimate 95% confidence intervals for estimated PM$_{2.5}$- and ozone-attributable mortality using error in the relative risk estimate only, though additional uncertainties arise from all other inputs to the health impact function, including concentrations, population, and mortality rates. We have attempted to bound our understanding of the simulated concentration changes for each scenario using a multi-model ensemble. All other sources of uncertainty are likely consistent across scenarios and therefore would not affect our results or conclusions about how air pollution-related mortality compares across emission scenarios. The exception may be for baseline mortality rates projected to future years: If our projection approach overestimates future mortality rates from the causes of death included in our study, we may overestimate the change in pollution-attributable mortality from emission changes, leading to exaggerated differences between scenarios. The opposite would be true if our projection approach underestimates mortality rates - the difference in air pollution-attributable mortality between emission scenarios would be muted.

Estimated air pollution-attributable health impacts may be underestimated because we only included PM$_{2.5}$ and ozone, though other air pollutants are affected by these emission scenarios and are linked with health outcomes[30,31] (e.g. NO$_2$ and pediatric asthma incidence). In addition, we only included several causes of mortality, and were not able to include a large range of morbidity endpoints that have been linked with air pollution in recent years, including dementia[32], chronic kidney disease[33], and adverse birth outcomes[34].
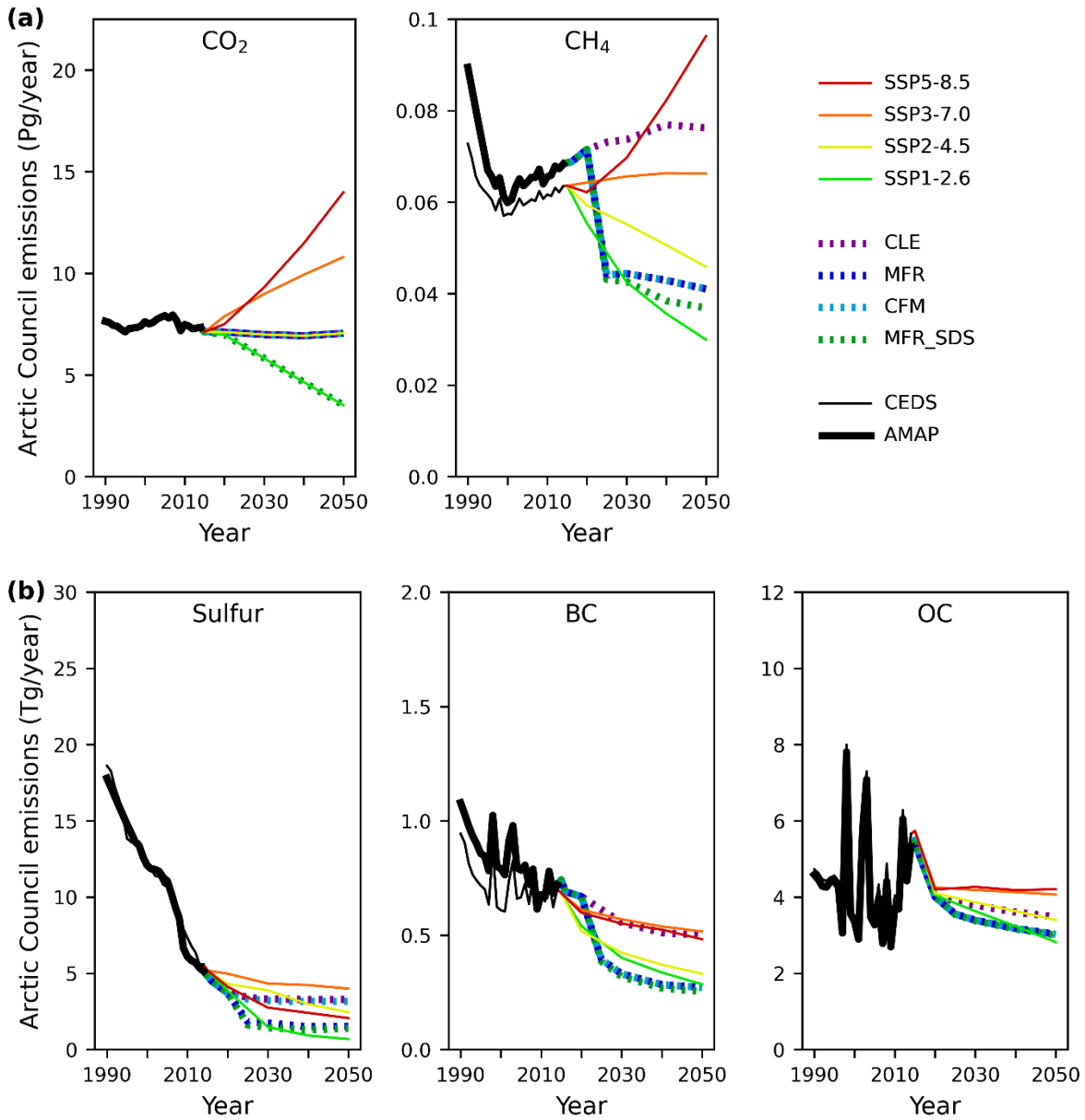
Supplementary Figure 1: Arctic Council countries annual mean greenhouse gas and particulate matter component emissions. a Historical carbon dioxide and methane emissions in two inventories (as indicated in the legend; thin black line: CEDS, thick black line: AMAP) and 8 future scenarios for 2015 to 2050 (colored lines and acronyms in legend; similar to Fig. 1). **b** Corresponding emissions of sulfur, black carbon (BC), and organic carbon (OC). Note that some of the scenarios are overlapping for some emitted species. The inter-annual variability in CH_4, BC, and OC emissions is largely due to wildfires.

Supplementary Figure 2: Arctic Council Asian Observer countries annual mean greenhouse gas and particulate matter component emissions. See Supplementary Figure 1 for details.
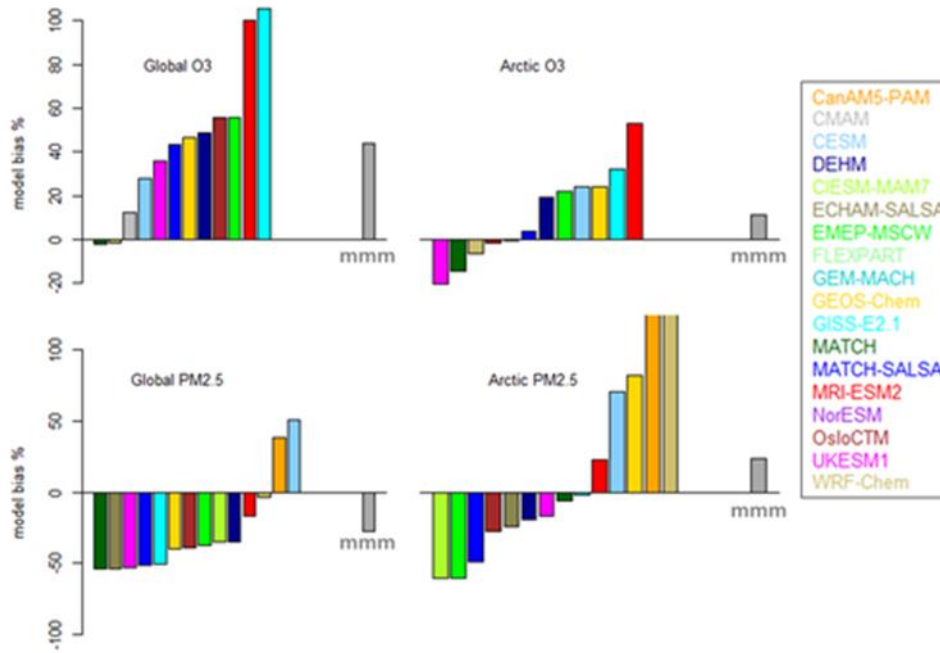
Supplementary Figure 3: Global and Arctic O$_3$ and PM$_{2.5}$ model relative biases and the multi-model mean (mmm) biases.
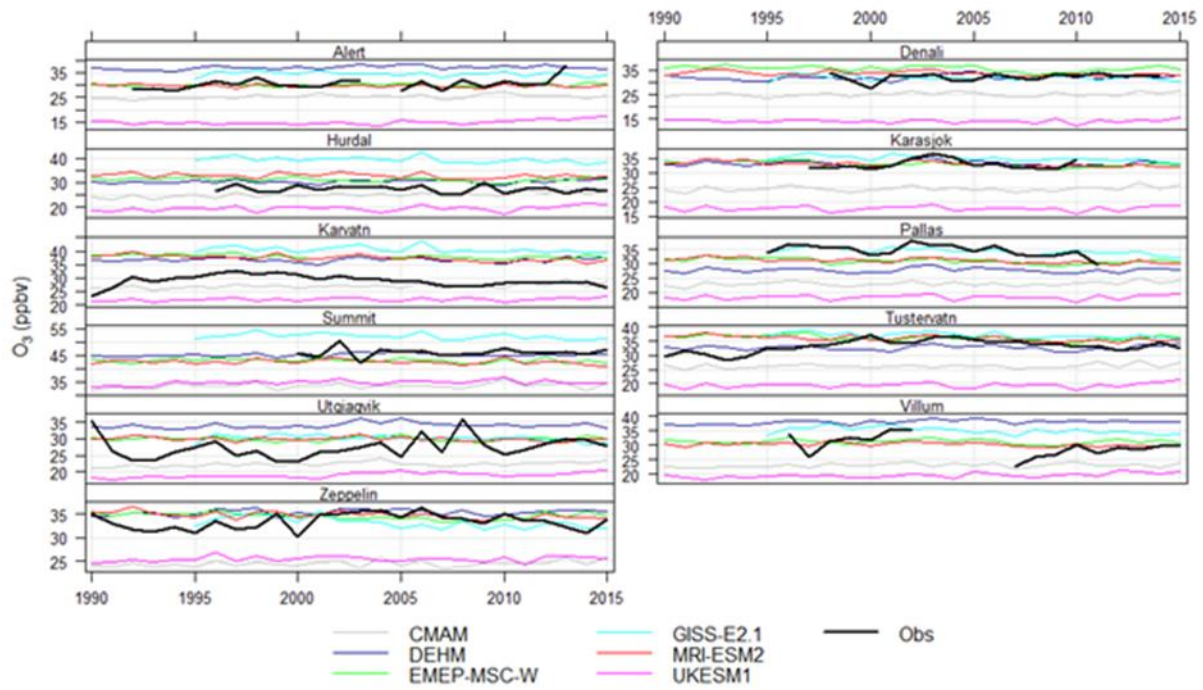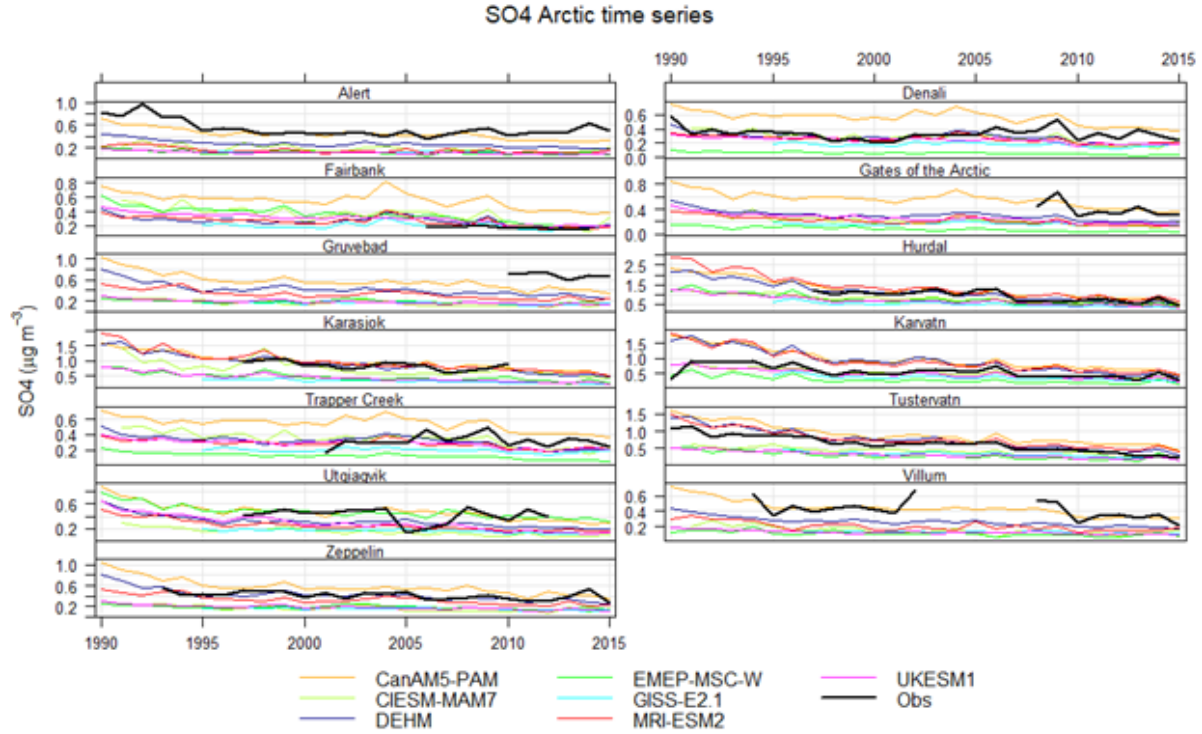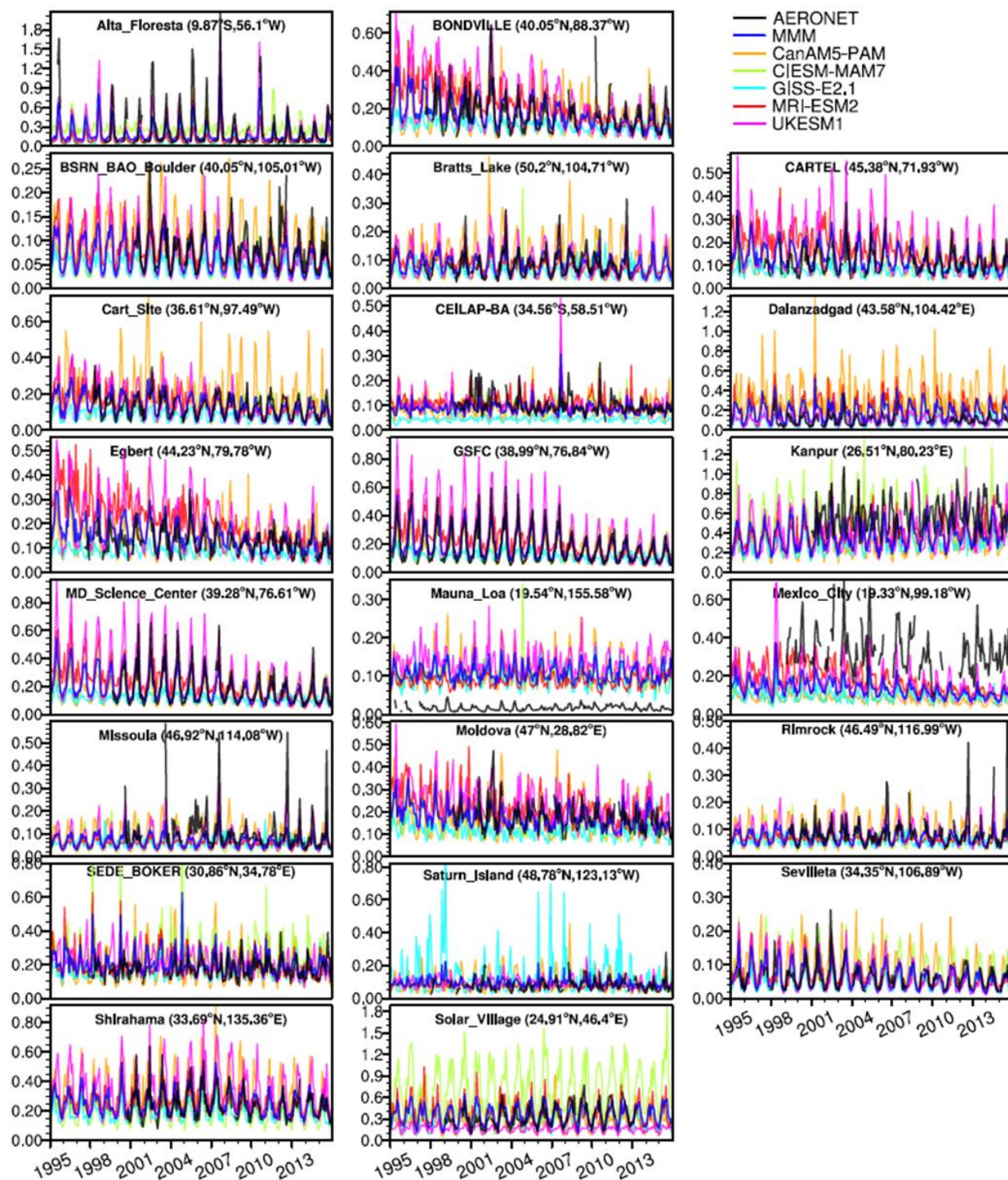
Supplementary Figure 4: Measured and modelled annual mean time series of surface-level O₃ at 11 Arctic ground stations.
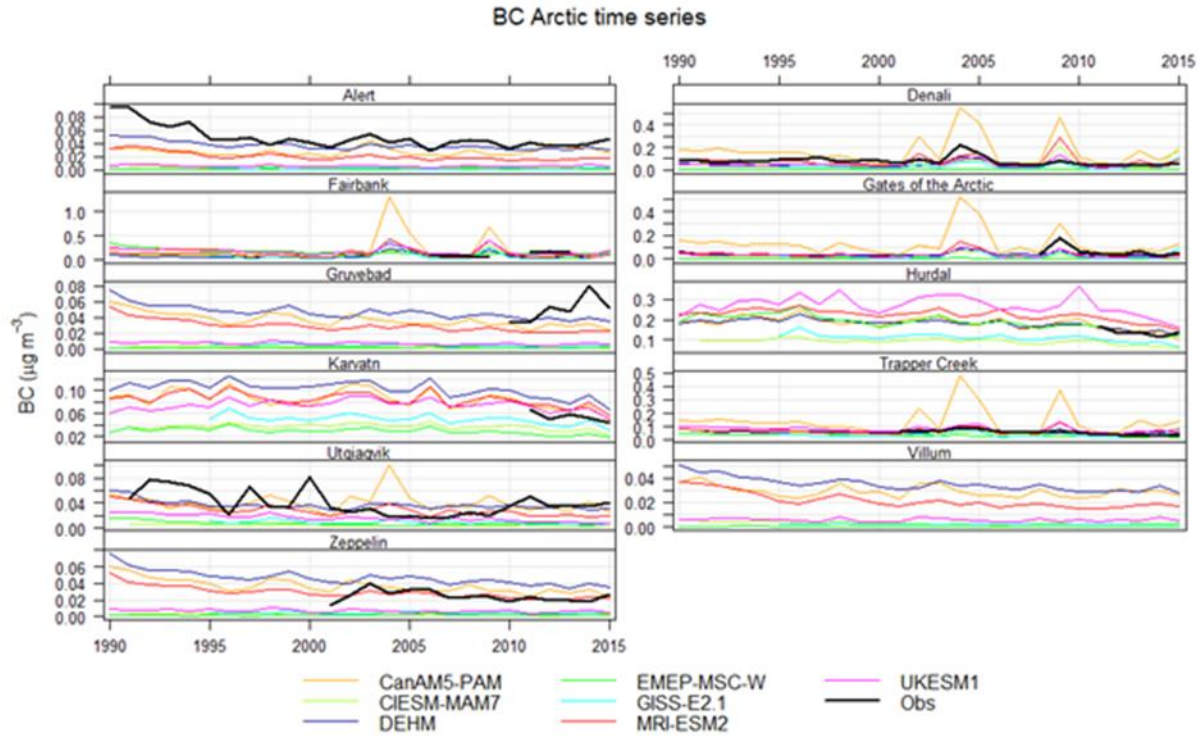
Supplementary Figure 5: Measured and modelled annual mean time series of surface-level BC at 11 Arctic ground stations.

Supplementary Figure 6: Measured and modelled annual mean time series of surface-level SO₄ at 13 Arctic locations.
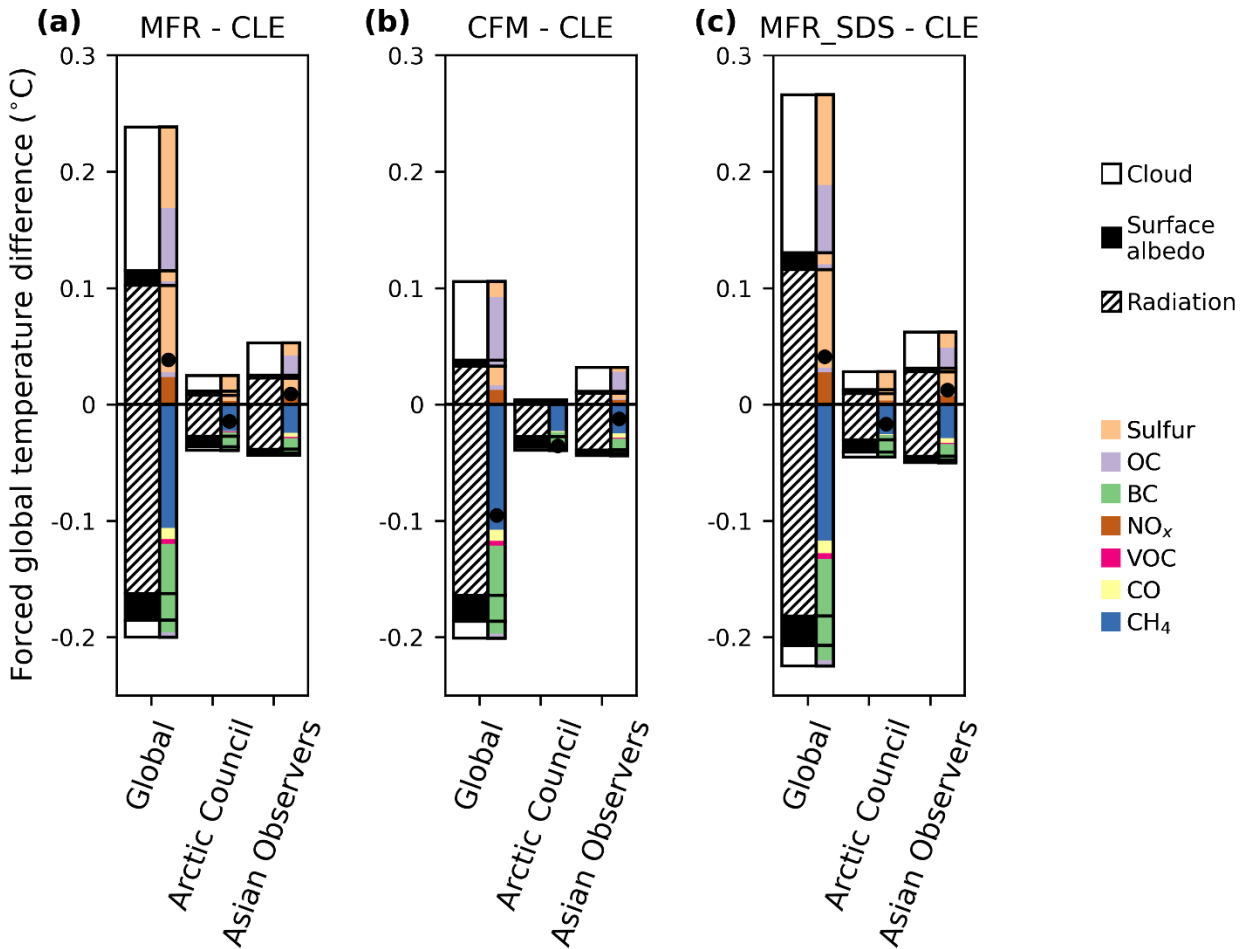
Supplementary Figure 7: Monthly mean AOD comparison between AERONET observations and model simulations. These AERONET stations were selected based on the availability of monthly mean data covering at least 60% of the total time-period, not necessarily consecutive months, between 1995 and 2015. MMM represents multi-model mean results. Note the different y-axis scale used for each station's results.
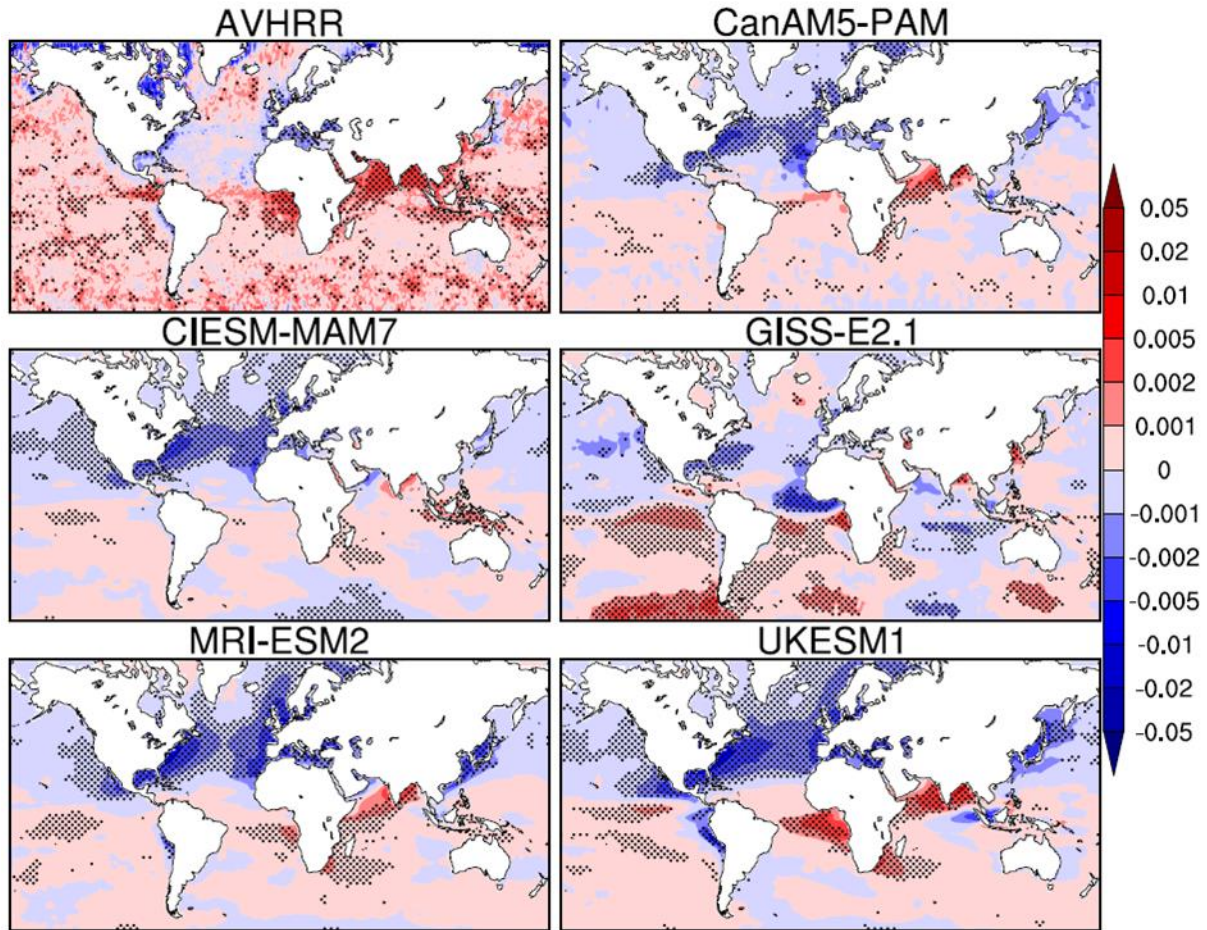
Supplementary Figure 8: The 1995 to 2015 AOD trend. Data from AVHRR (top left) and model simulations (rest of panels) are multiplied by a factor 10. Stippling represents regions where trends are statistically significant for 95% confidence level based on Student's t-test.
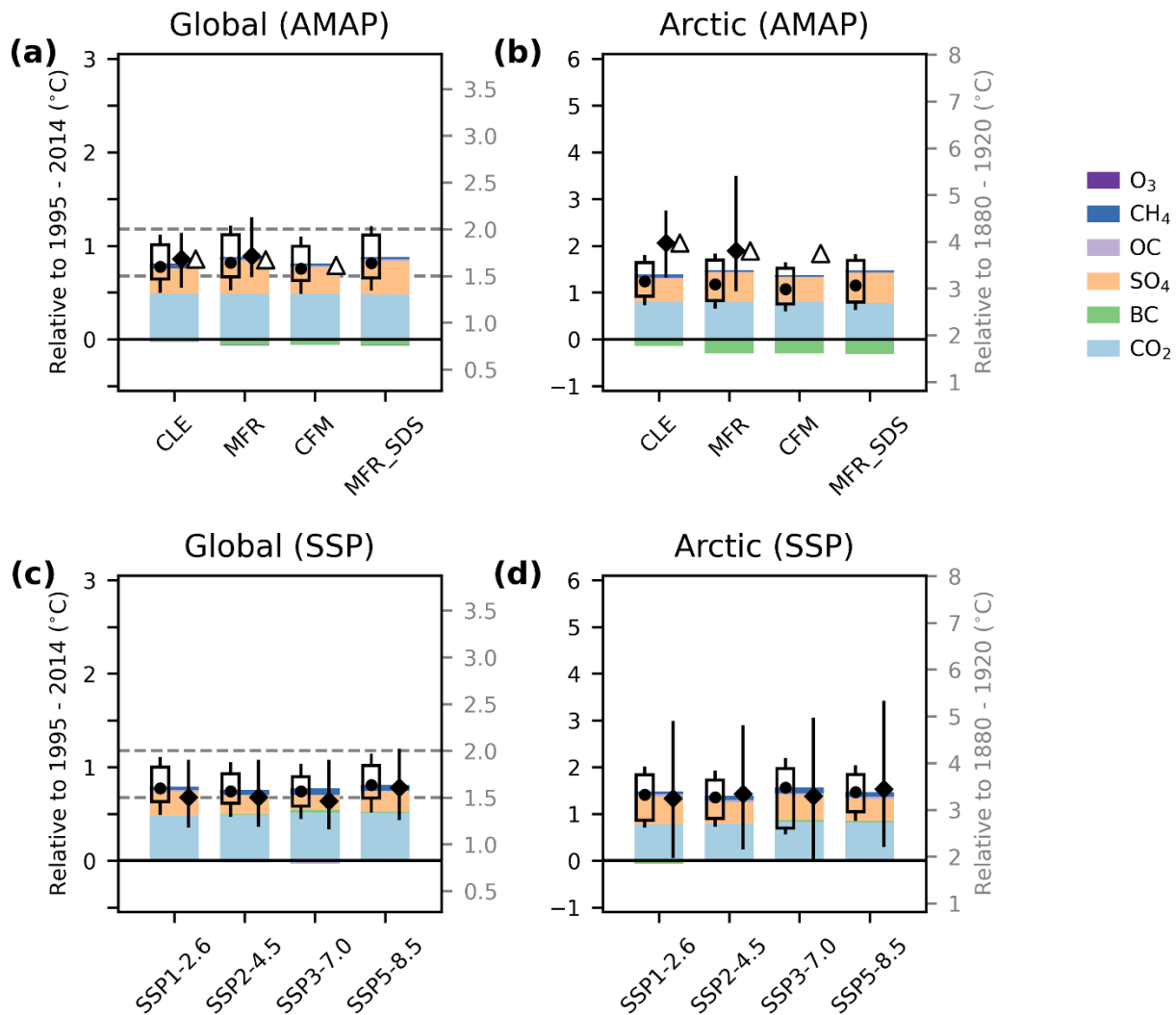
Supplementary Figure 9: Projected global and Arctic mean temperature changes in 2030. a and **b** are for Arctic Monitoring and Assessment Programme (AMAP) air pollution mitigation scenarios and **c** and **d** for Shared Socioeconomic Pathways (SSP) climate scenarios. Black diamonds refer to the multi-model median temperatures in Earth System Models (ESMs) for 2026 - 2035, relative to 1995 – 2014 (black font), where available. Warming relative to preindustrial conditions is also indicated (grey font, with 1.5 and 2 °C thresholds indicated by dashed lines in **a** and **c**). Color bars refer to the contributions from the individual changes in air pollutant and greenhouse gas emissions to forced temperature changes, based on emulator simulations (legend). Black bullets refer to the corresponding net changes. Results from the MRI-ESM2 are shown, the only ESM that ran the CFM scenario (triangles in **a** and **b**). Contributions from tropospheric ozone (O_3, less than 0.03 °C) and OC (less than 0.09 °C) are barely discernible but are included for the sake of completeness (AMAP scenarios only). 5-95% confidence intervals ($\pm$ 1.64 σ), resulting from uncertainties in all simulated processes, are indicated by black vertical lines (error bars). Confidence ranges due to radiative forcing uncertainty in the emulator are indicated by transparent rectangles. The emulator does not account for unforced natural variability and so confidence ranges are typically smaller than ESM confidence ranges.

Supplementary Figure 10: Impacts of changes in air pollution on future global climate.
Differences in forced global temperatures in 2050 are shown between Maximum Feasible
Reduction (**a**), Climate Forcing Mitigation (**b**), Sustainable Development activity (**c**) scenarios
and the Current LEgislation scenario. The differences are broken down into contributions of
global and regional emissions of chemically reactive species and the radiative forcing processes
that are associated with these emissions. 3 different radiative forcing processes are considered,
indicated by wide bars (hatched, black, and white; for interactions of air pollutants and CH_4 with
radiation, surface albedo, and clouds, respectively, see legend). Narrow colored bars refer to
emissions of 7 reactive species (see legend) from global sources and two regions (Arctic
Council, and Asian Arctic Council Observer countries: Japan, People's Republic of China,
Republic of India, Republic of Korea, Republic of Singapore). Black bullets refer to the net
temperature changes associated with global and regional emissions.
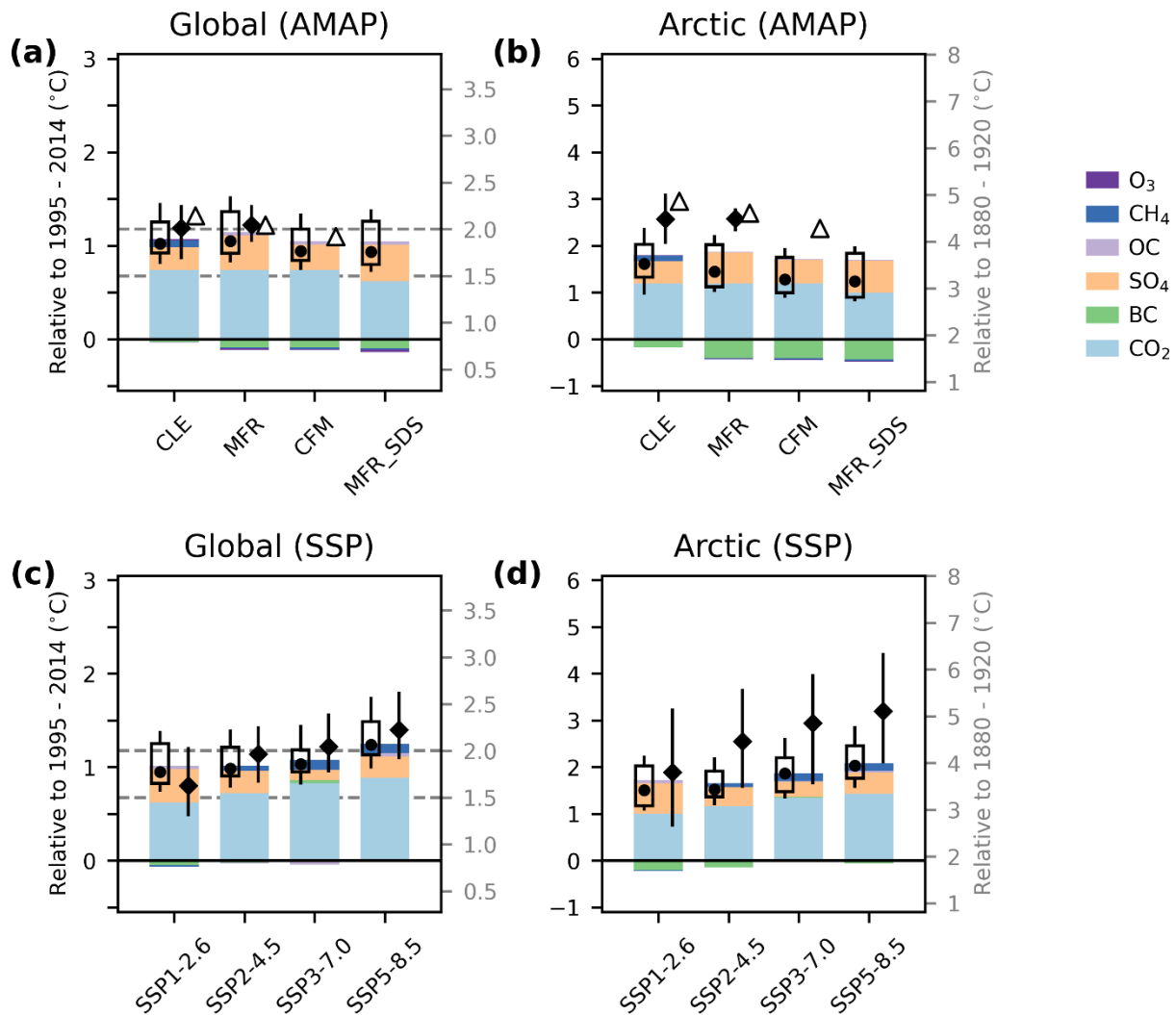
Supplementary Figure 11: Projected global and Arctic mean temperature changes in 2050.
a and **b** are for AMAP air pollution mitigation scenarios and **c** and **d** for Shared Socioeconomic Pathways (SSP) climate scenarios, using an ECS of 3 °C in the emulator and removing data from all ESMs that have an equilibrium climate sensitivity outside of the range from 2.5 to 4 °C. In comparison, the ECS for emulator simulation results in Fig. 2 is 3.7 °C and the ECS in the ESMs ranges from 1.8 to 5.6 °C. Black diamonds refer to the multi-model median temperatures in Earth System Models (ESMs) for 2046 - 2055, relative to 1995 – 2014 (black font), where available. Warming relative to preindustrial conditions is also indicated (grey font, with 1.5 and 2 °C thresholds indicated by dashed lines in **a** and **c**). Color bars refer to the contributions from the individual changes in air pollutant and greenhouse gas emissions to forced temperature changes, based on emulator simulations (legend). Black bullets refer to the corresponding net changes. Results from the MRI-ESM2 are shown, the only ESM that ran the CFM scenario (triangles in **a** and **b**). Contributions from tropospheric ozone (O_3, less than 0.03 °C) and OC (less than 0.09 °C) are barely discernible but are included for the sake of completeness (AMAP scenarios only). 5-95% confidence intervals ($\pm$ 1.64 σ), resulting from uncertainties in all simulated processes, are indicated by black vertical lines (error bars). Confidence ranges due to radiative forcing uncertainty in the emulator are indicated by transparent rectangles. The

emulator does not account for unforced natural variability and so confidence ranges are typically smaller than ESM confidence ranges.

Supplementary Acknowledgments

Supplementary References

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

2. Earth System Grid Federation, input4MIPs: Boundary Condition and Forcing Datasets for CMIP6, *Department of Energy Office of Science*, https://esgf-node.llnl.gov/projects/input4mips/ (2020).

3. Z. Klimont, K. Kupiainen, C. Heyes, P. Purohit, J. Cofala, P. Rafaj, J. Borken-Kleefeld, W. Schöpp, Global anthropogenic emissions of particulate matter including black carbon. *Atmos. Chem. Phys.* **17**, 8681–8723, https://doi.org/10.5194/acp-17-8681-2017 (2017).

4. L. Höglund-Isaksson, A. Gómez-Sanabria, Z. Klimont, P. Rafaj, W. Schöpp, Technical potentials and costs for reducing global anthropogenic methane emissions in the 2050 timeframe – results from the GAINS model. *Environmental Research Communications* **2**, 025004, https://doi.org/10.1088/2515-7620/ab7457 (2020).

5. International Institute for Applied Systems Analysis (IIASA), *ECLIPSE V6b*, global emission fields, https://previous.iiasa.ac.at/web/home/research/researchPrograms/air/ECLIPSEv6b.html (2020).

6. IEA, *World Energy Outlook 2018* (International Energy Agency, Paris, France, 2018).

7. S. M. Miller, A. M. Michalak, R. G. Detmers, O. P. Hasekamp, L. M. P. Bruhwiler, S. Schwietzke, China's coal mine methane regulations have not curbed growing emissions. *Nature Communications* **10**, 303 (2019).

8. J. Sheng, S. Song, Y. Zhang, R. G. Prinn, G. Janssens-Maenhout, Bottom-up estimates of coal mine methane emissions in China: a gridded inventory, emission factors, and trends *Environ. Sci. Technol.* **6**, 473–8 (2019).

9. M. Saunois, P. Brousquet, B. Poulter, A. Peregon, P. Ciais, et al, The global methane budget 2000-2012, *Earth Syst. Sci. Data* **8**, 697–751 (2016).

10. B. Zheng, D. Tong, M. Li, F. Liu, C. Hong, G. Geng, H. Li, X. Li, L. Peng, J. Qi, L. Yan, Y. Zhang, H. Zhao, Y. Zheng, K. He, Q. Zhang, Trends in China's anthropogenic

emissions since 2010 as the consequence of clean air actions. *Atmos. Chem. Phys.* **18,** 14095-14111 (2018).

11. U. Im, K. Tsigaridis, G. Faluvegi, P. L. Langen, J. P. French, R. Mahmood, R., M. A. Thomas, K. von Salzen, D. C. Thomas, C. H. Whaley, Z. Klimont, H. Skov, J. Brandt, Present and future aerosol impacts on Arctic climate change in the GISS-E2.1 Earth system model. *Atmos. Chem. Phys.* **21**, 10413–10438, https://doi.org/10.5194/acp-21-10413-2021 (2021).

12. E. E. McDuffie, S. J. Smith, P. O'Rourke, K. Tibrewal, C. Venkataraman, E. A. Marais, B. Zheng, M. Crippa, M. Brauer, R. V. Martin, A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970-2017): an application of the Community Emissions Data System (CEDS). *Earth Syst. Sci. Data* **12**, 3413-3442 (2020).

13. C. Wiedinmyer, R. J. Yokelson, B. Gullett, Global emissions of trace gases, particulate matter, and hazardous air pollutants from open burning of domestic waste, *Environ. Sci. Technol*. **48**, 16, 9523–9530 (2014).

14. Y. Kanaya, K. Yamaji, T. Miyakawa, F. Taketani, C. Zhu, Y. Choi, Y. Komazaki, K. Ikeda, Y. Kondo, Z. Klimont, Rapid reduction in black carbon emissions from China: evidence from 2009–2019 observations on Fukue Island, Japan. *Atmos. Chem. Phys.* **20**, 11, 6339-6356, (2020).

15. Y. Kanaya et al., Dominance of the residential sector in Chinese black carbon emissions as identified from downwind atmospheric observations during the COVID-19 pandemic. *Scientific Reports* **11**, 23378 (2021).

16. C. H. Whaley, R. Mahmood, K. von Salzen, B. Winter, S. Eckhardt, S. Arnold, S. Beagley, S. Becagli, R.-Y. Chien, J. Christensen, S. M. Damani, N. Evangeliou, G. Faluvegi, M. Flanner, J. Fu, M. Gauss, F. Giardi, W. Gong, J. L. Hjorth, L. Huang, U. Im, Y. Kanaya, Z. Klimont, T. Kühn, J. Langner, K. Law, L. Marelle, A. Massling, T. Onishi, N. Oshima, Y. Peng, D. Plummer, O. Popovicheva, L. Pozzoli, J.-C. Raut, M. Sand, L. N. Saunders, J. Schmale, S. Sharma, H. Skov, F. Taketani, M. Thomas, R. Traversi, K. Tsigaridis, S. Tsyro, S. Turnock, V. Vitale, M. Wang, K. A. Walker, D. Watson-Parris, T. Weiss-Gibbons, Model evaluation of short-lived climate forcers for the Arctic Monitoring and Assessment Programme: a multi-species, multi-model study. *Atmos. Chem. Phys.* **22**, 5775-5828 (2021).

17. R. Mahmood, K. von Salzen, M. Flanner, M. Sand, J. Langner, H. Wang, L. Huang, Seasonality of global and Arctic black carbon processes in the Arctic Monitoring and Assessment Programme models. *J. Geophys. Res.-Atmospheres* **121**, 12, 7100-7116, https://doi.org/10.1002/2016JD024849 (2016).

18. B. N. Holben, T.F. Eck, I. Slutsker, D. Tanre, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, A. Smirnov, AERONET - A federated instrument network and data archive for aerosol characterization. *Rem. Sens. Environ.* **66**, 1-16 (1998).

19. T. X.-P. Zhao, P. K. Chan, A. K. Heidinger, A global survey of the effect of cloud contamination on the aerosol optical thickness and its long-term trend derived from

operational AVHRR satellite observations. *J. Geophys. Res. Atmos.* **118**, 2849-2857, doi:10.1002/jgrd.50278 (2013).

20. X. Zhao, and NOAA CDR Program, NOAA Climate Data Record (CDR) of AVHRR Daily and Monthly Aerosol Optical Thickness (AOT) over Global Oceans, Version 3.0. NOAA National Centers for Environmental Information. [AOT_AVHRR_v03r00_monthly_avg] doi:10.7289/V5BZ642P [date accessed: Jan 30, 2018, last accessed on Feb 01, 2020] (2017).

21. N. Lenssen, G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy, and D. Zyss, Improvements in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos.* **124**, 6307-6326, doi:10.1029/2018JD029522 (2019).

22. GISTEMP Team, GISS Surface Temperature Analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. Dataset accessed 2020-02-14 at https://data.giss.nasa.gov/gistemp/ (2020).

23. J. Kelly, P. A. Makar, D. A. Plummer, Projections of mid-century summer air-quality for North America: effects of changes in climate and precursor emissions. *Atmos. Chem. Phys.* **12**, 5367–5390, https://doi.org/10.5194/acp-12-5367-2012 (2012).

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

25. A. van Donkelaar, R. V. Martin, M. Brauer, N. C. Hsu, R. A. Kahn, R. C. Levy, A. Lyapustin, A. M. Sayer, and D. M. Winker, Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science & Technology* **50**, 3762-3772 (2016).

26. S. T. Turnock, O. Wild, A. Sellar, A., F. M. O'Connor, 300 years of tropospheric ozone changes using CMIP6 scenarios with a parameterised approach. *Atmospheric Environment* **213**, 686-698, https://doi.org/10.1016/j.atmosenv.2019.07.001 (2019).

27. R. Van Dingenen, F. Dentener, M. Crippa, J. Leitao, E. Marmer, S. Rao, E. Solazzo, L. Valentini, TM5-FASST: a global atmospheric source–receptor model for rapid impact analysis of emission changes on air quality and short-lived climate pollutants. *Atmos. Chem. Phys.* **18**, 16173–16211, https://doi.org/10.5194/acp-18-16173-2018 (2018).

28. IHME (2019) data downloaded from https://cloud.ihme.washington.edu/index.php/s/PG0b3Xek3WVLO41 (last accessed October 2021)

29. B. Jones. B. C. O'Neill, Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environ. Res. Lett.* **11**, 084003, https://doi.org/10.1088/1748-9326/11/8/084003 (2016).

30. P. Achakulwisut, M. Brauer, P. Hystad, S.C. Anenberg, Global, national, and urban burdens of pediatric asthma incidence attributable to ambient NO2 pollution: estimates from global datasets. *The Lancet Planetary Health* **3**(4), e166-e178 (2019).

31. S. C. Anenberg, A. Mohegh, D.L. Goldberg, G.H. Kerr, M. Brauer, K. Burkart, P. Hystad, A. Larkin, S. Wozniak, L. Lamsal, Long-term trends in urban NO2 concentrations and associated pediatric asthma cases: estimates from global datasets. *The Lancet Planetary Health* **6**(1), e49-e58 (2022).

32. M. Ru, M. Brauer, J.-F. Lamarque, D. Shindell, Exploration of the global burden of dementia attributable to PM2.5: What do we know based on current evidence? *GeoHealth* **5**(5), doi: 10.1029/2020GH000356 (2021).

33. Y. Bo, J. R. Brook, C. Lin, L. Chang, C. Guo, Y. Zeng, Z. Yu, T. Tam, A.K.H. Lau, X.Q. Lao, Reduced ambient PM2.5 was associated with a decreased risk of chronic kidney disease: A longitudinal cohort study. *Environ. Sci. Technol.* **55**(10): 6876-6883, doi:10.1021/acs.est.1c00552 (2021).

34. GBD 2019 Risk Factor Collaborators, Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* **396**(10258), 1223-1249, doi:10.1016/S0140-6736(20)30752-2 (2020).