UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Role of Natural Variability vs Anthropogenic Forcing in Recent Climate Variations

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Earth and Planetary Sciences

by

Xueying Zhao

March 2023

Dissertation Committee:

Dr. Robert J. Allen, Chairperson
Dr. Wei Liu
Dr. David D. Oglesby

The Dissertation of Xueying Zhao is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor and my committee for their support. I am grateful to all the people that I met here.

To my parents for all the support.

ABSTRACT OF THE DISSERTATION

The Role of Natural Variability vs Anthropogenic Forcing in Recent Climate Variations

by

Xueying Zhao

Doctor of Philosophy, Graduate Program in Earth and Planetary Sciences
University of California, Riverside, March 2023
Dr. Robert J. Allen, Chairperson

Many of the observed changes are unprecedented, particularly since 1950s. Changes in the large-scale atmospheric circulations, e.g., Walker Circulation (WC) and tropical belt width, can be attributed to natural variability and anthropogenic forcing. This raises the question about the contribution of natural variability and anthropogenic forcing to observed changes. I used climate modelings to investigate the dominant driver of recent observed changes, focusing on the WC and tropical belt expansion.

The tropical belt has observed to widen over the past few decades. Natural variability, greenhouse gas, aerosol and stratospheric ozone depletion have been suggested to be important contributor to the tropical expansion. The net influence of natural variability and anthropogenic forcers on tropical belt width variations are complicated and contains large uncertainty, closely associated with aerosols. I used idealized PDRMIP simulations to investigate the response of tropical belt width to different forcers. Results show that absorbing black carbon (BC) aerosol drives tropical expansion and scattering sulfate aerosol drives contraction. Tropical belt expansion (contraction) is associated with an increase (decrease)

in extratropical static stability induced by absorbing (scattering) aerosol. Via linear attribution, result shows that BC might the largest driver of historical Northern Hemisphere tropical belt widening.

Climate models simulate a weakening of WC under the warming, however, I found WC intensified based on observations and multiple reanalyses since 1979. Atmosphere-only simulations driven by the real-world evolution of sea surface temperatures (SSTs) simulate the observed intensification of WC, whereas coupled ocean atmosphere simulations do not. Idealized simulations driven by the unforced component of SSTs yield significant WC strengthening, whereas negligible WC changes are simulated driven by the forced SST component. These findings suggest that the recent strengthening of the WC since 1979 is attributed to natural SST variability, particularly associated with a La Niña-like SST pattern.

Anthropogenic forcers−aerosols, are also generally considered to be the major cause of air pollution. Current air quality guidelines target man-made air pollutants, while the contributions of natural aerosols are of less importance. However, natural aerosols have strong geographic gradients and this suggests that spatially invariant air quality guidelines may handicap regions close to natural sources. Climate models are used to to construct a view of pre-industrial "pristine" air quality. Compared to the World Health Organization's globally uniform air quality thresholds, pristine sources of aerosols lead to poor air quality induced by natural (i.e., pristine) pollution alone, particularly near dusty regions, estimated nearly 1 billion people impacted. World Health Organization lowers fine particular annual mean threshold tin 2021. Based on the latest air quality guideline, estimated impacted

population rises to ∼3.6 billion. These findings imply that air quality guidelines established based on anthropogenic metrics unfairly bias countries close to natural sources.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Natural variations and anthropogenic forcers impact climate by altering the energy budget on Earth. These drivers perturb the energy system by influencing incoming radiative forcing. Positive radiative forcing leads to a warming and negative radiative forcing leads to a cooling. The observed average global surface temperature over 1880 to 2012 shows a positive trend (Stocker, 2014). This is associated with positive net anthropogenic radiative forcing. The total anthropogenic forcing 1750-2011 is estimated as is estimated as 2.3 (1.1−3.3) W/m^2 and it has increased more rapidly since 1970 than the previous decades, which related to greenhouse gas emissions (Stocker, 2014). Aerosol is another main contributor to the anthropogenic forcing. The radiative forcing from aerosols is estimated as -0.9 (-1.9−-0.1) W/m^2 over 1750-2011 (Stocker, 2014). Changes in large-scale climate variables (e.g., global mean temperature, the Walker Circulation, tropical belt displacement) have been reliably attributed to anthropogenic and natural forcings (Vecchi et al., 2006; Tao, Hu, & Liu, 2016; Grise, Davis, Staten, & Adam, 2018). Variations of large-scale circulations

have large impact on hydrological pattern over adjacent regions. It is important to better understand the contribution of natural variability and anthropogenic forcers to the climate systems.

The width of the tropical belt is linked to the Hadley cell circulation, with strong moist air ascent in tropical deep convection zones and dry air descent in the subtropics. Tropical belt expansion over the past few decades has been observed in multiple datasets (Hu & Fu, 2007; Lu, Vecchi, & Reichler, 2007; Seidel, Fu, Randel, & Reichler, 2008; S. M. Davis & Rosenlof, 2012; Birner, Davis, & Seidel, 2014; Lucas, Timbal, & Nguyen, 2014). This expansion is indicated by several metrics, including widening of the Hadley cell (Hu & Fu, 2007; Lu et al., 2007; Nguyen, Evans, Lucas, Smith, & Timbal, 2013) and poleward shifts in hydrological patterns and subtropical dry zones (Sousa et al., 2011; Scheff & Frierson, 2012; Cai, Cowan, & Thatcher, 2012; Staten, Lu, Grise, Davis, & Birner, 2018; Horinouchi, Matsumura, Ose, & Takayabu, 2019). Tropical expansions can be driven by internal (Garfinkel, Waugh, & Polvani, 2015; Grise et al., 2018) and natural (Lu, Chen, & Frierson, 2008; Grassi, Redaelli, Canziani, & Visconti, 2012; Nguyen et al., 2013; Tandon, Gerber, Sobel, & Polvani, 2013; Allen, Norris, & Kovilakam, 2014; Allen & Kovilakam, 2017a; Mantsis, Sherwood, Allen, & Shi, 2017; Amaya, Siler, Xie, & Miller, 2018a) variability, as well as anthropogenic drivers− greenhouse gases (Hu, Tao, & Liu, 2013; Tao et al., 2016; Grise & Polvani, 2016), stratospheric ozone depletion in the Southern Hamisphere, and absorbing aerosols in the Northern Hemisphere. The overall impact of anthropogenic drivers on tropical belt width variations are complex and have large uncertainties. This is particularly true for aerosols, where the 90% confidence range of aerosol effective radiative forcing (ERF)

is -0.4 to -2.0 W m^{-2} (Bellouin et al., 2019). Similarly, BC forcing from anthropogenic fossil and biofuel emissions possesses a relatively large uncertainty range of 0.05 to 0.80 W m^{-2} (Ramanathan & Carmichael, 2008; Koch et al., 2009; Bond, Doherty, Hahey, & et al., 2013; Myhre et al., 2013). The imprint of anthropogenic drivers on recent tropical expansion may even be indiscernible from natural drivers (Grise et al., 2019). The total influence of natural variability and anthropogenic forcers on tropical belt width variations are complicated and aerosol forcing introduce large uncertainties.Chapter 2 shows the work to better understand the effect of anthropogenic drivers on tropical belt width variations, with a focus on the role of aerosol.

The Walker Circulation (WC)−a large-scale zonal-vertical atmospheric circulation in the equatorial Pacific, has been suggested to weaken in the twentieth century due to anthropogenic forcing by climate models (Vecchi et al., 2006). However, a strengthening of the Walker Circulation is simulated by atmosphere-only climate models (AMIP) driven by observed SSTs over the late second half of the 20^{th} century (S. Ma & Zhou, 2016; Sohn, Yeh, Schmetz, & Song, 2013). In more recent decades, a stronger WC is also supported by cooling of the central and eastern tropical Pacific SSTs (a La Niña-like pattern) (Kosaka & Xie, 2013). This is likely related to that the observed radiative forcing has been increasing at a lower rate over the period from 1998 to 2011, compared to 1951 to 2011 (Stocker, 2014). Climate models simulate a surface warming trend larger than the observations over 1998 to 2011 (Stocker, 2014). The difference is likely due to natural internal climate variability, which diminishes the relevance of short trends for long-term climate change. It may also related to the forcing used by the models−an overestimate of the response to

increasing greenhouse gas and anthropogenic aerosol forcing. This discrepancy, in turn, leave the projected WC weakening in response to anthropogenic warming questionable and the dominant driver of the Walker Circulation is unclear. Chapter 3 discuss the work to investigate the contribution of multi-decadal SST variations on recent changes in the Walker Circulation.

The anthropogenic forcer−aerosols, can also originate from natural sources. Aerosols in the troposphere can either from anthropogenic or natural origins; stratospheric aerosols mostly from volcanic eruptions. Natural aerosol, like dust, is the most dominant aerosol species in the atmosphere, constituting 70% of the global aerosol mass burden and 25% of the aerosol optical depth (aerosol optical depth; Kinne et al., 2006). Atmospheric aerosol also contributes to air pollution, which have negative effects on human health. Air quality guideline is based on the measurements of ground-level concentrations of various air pollutants, including particulate matter (PM), ozone, nitrogen oxides and sulfur dioxide, such as the guideline from World Health Organization (WHO) (World Health Organization, 2006). Such air quality guideline, mainly target anthropogenic emissions, is used by many countries as reference to establish their own national standards and goals for air quality. However, the contribution of natural emitted pollutants are discounted and this bring bias, particularly over economically developing regions where natural aerosol levels can be high (e.g., Africa; Petkova, Jack, Volavka-Close, & Kinney, 2013). Contribution of natural (pristine) aerosols on present-day air quality has not been assessed. Chapter 5 discuss the work to quantify the impact of pristine aerosol background on current air quality metrics.

Chapter 2

Tropical Belt Width Proportionately More Sensitive to Aerosols Than Greenhouse Gases

2.1 Introduction

The width of the tropical belt is linked to the Hadley cell circulation, with strong moist air ascent in tropical deep convection zones and dry air descent in the subtropics. Tropical belt expansion over the past few decades has been observed in multiple datasets (Hu & Fu, 2007; Lu et al., 2007; Seidel et al., 2008; S. M. Davis & Rosenlof, 2012; Lucas et al., 2014). This expansion is indicated by several metrics, including widening of the Hadley cell (Hu & Fu, 2007; Nguyen et al., 2013) and poleward shifts in hydrological patterns and subtropical dry zones (Cai et al., 2012; Staten et al., 2018; Horinouchi et al., 2019).

In addition to internal (Garfinkel et al., 2015; Grise et al., 2018) and natural (Lu et al., 2008; Grassi et al., 2012; Nguyen et al., 2013; Tandon et al., 2013; Allen et al., 2014; Allen & Kovilakam, 2017a; Mantsis et al., 2017; Amaya et al., 2018a) variability, anthropogenic drivers are another important contributor. Tropical widening occurs in model simulations driven by GHGs (Hu et al., 2013; Tao et al., 2016; Grise & Polvani, 2016), which is associated with a poleward shift in the subtropical baroclinic instability zone (Lu et al., 2007). Stratospheric ozone depletion is important in the Southern Hemisphere (SH), particularly during austral summer (Polvani, Waugh, Correa, & Son, 2011; Min & Son, 2013; Waugh, Garfinkel, & Polvani, 2015). Absorbing aerosols, such as black carbon (BC), have been suggested as a cause of Northern Hemisphere (NH) tropical expansion (Allen, Sherwood, Norris, & Zender, 2012b; Allen et al., 2014; Kovilakam & Mahajan, 2015; Shen & Ming, 2018; Johnson, Haywood, & Hawcroft, 2019). BC warms the troposphere, particularly in the NH mid-latitudes where most emissions occur, which stabilizes the atmospheric column and results in a poleward shift of tropospheric jet streams (Allen, Sherwood, Norris, & Zender, 2012a; Allen et al., 2012b; Shen & Ming, 2018). Future decreases in scattering aerosol also induces mid-latitude tropospheric warming (primarily via aerosol-cloud interactions), leading to NH tropical widening through the 21st century (Allen & Sherwood, 2011; Allen & Ajoku, 2016).

The overall impact of anthropogenic drivers on tropical belt width variations are complex and have large uncertainties. This is particularly true for aerosols, where the 90% confidence range of aerosol effective radiative forcing (ERF) is -0.4 to -2.0 W m^{-2} (Bellouin et al., 2019). Similarly, BC forcing from anthropogenic fossil and biofuel emissions possesses

a relatively large uncertainty range of 0.05 to 0.80 W m^{-2} (Ramanathan & Carmichael, 2008; Koch et al., 2009; Bond et al., 2013; Myhre et al., 2013). The imprint of anthropogenic drivers on recent tropical expansion may even be indiscernible from natural drivers (Grise et al., 2019).

The goal of this work is to better understand the effect of anthropogenic drivers on tropical belt width variations, focused on the role of aerosols. To isolate the role of individual drivers, we utilize idealized simulations with very large single forcing perturbations in comprehensive coupled ocean-atmosphere models from Precipitation Driver and Response Model Intercomparison Project (PDRMIP). Although the forcings are idealized and their magnitude is not realistic, they allow a direct assessment of how individual drivers impact tropical belt width (e.g., N. A. Davis, Seidel, Birner, Davis, & Tilmes, 2016a). This is the first study that rigorously quantifies the aerosol impact−including individual aerosol species like BC and sulfate−on tropical belt width using multiple models. Moreover, this is the first study that quantifies the impact of regional (e.g., Europe, Asia) aerosol emissions on tropical belt width. This chapter is organized as follows: Methods and data are described in Section 2.2; Section 2.3 discusses the results and Section 2.4 summarizes the conclusions.

2.2 Method & Data

2.2.1 Model Simulations

Under the framework of PDRMIP (Myhre et al., 2017), we use the coupled global model simulations including the baseline simulation and a set of perturbation experiments. The baseline simulation is forced with all anthropogenic and natural climate forcing agents

at present-day (year 2000) levels (Samset et al., 2016). Global perturbation experiments (relative to present-day)−which represent very large perturbations, particularly for the aerosols−include a doubling of carbon dioxide (co2×2), 10 times black carbon concentration or emissions (bc×10), and 5 times sulfate concentrations or emissions (sul×5). Regional aerosol perturbation experiments include 10 times Asian black carbon concentration or emissions (bc×10 Asia), and 10 times sulfate concentrations or emissions in Europe (sul× 10 EU) and Asia (sul× 10 Asia) (L. Liu et al., 2018). The European (Asian) region is defined from $35°N - 70°$N and $10°W - 40°$E ($10°N - 50°$N and $60°E - 140°E$). An exception is HadGEM2-ES, which uses emissions from the year 1860 for its base run, and perturbation simulations are relative to the year 2000 (Stjern et al., 2017). HadGEM2-ES simulations will therefore feature larger increases in aerosols.

Nine models are used for the global perturbation experiments including CanESM2, HadGEM2-ES, HadGEM3, GISS-E2-R, MIROC-SPRINTARS, NCAR CESM1-CAM4, NCAR CESM1-CAM5, IPSL-CM5A and NorESM1 (Samset et al., 2016; Myhre et al., 2017). All models use a fully dynamical ocean model, except NCAR CESM1-CAM4 which uses a slab ocean model. Most models also performed regional perturbation experiments, except CanESM2 and HadGEM2. Although some models are concentration driven and some are emission driven, we don't find clear response differences between emission- and concentration-driven models. All simulations are interpolated to a 2.5° by 2.5° grid resolution using bilinear interpolation. We analyze the last 50-years of each simulation, when near-equilibrium is reached. For most experiments, models do not show significant trends in net top-of-atmosphere (TOA) radiative fluxes over the last 50 years. However, in the

case of co2×2, models will likely not have reached their equilibrium warming within the 100 years simulated here (Caldeira & Myhrvold, 2013; Samset et al., 2016).

We also use the corresponding PDRMIP fixed sea surface temperature (fSST) experiments to estimate the ERF, as well as the fast response (no surface temperature feedbacks) of the tropical belt width. ERF is estimated from the net TOA radiative fluxes (the sum of net longwave and shortwave fluxes) from the fSST simulations (perturbation minus baseline). This ERF definition is similar to (Forster et al., 2016), except SSTs here are based on the baseline climatology as opposed to the preindustrial. All fSST simulations were performed for at least fifteen years, so we analyze the last fifteen years of data for these calculations. The 90% confidence interval for a 15 year aerosol ERF is ~ 0.15 W m^{-2} (Forster et al., 2016). Annual global and hemispheric mean ERFs for individual models and experiments is included in Tables 2.1 and 2.2, and spatial ERF maps are including in Figure 2.1. As the coupled simulations yield the total climate response (slow surface temperature feedbacks and fast rapid adjustments), the fSST simulations yield the fast response (e.g., atmospheric heating).

Table 2.1: Annual mean effective radiative forcing (ERF) in the regional perturbation experiments for each individual model and the ensemble mean. Unit is W m^{-2}. n/a represents not available.

Model	bc×10 Asia			sul×10 Asia			sul×10 EU		
	Global	NH	SH	Global	NH	SH	Global	NH	SH
GISS-E2-R	0.18	0.49	-0.12	-0.74	-1.34	-0.14	-0.28	-0.53	-0.03
HadGEM3	0.19	0.45	-0.07	-0.69	-1.43	0.04	-0.25	-0.49	-0.004
IPSL-CM5A	0.004	0.13	-0.13	-0.57	-1.28	0.13	-0.31	-0.58	-0.05
MIROC-SPRINTARS	0.07	0.23	-0.09	-1.15	-2.08	-0.22	-0.41	-0.86	0.05
NCAR-CAM4	0.20	0.48	-0.08	-0.45	-1.13	0.22	n/a	n/a	n/a
NCAR-CAM5	0.40	0.72	0.09	-0.80	-1.50	-0.10	-0.47	-0.97	0.02
NorESM1	0.03	0.31	-0.25	-0.75	-1.39	-0.11	-0.40	-0.63	-0.16
Ensemble mean	0.15	0.40	-0.09	-0.74	-1.45	-0.02	-0.30	-0.68	-0.03

Table 2.2: Annual mean effective radiative forcing (ERF) in the global perturbation experiments for each individual model and the ensemble mean. Unit is W m^{-2}.

Model	bc×10			sul×5			co2×2		
	Global	NH	SH	Global	NH	SH	Global	NH	SH
CanESM2	1.54	2.30	0.79	-3.23	-4.66	-1.80	3.54	3.44	3.63
GISS-E2-R	1.24	1.89	0.60	-2.82	-3.63	-2.01	4.05	4.12	3.98
HadGEM2	1.43	2.20	0.66	-3.92	-6.29	-1.54	3.56	3.56	3.55
HadGEM3	0.68	1.18	0.19	-8.27	-8.19	-8.36	3.64	3.75	3.53
IPSL-CM5A	0.76	1.18	0.35	-2.75	-3.40	-2.10	3.32	3.33	3.31
MIROC-SPRINTARS	0.67	0.92	0.41	-2.73	-4.21	-1.25	3.69	3.63	3.75
NCAR-CAM4	0.76	1.33	0.19	-2.04	-2.86	-1.21	3.57	3.41	3.74
NCAR-CAM5	0.39	0.99	-0.20	-2.12	-3.11	-1.13	4.04	4.39	3.70
NorESM1	1.39	2.33	0.46	-3.70	-4.27	-3.14	3.50	3.38	3.63
Ensemble mean	0.98	1.59	0.38	-3.51	-4.51	-2.50	3.66	3.67	3.65

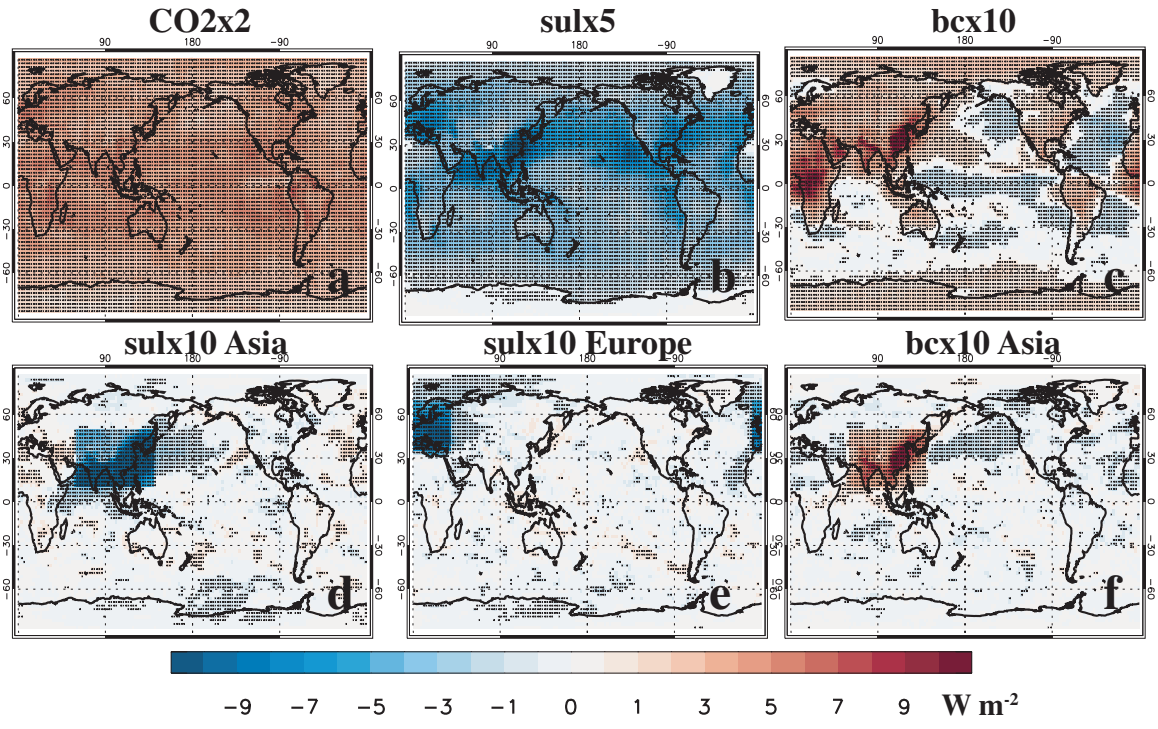

Figure 2.1: **Annual mean ensemble mean ERF**. (a) co2×2; (b) sul×5; (c) bc×10; (d) sul×10 Asia; (e) sul×10 Europe; and (f) bc×10 Asia. Changes significant at the 90% confidence level are denoted by black dots based on a t-test for the difference of means using the pooled variance. Units are W m^{-2}.

2.2.2 Tropical Edge Definition

The Tropical-width Diagnostics (TropD) package (Adam et al., 2018) is used to calculate the location of the tropical belt edge. The TropD package includes common tropical width metrics and standardizes the methodologies to diagnose the tropical belt edge by implementing simple mathematical methods. As recommended by N. Davis and Birner (2017) and Waugh et al. (2018), we focus on three metrics of tropical width: 1) the latitude where the zonal-mean meridional circulation (MMC) at 500 hPa becomes zero on the poleward side of the subtropical maxima, 2) the zero-crossing latitude where the zonal-mean near-surface easterly becomes mid-latitude westerly at the poleward side of the subtropical minimum in each hemisphere and equatorward of 60° latitude (UAS) and 3) the latitude where the zonal-mean precipitation minus evaporation becomes zero on the poleward side of the subtropical minimum (P−E). Changes in the latitude of the tropical belt edge are calculated as the difference between the perturbation simulation and the corresponding baseline simulation. Positive (negative) anomalies indicate tropical expansion (contraction). Note that GISS-E2-R didn't archive output to calculate the MMC metric. We have verified that similar results are obtained in the other two metrics when GISS-E2-R is removed from the the multi-model mean (not shown). We also note that recent analyses have generalized the MMC to the regional level by using the horizontally divergent wind (Schwendike et al., 2014; Staten, Grise, Davis, Karnauskas, & Davis, 2019). For conciseness, the quoted rates of expansion in the text are based on the average across the three metrics. In general, we get very similar results for all three metrics, but where exceptions exist they are noted. The figures show all metrics separately.

Statistical significance of tropical belt changes is examined by a standard two-tail t-test, accounting for the influence of serial correlation by using the effective sample size, $nyears(1-r_1)(1+r_1)^{-1}$, where $nyears$ is the number of years and r_1 is the lag-1 auto-correlation coefficient. The 90% confidence interval of ensemble mean response is also quantified as twice the standard error, $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the tropical width changes, and n is the number of models. We also quantify tropical width variations due to internal climate variability as the 90% confidence interval of tropical edge locations from the baseline simulation for each model. This is calculated by quantifying the boundary of the tropical edge in each hemisphere for each season and the annual mean for all years ($nyears$). The 90% confidence interval is based on a t-test, and estimated according to: $\overline{\Phi} \pm 1.68 \times \frac{\sigma'}{\sqrt{nyears}}$, where σ' is the standard deviation of tropical edge locations, $nyears$ is the number of tropical edge locations (i.e., number of years), $\overline{\Phi}$ is the mean tropical belt edge location, and 1.68 is the t-value with $nyears$-1 (49 for the coupled runs) degrees of freedom and a 0.05 probability.

Figure 2.2 shows individual model confidence intervals; the multi-model mean confidence interval ranges from 0.18° for MMC to 0.22° for P−E in the NH and from 0.08° for MMC to 0.11° for UAS in the SH. The smaller values in the SH may be due to the more zonally symmetric circulation and lack of land-sea contrasts. It may also be an artifact of finding the edge of the tropical belt, which is less well defined in the NH (Adam et al., 2018)−especially for JJA−which would also lead to enhanced variability in the NH.

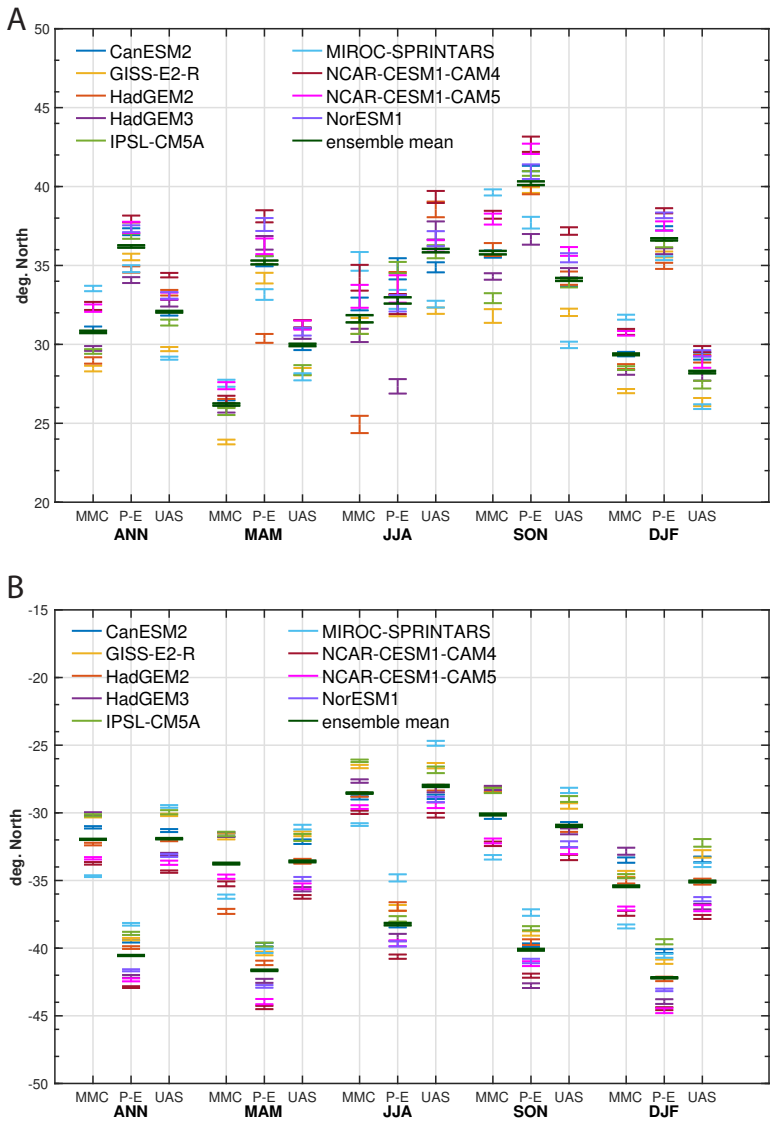

Figure 2.2: **Tropical belt edge location in the base experiment.** (A) Northern (°N) and (B) Southern (°S) Hemisphere annual and seasonal tropical belt edge locations in the base experiment for each individual model and ensemble mean based on the MMC, P−E and UAS metric. Errors bars show the 90% confidence interval of tropical belt edge location due to internal climate variability.

2.3 Results

The annual mean tropical belt edge response to anthropogenic drivers in the global perturbation experiments is shown in Figure 2.3. Consistent with prior studies (Tao et al., 2016; Grise & Polvani, 2016), co2×2 leads to tropical expansion. However, we find relatively weak tropical expansion in the NH, and much stronger and significant expansion in the SH. The annual mean ensemble mean poleward displacement of the SH tropical edge ($\sim 1^\circ \pm 0.19^\circ$) is about 5 times that of NH ($\sim 0.2^\circ \pm 0.15^\circ$). (Watt-Meyer, Frierson, & Fu, 2019) found twice as much GHG-induced tropical expansion in the SH, relative to the NH. GHG-forced tropical expansion is significant relative to internal climate variability in the SH, but not in the NH (gray shading in Fig. 1). The weaker rate of annual mean widening in the NH is partly due to tropical contraction during June-July-August (JJA) in all metrics (Figures 2.4-2.6). This result is consistent with prior analyses, where CMIP5 historical GHG and co2×4 simulations also yield NH tropical contraction during JJA (Tao et al., 2016; Grise & Polvani, 2016; Watt-Meyer et al., 2019). Further analysis is warranted, but this may be related to the JJA stationary wave response to CO_2 (Shaw, 2014; Shaw & Voigt, 2015). The largest NH widening trend due to co2×2 occurs in SON (N. A. Davis et al., 2016a; Grise & Polvani, 2016).

Similar to the single-model result from (Kovilakam & Mahajan, 2015), the bc×10 simulation yields significant tropical expansion in both hemispheres across all three metrics. The magnitude of annual mean ensemble mean NH tropical expansion ($\sim 0.96^\circ \pm 0.12^\circ$) is 2.3 times larger than in the SH ($\sim 0.41^\circ \pm 0.12^\circ$). This NH amplification is in contrast

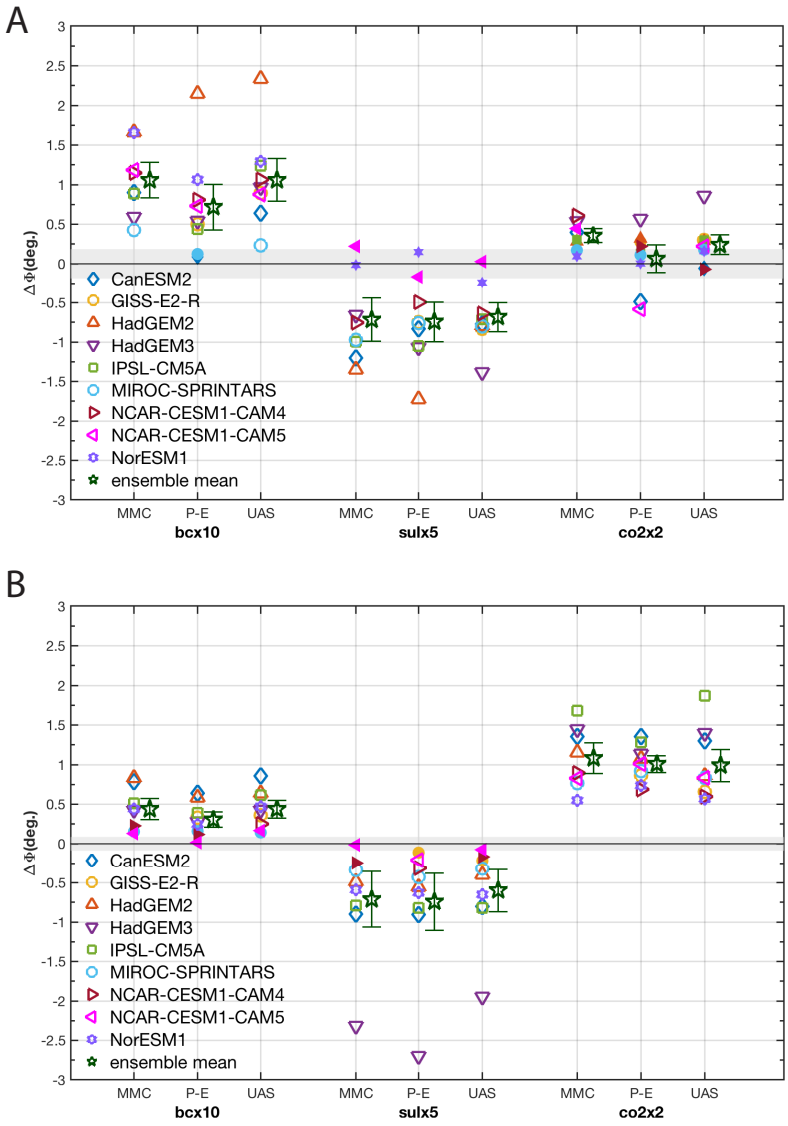

Figure 2.3: **Annual mean tropical belt edge response in global perturbation experiments.** (A) Northern and (B) Southern Hemisphere tropical belt edge change (° latitude) in response to bc×10, sul×5 and co2×2 based on the MMC, P-E and UAS metric for each individual model. Positive (negative) anomalies indicate tropical expansion (contraction). Unfilled model symbols indicate a response statistically significant at the 90% confidence level. The corresponding ensemble mean response is denoted by a dark green star and the error bar shows the approximate 90% confidence interval, estimated by $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the tropical width changes, and n is the number of models. Gray shading shows the 90% confidence interval of tropical belt edge displacements due to internal climate variability based on MMC metric (the 90% confidence intervals from the other two metrics are similar).

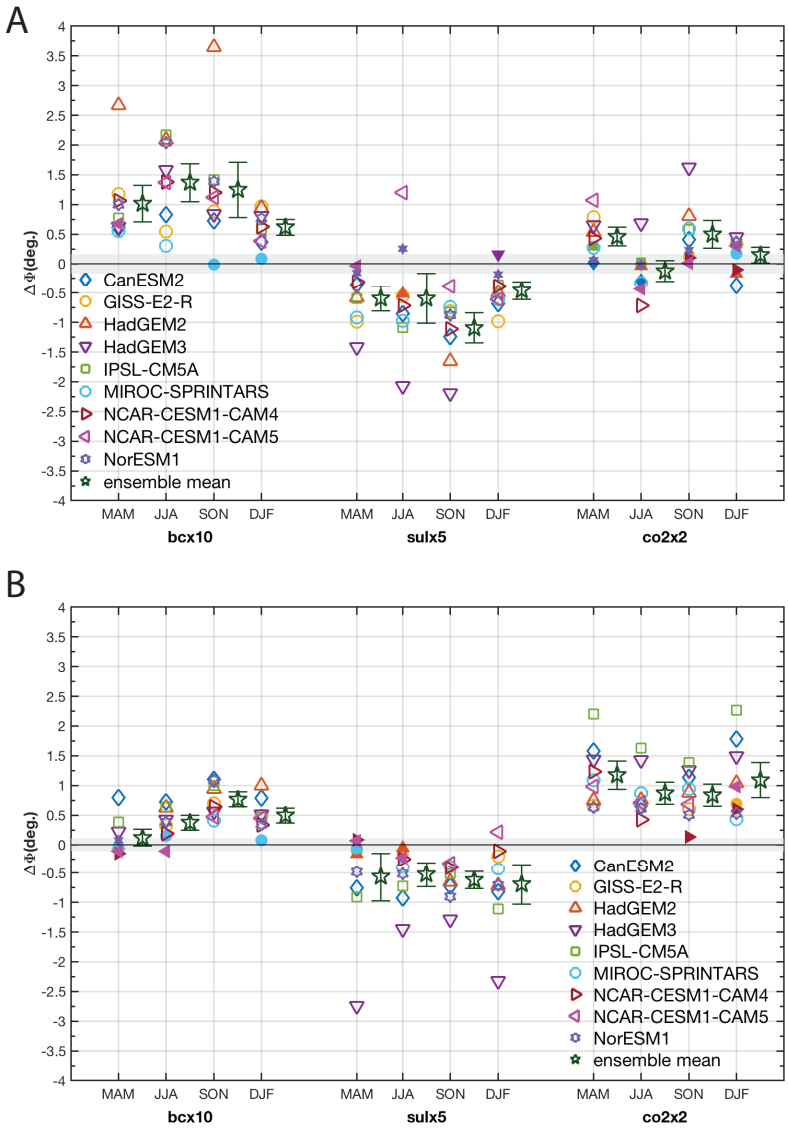

Figure 2.4: **Seasonal mean tropical belt edge response in global perturbation experiments based on UAS metric.** (A) Northern and (B) Southern Hemisphere tropical belt edge change (° latitude) in response to bc×10, sul×5 and co2×2 for each individual model. Positive (negative) anomalies indicate tropical expansion (contraction). Unfilled model symbols indicate a response statistically significant at the 90% confidence level. The corresponding ensemble mean response is denoted by a dark green star and the error bar shows the approximate 90% confidence interval, estimated by $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the tropical width changes, and n is the number of models. Grey shading shows the 90% confidence interval of tropical belt edge displacements due to internal climate variability based on UAS metric. Note the different y-axis between the two panels.

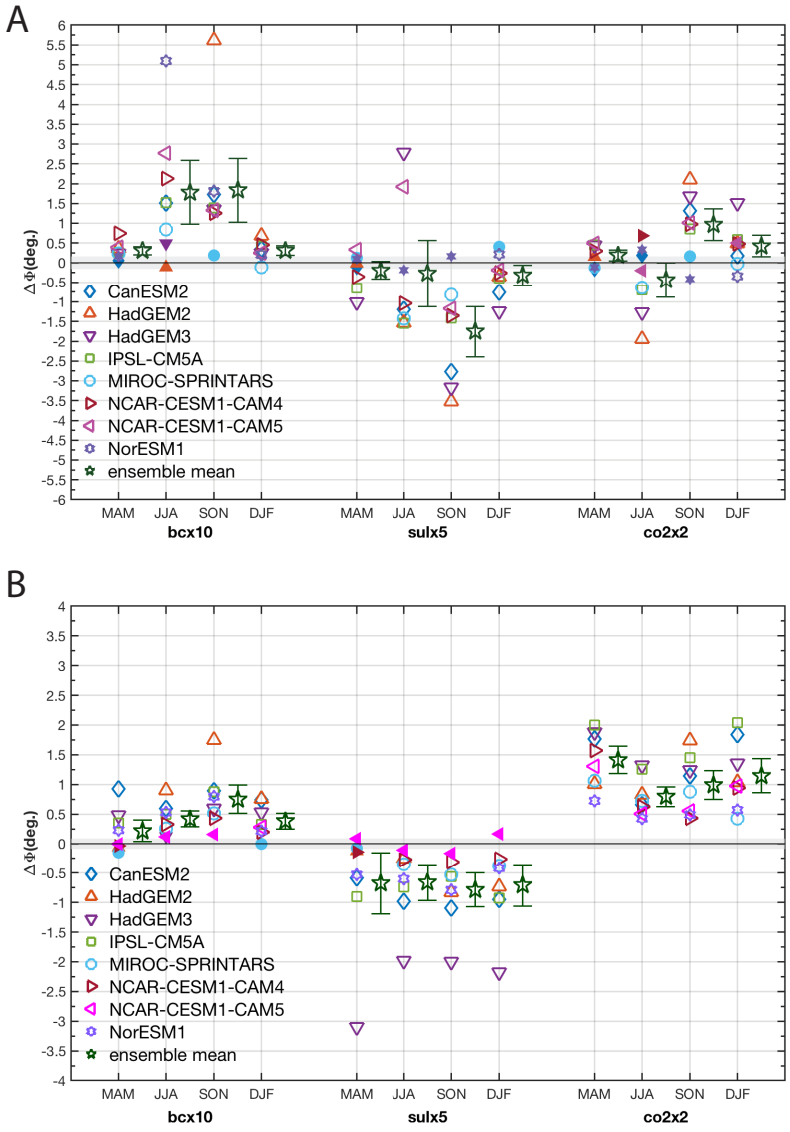

Figure 2.5: As in Figure 2.4 but based on MMC metric.

to co2×2 (where NH expansion is 20% as large as that in the SH), which is likely related to more BC burden in the NH, and consequently a larger hemispheric asymmetry in BC forcing (Table 2.1 and Figure 2.1). HadGEM2-ES generally yields the largest tropical expansion, especially in the NH, which is consistent with its large change in burden (Stjern et al., 2017) and relatively large BC ERF (Table 2.1). All seasons also show robust tropical expansion in response to bc×10 in both hemispheres (Figures 2.4-2.6), with maximum tropical widening occurring in the NH during JJA and September-October-November (SON) and in the SH during December-January-February (DJF) and SON. The large JJA BC response is interesting in light of the minimal JJA co2×2 expansion, which mainly results from the NH contraction in that season.

In contrast to BC, sul×5 drives significant annual mean tropical contraction in both hemispheres, at $\sim 0.69^\circ \pm 0.24^\circ$ for the NH and $\sim 0.66^\circ \pm 0.33^\circ$ for the SH. HadGEM3 yields the strongest tropical contraction, especially in the SH, consistent with its large sul×5 ERF (Table 2.1). Significant contraction also occurs in all seasons in both hemispheres except for JJA in the NH. Similar seasonal rates of contraction exist in the SH, while relatively larger seasonal variations exist in the NH, with maximum contraction in SON. The absence of NH contraction in JJA and the strongest contraction occurring in SON is analogous but opposite to the response to co2×2.

In terms of the regional perturbation simulations, bc×10 Asia yields significant tropical expansion (Figure 2.7). In the NH, the annual mean ensemble mean NH tropical

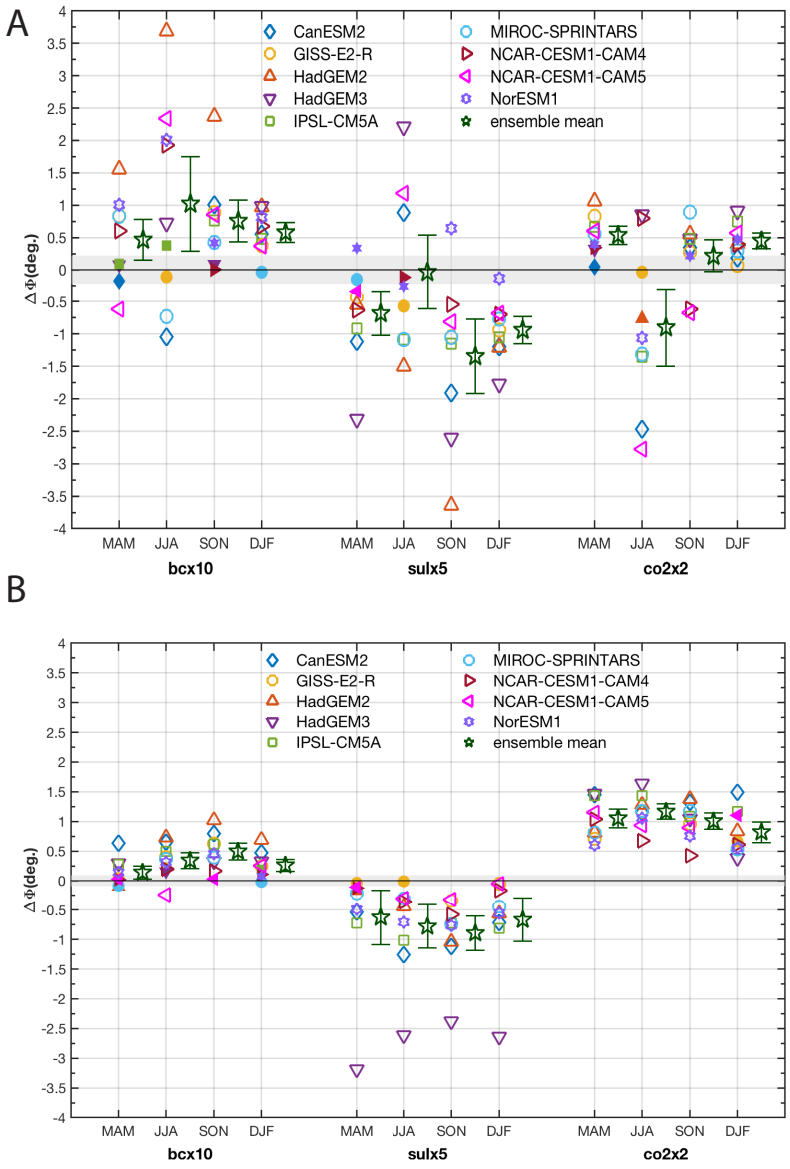

Figure 2.6: As in Figure 2.4 but based on P-E metric.

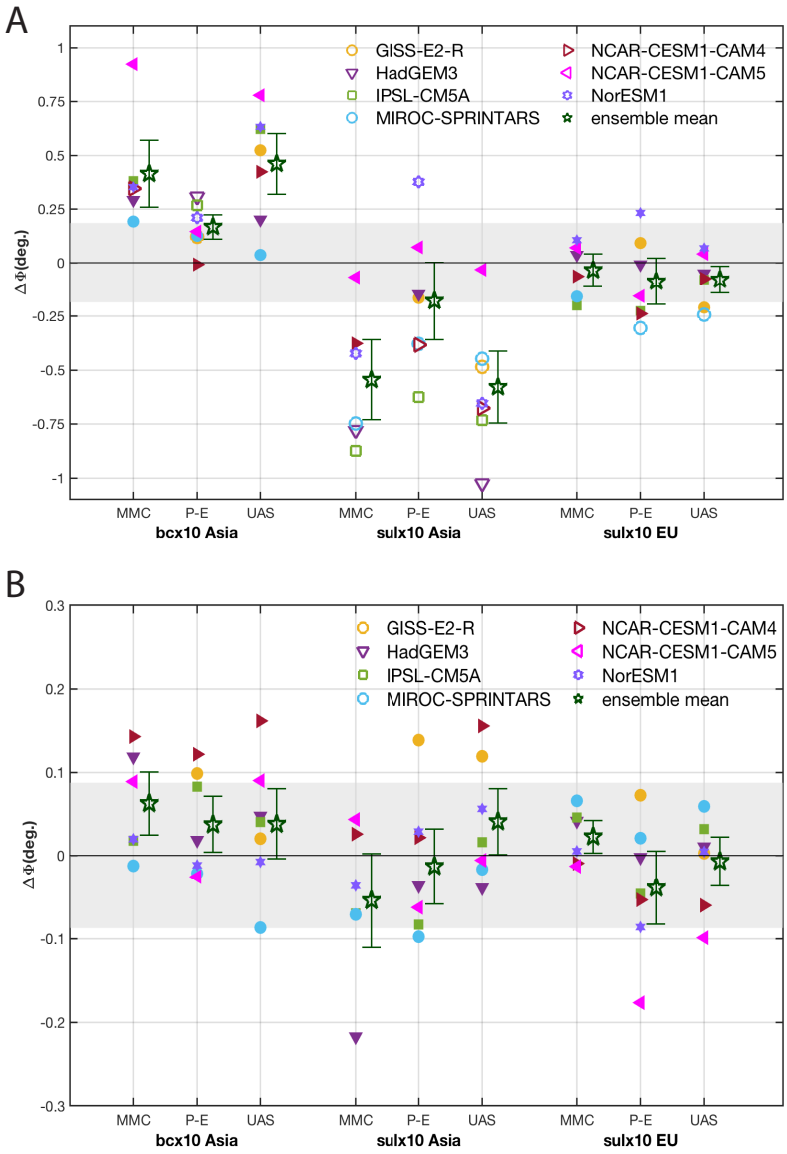

Figure 2.7: As in Figure 2.3 except for regional aerosol simulations including bc×10 Asia, sul×10 Asia and sul×10 EU. Note the different y-axis between the two panels.

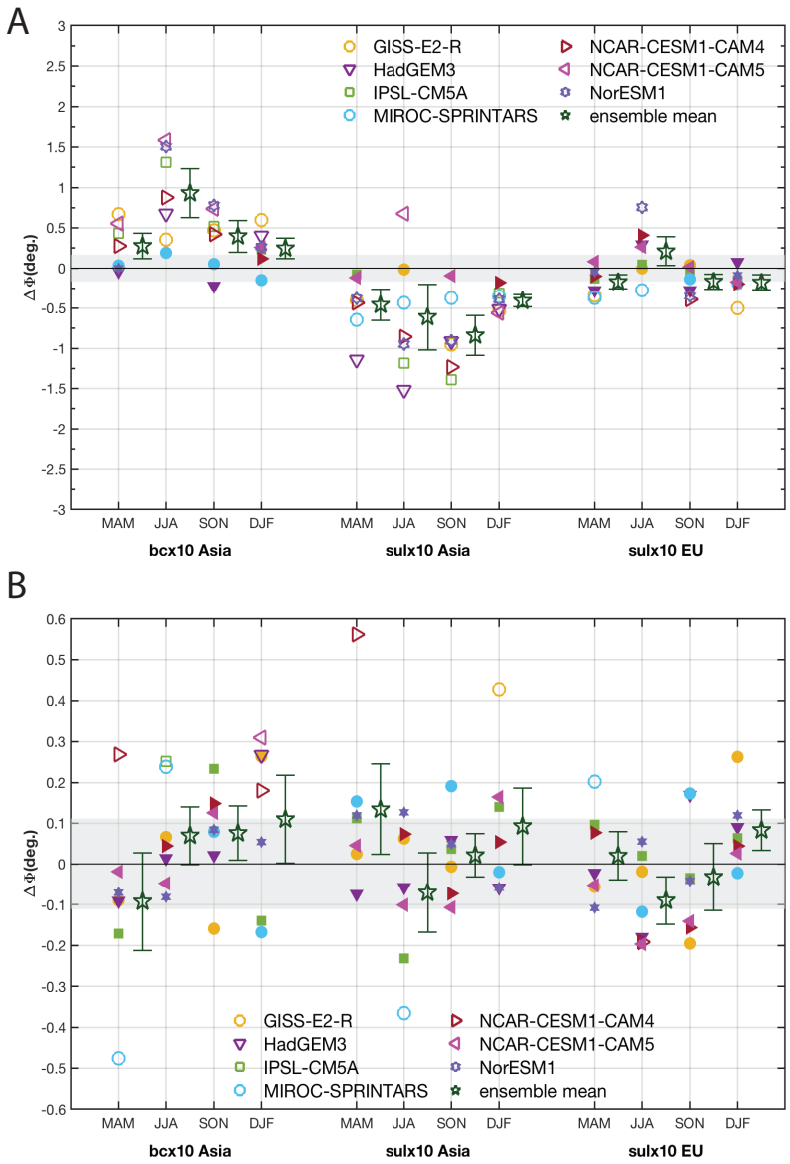

Figure 2.8: As in Figure 2.4 but for regional aerosol simulations including bc×10 Asia, sul×10 Asia and sul×10 EU. Note the different y-axis between the two panels.

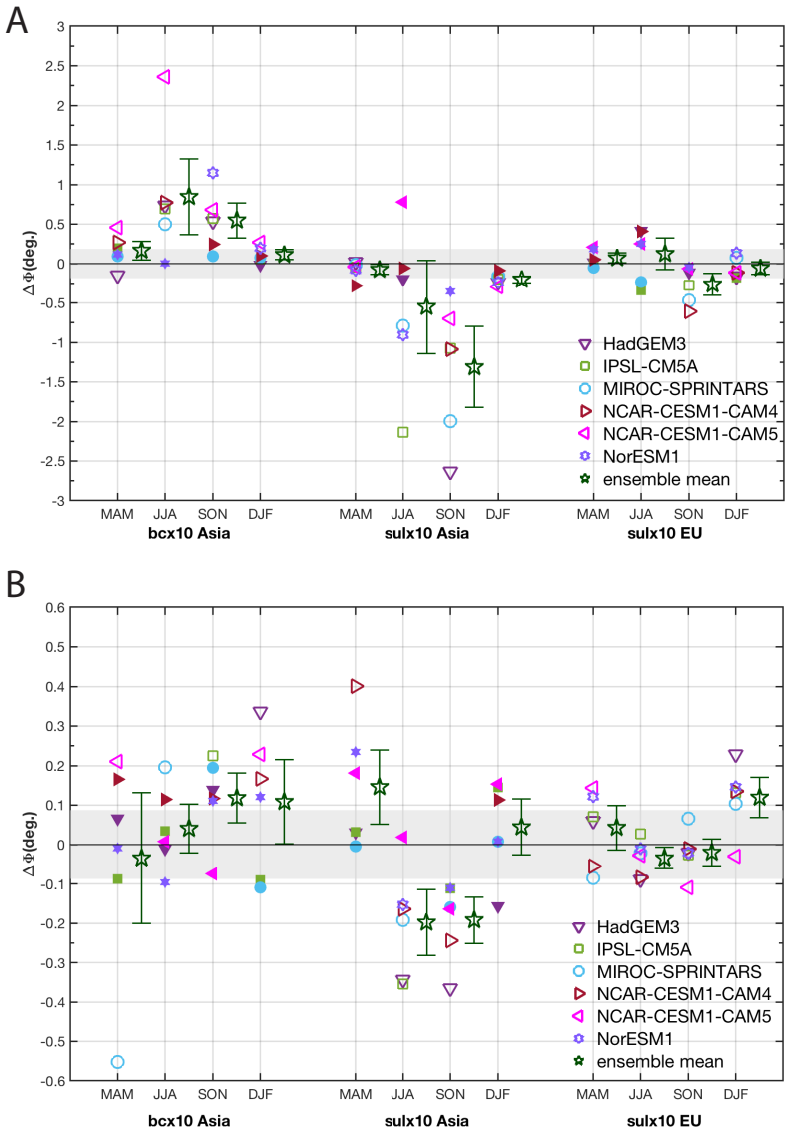

Figure 2.9: As in Figure 2.8 but based on MMC metric.

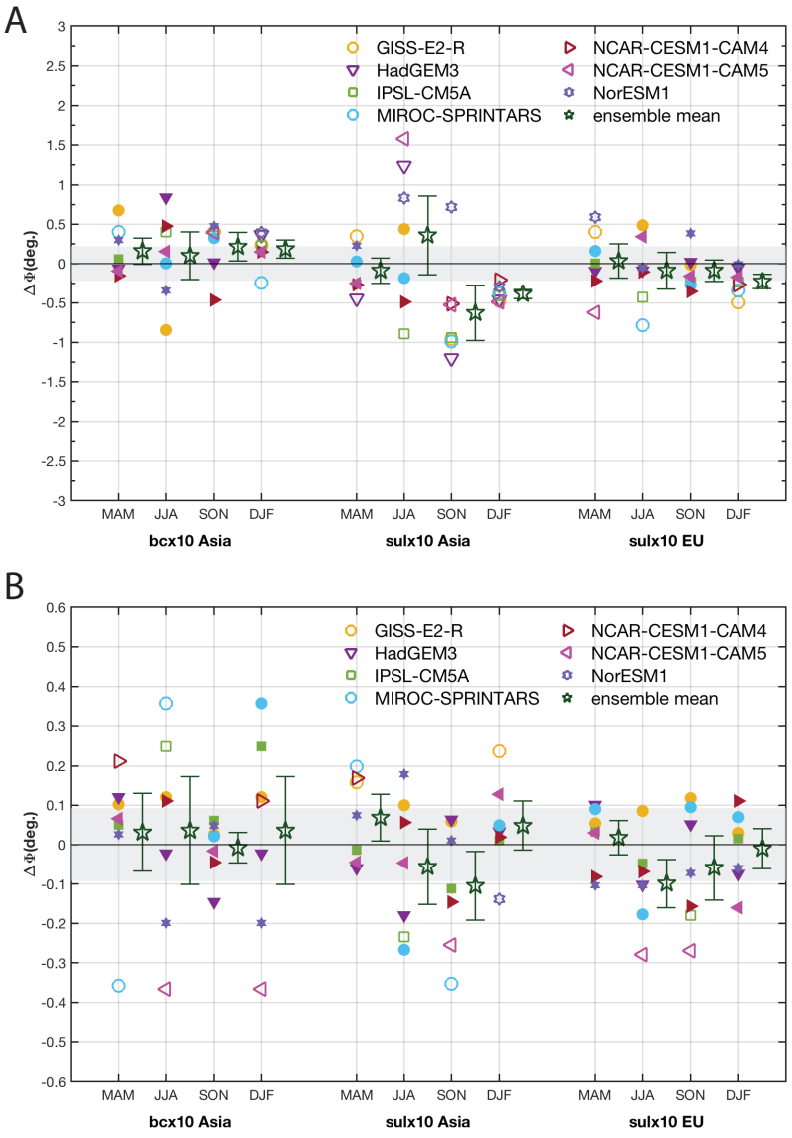

Figure 2.10: As in Figure 2.8 but based on P-E metric.

expansion is $\sim 0.36^\circ \pm 0.12^\circ$. We note that the P−E metric yields about half of the NH expansion (0.20°) as the other two metrics (0.45-0.48°). Despite bc×10 Asia's global (NH) ERF being only 15% (25%) of the corresponding ERF in bc×10, bc×10 Asia's NH tropical expansion is about 40% of the corresponding expansion in bc×10. This results suggests the importance of the forcing location, with Asian BC particularly efficient at widening the tropical belt. Maximum NH tropical expansion occurs in JJA (Figures 2.8-2.10). Significant tropical expansion also occurs in the SH, but it is much weaker at $\sim 0.05^\circ \pm 0.04^\circ$ (and within the 90% confidence interval due to internal climate variability).

The sul×10 Asia simulations yield significant NH contraction at $\sim 0.41^\circ \pm 0.18^\circ$, with maximum NH contraction in SON. Despite sul×10 Asia's global (NH) ERF being only 21% (32%) of the corresponding ERF in sul×5, sulx10 Asia's NH annual tropical contraction is 60% of the corresponding contraction in sulx5. This result again supports the importance of the location of the forcing to tropical width perturbations. The SH tropical belt edge response to Asian sulfate is not robust, and is inconsistent across metrics and seasons. The effect of European sulfate on tropical belt variations is also negligible, and not significantly different from internal climate variability. sul×10 Europe accounts for 9% (15%) of the sul×5 global (NH) ERF, and 10% of the NH tropical contraction. So unlike the Asian region, the tropical belt is not particularly sensitive to European sulfate emissions. As minimal NH contraction occurs under European sulfate, this implies the importance of sulfate from other regions (e.g., Asia and perhaps the U.S.). For SH contraction, the importance of sulfate from regions other than Europe and Asia is also implied from the large sul×5 SH contraction, but minimal contraction under sul×10 Asia and sul×10 EU.

All global (and regional) aerosol experiments show larger annual mean ensemble mean tropical edge displacements in the NH, as compared to the SH (particularly bc×10). This is related to the spatial heterogeneity of aerosols, with maximum loading and more importantly, ERF in the NH. For example, the annual mean ensemble mean bc×10 ERF is 1.59 W m^{-2} in the NH, compared to 0.38 W m^{-2} in the SH. By design, the regional aerosol experiments also have a much larger NH ERF, relative to the SH. However, for CO_2 forcing−which is uniform between hemispheres−much smaller tropical widening occurs in the NH. Watt-Meyer et al. (2019) ascribe this hemispheric asymmetry in CO_2 tropical expansion to a smaller sensitivity of the NH tropical edge to static stability changes. Of lesser importance is the pattern of the SST response and the CO_2 direct radiative effect.

The seasonal cycle of the NH and SH tropical edge response does not appear to be related to the seasonal cycle of global or hemispheric ERF. For example, bc×10 NH ERF is weakest in SON at 1.04 W m^{-2} when NH tropical expansion is relatively large (Figure 2.4-2.6). Similarly, bc×10 NH ERF is largest in MAM at 2.38 W m^{-2} when NH tropical expansion is relatively weak. We also note that latitudinal variations in the seasonal ERF are relatively small, and likely not an important contributor to the seasonal cycle of tropical edge displacements (not shown). These results reinforces the notion that there are certain seasons in which the tropical edge is more susceptible to latitudinal displacements (Watt-Meyer et al., 2019; Grise et al., 2018).

Figure 2.11 shows the ensemble mean annual mean ERF normalized tropical belt response to anthropogenic drivers. Aerosols generally drive larger NH tropical belt responses per ERF relative to GHGs. The NH tropical belt is most sensitive to BC (0.96° per W

m^{-2}), and in particular Asian BC (2.32° per W m^{-2}). Again, the quoted rates of tropical expansion represent the average over the three metrics. We note the P−E metric for bc×10 Asia yields a relatively smaller rate of expansion at 1.06° per W m^{-2}, versus ∼ 3° per W m^{-2} for the other two metrics. Compared to BC, the sensitivity of the NH tropical edge to sulfate is weaker. However, Asian emissions again yield the largest response (−0.55° per W m^{-2}) among the sulfate experiments. GHGs yield a much smaller normalized rate of NH tropical expansion at ∼0.06° per W m^{-2}.

In the SH, anthropogenic aerosols generally show smaller efficacy in driving tropical edge perturbations as compared to the NH. Global BC yields the largest SH efficacy at ∼ 0.41° per W m^{-2}, followed by Asian BC. Global sulfate yields a SH efficacy of −0.19° per W m^{-2}. In contrast to the aerosols, GHGs are more effective in perturbing the SH tropical edge (0.27° per W m^{-2}), as compared to the NH. Although not statistically significant, global BC and to some extent Asian BC still yield larger SH efficacies than co2×2. This suggests tropical width perturbations may be more responsive to the direct atmospheric heating by BC, as opposed to surface temperature driven feedbacks by CO_2 and SO_4. For both hemispheres, similar conclusions exist when we normalize by the change in global mean surface temperature, as opposed to ERF (Figure 2.12).

The larger displacement of the tropical belt edge in response to global aerosols, particularly BC, relative to GHGs suggests the potential importance of aerosols in perturbing the tropical belt, especially in the NH. Compared to global BC and sulfate forcing, the relatively larger NH response to the corresponding Asian aerosol indicates the location of the aerosol forcing is important. This is also supported by the larger NH tropical belt

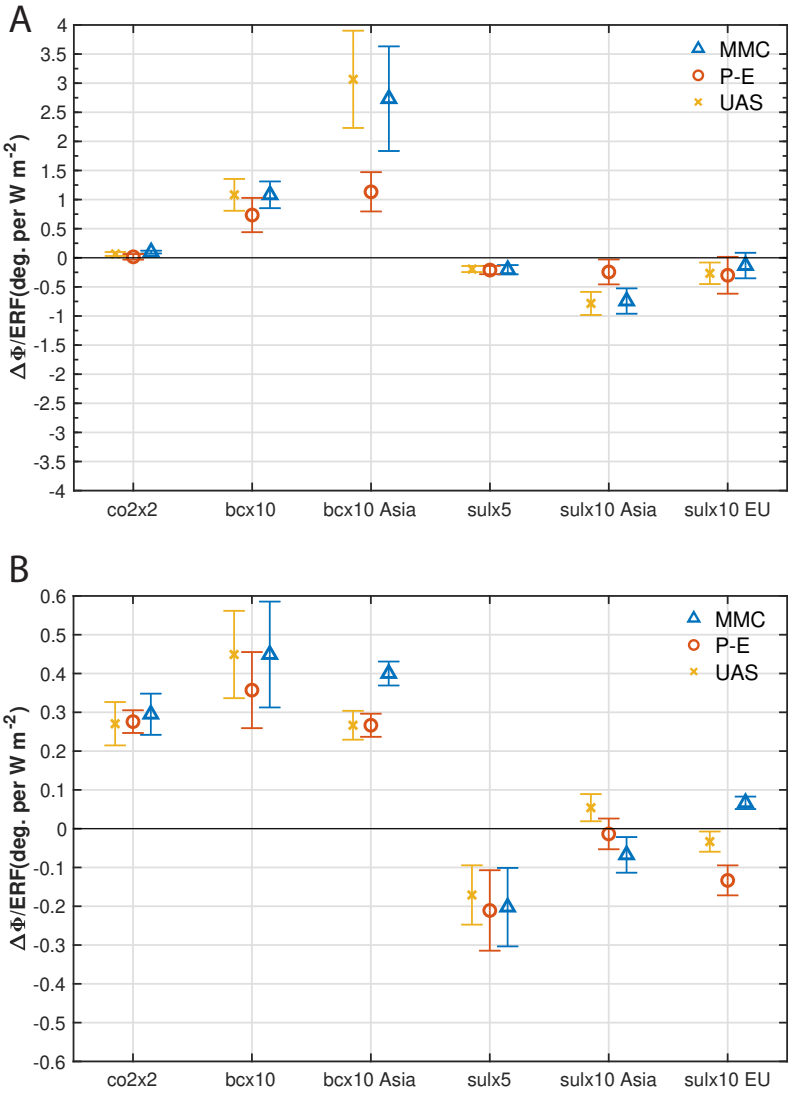

Figure 2.11: **Efficacy of anthropogenic drivers in perturbing the tropical belt edge.** (A) Northern and (B) Southern Hemisphere annual mean ensemble mean normalized tropical belt edge response to co2×2, bc×10, bc×10 Asia, sul×5, sul×10 Asia and sul×10 EU based on the MMC, P-E and UAS metric. The tropical belt edge response is normalized by the absolute value of the corresponding effective radiative forcing (ERF). Units are ° latitude per W m^{-2}. A positive (negative) response indicates tropical expansion (contraction). Note the different y-axis between the two panels. The 90% confidence interval is also included, estimated as $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the normalized tropical width changes, and n is the number of models.

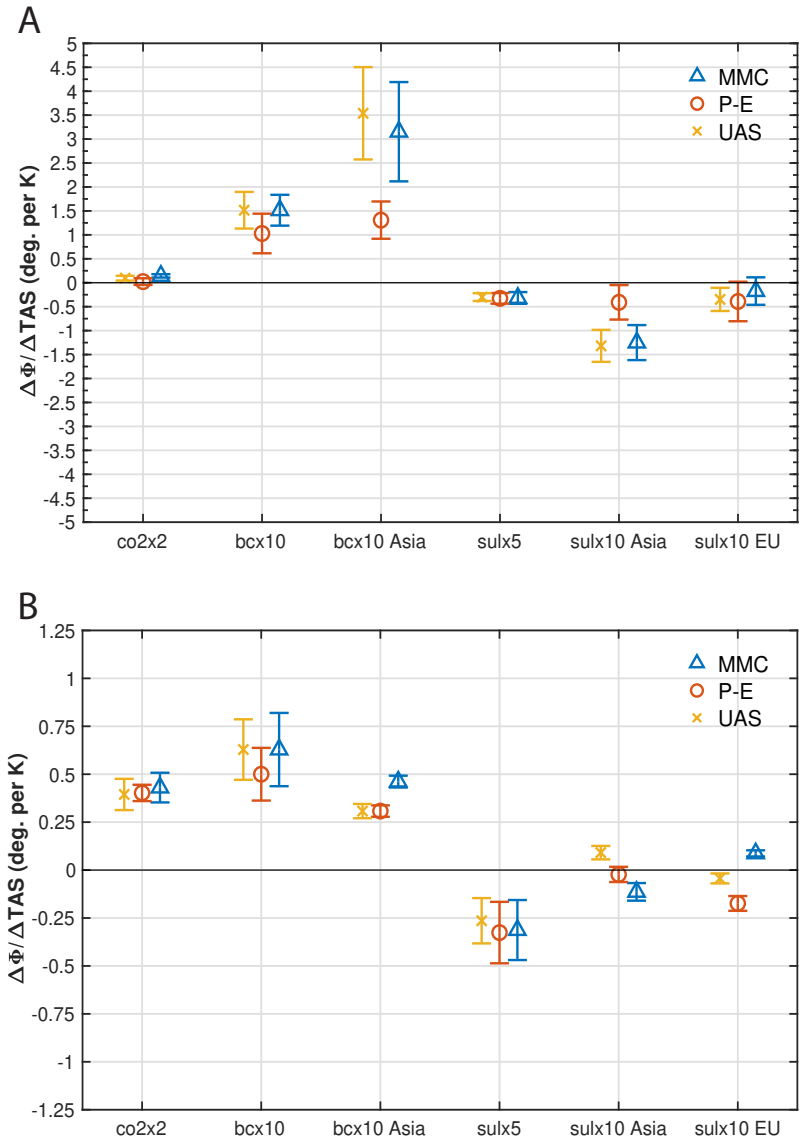

Figure 2.12: **Efficacy of anthropogenic drivers in perturbing the tropical belt edge.** (A) Northern and (B) Southern Hemisphere annual mean ensemble mean normalized tropical belt edge response to co2×2, bc×10, bc×10 Asia, sul×5, sul×10 Asia and sul×10 EU based on the MMC, P-E and UAS metric. The tropical belt edge response is normalized by the absolute value of the global mean surface temperature change. Units are ° latitude per K. A positive (negative) response indicates tropical expansion (contraction). The 90% confidence interval is also included, estimated as $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the normalized tropical width changes, and n is the number of models. Note the different y-axis between the two panels.

contraction in response to Asian sulfate, relative to European sulfate. These results imply that a forcing closer to the tropical belt edge is likely more effective in driving changes to tropical belt width, which is generally consistent with previous idealized simulations (Allen et al., 2012a).

Under one view, the width of the Hadley cell is determined by the poleward extent to which the angular momentum conservation continues until the resulting vertical shear becomes baroclinically unstable (Lu et al., 2007). Under such a scaling, the edge of the Hadley cell is sensitive to the gross static stability and the tropopause height near the poleward boundary of the circulation. Although this scaling is instructive, in reality the flow does not always conserve angular momentum (Schneider, 2006). We calculate the annual mean mid-latitude static stability, S as the potential temperature difference between the tropopause and surface ($\theta_{TP} - \theta_{SFC}$), averaged over 30-60° latitude, and the response (perturbation minus baseline) is normalized by the corresponding global ERF (absolute value in the case of sulfate). We estimate S from zonal mean temperature, as opposed to grid box by grid box values. Similar results are obtained when S is calculated over 30-60°, as well as from 20-40° and 30-40°.

The ensemble mean global ERF normalized tropical belt response is well correlated with the corresponding normalized mid-latitude ΔS (Figure 2.13). In the NH, correlations range from 0.97 to 0.98, depending on the metric. In the SH, the corresponding correlations range from 0.94 to 0.98. Similar, but somewhat weaker correlations are obtained without normalizing. Normalizing by the change in global mean surface temperature (as opposed to ERF) yields analogous conclusions (Figure 2.13). Comparing $\Delta\Phi$ versus ΔTAS (as opposed

to ΔS) yields similar correlations in the SH, but somewhat weaker correlations in the NH (Figure 2.14). Figure 2.13 shows that GHGs and absorbing aerosols drive an increase in the mid-latitude static stability, whereas scattering aerosols yield a decrease. The larger aerosol induced tropical belt response in the NH, relative to the SH, are consistent with larger perturbations to the NH extratropical static stability. Asian aerosols also lead to larger increases in the NH mid-latitude static stability, relative to the corresponding global forcing and GHGs. This supports our previous assertion that the location of the aerosol forcing is an important factor in the tropical belt width response.

The correlations between normalized (by global ERF) tropical belt response and static stability changes are mostly all positive and significant across experiments for individual models and metrics for the annual mean (Table 2.3). For example, 100% of the models yield a correlation of at least 0.8 (which is significant at the 90% confidence level) between the change in the MMC metric and static stability normalized by global ERF across experiments in the NH and SH. However, some models yield a negative correlation in the SH (e.g.,MIROC-SPRINTARS for UAS/P$-$E metrics). This is likely related to the tropical belt response in the regional forcing experiments, which is either weak or opposite of the other models (e.g., tropical contraction is simulated in the bc$\times$10 Asia experiment in MIROC-SPRINTARS). Similar results are generally obtained for the other two metrics, as well as when we normalize by the change in global mean surface temperature (as opposed to ERF). Weaker correlations occur based on non-normalized values (not shown).

Corresponding correlations for individual seasons are generally weaker, but more than 50% of the models yield a significant positive correlation ($r > 0.8$) for all seasons

(Table 2.4). JJA, and also to some extent SON, generally feature the largest percentage of models with significant correlations. For example, 63%/88% (75%/88%) of the models yield a JJA (SON) correlation of at least 0.8 based on the MMC metric for both the NH/SH. The smallest percentage of models with correlations larger than 0.8 occurs in MAM and DJF, particularly in the SH. Based on the MMC metric, 63% (75%) of the models yield a significant positive correlation in DJF (MAM) in the NH. The corresponding percentage of models in the SH is only 38% for DJF and 25% for MAM. Similar results are generally obtained with other metrics, as well as when we normalize by the change in global mean surface temperature (as opposed to ERF). Somewhat weaker results are obtained based on non-normalized values (not shown).

Table 2.5 shows the annual mean global ERF normalized correlations between $\Delta\Phi$ and ΔTAS. Correlations are also generally significant, but weaker than the corresponding correlations between $\Delta\Phi$ and ΔS (Supplementary Table 4). For example, based on the MMC metric, 75% (63%) of the model $\Delta\Phi$ and ΔTAS correlations are larger than 0.8 (90% significant) in the NH (SH). Based on $\Delta\Phi$ and ΔS, 100% of the correlations are larger than 0.8 in both hemispheres (Table 2.3).

Watt-Meyer et al. (2019) argue the hemispheric contrast in GHG-induced tropical expansion is related to weaker sensitivity to static stability changes in the NH. For non-normalized annual mean tropical edge displacements and ΔS across experiments, the ensemble-mean MMC-based tropical width sensitivity to ΔS is 0.19° with a model range of 0.07-0.29° K^{-1} in the NH. For P−E (UAS), the ensemble mean sensitivity is 0.14 with a range of 0.07-0.27° K^{-1} (0.20 with a range of 0.12-0.26° K^{-1}). In the SH, tropical width

sensitivity [$^\circ$ K^{-1}] to ΔS is 0.19 (0.07-0.24) for MMC, 0.20 (0.12-0.24) for P$-$E and 0.18 (0.11-0.23) for UAS. Thus, we obtain a similar tropical width sensitivity to ΔS in both hemispheres. Normalizing both quantities by the global mean ERF results in larger NH sensitivity (i.e., slope in Fig. 4), at 0.44 (0.2-0.8)$^\circ$ K^{-1} for the MMC metric; 0.21 (0.09-0.45)$^\circ$ K^{-1} for the P$-$E metric; and 0.48 (0.18-0.84)$^\circ$ K^{-1} for the UAS metric. In the SH, the corresponding sensitivities are 0.19 (0.08-0.26)$^\circ$ K^{-1} for MMC; 0.19 (0.12-1.61)$^\circ$ K^{-1} for P-E; and 0.17 (0.11-0.34)$^\circ$ K^{-1} for UAS.

The non-normalized ensemble mean tropical width sensitivity to ΔS varies across season. We find a stronger NH MMC-based tropical belt sensitivity to ΔS in JJA and relatively weaker sensitivities in the other seasons. For example, the sensitivity is 0.16° K^{-1} in JJA, and between -0.03 to 0.05° K^{-1} in the other three seasons. In the SH, the largest sensitivities occur in MAM, DJF and SON at 0.10, 0.08 and 0.07° K^{-1}, respectively. Weaker sensitivity occurs in JJA at 0.01° K^{-1}. Similar conclusions generally exist for the other two metrics, particularly the large NH sensitivity during JJA. Thus, across PDRMIP experiments, we do not find a weaker NH tropical belt sensitivity to static stability changes, as under CO_2 forcing alone (Watt-Meyer et al., 2019).

The correlations across models for given experiments are quite weak (Table 2.5) and generally not significant (so the sensitivities are also not significant). With 7-9 models (depending on the experiment), a correlation of at least 0.67 is required for significance at the 90% confidence level. For co2$\times$2, we find a NH non-normalized sensitivity based on the MMC metric of 0.06° K^{-1} and a SH sensitivity of 0.22° K^{-1} (Supplementary Table 8). Thus, there is a weaker tropical belt widening versus static stability sensitivity to CO_2

Table 2.3: Correlations between the annual mean change in tropical belt width ($\Delta\Phi$) and subtropical (30-60°) static stability (ΔS) across experiments for individual models and metrics. Both $\Delta\Phi$ and ΔS are normalized by the absolute value of the global mean ERF. Correlations significant at the 90% confidence level are denoted with bold font.

Model	MMC		UAS		P−E	
	NH	SH	NH	SH	NH	SH
CanESM2	0.93	0.93	0.84	0.90	0.86	0.96
GISS-E2-R	n/a	n/a	**0.98**	**0.80**	**0.82**	0.50
HadGEM2	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	0.95
HadGEM3	**0.92**	**0.85**	**0.94**	**0.90**	**0.90**	**0.84**
IPSL-CM5A	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**
MIROC-SPRINTARS	**0.99**	**0.99**	**0.88**	−0.73	**0.96**	−0.21
NCAR-CAM4	**0.95**	**0.95**	**0.98**	0.74	0.61	**0.93**
NCAR-CAM5	**0.96**	**0.96**	**0.96**	**0.96**	**0.85**	**0.84**
NorESM1	**0.99**	**0.95**	**0.99**	−0.39	**0.99**	−0.50

forcing in the NH, as compared to the SH. Similar results are obtained with the UAS metric (but not P−E). A similar result (larger SH sensitivity) generally exist for sul×5. Depending on the expansion metric, however, bc×10 tends to yield larger a NH sensitivity as compared to that in the SH. These results are generally applicable to normalized (by global ERF) sensitivities (Supplementary Table 9), which perhaps better shows CO_2 yields a weaker NH sensitivity relative to the SH, whereas the other experiments tend to yield the opposite (but again, this does depend on the metric in some cases). Thus, the $\Delta\Phi$ versus ΔS global ERF normalized correlations across model experiments−particularly for the annual mean−are also quite significant, and also generally larger than global ERF normalized $\Delta\Phi$ versus ΔTAS.

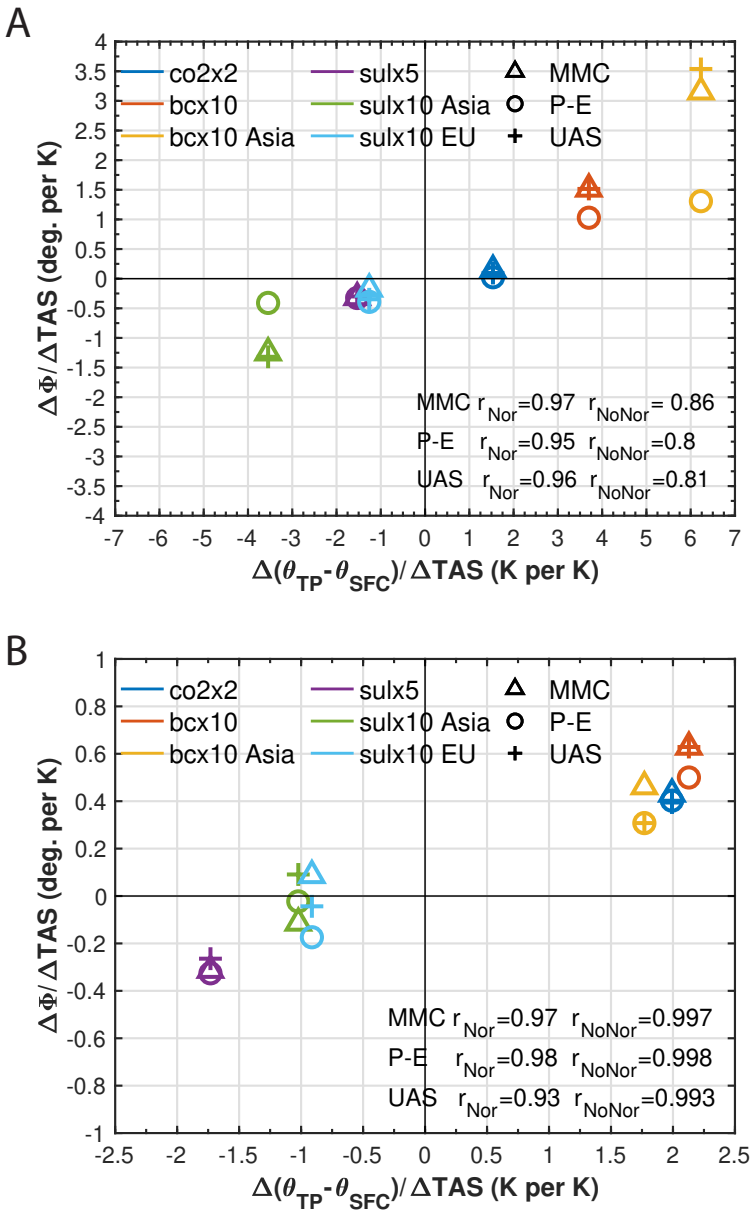

Figure 2.13: **Efficacy of tropical width perturbations versus normalized extratropical static stability.** Scatterplot of the annual mean ensemble mean (A) Northern and (B) Southern Hemisphere tropical belt edge response versus the corresponding subtropical static stability response. Both are normalized by the absolute value of the global mean surface temperature change. The tropical belt edge response is based on the MMC, P-E and UAS metric, as represented by the different symbols. A positive (negative) tropical belt edge response indicates tropical expansion (contraction). Also included are the correlations of normalized (r_{Nor}) and unnormalized (r_{NoNor}) tropical belt edge and extratropical static stability. Note the different axes between the two panels.

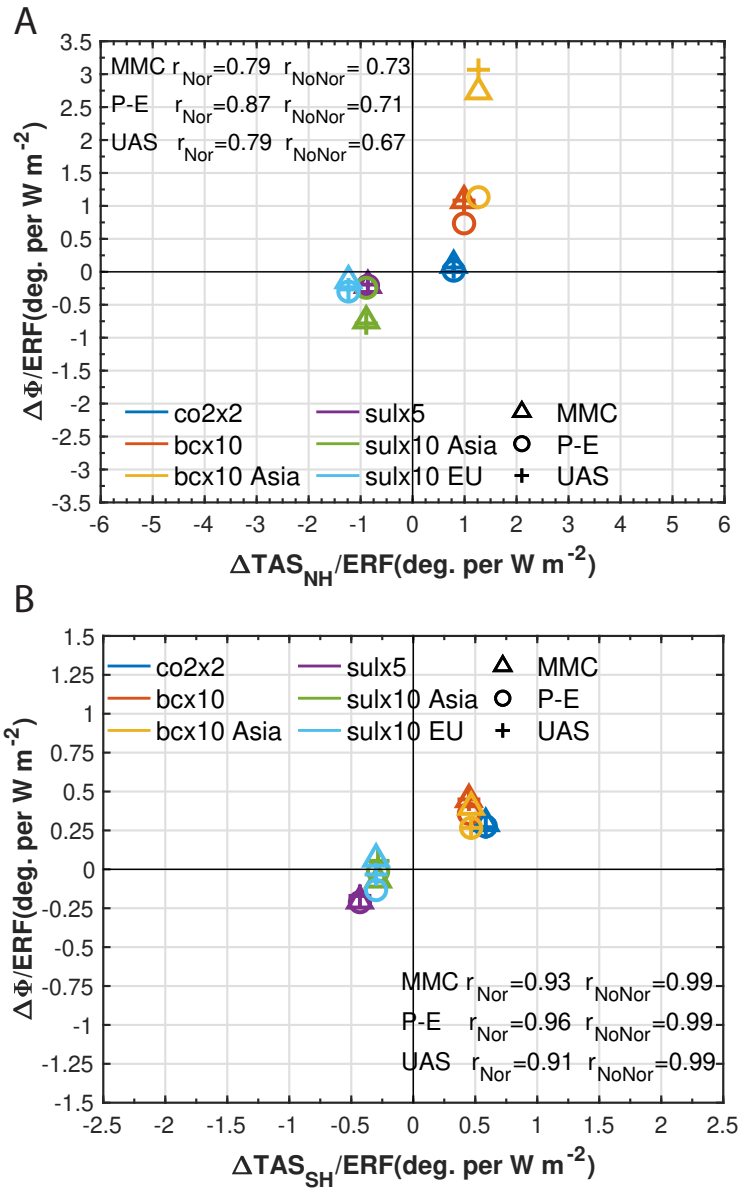

Figure 2.14: **Efficacy of tropical width perturbations versus normalized hemispheric surface temperature.** Scatterplot of the annual mean ensemble mean (A) Northern and (B) Southern Hemisphere tropical belt edge response versus the corresponding hemispheric surface temperature response. Both are normalized by the absolute value of the global mean ERF. The tropical belt edge response is based on the MMC, P-E and UAS metric, as represented by the different symbols. A positive (negative) tropical belt edge response indicates tropical expansion (contraction). Also included are the correlations of normalized (r_{Nor}) and unnormalized (r_{NoNor}) tropical belt edge and hemispheric surface temperature. Note the different axes between the two panels.

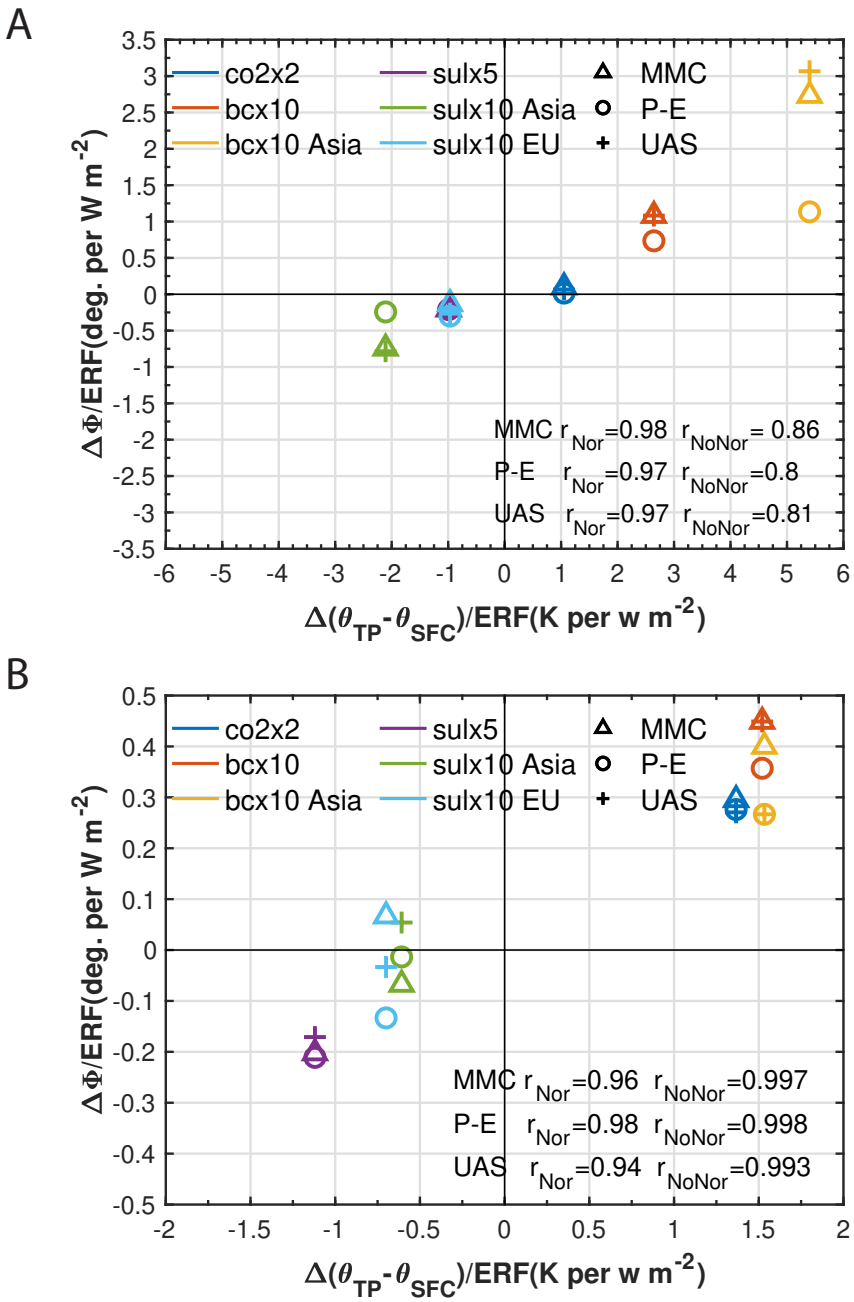

Figure 2.15: **Efficacy of tropical width perturbations versus normalized extratropical static stability.** Scatterplot of the annual mean ensemble mean (A) Northern and (B) Southern Hemisphere tropical belt edge response versus the corresponding subtropical static stability response. Both are normalized by the absolute value of the corresponding effective radiative forcing (ERF). The tropical belt edge response is based on the MMC, P-E and UAS metric, as represented by the different symbols. A positive (negative) tropical belt edge response indicates tropical expansion (contraction). Also included are the correlations of normalized (r_{Nor}) and unnormalized (r_{NoNor}) tropical belt edge and extratropical static stability. **Note the different x-axis and y-axis between the two panels.**

Table 2.4: Seasonal correlations between the change in the MMC-based tropical belt width ($\Delta\Phi$) and subtropical (30-60°) static stability (ΔS) across experiments for individual models. Both $\Delta\Phi$ and ΔS are normalized by the absolute value of the global mean ERF. Correlations significant at the 90% confidence level are denoted with bold font.

Model	NH MMC				SH MMC			
	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF
CanESM2	0.47	0.95	0.70	0.90	0.95	0.85	0.97	0.95
GISS-E2-R	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
HadGEM2	**0.99**	0.76	**0.99**	**0.99**	0.24	**0.99**	**0.99**	**0.99**
HadGEM3	0.01	0.73	**0.99**	**0.98**	**0.92**	**0.90**	**0.96**	−0.58
IPSL-CM5A	**0.90**	**0.98**	0.77	0.40	0.75	**0.94**	0.69	0.18
MIROC-SPRINTARS	**0.80**	**0.98**	**0.94**	**−0.84**	0.63	**0.98**	**0.99**	**−0.82**
NCAR-CAM4	**0.98**	**0.97**	**0.91**	**0.91**	0.21	**0.99**	**0.97**	0.75
NCAR-CAM5	**0.92**	**0.98**	**0.96**	**0.80**	0.59	**0.89**	**0.96**	0.21
NorESM1	**0.92**	0.26	**0.99**	**0.98**	0.63	−0.27	**0.81**	**0.85**

Table 2.5: Correlations between the annual mean change in tropical belt width ($\Delta\Phi$) and hemispheric surface temperature (ΔTAS) across experiments for individual models and metrics. Both $\Delta\Phi$ and ΔTAS are normalized by the absolute value of the global mean ERF. Correlations significant at the 90% confidence level are denoted with bold font.

Model	MMC		UAS		P−E	
	NH	SH	NH	SH	NH	SH
CanESM2	0.92	0.98	0.85	**0.99**	0.82	0.96
GISS-E2-R	n/a	n/a	0.17	0.78	0.27	0.61
HadGEM2	0.86	0.97	0.86	**0.99**	0.79	0.96
HadGEM3	0.69	0.77	0.75	**0.82**	0.77	**0.84**
IPSL-CM5A	**0.99**	**0.99**	**0.99**	**0.97**	**0.99**	**0.98**
MIROC-SPRINTARS	**0.98**	−0.27	**0.97**	−0.36	**0.87**	−0.81
NCAR-CAM4	**0.85**	0.60	0.70	0.56	**0.84**	0.29
NCAR-CAM5	0.73	**0.83**	0.65	**0.88**	0.73	**0.82**
NorESM1	**0.99**	**0.92**	**0.98**	−0.55	**0.99**	−0.47

Across drivers, ΔS come from different mechanisms. For CO_2 and sulfate, surface driven feedbacks are most important, whereas for BC atmospheric absorption is also important. This is supported by Figure 2.16 and Table 2.6, which shows the ensemble mean zonal mean annual ΔS for both the coupled and fixed SST (fSST) simulations for each of the six experiments. Also included is the corresponding ΔTAS from both coupled and fSST experiments, and the ERF. The coupled response yields the total climate response; the fSST response yields the fast response (e.g., atmospheric heating); and the difference yields the slow response (surface temperature feedbacks). For the global forcing experiments, the fast ΔS (from the fSST experiments) is a larger proportion of the total ΔS (from the coupled experiments) in bc×10, as compared to sul×5 and co2×2 (Figure 2.1 and 2.2; Table 2.6). Based on 30-60°N (global), the fast ΔS in bc×10 is 30% (15%) of the total ΔS; in co2×2, the fast ΔS is 9% (7%) of the total ΔS; and in sul×5 the fast ΔS is 14% (6%) of the total ΔS. Thus, the fast response (i.e., direct atmospheric heating) is about 2-3 times more important in the global BC simulation, relative to the global CO_2 and SO_4 simulations.

Similar results also exist for the regional forcing experiments. Based on 30-60°N (global), the fast ΔS in bc×10 Asia is 81% (69%) of the total ΔS; in sul×10 Asia, the fast ΔS is 33% (19%) of the total ΔS; and in sul×10 Europe, the fast ΔS is -14% (-14%) of the total ΔS. Reasons for the opposite signed ΔS in fSST and coupled sul×10 Europe simulations are unclear. Nonetheless, this result also shows that the fast response is 2-3 times important in bc×10 Asia relative to sul×10 Asia.

We note that similar, but somewhat weaker results exist for ΔTAS. Although fSST experiments do not allow the SSTs to respond to the forcing, land surfaces can. So fSST

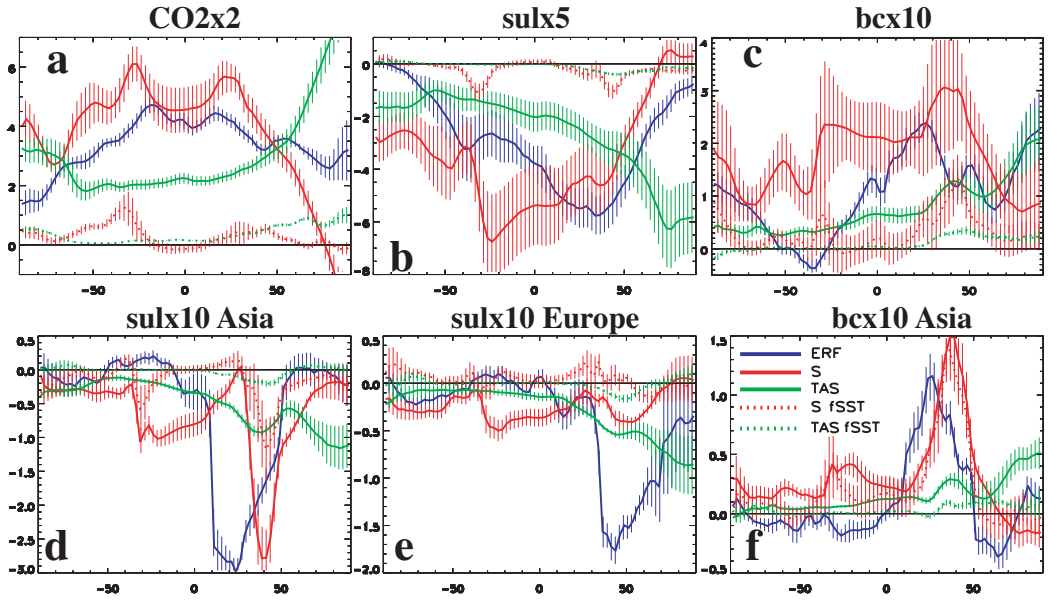

Figure 2.16: **Ensemble mean zonal mean annual responses.** Static stability (S; red) response for both the coupled (solid) and fixed SST (fSST; dashed) simulations for each of the six experiments, as well as the corresponding surface temperature (TAS; green) response, and the effective radiative forcing (ERF; blue). The coupled response yields the total climate response; the fSST response yields the fast response (e.g., atmospheric heating); and the difference yields the slow response (i.e., surface temperature feedbacks). S and TAS units are K; ERF units are W m^{-2}. Thin vertical cross hatching indicates the standard error of the response, estimated as $\frac{\sigma}{\sqrt{n}}$, where σ is the intermodel standard deviation and n is the number of models. Note that panels have a different scaling of the y-axis.

Table 2.6: Ensemble mean annual mean percentage of the total response (coupled simulations) due to the fast (fSST simulations) response for static stability (ΔS), surface temperature (ΔTAS) and tropical edge displacements ($\Delta\Phi$). ΔS and ΔTAS are shown for 30-60N, 30-60S and for the global mean; $\Delta\Phi$ is shown for the NH and SH. The very large SH $\Delta\Phi$ percentages for some experiments (e.g., sul×10 Europe) is due to a small (and not significant) $\Delta\Phi$ in the coupled run. Units are %.

Experiment	ΔS			ΔTAS			$\Delta\Phi$	
	30-60N	30-60S	Global	30-60N	30-60S	Global	NH	SH
Global Simulations								
co2×2	9	17	7	17	4	10	5	22
bc×10	30	15	15	25	3	12	52	54
sul×5	14	11	6	10	1	4	15	21
Regional Simulations								
bc×10 Asia	81	59	69	33	<1	23	69	−20
sul×10 Asia	33	29	19	17	<1	9	32	1000
sul×10 Europe	−14	-5	−14	21	<1	9	−86	1100

runs do not eliminate all of the surface temperature driven feedbacks. Thus, we are unable to completely separate the role of ΔTAS from ΔS. For the global forcing experiments, the fast ΔTAS is a larger proportion of the total ΔTAS in bc×10, as compared to sul×5 and co2×2. Based on 30-60°N (global), the fast ΔTAS in bc×10 is 25% (12%) of the total ΔTAS (Table 2.6); in co2×2, the fast ΔTAS is 17% (11%) of the total ΔTAS; and in 5×sul the fast ΔTAS is 10% (4%) of the total ΔTAS. Thus, the fast ΔTAS response is about 2-3 times more important in the global BC simulation, relative to the global CO_2 and SO_4 simulations.

Similar results also exist for the regional forcing experiments. Based on 30-60°N (global), the fast ΔTAS in bc×10 Asia is 33% (23%) of the total ΔTAS (Table 2.6); in sul×10 Asia, the fast ΔTAS is 17% (9%) of the total ΔTAS; and in sul×10 Europe, the fast ΔTAS is 21% (9%) of the total ΔTAS. The fast response is again 2-3 times important in bc×10 Asia relative to sul×10 Asia. Although it is difficult to separate ΔS from ΔTAS, these results suggest direct atmospheric heating by BC drives a large fraction of the ΔS.

We also calculate changes in tropical belt width from the fSST experiments. Although fewer significant responses occur, conclusions similar to those from the coupled simulations generally exist, including tropical expansion (contraction) under BC (SO_4)−particularly with Asian emissions−and weak NH tropical expansion under CO_2 (Figures 2.17 and 2.18). For the global simulations, we find that bc×10 fSST widening is a larger proportion of the total widening as compared to co2×2 and sul×5, especially in the NH (Table 2.6). bc×10 fSST widening is 52% and 54% of the total widening in the NH and SH, respectively. For co2×2 (sul×5) the corresponding percentages are 5% and 22% (15%

Table 2.7: Non-normalized correlations between the annual mean change in tropical belt width ($\Delta\Phi$) and subtropical (30-60°) static stability (ΔS) across models for individual experiments and metrics. Correlations larger than 0.67 are significant at the 90% confidence level (bold).

Experiment	MMC		UAS		P−E	
	NH	SH	NH	SH	NH	SH
co2×2	0.43	**0.96**	**0.70**	**0.90**	0.36	**0.76**
bc×10	0.57	**0.74**	**0.83**	0.49	**0.87**	0.63
sul×5	0.34	**0.98**	**0.86**	**0.97**	0.56	**0.98**
bc×10 Asia	0.28	**0.79**	0.43	0.46	−0.33	0.06
sul×10 Asia	**0.82**	**0.88**	**0.87**	**0.71**	0.43	0.62
sul×10 Europe	0.39	−0.01	0.43	0.26	0.30	0.27

and 21%). Similar results also exist for the regional experiments, particularly in the NH (where the signal is significant) for bc×10 Asia versus sul×10 Asia. The larger proportion of tropical widening in the fSST runs is consistent with the larger proportion of ΔS in the fSST runs for BC. This, in turn, supports the role of ΔS in BC induced tropical widening, and the role of direct atmospheric heating.

As previously mentioned, Watt-Meyer et al. (2019) argue the hemispheric contrast in GHG-induced tropical expansion is related to weaker sensitivity to static stability changes in the NH. For non-normalized annual mean tropical edge displacements and ΔS across experiments, the ensemble-mean MMC-based tropical width sensitivity to ΔS is 0.19° K^{-1} with a model range of 0.07-0.29° K^{-1} in the NH. In the SH, the MMC-based tropical width sensitivity to ΔS is 0.19 (0.07-0.24)° K^{-1}. Thus, we obtain a similar tropical width sensitivity to ΔS in both hemispheres. Similar conclusions generally exist with other metrics and seasons. However, across models, CO_2 does yield weaker NH sensitivity (Tables 2.7-2.9).

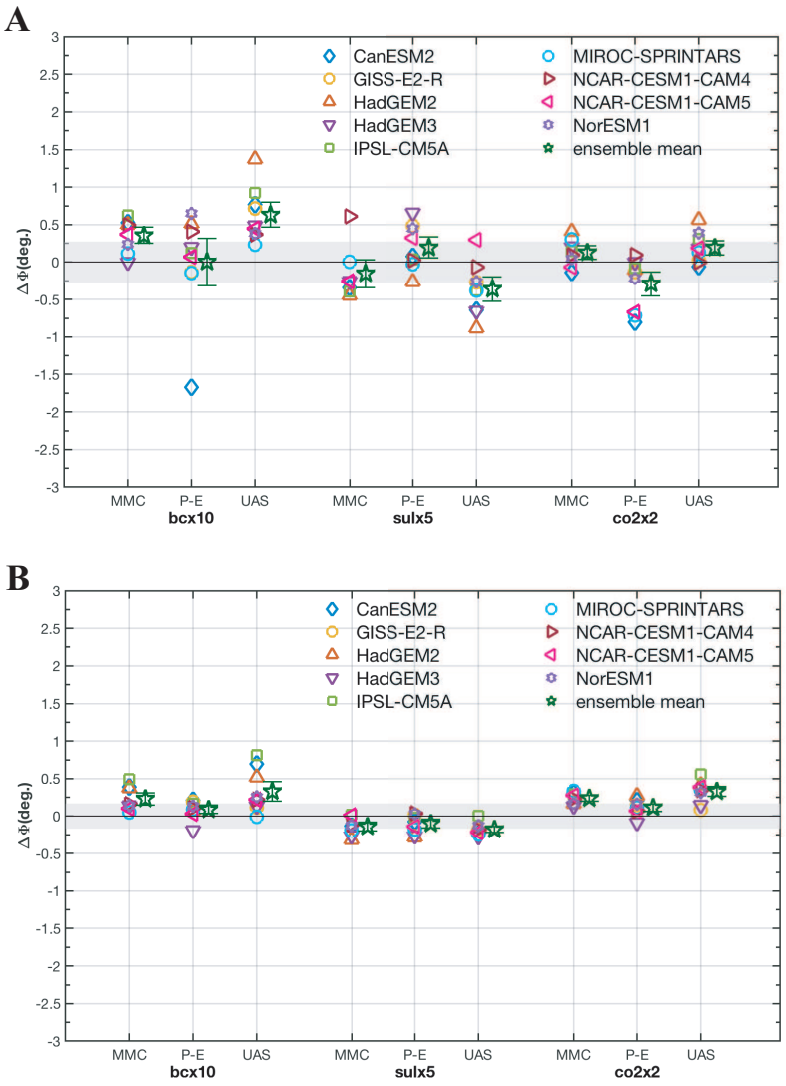

Figure 2.17: **Annual mean tropical belt edge response in global fSST perturbation experiments.** (A) Northern and (B) Southern Hemisphere tropical belt edge change (° latitude) in response to bc×10, sul×5 and co2×2 based on the MMC, P-E and UAS metric for each individual model. Positive (negative) anomalies indicate tropical expansion (contraction). Unfilled model symbols indicate a response statistically significant at the 90% confidence level. The corresponding ensemble mean response is denoted by a dark green star and the error bar shows the approximate 90% confidence interval, estimated by $2 \times \sigma/\sqrt{n}$, where σ is the intermodel standard deviation of the tropical width changes, and n is the number of models. Gray shading shows the 90% confidence interval of tropical belt edge displacements due to internal climate variability based on MMC metric (the 90% confidence intervals from the other two metrics are similar).

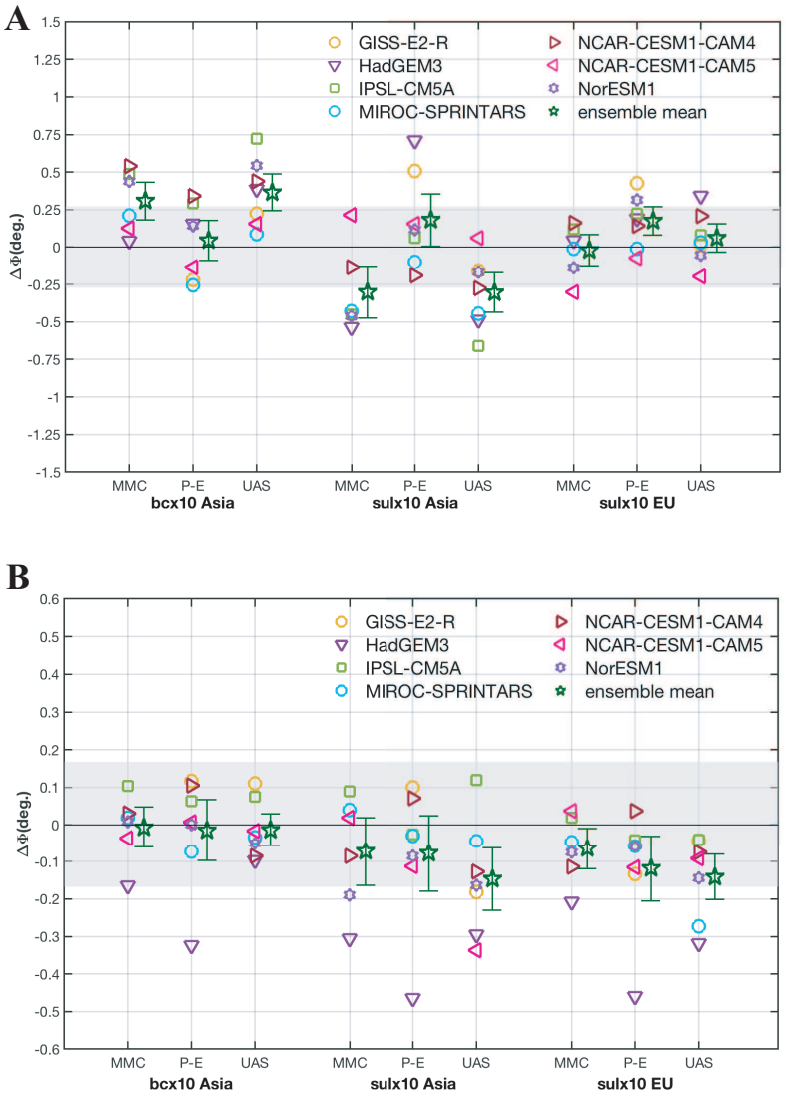

Figure 2.18: **Annual mean tropical belt edge response in regional fSST perturbation experiments.** As in Fig. 2.17 except for regional aerosol simulations including bc×10 Asia, sul×10 Asia and sul×10 EU. Note the different y-axis between the two panels.

Table 2.8: Non-normalized sensitivities between the annual mean change in tropical belt width ($\Delta\Phi$) and subtropical (30-60°) static stability (ΔS) across models for individual experiments and metrics. Units are ° latitude K^{-1}. Sensitivities significant at the 90% confidence level are denoted by bold font.

Experiment	MMC		UAS		P−E	
	NH	SH	NH	SH	NH	SH
co2×2	0.06	**0.22**	**0.15**	**0.23**	0.11	**0.10**
bc×10	0.09	**0.11**	**0.17**	0.07	**0.19**	0.08
sul×5	0.07	**0.20**	**0.14**	**0.17**	0.13	**0.23**
bc×10 Asia	0.17	**0.41**	0.27	0.28	−0.09	0.03
sul×10 Asia	**0.36**	**0.32**	**0.40**	**0.20**	0.22	0.20
sul×10 Europe	0.39	0.00	0.42	0.07	0.50	0.11

Table 2.9: ERF normalized sensitivities between the annual mean change in tropical belt width ($\Delta\Phi$) and subtropical (30-60°) static stability (ΔS) across models for individual experiments and metrics. Units are ° latitude K^{-1}. Sensitivities significant at the 90% confidence level are denoted by bold font.

Experiment	MMC		UAS		P−E	
	NH	SH	NH	SH	NH	SH
co2×2	0.05	**0.22**	0.13	**0.24**	0.10	0.12
bc×10	0.09	0.08	0.17	0.05	0.22	0.06
sul×5	0.36	**0.22**	0.22	0.20	0.27	**0.21**
bc×10 Asia	**0.90**	**0.19**	**1.49**	**0.45**	**0.63**	0.92
sul×10 Asia	0.27	0.32	**0.40**	0.12	0.26	0.17
sul×10 Europe	0.14	−0.04	0.30	0.02	0.14	0.02

2.4 Summary

With the usage of simulations from PDRMIP, the response of the tropical belt width to various anthropogenic drivers is quantified. Consistent with prior studies, results show that GHGs lead to tropical expansion, although relatively weak annual mean expansion in the NH due to tropical contraction during JJA are found. Absorbing aerosols lead to significant tropical expansion, whereas scattering aerosols lead to significant tropical contraction. Similar responses are generally found for the regional aerosol simulations, including tropical expansion (contraction)−particularly in the NH−in response to Asian BC (sulfate). Tropical belt width responses to European sulfate are relatively small and lack consistency across metrics. To account for the differing forcing magnitudes across PDRMIP experiments, tropical belt responses are normalized by the corresponding ERF. BC, especially Asian BC, is the most efficient driver of tropical widening, particularly in the NH. Furthermore, tropical belt width responses are highly correlated with changes in the midlatitude static stability, with absorbing (reflecting) aerosols driving increased (decreased) static stability and tropical expansion (contraction). The seasonal analysis suggests BC leads to maximum tropical widening in both hemispheres during summer and SON, which also corresponds to the observed maximum rates of tropical expansion (Hu & Fu, 2007; Allen & Kovilakam, 2017a; Grise et al., 2018). Although this seasonal correspondence further implies the importance of BC to observed tropical expansion, the NH tropical expansion in SON has been partially attributed to natural variability, particularly the PDO (Grise et al., 2018).

Chapter 3

Strengthening of the Walker Circulation in Recent Decades and the Role of Natural Sea Surface Temperature Variability

3.1 Introduction

The Walker Circulation (WC) is a large-scale zonal-vertical atmospheric circulation in the equatorial Pacific. Its variations are tied to the sea surface temperature (SST) gradient across the tropical Pacific. For example, during El Niño, the tropical Pacific SST gradient is weakened, as is the strength of the WC. In contrast, the tropical Pacific SST gradient and the WC are strengthened during La Niña. Similarly, changes in the sea level

pressure (SLP) gradient across tropical Pacific is an indicator of changes in the vertical motion of the WC, including the descending branch over cooler eastern Pacific and the ascending branch over the warm western Pacific. Variations in the strength and structure of the WC are closely tied to the southeast Asian monsoon (Wang, Liu, Kim, Webster, & Yim, 2012), precipitation in adjunct regions (Kosaka & Xie, 2013; J. Liu, Wang, Cane, Yim, & Lee, 2013), drying of eastern Africa (A. P. Williams & Funk, 2011) and fishery stocks related to nutrient supply brought by up-welling of cold water. Thus, understanding long-term changes in the WC are important for water resource management, ecosystems and agriculture.

The long-term temporal trend in WC strength is a popular and highly-debated subject. A weakened Walker Circulation in the twentieth century as response to greenhouse gas warming has been suggested by climate models (Held & Soden, 2006; Vecchi et al., 2006; Power & Smith, 2007; M. Collins et al., 2010; Power & Kociuba, 2011b, 2011a; Tokinaga, Xie, Deser, Kosaka, & Okumura, 2012), and it is supported by long-term sea-level pressure observations. This Walker Circulation weakening in the coupled climate models is due to anthropogenic forcing (Vecchi et al., 2006). Global temperature increases due to greenhouse gas emissions is expected to cause $\sim 7\%\ ^{\circ}C^{-1}$ increase in water vapor, based on the Clausius-Clapeyron relationship. However, increases in tropical precipitation occur at a slower rate ($\sim 3\%K^{-1}$; Held & Soden, 2006). This implies the tropical overturning circulation, which includes the WC and the easterly trade winds, slows down (Soden, Jackson, Ramaswamy, Schwarzkopf, & Huang, 2005). Furthermore, the Bjerknes feedback, a positive feedback between trade wind intensity and the zonal SST gradient, implies that the above changes will

lead to a reduced zonal SST gradient over the tropical Pacific and an El Niño-like response to warming. Such an SST pattern may lead to atmospheric dynamical changes, including enhanced divergence in the central and eastern tropical Pacific, a southeastward shift of North Pacific jet stream, and an increase in wintertime precipitation for California (Allen & Luptowitz, 2017).

In contrast to the above studies, strengthening of the Walker Circulation is simulated by atmosphere-only climate models (AMIP) driven by observed SSTs over the 20^{th} century (Meng et al., 2012; Sandeep, Stordal, Sardeshmukh, & Compo, 2014), particularly since the late second half (S. Ma & Zhou, 2016; Sohn et al., 2013). This strengthening is further supported by various observational data sets (Sohn et al., 2013; Karnauskas, Seager, Kaplan, Kushnir, & Cane, 2009; A. Solomon & Newman, 2012; L'Heureux, Lee, & Lyon, 2013). In more recent decades, a stronger WC is also supported by cooling of the central and eastern tropical Pacific SSTs (a La Niña-like pattern) (Kosaka & Xie, 2013). Observed and unprecedentedly strengthening of the easterly trade winds (England et al., 2014; M. Watanabe et al., 2014) and cooling in the eastern equatorial Pacific (Kosaka & Xie, 2013) contribute to an acceleration of tropical atmospheric overturning−opposite what coupled climate models suggest. The abnormally strong wind stress is found to be regulated by internal climate variability in the 2000s, and the east Pacific cooling is dominated by decadal variability, suppressing the effect of anthropogenic forcing (M. Watanabe et al., 2014).

Projected WC weakening in response to anthropogenic warming is found to be doubtful due to lack of consistent weakening in coupled climate models (Plesca, Grützun, &

Buehler, 2018). No conclusive projection for El Niño and the Southern Oscillation (ENSO) exists in global climate models (M. Collins et al., 2010; Cane, 2005; DiNezio, Clement, & Vecchi, 2010). Thus, whether a La Niña-like (related to WC strengthening) or a El Niño-like (related to WC weakening) tropical Pacific SST pattern will dominate this century is undetermined. In turn, the dominant driver of the Walker Circulation, whether anthropogenic emissions or natural variability, is not well constrained.

The aim of this study is to quantify the contribution of multi-decadal SST variations on recent changes in the Walker Circulation. We investigate the imprints of both anthropogenic warming and decadal internal climate variability on real world SSTs directly, instead of changes in surface air temperature forced by wind stress (M. Watanabe et al., 2014). Idealized climate model experiments are conducted with the Community Atmosphere Model version 5 (CAM5) (Neale et al., 2010) to quantify the effects of forced versus unforced SST evolution. The response to the forced signal is estimated based on coupled climate model simulations. This chapter is organized as follows: Data and Methods are described in Section 3.2; Section 3.3 discusses the results and Section 3.4 summarizes the conclusions.

3.2 Data & Methods

3.2.1 Trend

The strength of the Walker Circulation is estimated by trends of the east minus west SLP anomaly differences across the tropical Pacific, which is associated with vertical motions of the WC. The described region over the tropical Pacific is 4.74°S−4.74°N in

latitude, and 128.39°E− 151.05°E and 211.47°E−231.61°E in longitude for the western and eastern tropical Pacific edge, respectively. We also examined additional regions, which yield similar results (Table 3.1). Anomalies are calculated by removing the long term monthly mean at each grid point. Trend significance for the time series and spatial plots is examined by a standard two-tail student t-test, accounting for the influence of serial correlation by using the effective sample size, $n(1 - r_1)(1 + r_1)^{-1}$, where n is the number of years and r_1 is the lag-1 auto-correlation coefficient. Trend uncertainty will be estimated as twice the standard error, $2 \times \sigma/\sqrt{n}$, where σ is the standard deviation of the trends, and n is the number of model realizations.

3.2.2 Observation and Reanalysis Data

Observation-based WC trends are estimated from the Met Office Hadley Centre's Sea Level Pressure version 2 (HadSLP2; Allan & Ansell, 2006), NCEP/National Center for Atmospheric Research Reanalysis (R1; Kalnay et al., 1996), ERA-Interim (Dee et al., 2011), MERRA2 (Gelaro, McCarty, Suárez, et al., 2017), CFSR (Saha et al., 2010) and Japan Reanalysis (JRA-55; Kobayashi et al., 2015). WC trends are estimated from 1979 to 2014, except JRA-55 which ends in 2013, and MERRA2 which starts in 1980. Since the spatial resolution varies among our data sets, we unify all to a $1.9° \times 2.5°$ resolution−same as the Community Atmosphere Model version 5 (CAM5) model−before commencement of the analysis. Observed SST trends are estimated from Hadley Centre SST data set version 3 (HadSST3; Kennedy, Rayner, Smith, Parker, & Saunby, 2011a, 2011b).

Monthly mean data from the Coupled Model Intercomparison Projection version 5 (CMIP5; Taylor, Stouffer, & Meehl, 2012a) archive is utilized, including 22 atmosphere-

Table 3.1: **1979-2014 Walker Circulation trends based on alternative western and eastern tropical Pacific regions.** All regions indicate an intensification of the Walker Circulation. Bold coordinates indicates the region emphasized in this study. All trends are significant at 99% confidence level, based on a standard t−test. Trend units are hPa century^{-1}.

Latitude	Western Tropical Pacific	Eastern Tropical Pacific	Trend
	128.39°E-151.05°E	**211.47°E-231.61°E**	6.08
	123.36°E-156.08°E	206.43°E-236.64°E	6.02
4.74°S-4.74°N	118.32°E-161.12°E	201.40°E-241.68°E	5.93
	113.29°E-151.05°E	211.47°E-246.71°E	5.80
	108.25°E-151.05°E	211.47°E-251.75°E	5.51
	128.39°E-151.05°E	211.47°E-231.61°E	5.59
	123.36°E-156.08°E	206.43°E-236.64°E	5.48
10.42°S-10.42°N	118.32°E-161.12°E	201.40°E-241.68°E	5.33
	113.29°E-151.05°E	211.47°E-246.71°E	5.32
	108.25∘E-151.05°E	211.47°E-251.75°E	5.13

only (AMIP) models and coupled ocean atmosphere (CMIP) models. Table 3.2 lists the 22 models used in this study. Both CMIP and AMIP simulations use identical external forcings, including temporal evolution of greenhouse gases, anthropogenic aerosols, solar insolation and volcanic aerosols. AMIP simulations are driven by the observed evolution of SSTs and sea ice. CMIP5 simulations nominally end in 2005. Representative Concentration Pathway 4.5 (RCP4.5) is used to extend the CMIP5 simulations through 2008. Only models from the CMIP5 archive that contain both AMIP and CMIP simulations are utilized. Furthermore, the same number of realizations for each AMIP and CMIP experiment is used. Since most AMIP simulations end in 2008, our analysis of CMIP5 simulations is focused on the 1979-2008 time period. Other analyses, however, are based on a longer time period, from 1979-2014.

Analogous sets of AMIP-type simulations are conducted with CAM5. CAM5 AMIP simulations are based on 10 ensemble members with varying initial conditions, by

Table 3.2: **Coupled Model Intercomparison Project version 5 models (CMIP5) and number of simulations used for each experiment.** Experiments include coupled atmosphere simulations (CMIP) and atmosphere-only (AMIP) simulations. Both are driven by identical forcing, including greenhouse gases, aerosols, ozone solar variability and volcanic eruptions. AMIP simulations are also driven by the real world evolution of sea surface temperatures and sea ice. All simulations span 1979-2008.

Institution	Model	CMIP	AMIP
CSIRO and Bureau of Meterology	ACCESS1.0	1	1
	ACCESS1.3	1	1
Beijing Climate Center	BCC-CSM1.1	3	3
	BCC-CSM1.1(m)	3	3
GCESS, Beijing Normal University	BNU-ESM	1	1
National Center for Atmospheric Research	CCSM4	5	5
CNRM/CERFACS	CNRM-CM5	1	1
CSIRO, Industrial Research Organization and QCCCE	CSIRO-Mk3.6.0	10	10
EC-EARTH consortium	EC-EARTH	1	1
LASG, IAP, Chinese Academy of Sciences and CESS	FGPALS-g2	1	1
NASA Goddard Institute for Space Studies	GISS-E2-R	6	6
Institute for Numerical Mathematics	INM-CM4.0	1	1
Institut Pierre-Simon Laplace	IPSL-CM5A-LR	4	4
	IPSL-CM5A-MR	1	1
	IPSL-CM5B-LR	1	1
JAMEST, AORI, and NIES	MIROC-ESM	1	1
	MIROC5	2	2
Max Planck Institute for Meteorology	MP1-ESM-LR	3	3
	MP1-ESM-ME	3	3
Meteorological Research Institute	MRI-CGCM3	2	2
Norwegian Climate Center	NorESM1-M	3	3

applying a random surface temperature perturbation. Simulations are initiating from 1970 to allow initial state of the atmosphere to diverge as a result of internal atmospheric variability. Ensemble members are extended to 2014 using RCP 4.5 time varying forcing, including solar radiation, estimated concentrations of greenhouse gases, volcanic aerosols, ozone and primary emissions of sulfur dioxide and black and organic carbon.

We also analyze 40 CMIP simulations from Community Earth System Model Large Ensemble Project (CESM LENS) models (Kay et al., 2015). CESM LENS simulations, referred as CAM5 CMIP in this manuscript, were downloaded from the Earth System Grid at the National Center fro Atmospheric Research (NCAR). Theses simulations were extended from 2005 to 2014 using RCP 8.5−the only available future pathway. CAM5 CMIP simulations feature identical forcing, but different initial conditions. Both CAM5 AMIP and CMIP simulations allow assessment of natural climate variability. Note that ensemble mean averages out the internal variability and hence represents the forced signal.

3.2.3 Idealized CAM5 Experiment

It is assumed that the observed SST evolution is composed of two components, including an unforced (UFSST) and forced (FSST) component. The unforced component is due to the natural variability of SSTs (e.g., El Niño and La Niña), and the forced component is due to external forcing (i.e., anthropogenic emissions). The forced component is obtained from the ensemble mean SSTs from the CMIP5 coupled ocean atmosphere models. Solar radiation and volcanic aerosols are considered as external factors. The unforced component is obtained by removing the forced SST trend from the observed SST at each grid point using linear regression. To do so, the trend of monthly FSST is estimated by taking a least

squares trend at each grid point, multiplying this slope by the corresponding year, and then subtracting this SST estimate from the observed SST. We perform CAM5 experiments using forced (CAM5 FSST) and unforced (CAM5 UFSST) SST fields. CAM5 FSST and CAM5 UFSST show the role of SST evolution on variability of the Walker Circulation, including the relative roles of externally forced and unforced SST evolution.

The occurrence of a more El Niño-like mean state of the tropical Pacific is a likely response to anthropogenic warming (Cai et al., 2015). However, coupled climate models may be deficient in their tropical response to warming (Karnauskas et al., 2009; Fyfe, Gillett, & Zwiers, 2013). Uncertainty lies in whether the tropical response will resemble a more La Niña-like SST pattern (as has occurred since ∼1979), or a more El Niño-like pattern. Changes in the tropical Pacific mean state depend on the relative strength of amplifying effects associated with ocean energy balance (including an increase in SST gradient and stronger ENSO events), and damping effects associated with atmospheric dynamics (including weakening zonal winds, upwelling and weaker ENSO events; M. Collins et al., 2010).

3.3 Results

Based on HadSLP2 observations, the WC strengthens from 1979-2014, which is statistically significant at the 99% confidence level based on a standard t-test (Figure 3.1A). Observed eastern tropical Pacific SLP yields positive trends, while negative trends prevail over the western tropical Pacific (Figure 3.1B). Over the same time period, a La Niña-like SST trend pattern exists (Figure 3.1C). This result is robust to the eastern and western

tropical Pacific box definitions (Table 3.1). Observations over a shorter period, 1979-2008, also depicts WC strengthening, but not as strong as that over longer period (4.65 versus 6.08 hPa $century^{-1}$; Table 3.3 & Figure 3.2). These observational results are further supported by reanalyses data, which also show intensification of the WC (Figure 3.1), though with magnitude differences. Furthermore, only trends based on R1 and JRA55 are significant. The spatial pattern of observed SLP trends is also captured by reanalyses (not shown). A shorter time period (1979-2008) also yields positive trends of the SLP gradient, with only HadSLP2 yielding a significant WC intensification (Table 3.3). 1979-2008 trends of the SLP gradient are generally weaker than those over the longer, 1979-2014 time period (except MERRA2). This is likely related to a strong La Niña event during late 2007 to early 2008, leading to a larger SLP gradient over the tropical Pacific.

Figure 3.3 shows the 1979-2014 ensemble mean simulated change in the strength of the WC in CAM5 AMIP and CESM LENS (referred as CAM5 CMIP) simulations. CAM5 AMIP yields a positive trend but weaker than that from observation, with ensemble mean of 2.3 hPa $century^{-1}$ significant at the 90% confidence level (Table 3.3 & Figure 3.3A). Trends of SLP gradient estimated from CAM5 AMIP realizations ranges from 2.15 to 2.59 hPa century^{-1} (Figure 3.4A). The corresponding SLP trend pattern is also consistent with that from observation and reanalyses (Figure 3.3C). In contrast, CAM5 CMIP simulations yield a general increase in SLP over most of the tropical Pacific (Figure 3.3D), and a weakly negative ensemble mean WC trend (Figure 3.3B). CAM5 CMIP also yields a larger range of trends (-3 to 1 hPa century^{-1}) than CAM5 AMIP, and there is no consensus on the

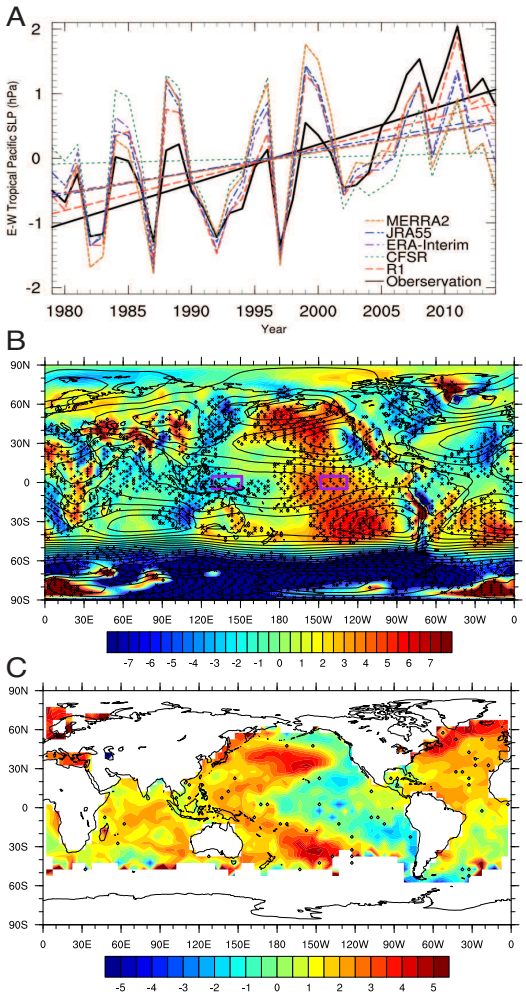

Figure 3.1: **1979-2014 trends of tropical Pacific SLP gradient, SST and SLP pattern from observation and reanalyses.** (A) Time series of east (4.74°S−4.74°N, 211.47°E−231.61°E) minus west (4.74°S−4.74°N, 128.39°E− 151.05°E) tropical Pacific SLP. Black lines represent HadSLP2; red lines represent NCEP/NCAR (R1) reanalysis; green lines represent CFSR reanalysis; purple lines represent ERA-Interim reanalysis; blue lines represent JRA55 (which ends in 2013) and orange lines represent MERRA2 reanalysis (which starts in 1980). Also included is the least squares linear trend corresponding to each data set. (B) Observed SLP trend (hPa $century^{-1}$) pattern. Thin black lines represent climatological SLP (hPa). Purple boxes denote the western and eastern tropical Pacific regions used in (A). (C) Observed SST trend ($°C\ century^{-1}$) pattern based on HadSST3. Symbols in (B, C) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation. Warm (cold) colors represent positive (negative) trends.

Table 3.3: **Walker Circulation trends based on multiple data sets and model simulations.** Trends in 1979-2014 and 1979-2008 are based on observations (HadSLP2), reanalyses including R1(NCAR/NCEP), JRA55, CFSR, ERA-Interim and MERRA2, and atmosphere-only (AMIP) and coupled ocean atmosphere (CMIP) simulations from CAM5 and CMIP5 (only available in 1979-2008). Idealized CAM5 unforced (UFFST) and forced (FSST) trends are also included in both periods. Symbols represent trend significance at the 90 % (+), 95% (diamond) or 99% (star) confidence level, based on a standard t-test. Trend units are hPa century^{-1}. Fields with "-" represents unavailable because not all CMIP/AMIP simulations from CMIP5 archive extend through 2014.

Data	1979-2014	1979-2008
HadSLP2	$6.08^{\star}$	$4.65^{\star}$
R1	$4.89^{\star}$	3.8
JRA55	$3.51^{\star}$	3.35
CFSR	0.50	0.38
ERA-Interim	3.21	3.00
MERRA2	3.20	4.87
CMIP5 AMIP	-	3.1
CMIP5 CMIP	-	-0.15
CAM5 AMIP	2.3^{+}	2.32
CAM5 CMIP (CESM LENS)	-0.32	-0.16
CAM5 UFSST	$2.59^{\diamond}$	2.54
CAM5 FSST	-0.21	-0.09

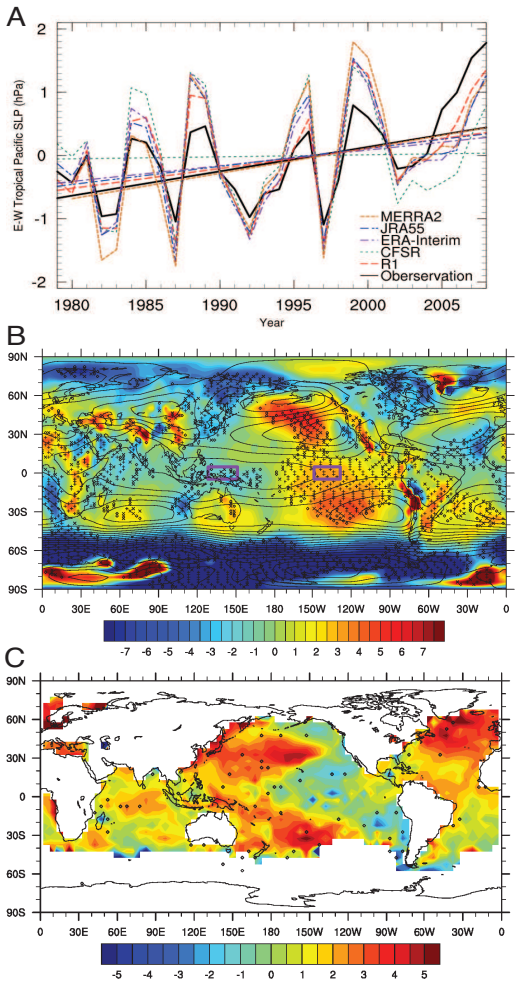

Figure 3.2: **1979-2008 trends of tropical Pacific SLP gradient, SST and SLP pattern from observation and reanalyses.** (A) Time series of east (4.74°S−4.74°N, 211.47°E−231.61°E) minus west (4.74°S−4.74°N, 128.39°E− 151.05°E) tropical Pacific SLP. Black lines represent HadSLP2; red lines represent NCEP/NCAR (R1) reanalysis; green lines represent CFSR reanalysis; purple lines represent ERA-Interim reanalysis; blue lines represent JRA55 (which ends in 2013) and orange lines represent MERRA2 reanalysis (which starts in 1980). Also included is the least squares linear trend corresponding to each data set. (B) Observed SLP trend (hPa $century^{-1}$) pattern. Thin black lines represent climatological SLP (hPa). Purple boxes denote the western and eastern tropical Pacific regions used in (A). (C) Observed SST trend ($°C\ century^{-1}$) pattern based on HadSST3. Symbols in (B, C) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation. Warm (cold) colors represent positive (negative) trends.

sign of the SLP gradient trend derived from CAM5 CMIP simulations (Figure 3.4A). About two thirds of the CAM5 CMIP simulations yield WC weakening, and one third yield WC strengthening. Individual trends from CAM5 CMIP realizations largely underestimate those from observation and reanalyses, except CFSR. Thus, we find robust results across CAM5 AMIP simulations that are consistent with observations. This implies the real-world evolution of SSTs is important to the observed strengthening of the WC. HadSLP2 observation shows that WC weakening dominants the bulk of the 20^{th} century (1920-2000; Figure 3.5). This weakening is captured by the CAM5 CMIP (CESM LENS) simulations, with a robust ensemble mean trend at -0.25 hPa $century^{-1}$. This is in an agreement with previous work arguing the long-term WC weakening is a greenhouse gas (GHG) forced response (Vecchi et al., 2006; Power & Kociuba, 2011b). This also implies that the effect of GHG warming is not the major cause of recent WC strengthening, which may be counteracted by the natural variability.

We also conducted the same analysis with CAM5 simulations over the shorter period−1979 to 2008, and similar results are obtained (Figure 3.6). WC strengthening is simulated from CAM5 AMIP, with an ensemble mean trend of 2.32 hPa century^{-1}, nearly the same as that over the longer period (Table 3.3 & Figure 3.6A). The SLP trend patterns in CAM5 AMIP also resembles a La Niña-like pattern, consistent with that from observations (Figure 3.6C). CAM5 CMIP, however, yields a much smaller and negative ensemble mean trend at -0.16 hPa century^{-1}(Table 3.3, Figure 3.6B & D). A wider distribution of CAM5 CMIP5 trends also exists, ranging from -2.75 to 2.5 hPa century^{-1}. 45% of CAM5 CMIP

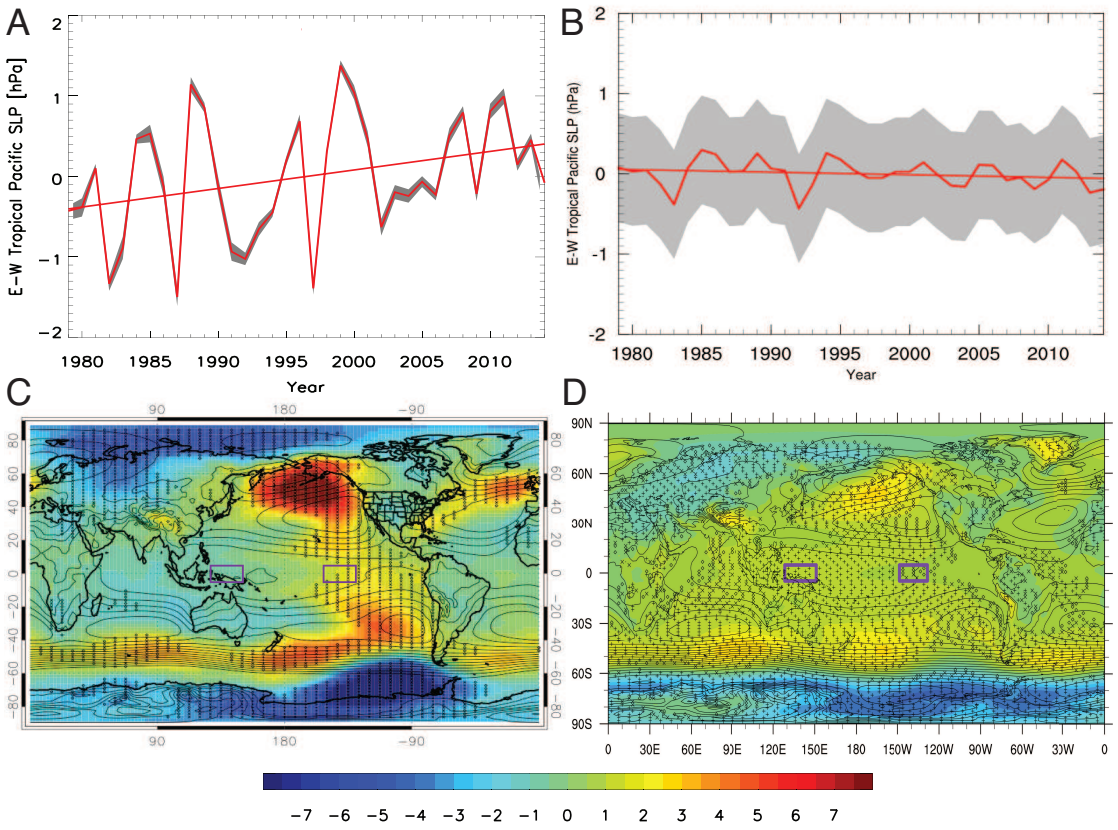

Figure 3.3: **1979-2014 CAM5 AMIP and CAM5 CMIP (CESM LENS) ensemble mean SLP trends.** Left panel (A, C) CAM5 AMIP and right panel (B, D) CAM5 CMIP. (A, B) time series of east minus west tropical Pacific SLP gradient. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set. (C, D) SLP trend (hPa $century^{-1}$) pattern. Purple boxes denote the western and eastern tropical Pacific regions used in (A, B). Symbols in (C, D) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation, and thin black lines represent climatological SLP (hPa). Warm (cold) colors represent positive (negative) SLP trends.

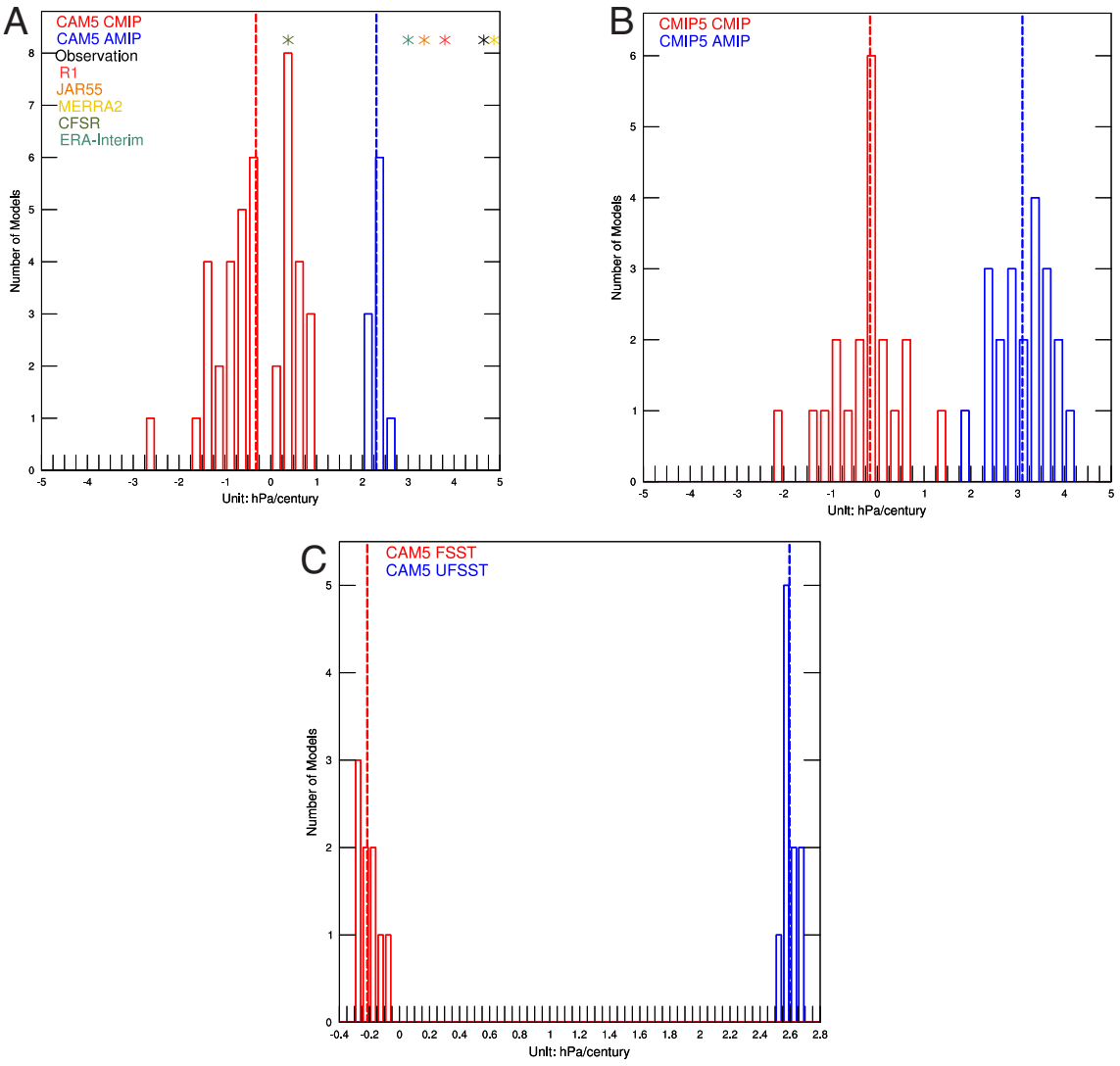

Figure 3.4: **Histogram of tropical Pacific SLP gradient trends.** Trends based on (A) 1979-2014 CAM5 simulations, (B) 1979-2008 CMIP5 simulations and (C) 1979-2014 CAM5 UFSST and FFST simulations. Blue bars represent trends from CMIP simulations in (A, B) and unforced SST component (UFFST) in (C). Red bars represent trends from AMIP simulations in (A, B) and forced SST component (FSST) in (C). Dash lines represent trends from ensemble mean of AMIP/UFSST (blue) and CMIP/FSST (red) simulations. In (A), stars represent trends from observation (black), NCEP/NCAR (R1) (red), JRA55 (orange), MERRA2 (gold), CFSR (olivegreen) and ERA-Interim (light green) reanalysis. Unis is hPa $century^{-1}$.

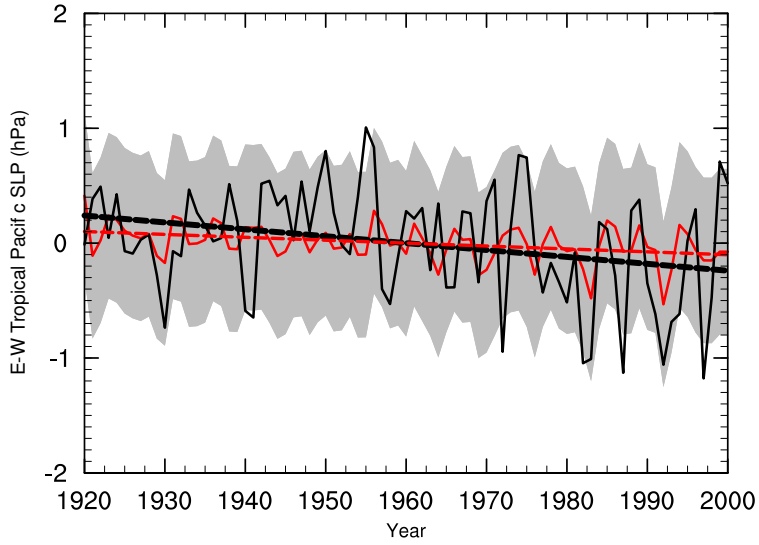

Figure 3.5: **1920-2000 trends of tropical Pacific SLP gradient from observation and CAM5 CMIP (CESM LENS) ensemble mean.** Time series of east minus west tropical Pacific SLP. Black lines represent HadSLP2; red lines represent CAM5 CMIP (CESM LENS) ensemble mean. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set.

realizations yield positive trends and 55% yield negative trends (Figure 3.7A). Thus, similar to observations, CAM5 simulations yield similar results over both time periods, although WC strengthening is less statistically significant over the 1979-2008 time period.

Results based on 1979-2008 CAM5 simulations are further supported by those from CMIP5 (Figure 3.8). CMIP5 AMIP yields a positive ensemble mean trend nearly as large as that from observations, at 3.1 hPa century^{-1}, although it is not significant at the 90% confidence level (Table 3.3 , Figure 3.8A & C). All CMIP5 AMIP simulations yield WC strengthening, ranging from 1.75 to 4.25 hPa century^{-1}, which agrees well with the observed trends (Figure 3.8B). In contrast, a small negative CMIP5 CMIP ensemble mean trend is obtained at -0.15 hPa century^{-1} (Table 3.3 , Figure 3.8B & D) and 73% of these realizations depict WC weakening (Figure 3.8B). Therefore, similar to CAM5 AMIP, CMIP5 AMIP simulates WC strengthening that agrees with observations, while CAM5 CMIP and CMIP5 CMIP do not. Robust results from both CAM5 and CMIP5 simulations over the shorter time period provide additional evidence that WC intensification is related to real-world SST variations, rather than anthropogenic emissions.

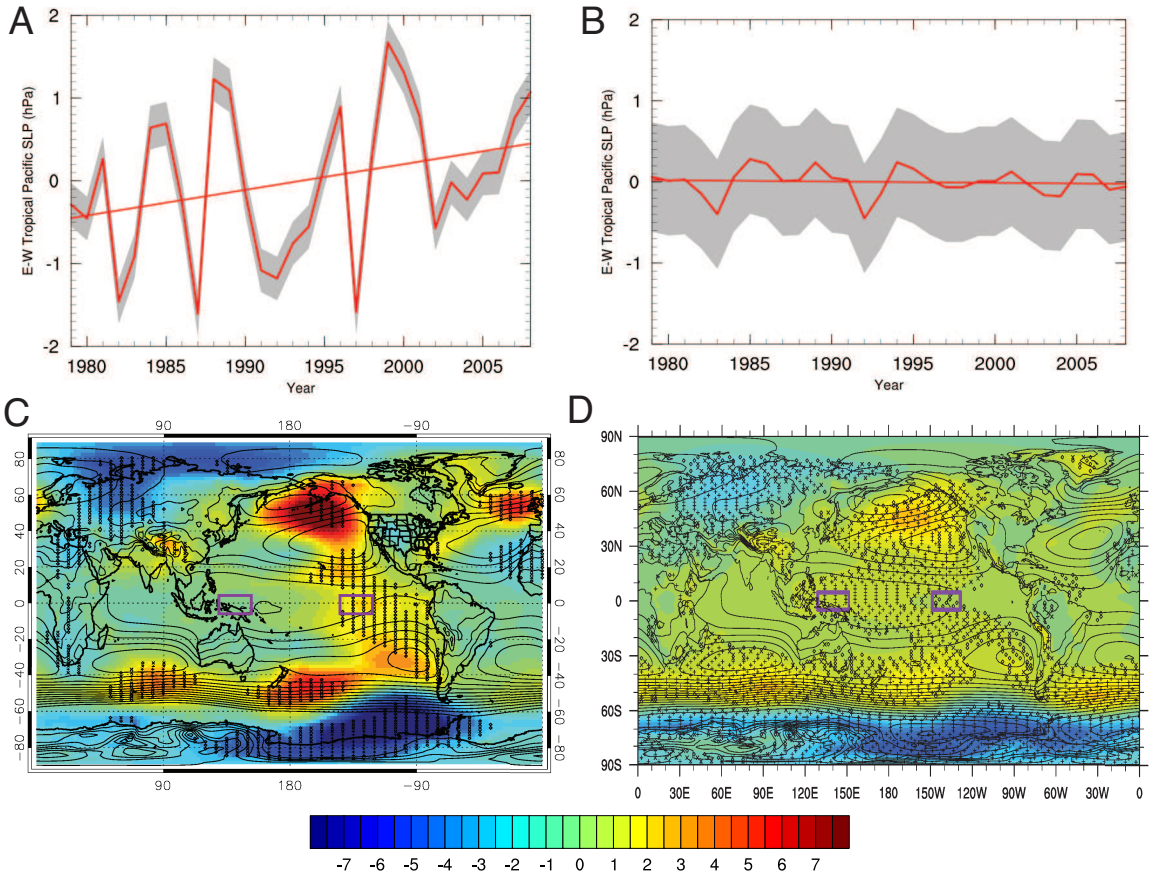

Figure 3.6: **1979-2008 CAM5 AMIP and CAM5 CMIP (CESM LENS) ensemble mean SLP trends.** Left panel (A, C) CAM5 AMIP and right panel (B, D) CAM5 CMIP. (A, B) time series of east minus west tropical Pacific SLP gradient. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set. (C, D) SLP trend (hPa $century^{-1}$) pattern. Purple boxes denote the western and eastern tropical Pacific regions used in (A, B). Symbols in (C, D) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation, and thin black lines represent climatological SLP (hPa). Warm (cold) colors represent positive (negative) SLP trends.

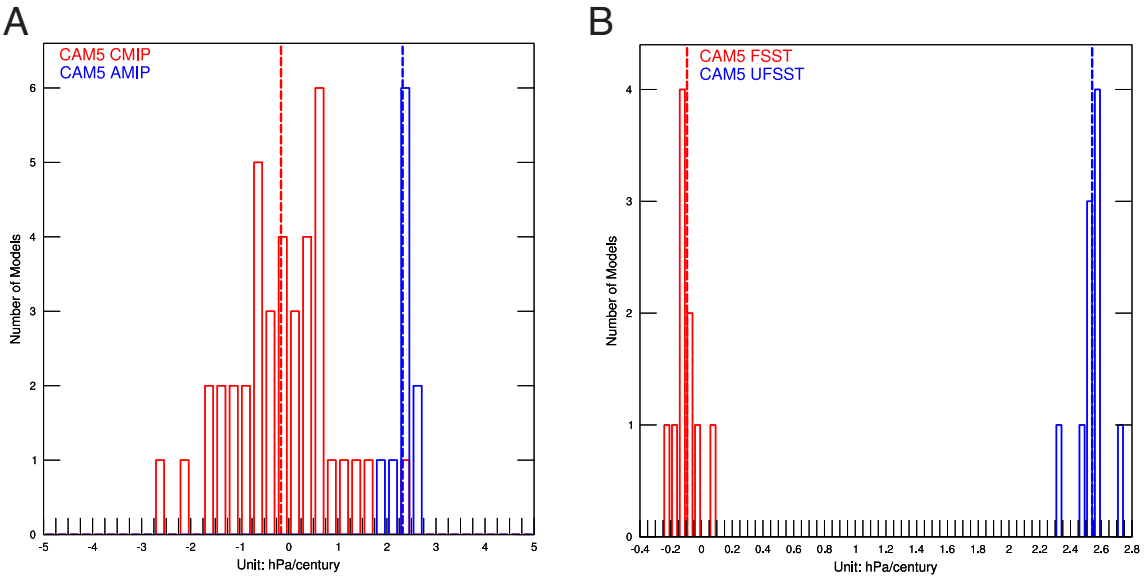

Figure 3.7: **1979-2008 Histogram of tropical Pacific SLP gradient trends.** 1979-2008 trends from (A) CAM5 simulations and (B) CAM5 UFSST and FFST simulations. Blue bars represent trends from CMIP simulations in (A) and CAM5 simulations driven by the unforced SST component (UFFST) in (B). Red bars represent trends from AMIP simulations in (A) and CAM5 driven by the forced SST component (FSST) in (B). Dashed lines represents the corresponding ensemble mean trend. Unit is hPa $century^{-1}$.

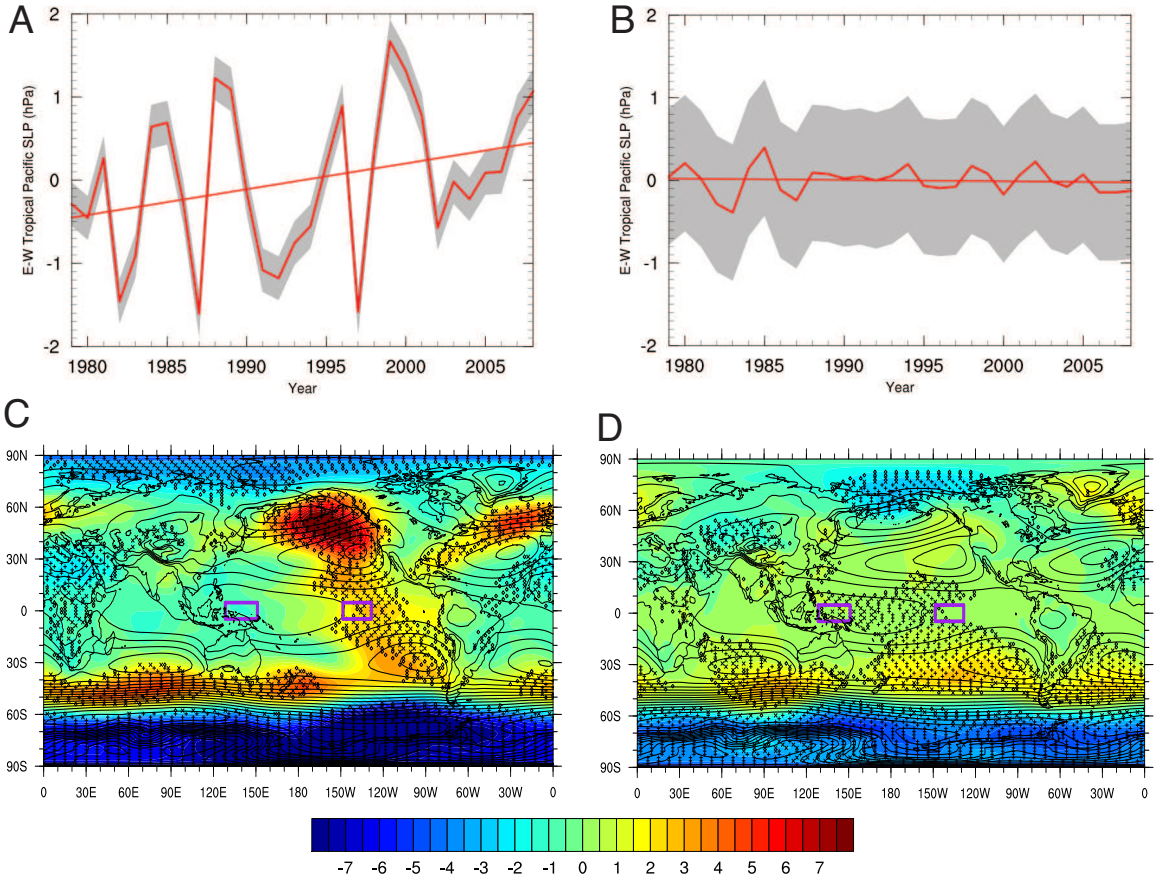

Figure 3.8: **1979-2008 CMIP5 AMIP and CMIP ensemble mean SLP trends.** Left panel (A, C) CMIP5 AMIP and right panel (B, D) CMIP5 CMIP. (A, B) time series of east minus west tropical Pacific SLP for. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set. (C, D) SLP trend (hPa $century^{-1}$) pattern. Purple boxes denote the western and eastern tropical Pacific regions used in (A, B). Symbols in (C, D) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation, and thin black lines represent climatological SLP (hPa). Warm (cold) colors represent positive (negative) SLP trends.

The real-world evolution of SSTs is composed of an unforced and a forced component. The unforced SSTs are controlled by internal variability of SSTs, such as El Niño Southern Oscillation (ENSO). The forced SSTs are driven by external forcing due to anthropogenic emissions, volcanic eruptions and solar radiation variations. We assume that the observed SSTs can be decomposed into a forced (FSST) and unforced (UFSST) component. 1979-2014 SLP changes based on CAM5 using the unforced and forced SSTs and are shown in Figure 3.9 (Figure 3.10 shows 1979-2008). The CAM5 UFSST ensemble mean yields a positive SLP gradient trend of $\sim$2.5 hPa century^{-1} over both time periods (Table 3.3). From 1979-2014, this trend is significant at the 95% significance level (Table 3.3). Trends simulated from CAM5 UFSST fall in a narrow range of 2.34 to 2.7 hPa century^{-1} for both periods (Figure 3.4C & Figure 3.7B). The CAM5 UFSST ensemble mean trend pattern resembles the observations over the tropical Pacific, with positive trends dominating in the eastern tropical Pacific and negative trends prevailing in the western tropical Pacific (Figure 3.10C & Figure 3.10C). In contrast, the CAM5 FSST ensemble mean yields weakly negative and insignificant SLP gradient trends of -0.21 hPa century^{-1} for 1979-2014 and -0.09 hPa century^{-1} for 1979-2008 (Table 3.3 , Figure 3.10C & Figure 3.10C). All CAM5 FSST SLP gradient trends yield WC strengthening and exhibit a relatively wider distribution than those based on CAM5 UFSST, with a range of -0.3 to -0.09 hPa century^{-1} for both periods (Figure 3.4C & Figure 3.7B). The corresponding 1979-2014 SST trend pattern from these simulations were examined by (Figure 3.11; Allen & Kovilakam, 2017b). A cold ENSO-like SST trend pattern is associated with UFSST (Figure 3.11B), consistent with the observed SST trend pattern (Figure 2.1C & Figure 3.11A). However, a spatially uniform

warming pattern is produced by FSST (Figure 3.11C). Thus, CAM5 UFSST captures the observed WC intensification. It suggests the recent intensification of the Walker Circulation is largely due to natural SST variations, primarily due to a La Niña-like SST pattern.

3.4 Summary

Through observations, multiple reanalyses and climate model simulations, we have showed the importance of real-world SST evolution to recent intensification of the Walker Circulation. Since 1979, observations and reanalyses yield positive SLP gradient trends over the equatorial Pacific, suggesting an intensified Walker Circulation. CAM5 atmosphere-only (AMIP) simulations also simulate significant strengthening of the SLP gradient over tropical Pacific, while CAM5 coupled atmosphere-ocean (CMIP) models yield relatively small, negative ensemble mean trends. A large range of of trends−including both positive and negative−is also found with CAM5 CMIP simulations. CMP5 AMIP and CMIP5 CMIP simulations yield similar results over a shorter time period (1979-2008), with CMIP5 AMIP simulations capturing the observed strengthening of the Walker Circulation, though not statistically significant. These results imply that the real-world evolution of SSTs is important for the observed intensification of the Walker Circulation. Note that the observed WC trend is not captured by either AMIP or CMIP simulations. Furthermore, the range of WC trends from AMIP simulations does not overlap that from CMIP. This suggests that models, particularly the coupled ocean atmosphere models, tend to underestimate the strength of tropical interdecadal SST variability. Given a sufficiently large ensemble of CMIP simulations, the observed value should at least be within the tail of the simulated distribution,

but this is not the case. Therefore, CMIP may be able to accurately simulate the long-term mean response to anthropogenic perturbations, but may not be able to simulate the strength of the interdecadal variability in tropical SST and WC intensity. Underestimating decadal variability may produce a response with a narrow spread around the mean value and a potential underestimation of extremes at the regional scale.

Idealized experiments were conducted with CAM5, to isolate the forced and unforced component of real-world SST revolution. Robust results are obtained from simulations driven by the unforced SSTs, where a significant strengthening of the Walker Circulation is obtained. UFSST resembles the observed SSTs, and in particular, the La Niña-like SST pattern over the same time period (Allen & Kovilakam, 2017b). This further implies the unforced component of real-world SST evolution is the dominant driver of the recent intensification of the Walker Circulation. Since 1979, natural SST variability has played a larger role in long-term evolution of the WC than anthropogenic emissions, contrary to that over the entire 20^{th} century.

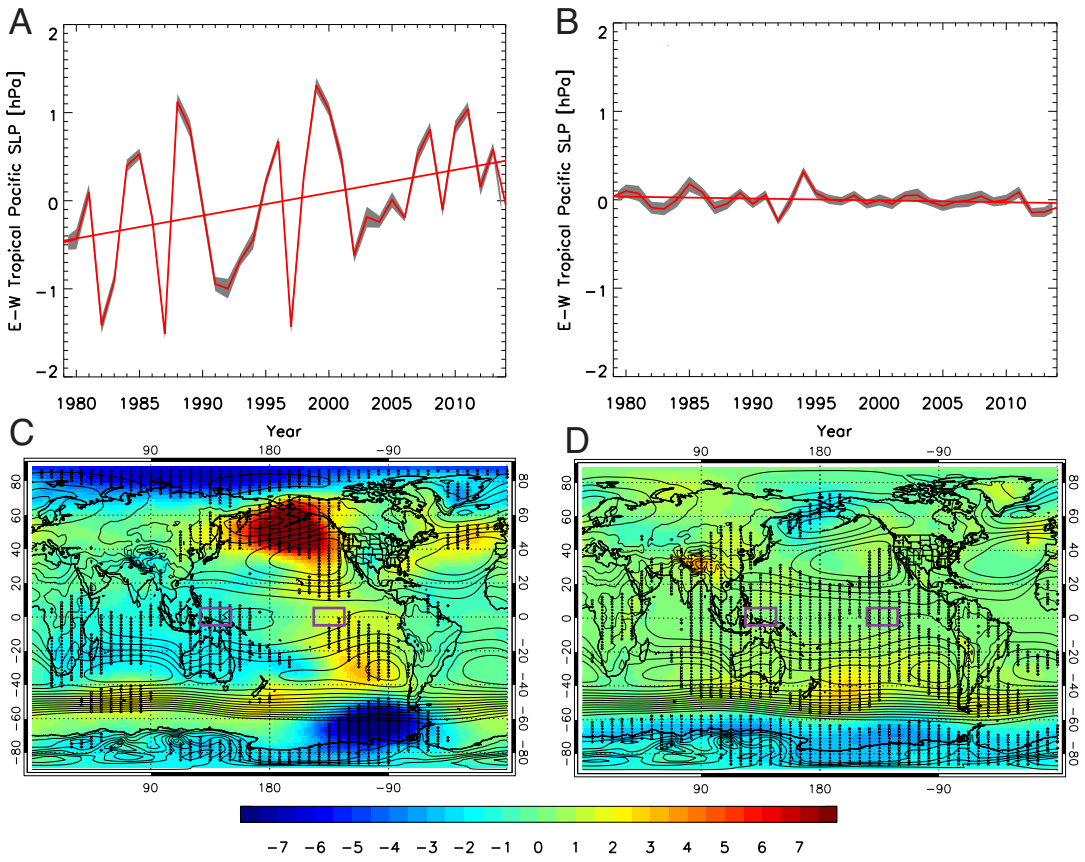

Figure 3.9: **1979-2014 ensemble mean SLP trends from idealized AMIP-type experiment with CAM5 models.** Left panel (A, C) CAM5 UFSST and right panel (B ,D) CAM5 FSST. (A, B) time series of east minus west tropical Pacific SLP, including the linear trend. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set. (C, D) SLP trend (hPa $century^{-1}$) pattern. Purple boxes denote the western and eastern tropical Pacific regions used in (A, B). Symbols in (C, D) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation, and thin black lines represent climatological SLP (hPa). Warm (cold) colors represent positive (negative) SLP trends.

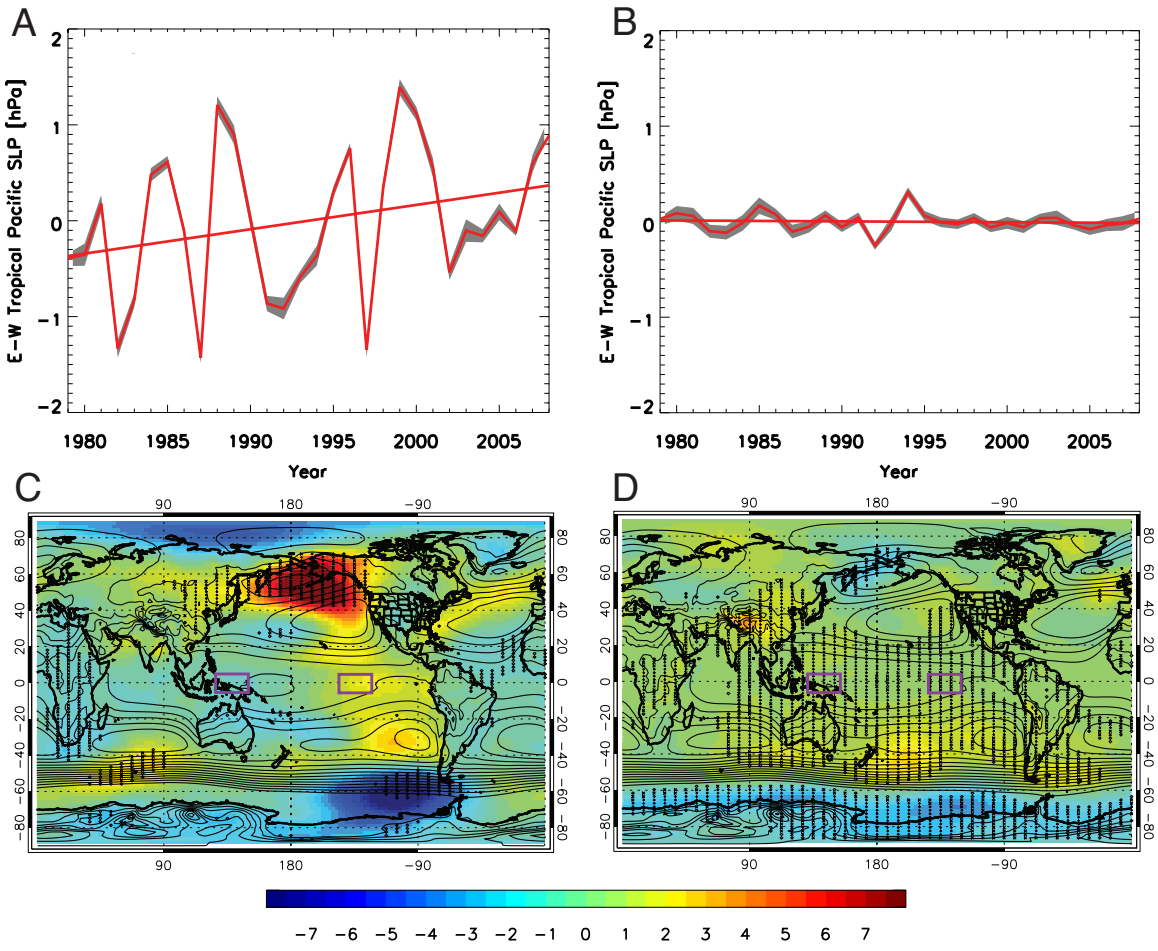

Figure 3.10: **1979-2008 ensemble mean SLP trends from idealized AMIP-type experiment with CAM5.** Left panel (A, C) CAM5 UFSST and right panel (B, D) CAM5 FSST. (A, B) time series of east minus west tropical Pacific SLP, including the linear trend. Gray shading represents uncertainty across realizations, estimated as twice the standard error. Also included is the least squares linear trend corresponding to each data set. (C, D) SLP trend (hPa $century^{-1}$) pattern. Purple boxes denote the western and eastern tropical Pacific regions used in (A, B). Symbols in (C, D) represent trend significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation, and thin black lines represent climatological SLP (hPa). Warm (cold) colors represent positive (negative) SLP trends.

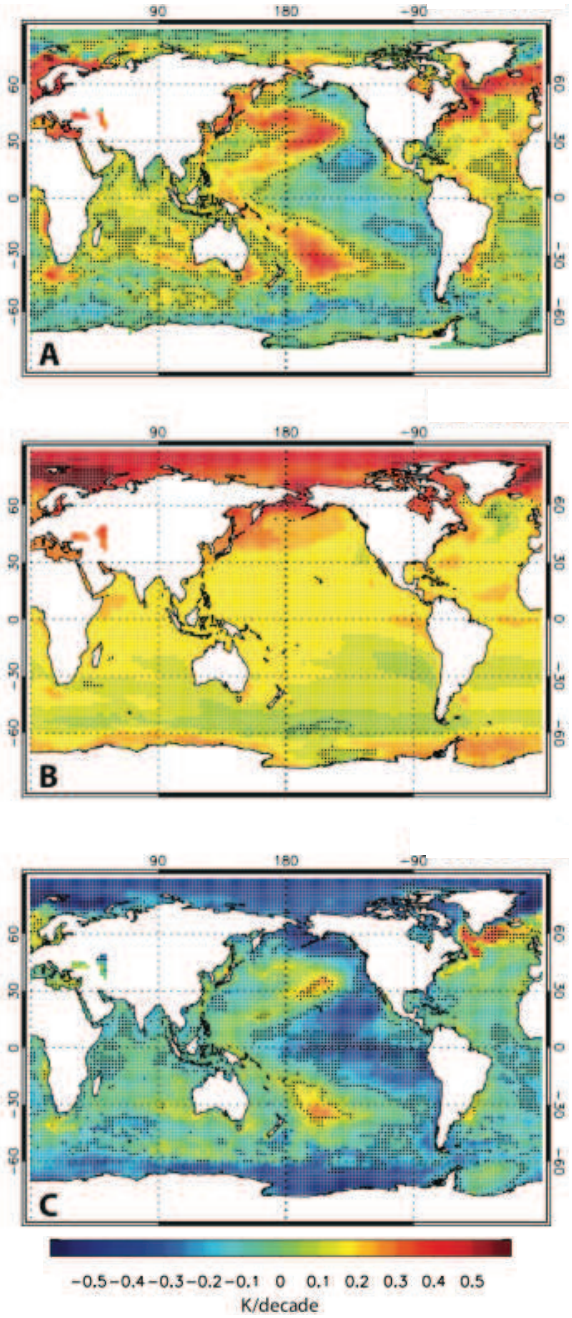

Figure 3.11: **1979-2014 obseved SST trends.** (A) Hadley Centre observed SST trends; (B) the forced component, estimated from the ensemble mean of the CMIP5 20^{th} century all forcing experiments, combined with RCP4.5; (C) the unforced component, which is estimated by removing the forced SST trend from the observed SST at each grid point. Trend symbols represent significance at the 90 % (diamond), 95% (X) or 99% (+) confidence level, accounting for autocorrelation. Units are K decade^{-1}.

Chapter 4

An Implicit Air Quality Bias due to the State of Pristine Aerosol

4.1 Introduction

Clean air is a key trans-boundary resource needed to sustain ecosystems and human societies, yet air pollution remains a common threat to human health (Shaddick, Thomas, Mudu, Ruggeri, & Gumy, 2020). Epidemiological studies have demonstrated that exposure to air pollutants over both short and long time scales is associated with disease, including cardiovascular disease and lung cancer (Brunekreef & Holgate, 2002; Pope III & Dockery, 2006; Apte, Marshall, Cohen, & Brauer, 2015; World Health Organization, 2016; Burnett et al., 2018; Landrigan et al., 2018). Generally, air quality is quantified by measurements and monitoring of the surface concentrations of various air pollutants, including particulate matter (PM), ozone, nitrogen oxides and sulfur dioxide. The World Health Organization

(WHO) has established air quality guidelines based on such ground-level air pollutant monitoring coupled with epidemiological studies (World Health Organization, 2006). The end results are spatially uniform thresholds recommended for specific air pollutants, like $PM_{2.5}$-fine particulate matter with diameter less than 2.5 µm, for which the WHO recommends an annual mean $PM_{2.5} \leq 10\,\mu g\ m^{-3}$ (World Health Organization, 2006). Moreover, many countries look to the WHO recommendations in order to establish their own national standards and goals for air quality. For example, the European Environment Agency (EEA) recommends an annual mean $PM_{2.5}$ threshold of $25\,\mu g\ m^{-3}$ (European Council Directive, 2008), the US Environmental Protection Agency (EPA) uses a threshold of $12\,\mu g\ m^{-3}$ (US EPA, 2016), and China uses a threshold of $35\,\mu g\ m^{-3}$ everywhere except locations that qualify for special protections, such as national parks (Ambient Air Quality Standards, GB 3095-2012, 2012). Thus, the more stringent WHO recommendations serve as a benchmark for improving air quality via anthropogenic emission reductions, regulations, and abatement measures. However, none of these air quality metrics directly accounts for the background of naturally occurring pollutants.

Much of the science (i.e., ground monitoring and epidemiological studies) behind establishing air quality metrics (i.e., WHO air quality guidelines) are or were exclusively conducted in economically developed regions where natural aerosol levels are generally low. Since industrialization the composition of atmospheric aerosols and their sources have changed significantly, particularly over economically developed regions, from the dominance of natural (pristine) aerosols during the pre-industrial era (with minimal contribution from early human activities) to a larger contribution from anthropogenic aerosols during the in-

dustrial era (Hoesly et al., 2018). Now, over economically developed regions, $PM_{2.5}$ and other pollutants largely originate from combustion sources, including from the transportation (e.g., ships, trucks, cars) and energy sectors (e.g., coal and wood burning; World Health Organization, 2006). While this implicitly suggests that human health risks from polluted air are primarily associated with anthropogenic aerosols, air quality guidelines generally discount any contributions from natural aerosols, including dust and sea salt. However, air pollution can be or is worse over economically developing regions where natural aerosol levels can be high (e.g., Africa; Petkova et al., 2013), and anthropogenic pollution may convolve to disguise the contribution of pristine aerosol to air quality. In fact, natural aerosol, like dust, is the most dominant aerosol species in the atmosphere, constituting 70% of the global aerosol mass burden and 25% of the aerosol optical depth (AOD; Kinne et al., 2006). Given that current air quality guidelines mainly target anthropogenic emissions, natural contributions to air pollution may remain overlooked.

The impacts of human emissions since industrialization on the Earth system have been well recognized in climate research. Most prior work has focused on assessing changes in anthropogenic aerosol radiative forcing over the industrial era by quantifying pre-industrial aerosol radiative properties (Carslaw et al., 2013; Hamilton et al., 2014; Carslaw et al., 2017; S. Solomon et al., 2007; Stocker, 2014; Watson-Parris et al., 2020). No effort has attempted to assess the impact of pristine aerosols on present-day air quality. Large community-wide modeling initiatives like the Coupled Model Intercomparison Project (CMIP) run standard pre-industrial control (PIC) simulations to model pre-industrial atmospheric conditions. Here we use model output from CMIP6 (Eyring et al., 2016) and CMIP5 (Taylor, Stouffer,

& Meehl, 2012b) to construct a view of pristine air quality, specifically $PM_{2.5}$, and assess its contribution to present-day air quality. Direct $PM_{2.5}$ observations during the pre-industrial period are unavailable. Thus, in order to assess the quality of PIC simulations, models' representations of present-day (1980-2014) $PM_{2.5}$ are examined relative to a suite of observational data sets and reanalyses. The findings offer insight into how the pristine aerosol background might inform current air quality metrics. This chapter is organized as follows: Data and Methods are described in Section 4.2; Section 4.3 discusses the results and Section 4.4 summarizes the conclusions.

4.2 Data

Constructing an understanding of pre-industrial air quality utilizing the tools of CMIP is a complex exercise that must be constrained by assessments of the models' efficacy over time periods during which observational data exists. Assessing the impact on populations involves another layer of geographically tagged data. Thus, in advance of detailing the air quality aspects, short descriptions of the data sources and utilization methods are given with significant details added to the Supplementary Materials for interested readers.

4.2.1 Model Data

Monthly data from the PIC simulations generated by 12 coupled ocean-atmosphere CMIP6 models are used Table 4.2 (Eyring et al., 2016). Most models include interactive tropospheric chemistry and aerosol schemes. Data at the lowest model level is extracted to represent surface conditions. The first 50 years of the PIC simulations (which is enough

to account for decadal variability) are used to generate global maps of pristine aerosol. Forcing levels of anthropogenic aerosol, greenhouse gases and solar radiation are fixed to 1850 levels. Historical (HIS) simulations from the same CMIP6 models are used to examine their ability to simulate present-day (1980-2014) $PM_{2.5}$. Anthropogenic (greenhouse gas and anthropogenic aerosol) and natural (solar irradiance and volcano activities) forcings over the historical periods are used.

Of the 12 CMIP6 models, five individually−GFDL-ESM4, MIROC-ES2L, NorESM2-LM, GISS-E2-1-G and MRI-ESM2-0, directly archive output $PM_{2.5}$ using their own model-specific methodology. They all contain black carbon (BC), sulfate (SO_4), organic matter (OM), dust (DU) and sea salt (SS). Archived $PM_{2.5}$ from GFDL-ESM4 and GISS-E2-1-G also include nitrate and ammonium.

In GFDL-ESM4 (Krasting et al., 2018; Zhao et al., 2018a, 2018b; Held et al., 2019), all aerosols are lognormally distributed, except dust and SS, which include five size bins from 0.1 to 20 µm in diameter. Emitted dust is discretized based on the size of tracers with a constant proportion in each bin (in diameter, 0.1-2 µm, 5%; 2-4 µm,15%; 4-6 µm, 30%; 6-12 µm, 27%; and 12-20 µm, 23%)(Zhao et al., 2018b). $PM_{2.5}$ is estimated as: 0.96×SOA + DU1 + 0.25×DU2 + 0.97×SO4 + SS1 + SS2 + 0.167×SS3 + 0.995×(BCPHOB + BCPHIL) + 0.96×(OMPHOB + OMPHIL) + 0.954×NO3 + 0.973×NH4, where SOA is secondary organic aerosol; PHOB refers to hydrophobic and PHIL refers to hydrophilic. Numbers after aerosol species refer to the size bin (e.g., DU1 refers to the smallest dust size bin).

Table 4.1: **CMIP6 models.**

Model	Resolution	Aerosol scheme	Chemistry scheme	References
BCC-ESM1 (ESM)	2.813° × 2.813°	Mass-based aerosol scheme. Prescribed stratospheric aerosols.	CAM-Chem (based on MOZART). Tropospheric chemistry only. Prescribed DMS emissions.	(Wu et al., 2019, 2020)
CESM2-WACCM (ESM)	0.9° × 1.25 °	Four-mode version of the Modal Aerosol Module (MAM4) with a more comprehensive SOA approach using the Volatility Basis Set scheme (SOA-VBS)	MOZART-based TSMLT1, covering troposphere, stratosphere, mesosphere and lower thermosphere	(Tilmes et al., 2019; Danabasoglu et al., 2020; Emmons et al., 2020)
CNRM-ESM2-1 (ESM)	1.4° × 1.4 °	Mass-based TACTIC_v2scheme for tropospheric aerosols.	No representation of lower tropospheric (below 560 hPa) chemistry.	(Séférian et al., 2019; Michou et al., 2020)
GFDL-CM4 (ESM)	1.0° × 1.25 °	Bulk mass-based scheme	Interactive troposphere-stratosphere with simplified chemical production of sulfate	(Held et al., 2019), (Zhao et al., 2018a, 2018b)
GFDL-ESM4 (ESM)	1.0° × 1.25 °	Bulk mass-based scheme	Interactive troposphere-stratosphere	(Krasting et al., 2018; Held et al., 2019; Zhao et al., 2018a, 2018b)
GISS-E2-1-G*(GCM)	2.0° × 2.5 °	Mass based One-Moment Aerosol (OMA) scheme	Coupled stratosphere-troposphere and Carbon Bond Mechanism 4 (CBM-4) chemical mechanism	(Rind et al., 2020; Kelley et al., 2020)
HadGEM3-GC31-LL (ESM)	1.25° × 1.875 °	GLOMAP-Mode microphysical aerosol scheme with 2-moment, 5-mode simulating both mass and number. Mass-based CLASSIC bin scheme for dust.	UKCA simplified stratosphere-troposphere chemistry scheme with prescribed monthly-mean climatologies of oxidants.	(K. D. Williams et al., 2018; Mulcahy et al., 2020)
MRI-ESM2-0 (ESM)	1.86° × 1.875 °	MASINGAR mk-2r4c	MRI Chemistry Climate Model version 2.1 (MRI-CCM2.1) covering troposphere and middle atmosphere.	(Yukimoto et al., 2019)
MIROC-ES2L (ESM)	2.813° × 2.813°	SPRINTAS simulating tropospheric aerosol mass mixing ratios and optical properties.	No atmospheric chemistry scheme, only sulfur chemistry in aerosol scheme.	(Hajima et al., 2020)
UKESM1-0-LL (ESM)	1.25° × 1.875 °	GLOMAP-Mode microphysical aerosol scheme with 2-moment, 5-mode simulating both mass and number. Mass-based CLASSIC bin scheme for dust.	UKCA full stratosphere-troposphere chemistry scheme. Interactive photolysis	(Sellar et al., 2019; Mulcahy et al., 2020; Archibald et al., 2020)

Table 4.1: **CMIP6 models (cont.)**

Model	Resolution	Aerosol scheme	Chemistry scheme	References
NorESM2-MM (ESM)	0.9° × 1.25 °	CAM5-Oslo with 4 modes based on MAM in CAM5	Simplified chemistry for use in aerosol scheme. Other fields prescribed.	(Seland et al., 2020)
NorESM2-LM (ESM)	1.9° × 2.5 °	CAM5-Oslo with 4 modes based on MAM in CAM5	Simplified chemistry for use in aerosol scheme. Other fields prescribed.	(Seland et al., 2020)

⋆: We used physics version 3 simulations.

For GISS-E2-1-G, the physics version 3 model includes an interactive mass-based One-Moment Aerosol (OMA)-version module (Rind et al., 2020; Kelley et al., 2020). In the OMA module, SS has two distinct size classes-fine mode with dry radius from 0.1 to 1µm with dry effective radius of 0.44 µm, and coarse mode with dry radius from 1 to 4 µm with dry effective radius of 1.7 µm (Tsigaridis, Koch, & Menon, 2013). For dust, a dust module with 5 classes is used, including a clay and 4 silt classes, with diameters ranging from 0.1 to 32 µm (clay: 0.1-2 µm where clay dust mass makes up 8% of all dust mass; silt class 1: 2-4 µm). $PM_{2.5}$ is estimated as: 0.96×SOA + DU1 + 0.25×DU2 + 0.97×SO4 + SS1 + 0.25×SS2 + 0.995×(BCPHOB + BCPHIL) + 0.96×OM + 0.954×NO3 + 0.973×NH4.

For MRI-ESM2-0 (Yukimoto et al., 2019), SS and dust are discretized into 10 bins ranging from 0.1 to 20 µm in diameter and the first 6 bins (from 0.1 to 2.16 µm in diameter) mostly cover $PM_{2.5}$. Size distributionS of other aerosols are assumed to be lognormal. Sulfate is assumed to be sulfate nitrate $(NH_4)_2SO_4$ and the mass is scaled based on molecular weight. $PM_{2.5}$ is estimated as: (132.1369/96.06)×SO_4 + BC + OA + SS1 + SS2 + SS3 + 0.988×SS4 + 0.901×SS5 + 0.387×SS6 + DU1 + DU2 + DU3 + 0.988×DU4 + 0.901×DU5 + 0.387×DU6.

For MIROC-ES2L (Hajima et al., 2020), emitted dust is discretized into 10 bins from 0.1 to 10 µm in radius and the first 6 bins (from 0.1 to 3.16µm in radius) include $PM_{2.5}$ (Takemura et al., 2000). Sea salt is partitioned into 10 effective radii from 0.05 to 10 µm and lognormally distributed with a standard deviation of 2.0. BC is hydrophobic and sulfate is hydrophilic. $PM_{2.5}$ is estimated as: $0.97\times SO_4$ + BC + OA + SS1 + SS2 + SS3 + $0.988\times$SS4 + $0.901\times$SS5 + $0.387\times$SS6 + DU1 + DU2 + DU3 + $0.988\times$DU4 + $0.901\times$DU5 + $0.387\times$DU6.

For NorESM2-LM (Seland et al., 2020), $PM_{2.5}$ is calculated online during the simulation. The initial size distributions of all aerosols are lognormal. The mixture process e.g., condensation and coagulation leads to changes in compositions and shapes of the background modes, leading to the size distributions that are no longer lognormal(Kirkevåg et al., 2018). The fraction of each mode to PM2.5 is not determinable.

Monthly surface concentrations from PIC and HIS simulations from 10 coupled ocean-atmosphere CMIP5 models (Taylor et al., 2012b) are also analyzed (Table 4.2). CMIP5 historical simulations are extended to 2014 following the Representative Carbon Pathway (RCP) 8.5. CMIP5 CSIRO-Mk3.6 doesn't archive sea salt. CMIP5 pre-industrial control simulations use aerosol emissions fixed at 1850-year forcing levels. Most models also use 1850-year forcing levels of anthropogenic aerosol, greenhouse gas and solar radiation. Only GFDL-CM3 and GFDL-ESM2M use 1860-year forcing levels. Climate conditions and aerosol emissions are similar between 1850 and 1860 (Lamarque et al., 2010).

The surface mass concentration ($[sconc]$) of $PM_{2.5}$ and other aerosol species in all CMIP6 models is derived from the corresponding archived mass mixing ratios (mmr), using

the equation $[sconc] = \frac{M_{air}P}{RT} \times mmr$, where M_{air} is the average molar mass of air defined as $28.96\,\text{g mol}^{-1}$ for dry air, R is the gas constant $8.314\,\text{m}^3\text{Pa K}^{-1}\,\text{mol}^{-1}$, P is surface pressure and T is surface temperature.

All model analyses are based on multi-model mean (MMM) annual means. To weigh each model equally in the MMM, model means are calculated by averaging over all realizations for a given model. The number of realizations for each model varies and are described in Table 4.2 for CMIP6 and Table 4.2 for CMIP5. All data are spatially interpolated to $1^\circ \times 1^\circ$ resolution using bilinear interpolation. The climatological difference in the MMM $PM_{2.5}$ relative to observations and reanalyses is computed, as is the 90% confidence interval of the difference, which is estimated by $\frac{t \times SD}{\sqrt{n-1}}$ where SD is the standard deviation of the differences across models, n is number of models and t is the corresponding t-value at a 90% significance level with $(n-1)$ degrees of freedom. We also estimate the model agreement on the sign of the climatological difference of MMM $PM_{2.5}$ relative to observations and reanalyses. The model agreement on the sign of the difference at each grid box is estimated by the percentage of the number of models that yield a positive (model overestimate) or negative (model underestimate) difference. For instance, a 70% model agreement on model overestimation of annual mean $PM_{2.5}$ relative to MERRA$-$2 means that 70% of the individual CMIP6 models overestimate annual mean $PM_{2.5}$ relative to MERRA$-$2. A percentage larger than 66% ($\geq$ 8 models in CMIP6 and $\geq$ 7 models in CMIP5) represents strong model agreement.

Table 4.2: **CMIP5 models.**

Model	Resolution	Aerosol scheme	Chemistry scheme	References
CanESM2 (ESM)	2.8°× 2.8°	Interactive	Interactive aerosol chemistry (gas-phase and aqueous-phase chemistry)	(Arora et al., 2011)
CSIRO-Mk3.6 (GCM)	1.875°× 1.875°	Mass-based interactive aerosol scheme	Not implemented	(Rotstayn et al., 2012)
GFDL-CM3 (ESM)	2.0°× 2.5°	Bulk mass-based scheme	A fully coupled stratosphere-troposphere chemistry scheme	(Donner et al., 2011)
GFDL-ESM2M (ESM)	2.813°× 2.813°	Semi-interactive mass-based scheme	Not implemented	(Dunne et al., 2012)
GISS-E2-H⋆ (GCM)	2.0°× 2.5°	Interactive stratosphere-troposphere	G-PUCCINI, fully interactive for p3	(Schmidt et al., 2006)
IPSL-CM5A-LR (ESM)	1.9°× 3.75°	Semi-interactive troposphere	Not implemented	(Dufresne et al., 2013)
IPSL-CM5A-MR (ESM)	1.25°× 2.5°	Semi-interactive troposphere	Not implemented	(Dufresne et al., 2013)
IPSL-CM5B-LR (ESM)	1.9°× 3.75°	Semi-interactive troposphere	Not implemented	(Dufresne et al., 2013)
MIROC-ESM (ESM)	1.4°× 1.8134°	SPRINTARS	Not implemented	(S. Watanabe et al., 2011)
MIROC-ESM-CHEM (ESM)	2.8°× 2.8 °	SPRINTARS	CHASER	(S. Watanabe et al., 2011)

⋆: We used physics version 3 simulations.

4.2.2 Reanalyses and Observational Data Sets

To evaluate the CMIP model skill in simulating present-day $PM_{2.5}$, several reanalyses and observational data sets at the surface level are used. Reanalyses include 1980-2014 monthly MERRA−2 products (Gelaro, McCarty, Suárez, et al., 2017) and 2003-2014 monthly $PM_{2.5}$ from the Copernicus Atmosphere Monitoring Service (CAMS; Inness et al.,

2019; Akritidis et al., 2020). Observational data sets include 2000-2009 ground-based $PM_{2.5}$ measurements from the Global Aerosol Synthesis and Science Project (GASSP; Reddington et al., 2017) and 1998-2014 satellite based $PM_{2.5}$ estimates (van Donkelaar, Martin, Brauer, & Boys, 2015; Van Donkelaar et al., 2016). CAMS reanalysis and both observational data sets archive $PM_{2.5}$ directly. A summary of reanalyses and observational data sets is given in Table 4.3.

$PM_{2.5}$ from both observational data sets includes black carbon, organic aerosol, sulfate, sea salt, mineral dust, ammonium and nitrate. In contrast, archived $PM_{2.5}$ from CAMS and $PM_{2.5}$ extracted from MERRA-2 (which doesn't archive $PM_{2.5}$ and is estimated following the reference method discussed in Section "Approximation method of $PM_{2.5}$") lack ammonium and nitrate. For satellite based $PM_{2.5}$ estimates, two types of $PM_{2.5}$ estimates are available: one that includes all aerosol species$-$anthropogenic and natural (sea salt and dust), and one that includes only anthropogenic species. We estimate the fine dust by subtracting $PM_{2.5}$ estimates without natural aerosols from the $PM_{2.5}$ estimates with all aerosol species. This may lead to overestimation of continental fine dust, particularly over coastal regions. Our implicit assumption is that fine dust is the dominant natural aerosol species over land. The satellite based estimates contain further caveats which involve underestimation of surface $PM_{2.5}$, for example over bright continental surfaces, as a result of errors and limitations in satellite retrieved AOD products (van Donkelaar et al., 2015).

For all model-data comparisons, we use the time period of the data (either reanalyses or observational data). $PM_{2.5}$ from the two reanalyses and the satellite based estimates are regridded using bilinear interpolation to $1^\circ \times 1^\circ$ (as with CMIP5/6). For GASSP, the

Table 4.3: **Reanalyses and Observations.** Global fine dust factor is estimated as the spatial and temporal average of the archived fine dust in MERRA−2 and CAMS divided by the corresponding archived total dust. Global land fine dust factor is given in the parenthesis."n/a" denotes unavailable.

Data	Temporal Coverage	Aerosol mechanism	Aerosol module	Archived Species	Fine dust fraction
MERRA−2	1980-2014	Integrated multiple satellite AOD measurements (AVHRR, MODIS, MISR) and ground-based AERONET measurements, assimilated by the GEOS-5 model	GOCART aerosol module. DU is partitioned into five size bins from 0.1 to 10 µm in particle radius. SS is simulated in five size bins from 0.03 to 10 µm in radius. Organic matter (OM) and BC are included as both hydrophobic and hydrophilic components.	BC, SO4, OM, SS, fine SS, DU, fine DU	0.32 (0.3)
CAMS	2003-2014	Assimilated by IFS(CB05) model, integrated multiple satellite AOD measurements (MODIS, Envisat, Aura, Metop-A/B, NOAA-14/16/17/18/19).	Hybrid bulk-bin aerosol scheme. DU is partitioned into three radius ranges, including 0.03-0.55µm (DU1), 0.55-0.9µm (DU2) and 0.9-20µm (DU3). Archived fine dust is estimated as (DU1+ DU2). SS is partitioned into three size ranges, including 0.03-0.5µm (SS1), 0.5-5µm and 5-20µm (SS2) in particle radius. Archived fine SS is estimated as (SS1+0.5×SS2)/4.3 (dividing a factor of 4.3 is to convert to dry mass). OM and BC are included as two types-hydrophobic and hydrophilic.	BC, SO4, OM, SS1, SS2, DU1, DU2, DU2 and $PM_{2.5}$	0.8 (0.72)
Satellite based surface $PM_{2.5}$ estimates	1980-2014	Integrated from MODIS satellite AOD products, assimilated by GEOS-CHEM, constrained by observations (CALIPSO, AERONET and ground measurement from GDB)	n/a	$PM_{2.5}$ with aerosol species from anthropogenic (ammonium sulfate, ammonium nitrate, primary and secondary OA and BC) and natural (SS and DU) sources and $PM_{2.5}$ with aerosol species from anthropogenic sources	n/a

Table 4.3: **Reanalyses and Observations (cont.)**

Data	Temporal Coverage	Aerosol mechanism	Aerosol module	Archived Species	Fine dust fraction
Ground-based measurements (GASSP)	2000-2009	Ground-station measurements collected from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network, National Air Pollution Surveillance Program (NAPS) in North America, the European Monitoring and Evaluation Program (EMEP) and Asia-Pacific Aerosol Database (A-PAD)	n/a	$PM_{2.5}$	n/a

sites aren't distributed evenly in space and are clustered over some areas, and thus model data is interpolated onto the locations of sites.

4.2.3 Population Data

The human population impacted by pristine aerosol pollution is estimated using the Gridded Population of the World Version (GPWv4) with raster cell (Center for International Earth Science Information Network-CIESIN-Columbia University, 2018). It provides estimates of the population on fine grid resolutions of $1^\circ \times 1^\circ$ with fixed latitudinal and longitudinal distance at ≈110 km. This data is re-gridded to match our spatial coordinates. At the time it was accessed (January 2020), the most up-to-date world population data available was from 2015. The total world population estimated by GPWv4 in 2015 was about 7.32 billion.

4.3 Approximation Method of $PM_{2.5}$

Unfortunately utilizing CMIP model output to map air quality parameters is complicated because most models do not directly archive all of the fields of interest and those that do use varying methodologies, as well as different aerosol species. For this study $PM_{2.5}$ across all models is calculated using an approximation formula adapted from Fiore et al. (2012) and Silva et al. (2013),

$$PM_{2.5} = BC + SO_4 + OA + 0.25 \times SS + 0.1 \times DU \tag{4.1}$$

that we refer to as the reference method. In the reference method, BC is black carbon; SO_4 is sulphate; OA is organic aerosol (including primary organic aerosol and secondary organic aerosol if available); SS is sea salt; DU is dust. BC, SO_4 and OA are assumed to be $\leq 2.5\,\mu m$ in diameter. We exclude nitrate and ammonium since most models do not archive these two aerosol species. The reference method assumes that 1/10 (1/4) of the total dust (sea salt) is within the $2.5\,\mu m$ size range, which is referred to as fine dust (fine sea salt), although this fraction is likely dependent on the aerosol schemes and size distributions of each model. For the CMIP6 CNRM-ESM2-1, we obtain abnormally large fine sea salt concentrations following the reference method. To account for the large sea salt size range of up to $20\,\mu m$, a much smaller factor of ≈ 0.01 is used for this model (cf., Allen et al., 2020).

For these five models that directly archive $PM_{2.5}$ as mentioned before, the reference method underestimates present-day archived $PM_{2.5}$ over land (Figure 4.1). The global land area-weighted average underestimation is $-6.65\,\mu g\ m^{-3}$, with a 90% uncertainty in-

terval of $\pm 4.16\,\mu g\ m^{-3}$. Over dusty regions−which include Africa, the Middle East, eastern and central Asia, northwest China (Taklamakan desert), southern Mongolia (Gobi desert), and northwestern India (red-boxed areas in Figure 4.2a, which occupy 13.83% of global land)−the underestimation is larger at -16.27$\pm$10.32 $\mu g\ m^{-3}$. The underestimation bias likely emerges for several reasons, including the aforementioned aerosol species variability (e.g., some models include nitrate and ammonium aerosol, but this is not included in the reference method) and different contributions from fine aerosols within the $PM_{2.5}$ size range−most notably dust.

Given that dust is the major component of $PM_{2.5}$ over land, we further investigate the credibility of the assumed dust contribution in the reference method. We compute a temporally and spatially invariant fine dust factor as a time averaged, area-weighted mean of the archived fine dust (that is dust aerosol within the $PM_{2.5}$ size range) divided by the archived total dust in two reanalyses (Table 4.3). The resulting fine dust factor is 0.32 over the globe and 0.30 over global land in MERRA−2 (Figure 4.3; Table 4.3). For CAMS, the corresponding fine dust factor is 0.80 over the globe and 0.72 over global land (Table 4.3). Thus, the fine dust factor extracted from both reanalyses is significantly larger than the 0.1 assumed in the reference method.

According to the IPCC fourth Assessment (AR4), clay-sized dust (with diameter less than $2\,\mu m$) emissions (350 Tg yr^{-1}) account for ≈20% of total dust emission (1800 Tg yr^{-1}; S. Solomon et al., 2007). Likewise, observed fine dust AOD, which describes the contribution to scattering and absorption of radiation by fine dust, is fractionally ≈0.2 relative

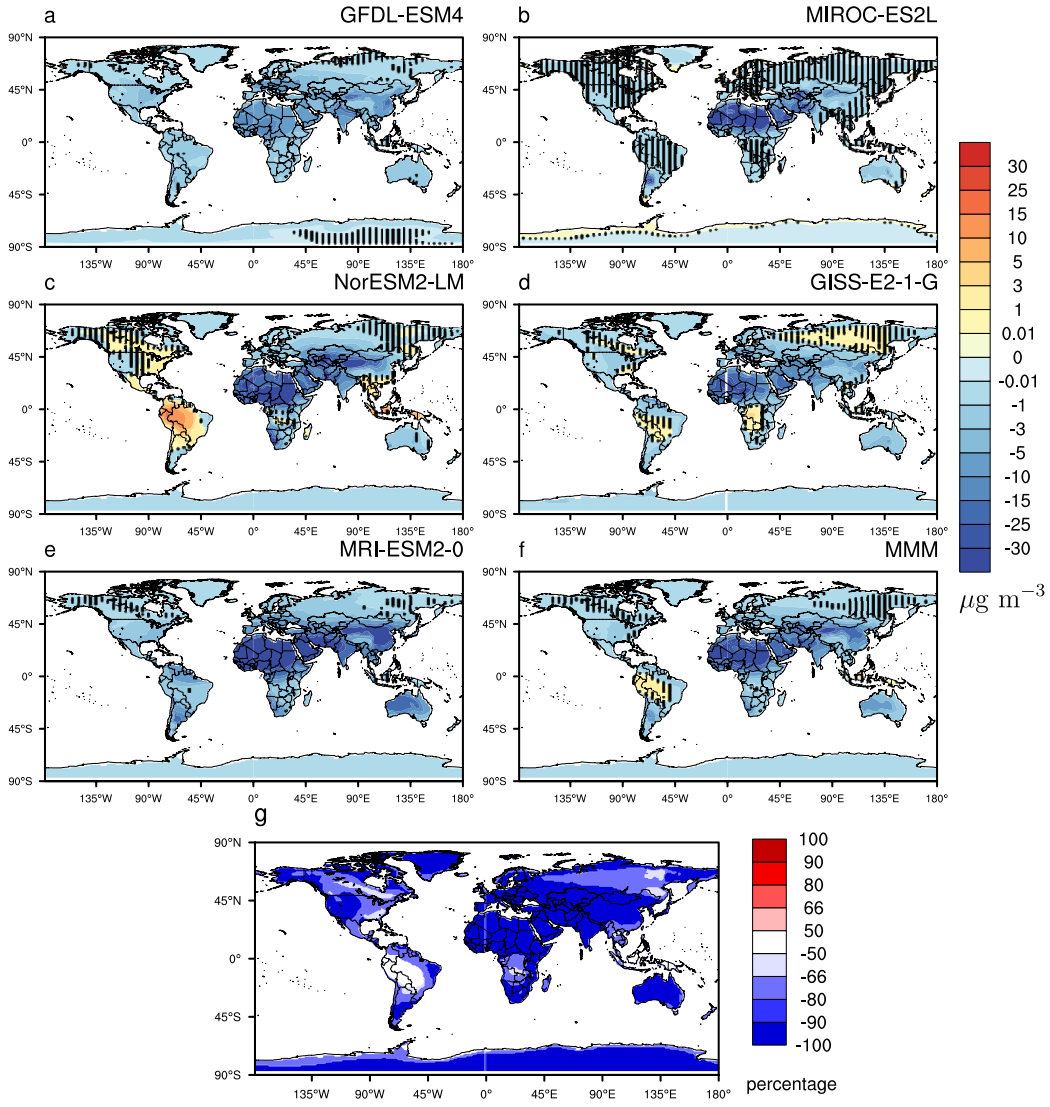

Figure 4.1: **Evaluation of $PM_{2.5}$ approximation method using CMIP6 models.** 1980-2014 climatological $PM_{2.5}$ difference (the estimated $PM_{2.5}$ following the reference method minus the archived $PM_{2.5}$) in (a) GFDL-ESM4, (b) MIROC-ES2L, (c) NorESM2-LM, (d) GISS-E2-1-G, (e) MRI-ESM2-0 and (f) multi-model mean (MMM). Non-stippled differences in (a, b, c, d, e, f) are significant at the 90% confidence level, based on a standard t-test. Units in (a, b, c, d, e, f) are μg m^{-3}. (g) shows model agreement (units of percentage) on the sign of differences in (f). Red shading represents model overestimation of $PM_{2.5}$ or fine dust relative to that archived, and blue shading represents underestimation.

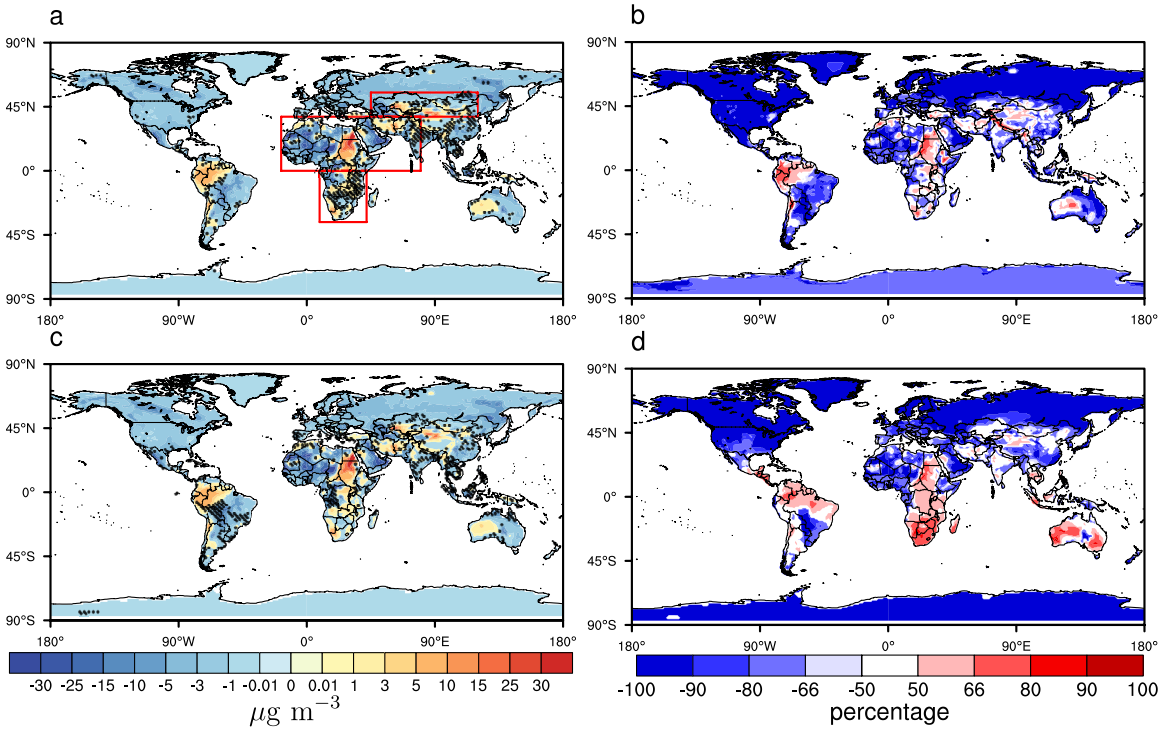

Figure 4.2: **Present-day evaluation of CMIP6 $PM_{2.5}$ and fine dust relative to MERRA−2.** 1980-2014 climatological mean difference (CMIP6 MMM − MERRA−2) for (a) $PM_{2.5}$ and (c) fine dust. Both quantities are estimated using the reference method. (b, d) shows CMIP6 model agreement (units of percentage) on the sign of differences in (a, c), respectively. Red shading represents model overestimation of $PM_{2.5}$ or fine dust relative to MERRA−2, and blue shading represents underestimation. Non-stippled in (a, c) are significant at the 90% confidence level based on a standard t-test. Units in (a, c) are μg m^{-3}. Red-boxed land areas in (a) indicate dusty regions, which include 13.83% of global land area.

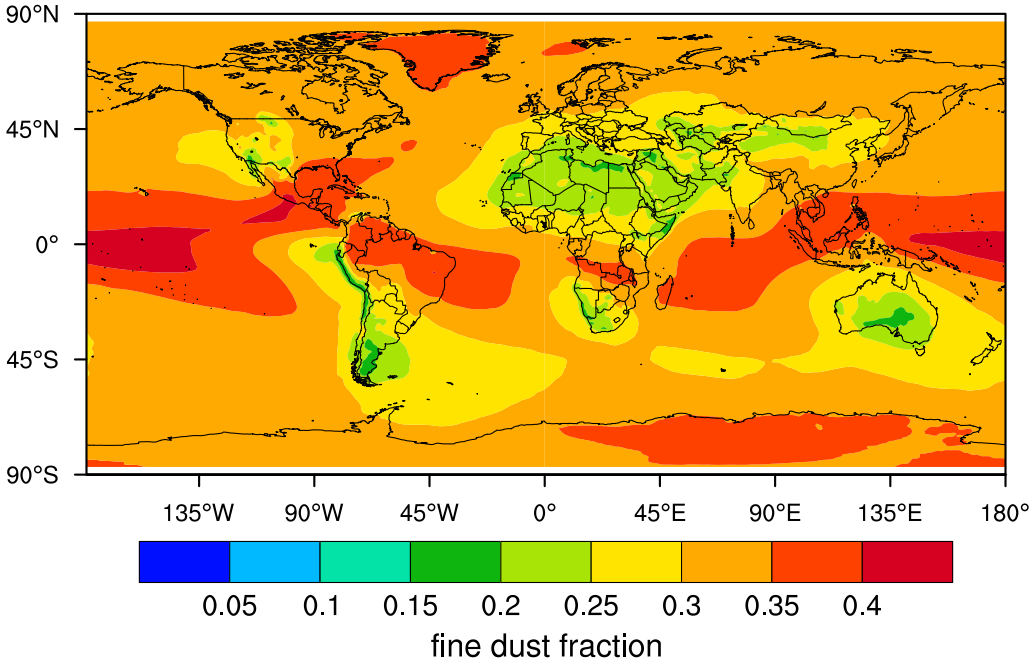

Figure 4.3: **Present-day MERRA−2 fine dust fraction.** 1980-2014 climatological fine dust fraction estimated using archived fine dust divided by total dust in MERRA−2.

to total dust over dust source regions (Chung et al., 2016). Observationally constrained dust size distributions and theoretical analyses assert that the clay-sized dust fraction is 0.043 (with a 95% confidence interval from 0.035 to 0.057; Kok, 2011; Kok et al., 2017). This evidence implicitly suggests that the reference method's fine dust factor of 0.1 falls within a reasonable range. Given that models tend to overestimate fine dust (Kok, 2011; Kok et al., 2017; Adebiyi & Kok, 2020) and that both reanalyses are model-based, it is likely that MERRA−2 and CAMS have too large fine dust factors. Furthermore, if CMIP5/6 models underestimate total dust (and this also applies to the surface level; Adebiyi & Kok, 2020), and a fine dust factor of 0.1 is a reasonable approximation as previously argued, this suggests fine dust computed with the reference method likely represents an underestimation. Overall, we conclude that the reference method with a fine dust factor of 0.1 yields a suitable and conservative representation of climatological $PM_{2.5}$.

4.4 Results

4.4.1 Analysis of Present-day $PM_{2.5}$

First we evaluate models' credibility in simulating present-day $PM_{2.5}$ relative to multiple data sets. Analysis of model results compared to two reanalyses (MERRA−2 and CAMS) is discussed in Model versus Reanalyses Section, and the comparison with the two observational data sets (satellite based estimates round-based measurements GASSP) is described in Model versus Observational Data Sets Section.

Model Versus Reanalyses Relative to MERRA−2, over most continents the CMIP6 MMM underestimates $PM_{2.5}$ (Figure 4.2a). $PM_{2.5}$, as well as fine dust, from models and MERRA−2 reanalysis are estimated here using the reference method. CMIP6 underestimation of $PM_{2.5}$ also exhibits a high degree of model agreement (Figure 4.2b). We notice that there are some limited regions where the CMIP6 MMM overestimates $PM_{2.5}$ relative to MERRA−2 but with low model agreement. Table 4.4 shows that global land area-weighted average underestimation is -2.89±0.15 µg m^{-3}. This underestimation bias is even larger over dusty regions (red-boxed areas in Figure 4.2a), with a regional land area-weighted average difference of -3.39±2.55 µg m^{-3} (Table 4.4). This implies the importance of fine dust to the $PM_{2.5}$ bias. Consistently, the CMIP6 MMM underestimates fine dust relative to MERRA−2, with global land and dusty region area-weighted biases of -0.87±0.85 µg m^{-3} and -2.41±2.28 µg m^{-3}, respectively (Figure 4.2c & Table 4.4). Underestimation of fine dust is also a robust result, with over 90% model agreement in most locations (Figure 4.2d).

We further evaluate CMIP6 fine dust from the reference method relative to the archived MERRA−2 fine dust. We obtain a larger CMIP6 MMM fine dust underestimation over almost all land with a global land area-weighted average difference of -7.21±0.85 µg m^{-3} (Table 4.4 & Figure 4a), with a high degree of model agreement (Figure 4b). As expected, a larger bias exists over dusty regions, where the regional land area-weighted average difference is -18.78±2.28 µg m^{-3} (Table 4.4). As previously discussed (Section "Approximation method of $PM_{2.5}$"), the fine dust factor extracted from MERRA−2 is ≈ 3 times as large as the reference method's 0.1 fine dust factor. This helps explain the larger fine dust underestimation bias when archived MERRA−2 fine dust is used, as opposed to that based

on the reference method.

Compared to CAMS, the CMIP6 MMM underestimates $PM_{2.5}$ and fine dust, with global land area-weighted average differences of -15.60±1.18 µg m^{-3} and -8.20±0.84 µg m^{-3}, respectively (Table 4.4 & Figure 4.5). We also find larger underestimation of $PM_{2.5}$ and fine dust over dusty regions, with corresponding regional land area-weighted average biases of -27.63±2.56 µg m^{-3} and -20.57±2.25 µg m^{-3}, respectively. The CMIP6 MMM underestimation of $PM_{2.5}$ and fine dust over global land and dusty regions relative to CAMS is larger than that relative to MERRA−2. As discussed in Section "Approximation method of $PM_{2.5}$", this is supported by the much larger fine dust factor (0.72 over global land) in CAMS relative to MERRA−2 (0.30 over global land).

Additionally, CMIP6 archived $PM_{2.5}$ is compared to that in CAMS, as well as MERRA−2 $PM_{2.5}$ based on the reference method (Table 4.4), and we also find that models underestimate present-day $PM_{2.5}$. This further supports the assertion that models tend to underestimate $PM_{2.5}$.

Similar results are obtained from an identical analysis using 10 CMIP5 models (Table 4.4 & Figure 4.6). Moreover, relative to the CMIP6 MMM underestimation, the CMIP5 MMM generally yields larger underestimations of present-day $PM_{2.5}$ and fine dust when compared to all data sets (Table 4.4). A prior study has shown that CMIP5 models underestimate dust emissions and transport processes (Evan, Flamant, Fiedler, & Doherty, 2014), consistent with our results. Furthermore, CMIP5 models include less developed aerosol representations than CMIP6 models (Tables 4.2 & 4.2).

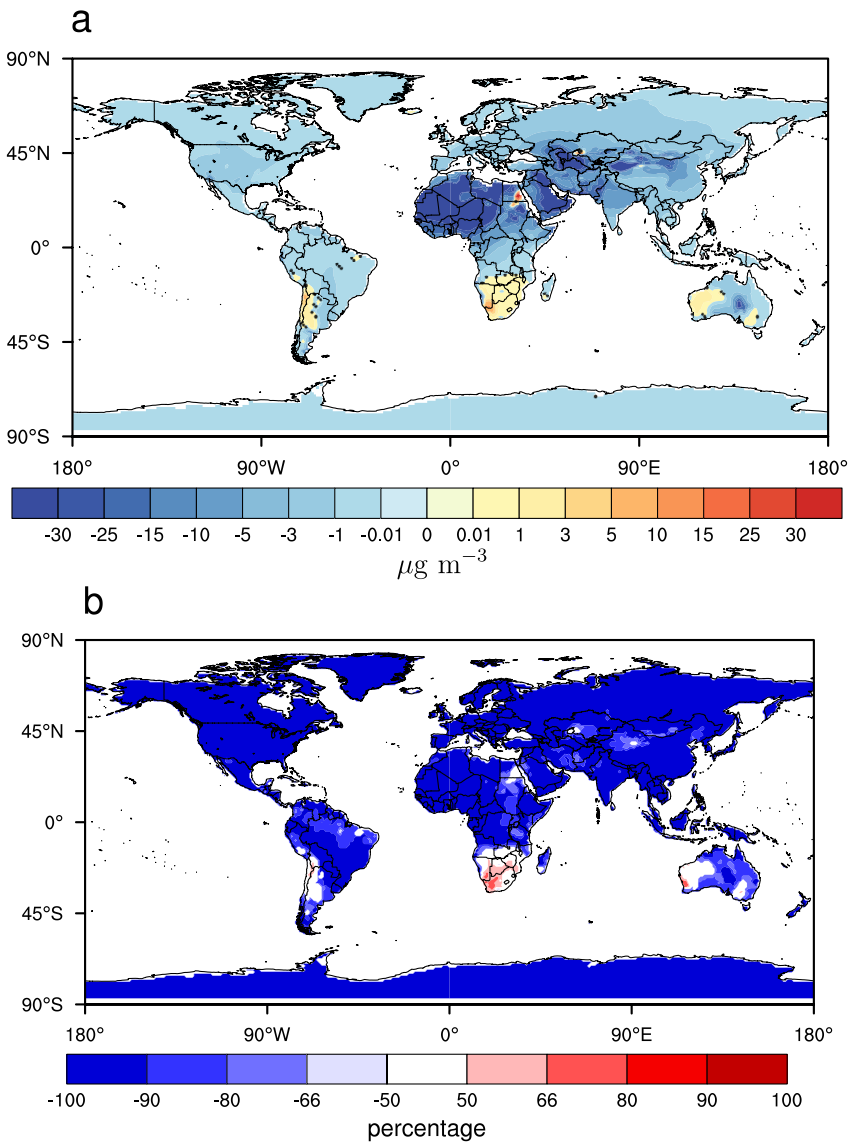

Figure 4.4: **Present-day evaluation of fine dust estimated from CMIP6 relative to archived fine dust in MERRA−2.** (a) 1980-2014 climatological mean difference (CMIP6 MMM − MERRA−2) for fine dust. Fine dust in CMIP6 models is estimated using the reference method. From MERRA−2 archived fine dust is used. (b) Shows CMIP6 model agreement (units: percentage) on the sign of differences in (a). Red shading represents model overestimation of fine dust relative to MERRA−2, and blue shading represents underestimation. Non-stippled in (a, c) are significant at the 90% confidence level based on a standard t-test. Units in (a) are μg m^{-3}.

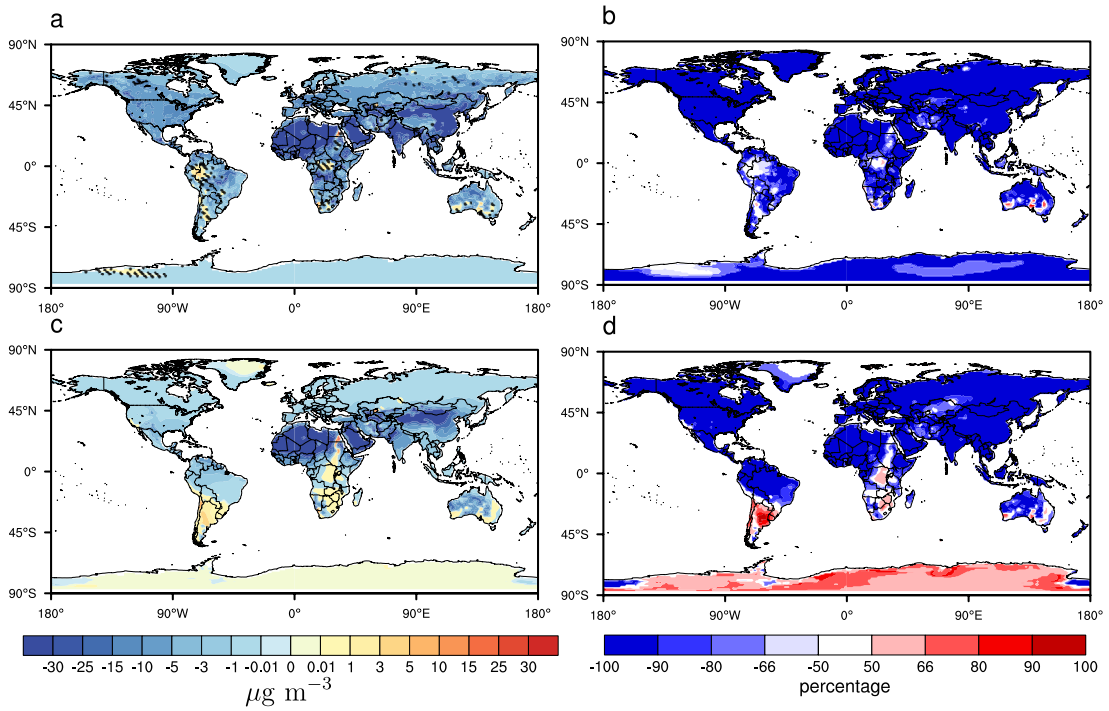

Figure 4.5: **Evaluation of present-day CMIP6 $PM_{2.5}$ relative to CAMS reanalysis.** 2003-2014 climatological mean difference (CMIP6 MMM-CAMS) for (a) $PM_{2.5}$ and (c) fine dust. Both quantities in models are estimated using the reference method. (b, d) shows CMIP6 model agreement (units of percentage) on the sign of the differences in (a, c), respectively. Red shading represents model overestimation of $PM_{2.5}$ or fine dust relative to CAMS, and blue shading represents underestimation. Non-stippled differences in (a, c) are significant at the 90% confidence level based on a standard t-test. Units in (a, c) are μg m^{-3}.

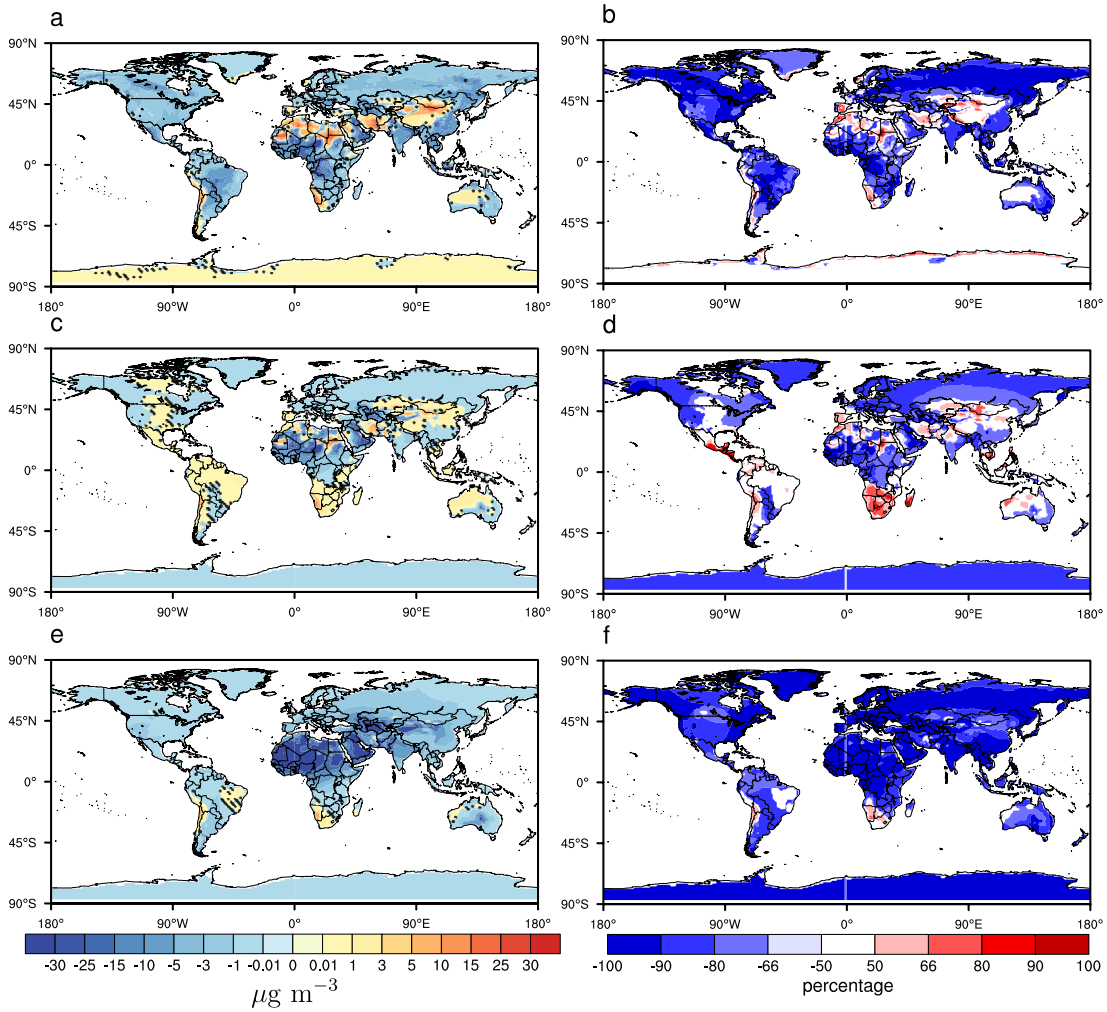

Figure 4.6: **Present-day evaluation of CMIP5 $PM_{2.5}$ and fine dust relative to MERRA−2.** 1980-2014 climatological mean difference (CMIP5 MMM − MERRA−2) for (a) $PM_{2.5}$ and (c) fine dust. Both quantities are estimated using the reference method. (e) is analogous to (c), but MERRA−2 is based on archived (actual) fine dust. (b, d, f) shows CMIP6 model agreement (units of percentage) on the sign of differences in (a, c, e), respectively. Red shading represents model overestimation of $PM_{2.5}$ or fine dust relative to MERRA−2, and blue shading represents underestimation. Non-stippled differences in (a, c, e) are significant at the 90% confidence level based on a standard t-test. Units in (a, c, e) are µg m^{-3}.

Model Versus Observational Data Sets When compared with the two observational data sets, including 1998-2014 satellite based $PM_{2.5}$ estimates (Figure 4.7) and 2000-2009 GASSP ground-based measurements (Figure 4.8), the models also generally underestimate present-day $PM_{2.5}$ and fine dust and the underestimation is generally larger over dusty regions, supporting the results from MERRA-2 and CAMS (Table 4.4). However, there are other uncertainties that may also contribute to model-measurement differences. For example, $PM_{2.5}$ from the two observational data sets include nitrate and ammonium aerosols while models' $PM_{2.5}$ from the reference method do not.

When comparing archived $PM_{2.5}$ from CMIP6 with two observational data sets (Table 4.4), we find that models underestimate present-day $PM_{2.5}$ relative to GASSP but overestimate $PM_{2.5}$ relative to satellite based estimates. This is likely due to systematic errors and limitations in satellite retrieved AOD products that tend to underestimate surface $PM_{2.5}$ over bright continental surfaces (van Donkelaar et al., 2015). The analogous CMIP5 analysis also shows that models underestimate $PM_{2.5}$ (Table 4.4).

A better assessment of the uncertainties associated with the reference method will likely require measurements from more monitoring networks with a wider temporal and spatial coverage, especially over poorly monitored dusty regions. Other factors may also contribute to disparities between model and reanalyses/observational data sets. Models have deficiencies in their representations of aerosols and aerosol related processes (Ciarelli et al., 2016; Kok, 2011; Solazzo et al., 2017; Kok et al., 2017; Glotfelty, He, & Zhang,

Table 4.4: **Present-day evaluation of CMIP5/6 $PM_{2.5}$ and fine dust relative to multiple data sets.** Climatological area-weighted mean difference (CMIP5/6 MMM − data set) over global land and dusty regions for $PM_{2.5}$ and fine dust. Data sets include MERRA−2 and CAMS reanalysis and two observational data sets, including a satellite based estimate (SAT) and ground-based measurements (GASSP). Dusty regions are denoted by the red-boxed areas in Figure 4.2a. The corresponding 90% uncertainty interval of the difference is also included. $PM_{2.5}$ or fine dust from the model or a data set is estimated using the reference method. Otherwise, $PM_{2.5}$ or fine dust "directly" obtained is denoted as "archived". Units are μg m^{-3}. "n/a" denotes unavailable.

		CMIP6		CMIP6 archived		CMIP5	
		Global land	Dusty region	Global land	Dusty region	Global land	Dusty region
MERRA−2	$PM_{2.5}$	-2.89±0.15	-3.39±2.55	-2.77±5.65	-4.42±12.75	-3.53±3.25	-3.91±2.64
	fine dust	-0.87±0.85	-2.41±2.28	n/a	n/a	-1.35±0.92	-3.90±2.01
	archived fine dust	-7.21±0.85	-18.78±2.28	n/a	n/a	-7.69±0.93	-20.27±2.01
CAMS	archived $PM_{2.5}$	-15.60±1.18	-27.63±2.56	-9.51±5.55	-12.66±12.38	-16.62±3.25	-28.57±7.49
	archived fine dust	-8.20±0.84	-20.57±2.25	n/a	n/a	-8.66±0.94	-21.99±2.04
SAT	archived $PM_{2.5}$	-6.01±1.35	-11.19±2.57	0.95±6.43	3.96±12.61	-7.16±3.69	-12.13±7.43
	archived fine dust	-4.03±0.97	-8.79±2.27	n/a	n/a	-4.61±1.07	-10.30±2.01
GASSP	archived $PM_{2.5}$	-3.77±0.54	n/a	-1.28±4.04	n/a	-1.77±2.29	n/a

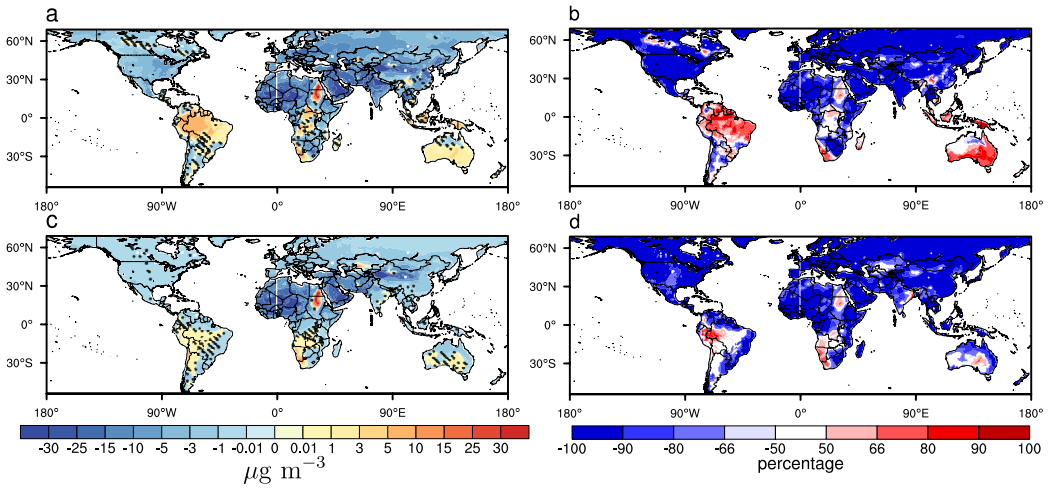

Figure 4.7: **Evaluation of present-day CMIP6 MMM $PM_{2.5}$ relative to satellite based (SAT) estimates.** 1980-2014 climatological mean difference (CMIP6 MMM-SAT) for (a) $PM_{2.5}$ and (c) fine dust. Both quantities in models are estimated using the reference method. (b, d) shows CMIP6 model agreement (units of percentage) on the sign of the differences in (a, c, e), respectively. Red shading represents model overestimation of $PM_{2.5}$ or fine dust relative to SAT estimates, and blue shading represents underestimation. Differences in (a, c) are significant at the 90% confidence level based on a standard t-test. Units in (a, c) are μg m^{-3}.

2017), as well as with their simulated meteorological fields (McNider & Pour-Biazar, 2020). Model underestimation of $PM_{2.5}$ may also result from their inability to capture complex local emission sources due to coarse resolution (Glotfelty et al., 2017).

4.4.2 Estimation of Pristine $PM_{2.5}$

The evaluation of the CMIP results in the present-day yields confidence that they are representative of conservative estimates for $PM_{2.5}$ (based on the reference method and models' underestimation of present-day $PM_{2.5}$), and thus the PIC simulations can be used to generate a baseline understanding of pristine conditions. The CMIP6 PIC MMM shows that pristine aerosol levels greater than 10 μg m^{-3}, the WHO recommended annual mean $PM_{2.5}$ threshold, are present over large areas, particularly in dusty regions

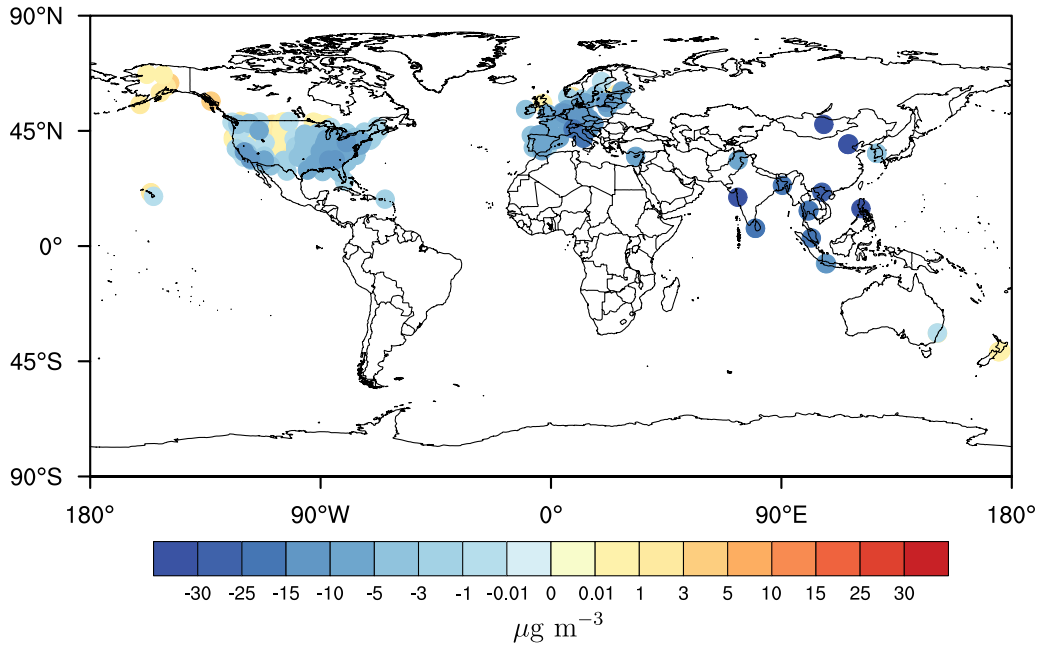

Figure 4.8: **Evaluation of present-day CMIP6 MMM $PM_{2.5}$ relative to GASSP ground-based measurements.** 2000-2009 climatological mean difference (CMIP6 MMM - GASSP) for $PM_{2.5}$. CMIP6 MMM $PM_{2.5}$ is estimated using the reference method. Units are µg m^{-3}.

Table 4.5: **Pristine $PM_{2.5}$ based on CMIP5/6.** Pre-industrial climatological area-weighted $PM_{2.5}$ based on the CMIP5/6 MMM. Pristine $PM_{2.5}$ from CMIP5/6 is estimated using the reference method. Pristine $PM_{2.5}$ from the five CMIP6 models that explicitly archive $PM_{2.5}$ is denoted as "CMIP6 archived". The corresponding 90% uncertainty interval of the pristine $PM_{2.5}$ is also included. Dusty regions are the red-boxed regions in Figure 4.2a. Stippled dusty regions are where the MMM pristine $PM_{2.5}$ exceeds the WHO threshold ($10\,\mu g\ m^{-3}$) and at least 2/3 of the models agree (Figure 4.10b). Population within stippled dusty regions, as well as in regions where only the MMM pristine $PM_{2.5}$ exceeds $10\,\mu g\ m^{-3}$ is also included. The corresponding percentage of the total world population over those regions is given in parentheses. Units of pristine $PM_{2.5}$ are $\mu g\ m^{-3}$. Units of population are billions of people.

		CMIP6	CMIP6 archived	CMIP5
$PM_{2.5}$	Global land	7.06±0.97	12.78±5.28	5.81±1.25
	Dusty regions	14.08±2.30	28.72±12.71	11.06±2.45
	Stippled dusty regions	24.41±2.49	44.03±3.73	23.36±2.65
Population	MMM $> 10\,\mu g\ m^{-3}$	1.72(23.49%)	3.49(47.72%)	0.71(9.70%)
	Stippled dusty regions	0.97 (13.28%)	2.12 (28.95%)	0.16 (2.23%)

(Figure 4.9a). Over these dusty areas, the global land area-weighted average annual mean $PM_{2.5}$ is 14.08±2.30 $\mu g\ m^{-3}$ (Table 4.5). We note that there is also a relatively large corresponding inter-model variance over these dusty regions−highlighting model structural uncertainty−ranging from 8.44 to 24.25 $\mu g\ m^{-3}$(Figure 4.9b).

In Figure 4.10a the percentage of models that simulate pristine $PM_{2.5}$ greater than $10\,\mu g\ m^{-3}$ is mapped. More than two thirds of CMIP6 models characterize the annual mean of pristine aerosol levels as greater than $10\,\mu g\ m^{-3}$ over most dusty regions−which we refer to as "stippled dusty regions"−including northern and central Africa, parts of the Middle East, central Asia, and the Indo-Gangetic plain, and through northwest China (the Taklamakan desert; Figure 4.10b). The land area-weighted average annual mean $PM_{2.5}$ over the stippled dusty regions is 24.41±2.49 $\mu g\ m^{-3}$ (Table 4.5), which is more than two times the WHO threshold. Therefore, relative to the WHO air quality guidelines, these

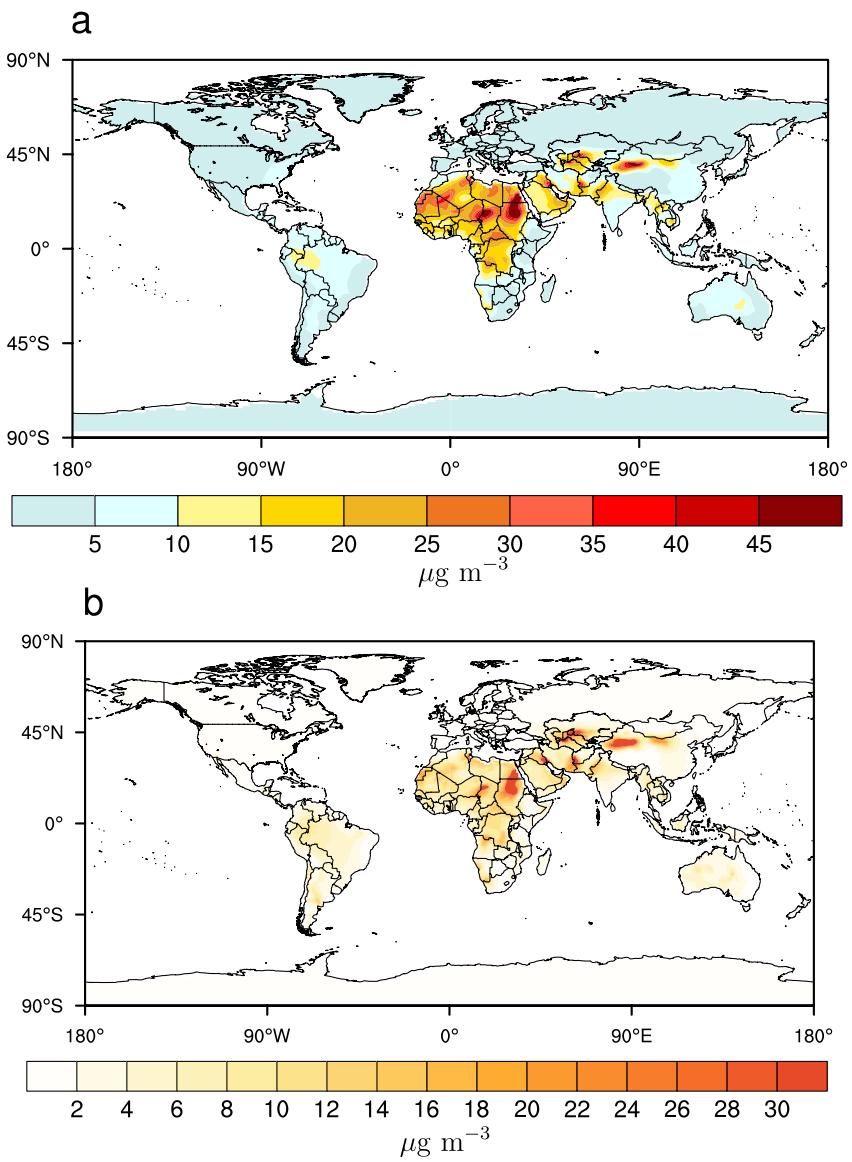

Figure 4.9: **Global map of pristine $PM_{2.5}$ based on CMIP6 models.** (a) Pre-industrial climatological $PM_{2.5}$ based on the CMIP6 MMM; (b) standard deviation of PI $PM_{2.5}$ across models. Units in (a, b) are µg m^{-3}.

regions already possess poor air quality without contributions from anthropogenic emissions. Collectively these regions represent a human population of roughly 0.97 billion, or about 13.28% of the world's population (Figure 4.10b & Table 4.5). This impacted population increases to ≈1.72 billion (23.49% of global population) when we only consider where the MMM pristine $PM_{2.5}$ exceeds the WHO threshold (i.e., without the 2/3 model agreement criterion; Table 4.5).

Similar results are obtained using the archived pristine $PM_{2.5}$ produced by the five CMIP6 models that have explicit $PM_{2.5}$ products (Figure 4.11). Moreover, consistent with the view that our methodology (i.e., the reference method) produces conservative estimates, the archived pristine $PM_{2.5}$ in fact yields broader stippled dusty regions where the MMM $PM_{2.5}$ exceeds the WHO threshold with at least 2/3 model agreement. The corresponding stippled dusty region's average annual mean $PM_{2.5}$ is $44.03 \pm 3.73\,\mu g\ m^{-3}$−which is more than four times the WHO threshold−and it suggests ≈ 2.12 billion people (28.95% of the world's population) are impacted (Table 4.5). If we only require the MMM pristine $PM_{2.5}$ to exceed the WHO threshold, this impacted population increases to ≈3.49 billion (47.72% of global population). Results from CMIP5 yield similar conclusions, but with less overall stippled dusty regions. Here, the area-weighted average annual mean $PM_{2.5}$ over stippled dusty regions is $23.36 \pm 2.65\,\mu g\ m^{-3}$, affecting about 0.16 billion people (2.23% of the global population; Table 4.5 & Figure 4.12). Similarly, this impacted population increases to ≈0.71 billion (9.70% of global population) when we only require the MMM pristine $PM_{2.5}$ to exceed the WHO threshold.

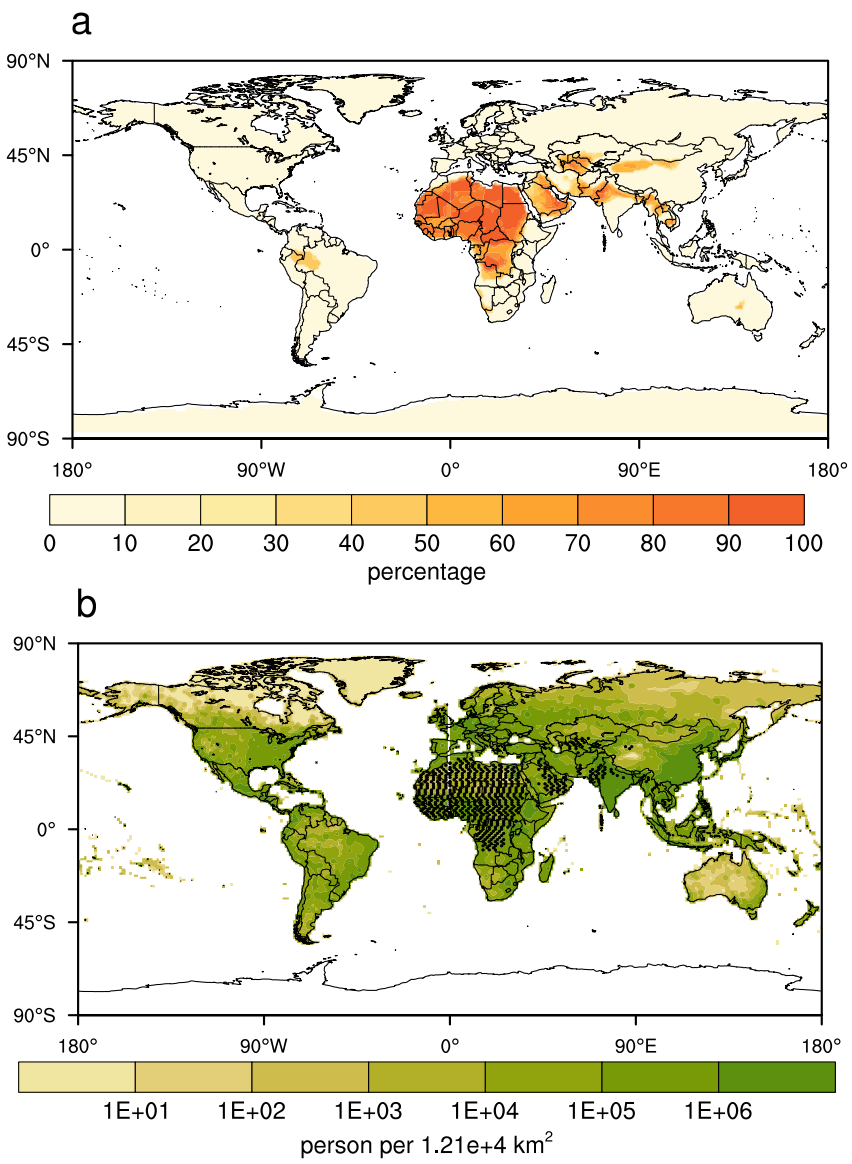

Figure 4.10: **Model agreement on simulating pristine aerosol pollution and world gridded population.** (a) Percentage of models that simulate a pristine $PM_{2.5}$ level that exceeds the WHO annual mean threshold for poor air quality ($10\,\mu g\ m^{-3}$). (b) World gridded population. Unit in (b) is person per 12 100 km^2. Stippling in (b) represents where the CMIP6 MMM $PM_{2.5}$ exceeds the WHO threshold (i.e., Figure 4.9a) and at least two thirds ($\geq 66\%$) of the models agree.

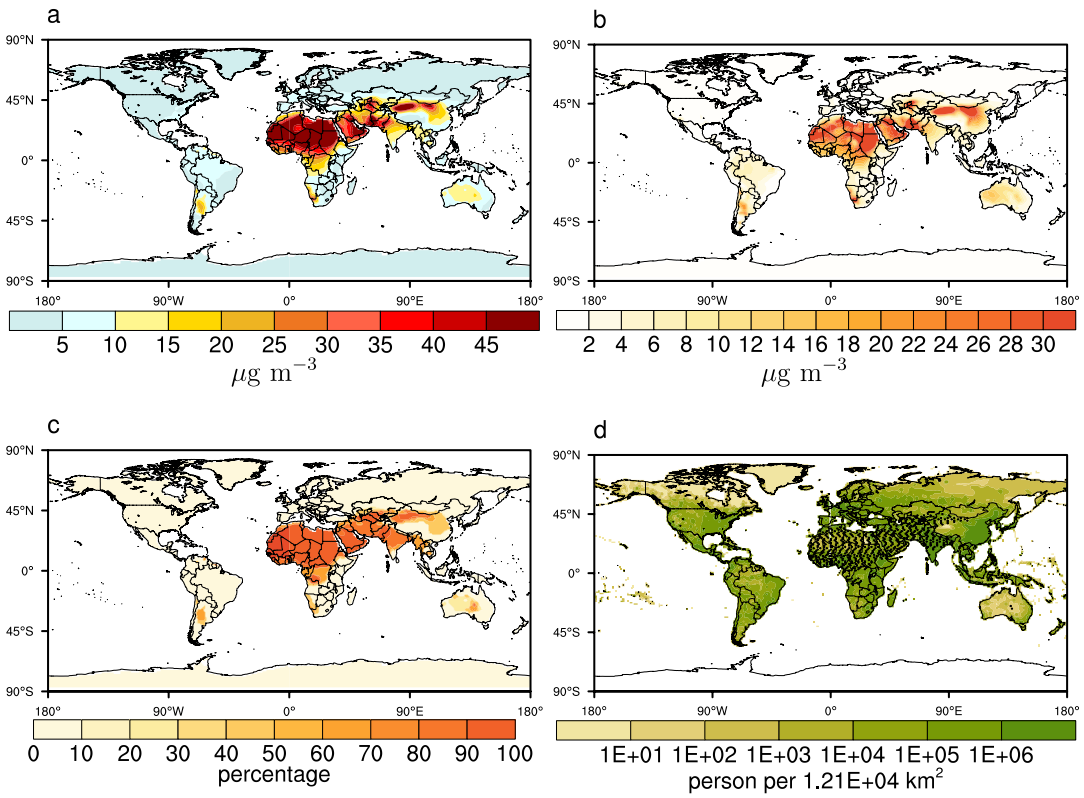

Figure 4.11: **Global map of pristine $PM_{2.5}$ based on CMIP6 archived $PM_{2.5}$.** (a) Pre-industrial (PIC) climatological $PM_{2.5}$ based on MMM archived $PM_{2.5}$ in five CMIP6 models, including GFDL-ESM4, MIROC-ES2L, NorESM2-LM, GISS-E2-1-G and MRI-ESM2-0; (b) standard deviation of PIC $PM_{2.5}$ across models; (c) the percentage of models that simulate a pristine $PM_{2.5}$ level that exceeds the WHO threshold for poor air quality (10 µg m^{-3}); and (d) world gridded population. Stippling in (d) represents MMM $PM_{2.5}$ in (a) exceeds the WHO threshold and at least two thirds ($\geq$ 66%) model agreement in (c). Units in (a, b) are µg m^{-3} and units in (d) are person per 12 100 km^2.

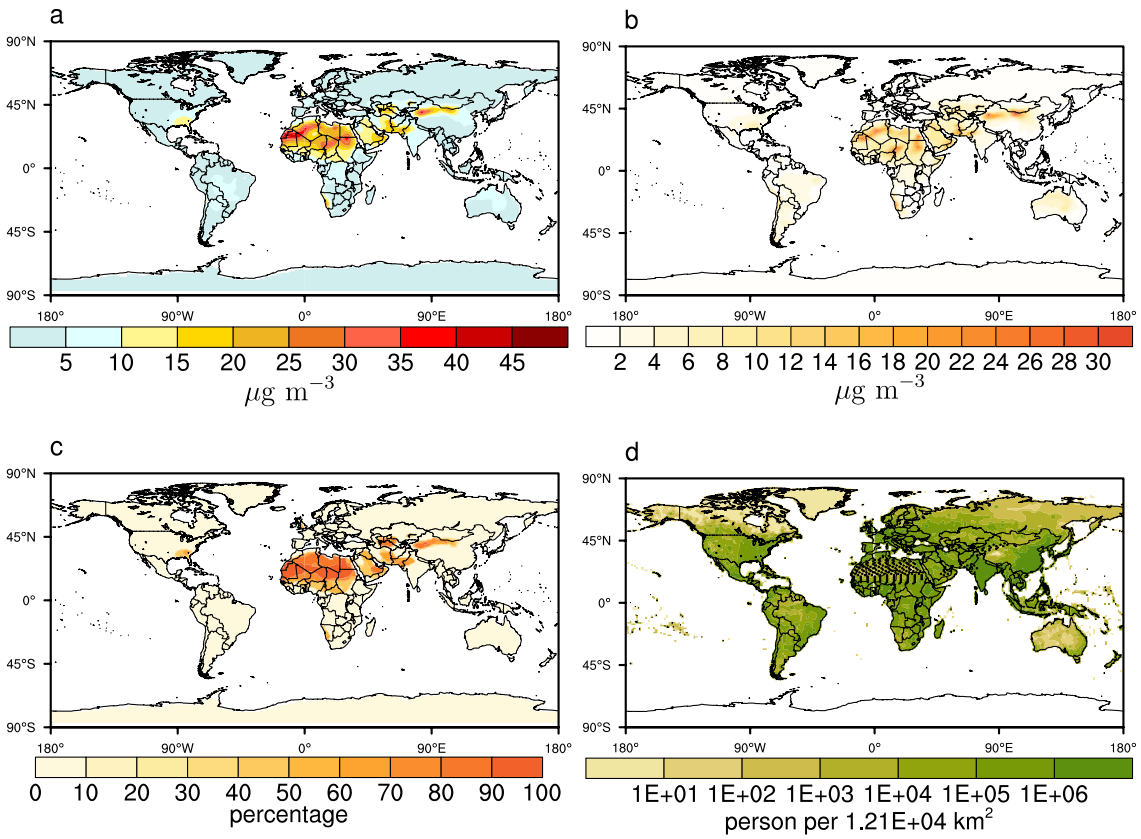

Figure 4.12: **Global map of pristine $PM_{2.5}$ based on CMIP5 models.** (a) Pre-industrial (PIC) climatological $PM_{2.5}$ based on the CMIP5 MMM using the reference method; (b) corresponding standard deviation across models of PIC $PM_{2.5}$ across models; and (c) the percentage of models that simulate a pristine $PM_{2.5}$ level that exceeds the WHO threshold for poor air quality (10 µg m^{-3}); and (d) world gridded population. Stippling in (d) represents MMM $PM_{2.5}$ in (a) exceeds the WHO threshold and at least two thirds ($\geq$ 66%) model agreement in (c). Units in (a, b) are µg m^{-3} and in (d) is person per 12100 km^2.

4.5 Summary

Result show that parts of Africa and Asia would have poor air quality with minimal human contributions is complicated by the fact that climate change has likely led to increasing emissions and concentrations of natural aerosol components, including dust (N. Zeng, 2003; Romm, 2011; Dai, 2011; Allen, Landuyt, & Rumbold, 2016; Allen, Hassan, Randles, & Su, 2019). Furthermore, anthropogenic land use change has also increased dust emissions (Ginoux, Prospero, Gill, Hsu, & Zhao, 2012). To estimate how natural emissions have changed from the pre-industrial to the present day, we compare pristine conditions with the present-day (1980-2014). In CMIP6 we observe that most land including Africa experiences a small increase in annual mean fine dust levels with a high degree of model agreement. However, some regions, including dusty Asian regions, experience a small decrease (Figure 4.13). Global land area-weighted average annual mean fine dust increases about $0.08 \pm 0.08\,\mu g\ m^{-3}$, and fine dust over dusty regions shows a larger increase at $0.26 \pm 0.25\,\mu g\ m^{-3}$. Similar results are obtained from CMIP5, wherein global land area-weighted average annual mean fine dust increases $0.005 \pm 0.007\,\mu g\ m^{-3}$ ($0.004 \pm 0.003\,\mu g\ m^{-3}$ over dusty regions). Similar analyses with fine sea salt, BVOC and DMS emissions show small changes since the pre-industrial (e.g., global land area-weighted average annual mean fine sea salt increases by $0.04 \pm 0.05\,\mu g\ m^{-3}$ and increases by $0.03 \pm 0.04\,\mu g\ m^{-3}$ over dusty regions in CMIP6). Overall, this evidence suggests that natural emission changes, particularly over dusty regions, are negligible from the pre-industrial to the early 21^{st} century. This means that natural emissions of aerosol species in our pristine $PM_{2.5}$ baseline remain nearly stationary over dusty regions from the pre-industrial to present day. Changes in emissions

of precursor gases (e.g., BVOC and SOA from forests) and other natural aerosol species currently are not quantified within CMIP5/6 models. Also, it is possible that continued climate change (i.e., more global warming) will exacerbate poor regional air quality due to increases in natural aerosols.

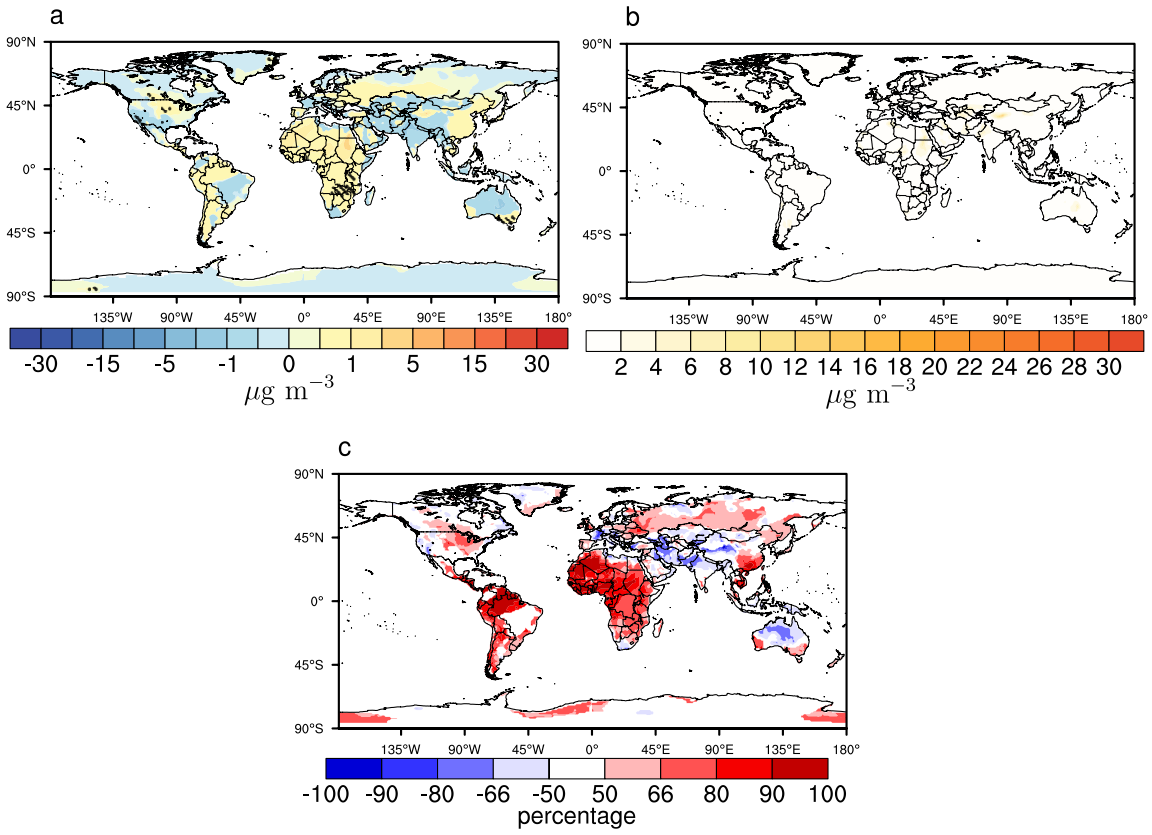

Figure 4.13: **Difference in fine dust between present-day and pristine conditions in CMIP6 models.** (a) Climatological CMIP6 MMM fine dust difference (present-day minus pristine conditions). Fine dust is estimated using the reference method. (b) Standard deviation of fine dust difference across CMIP6 models; and (c) CMIP6 model agreement (units of percentage) on the sign of differences in (a). Red shading represents increase in annual mean fine dust over historical period, and blue shading represents decrease. Units in (a, b) are μg m^{-3}.

Chapter 5

Conclusions & Discussions

5.1 Tropical Belt Width Proportionately More Sensitive to Aerosols Than Greenhouse Gases

In the chapter 2, I discussed the work to investigate the anthropogenic forcings role of the observed tropical belt expansion using idealized PDRMIP simulations. Results show that absorbing black carbon aerosol drives tropical expansion and scattering sulfate aerosol drives contraction. Black carbon, especially from Asia, is more efficient per unit radiative forcing than GHGs in driving tropical expansion, particularly in the Northern Hemisphere. Tropical belt expansion (contraction) is associated with an increase (decrease) in extratropical static stability induced by absorbing (scattering) aerosol.

It is difficult to use the PDRMIP idealized simulations to attribute the observed tropical belt expansion to anthropogenic forcings, particularly in the case of aerosols, which have changed spatially and temporally over recent decades (Hoesly et al., 2018). Owing

to the step change perturbations, PDRMIP integrations may be expected to yield a larger response relative (per unit forcing) to the real-word situation, since the system has had more time to equilibrate. Natural variability must also be considered, as it likely has played a crucial role in recent tropical expansion, while also disguising the fingerprint of anthropogenic forcing (Staten et al., 2018). Nonetheless, to illustrate the impact of these drivers on tropical widening in response to real-world changes in emissions, we follow (Tang et al., 2018) and approximate the impact of GHGs, BC and sulfate on tropical belt width by scaling the tropical belt response for each individual forcing and hemisphere according to: $\Delta\Phi_{scaled} = \Delta\Phi \times (\mathrm{ERF}_{1750-2011}/\mathrm{ERF}_{PDRMIP})$. We use a 1750-2011 GHG forcing of 2.83 W m^{-2} (90% confidence range of 2.54−3.12 W m^{-2}); a BC forcing of 0.40 W m^{-2} (0.05 to 0.80 W m^{-2}); and a sulfate forcing of −0.40 W m^{-2} (−0.60 to −0.20 W m^{-2}) (Myhre et al., 2013). Including aerosol-cloud interactions likely doubles the sulfate forcing (multi-model mean total sulfate forcing of −0.89 W m^{-2} with a range from −0.34 to −1.62 W m^{-2}) (Boucher et al., 2013). Furthermore, BC thermodynamic effects on clouds is not accounted for here (e.g., Allen, Amiri-Farahani, et al., 2019). Using these forcings and the ERF normalized tropical edge response for each driver and hemisphere (and assuming linearity to ERF), we estimate 1750-2011 NH tropical widening of 0.17° (0.15 to 0.19°) for GHGs; 0.38° (0.05 to 0.77°) for BC; and −0.1° (−0.15 to −0.05°) for sulfate. In the SH, we estimate 0.76° (0.69 to 0.84°) for GHGs; 0.16° (0.02 to 0.33°) for BC; and −0.08° (−0.11 to −0.04°) for sulfate. Cross-checking the GHG results with CMIP5 GHG-only tropical widening trends over a similar (but somewhat different) time period of 1850-2005 yields tropical widening of 0.23° in the NH and 0.70° in the SH (Allen & Ajoku, 2016). These values compare well

to the above GHG inferred estimates. Although this approach represents an approximation and possesses several caveats, it suggests that among the drivers, BC has likely caused a substantial portion NH tropical widening over the historical time period, though there is large uncertainty associated with BC forcing, with a 90% confidence range of 0.05 to 0.80 W m^{-2} (Ramanathan & Carmichael, 2008; Koch et al., 2009; Bond et al., 2013; Myhre et al., 2013). To the extent that BC and sulfate (or other reflecting aerosols) are co-emitted, however, this will partially offset the BC effect, especially when aerosol-cloud radiative effects are accounted for. Even so, the net effect of BC and sulfate on the NH tropical edge is similar to GHGs at 0.18° and 0.17°, respectively.

Considering the global plan for sustainable development goal proposed by United Nation, emission reductions in greenhouse gases, aerosols and their precursors are anticipated over the next few decades. Anthropogenic emissions would reduce gradually over time. The proportions of greenhouse gases and aerosols would fluctuate correspondingly. Future studies can focus on tropical belt width response to projected emission reductions. The project is focused on investigate the possible role of individual forcer in perturbing tropical belt width. Contribution of realistic single forcing, as well as regional emissions, on historic tropical belt width variations would be another interesting topic to study.

5.2 Strengthening of the Walker Circulation in Recent Decades and the Role of Natural Sea Surface Temperature Variability

In response to global warming, climate model simulations suggest a weakening of the Walker Circulation (WC), which is supported by long-term sea level pressure observations over the 20^{th} century. As shown in the chapter 3, the observations and multiple reanalyses yield the opposite trend from 1979 to present−a WC intensification. Atmosphere-only simulations driven by the real-world evolution of sea surface temperatures (SSTs) simulate this observed intensification, whereas coupled ocean atmosphere simulations do not. Thus, the recent WC intensification is related to real-world SST evolution. Assuming the multi-model mean SSTs from 20^{th} century coupled climate model simulations accurately represent the externally forced response, the observed SSTs can be decomposed into a forced and an unforced component. Idealized Community Atmosphere Model version 5 (CAM5) simulations driven by the unforced component of SSTs yield significant WC strengthening, whereas negligible WC changes occur when driven by the forced component of SSTs. Although coupled climate models may be deficient in their tropical response to anthropogenic warming, our results suggest natural SST variability, and in particular a La Niña-like SST pattern, is primarily responsible for the strengthening of the WC since 1979.

These results are consistent with previous studies, accompanied by a "faux pause" in global mean surface temperature increase (Kosaka & Xie, 2013). A negative phase of Pacific Decadal Oscillation (PDO) and more La Niña events is suggested to be an important

cause of the recent warming hiatus (Trenberth & Fasullo, 2013). Radiative effects from GHGs are superimposed with natural variability over this period. In recent decades, natural variability related to the observed La Niña-like SST pattern is found to be exceptionally strong, as compared to CMIP5 control runs (Amaya, Siler, Xie, & Miller, 2018b). We also note that as the PDO transitions to the opposite phase, it is likely that the strengthening of the Walker Circulation will weaken in the next decade. Furthermore, as GHGs continue to accumulate in the atmosphere, the forced signal will become more important through this century, implying weakening of the Walker Circulation. Regardless, future changes in the strength of the Walker Circulation will be heavily modulated by natural SST variability related to ENSO/PDO cycles.

However, studies have suggested that models may not accurately simulate the tropical response to warming (Stocker, 2014)−in some models, models simulate an overestimation response to greenhouse gases and/or other anthropogenic forcing. The observed cooling in the central and eastern tropical Pacific in recent decades cannot be well captured by coupled models (Kociuba & Power, 2015). If the true forced SST signal is more La Niña-like, a more positive SLP gradient trend would be estimated from CAM5 FSST. Therefore, a smaller positive trend of SLP gradient over equatorial Pacific would be simulated by CAM5 UFSST, and the intensification of the Walker Circulation by natural SST variations would be smaller. Study nudged observed central to eastern tropical Pacific sea surface temperature in climate model is able to reproduce the observed La Niña-like pattern, including the intensified Walker Circulation, highlighting the importance of decadal natural variability, associated a La Niña-like cooling (Kosaka & Xie, 2013). It seems to

suggest western tropical Pacific warms more than the east. It would be interesting to study the sensitivity of eastern vs western tropical pacific to different warming pattern.

5.3 An Implicit Air Quality Bias due to the State of Pristine Aerosol

In the chapter 4, climate models are used to construct a view of pre-industrial "pristine" air quality, including fine particulate matter with diameters less than 2.5 µm ($PM_{2.5}$). Under pristine conditions, $PM_{2.5}$ levels over regions in geographic proximity to dust sources, including parts of Africa and Asia, exceed World Health Organization air quality guidelines. We estimate that this pristine air pollution, which is unassociated with human activities, impacts up to about one billion people globally. The results show that natural aerosols, with strong geographic gradients, can lead to poor air quality over regions close to sources, and that in many areas no amount of anthropogenic emission reductions will result in clean air.

It is clear that our state-of-the-art climate models are tools that can also be used to provide us with a conservative picture of pre-industrial air quality, and that it is highly likely that in many parts of the world that "pristine" air would already be naturally polluted according to current standards. Given that today more than a billion people live in these areas, this presents a fairness and regulatory dilemma, and leads to many further questions. The results shed light on the shortcomings of establishing metrics that do not account for natural geospatial diversity, and also point to the potential importance of understanding and differentiating between natural and anthropogenic sources of air pollution. How toxic aerosol

species are to human health is related to exposure concentrations and duration (Brook et al., 2010). Although many toxic pollutants come from human emissions, studies suggest that long-term exposure to desert dust can also have adverse health effects (Giannadaki, Pozzer, & Lelieveld, 2014; De Longueville, Hountondji, Henry, & Ozer, 2010; Goudie, 2014). That said, in dusty regions there may in fact be harmful pollutants that warrant public awareness and/or regulation and do originate from human activities. In these cases, if natural levels of $PM_{2.5}$ exceed the best-practices standards, how should affected countries evaluate and/or strive to improve their air quality? This rhetorical question is particularly prescient for those countries that would like to achieve the Sustainable Development Goals (SDGs) as enumerated in the 2030 Agenda for Sustainable Development adopted by the United Nations in 2015. For example, 101 out of 247 SDGs indicators are environment related (Y. Zeng et al., 2020); household and ambient air pollution induced mortality rate is the measurement metric included in SDG Goal 3 (United Nations, 2015). In 2016, 50-89% of the total air-pollution related deaths over dusty Asian and African regions were due to ambient air pollution (United Nations, 2020). Our results suggest this is largely due to pristine aerosol. PM air pollution including $PM_{2.5}$ is a major contributor to African premature death, and deaths induced by natural aerosol (i.e., desert dust) dominate, surpassing deaths related to anthropogenic emissions (Bauer, Im, Mezuman, & Gao, 2019). Furthermore, small changes in dust concentration can have large changes in health outcomes. Over sub-Saharan Africa, every $10\,\mu g\ m^{-3}$ increase in $PM_{2.5}$ induced by imported Saharan dust can lead to a 24% (95% confidence range 10% to 35%) increase in infant mortality (Heft-Neal, Burney, Bendavid, Voss, & Burke, 2020). This means that pristine air pollution may be a

crucial concern for local communities with respect to health over these regions. Previous study points out the concentration and exposure duration to particular matters seem to be two important factors in causing an effect on human health. WHO air quality guideline mainly targets anthropogenic sourced particular matter. Whether one particular matter from anthropogenic source is more toxic than it from a natural source remains uncertain. It would be an interesting topic for future medical study.

Also, our findings support the need for increased access to measurements and monitoring within these regions (as would contribute to SDG 17 "Partnership for the Goals" and many targets therein), such that improved metrics, that perhaps identify particular anthropogenic tracers, can be implemented into both air quality monitoring and targeting. It must be noted that less economically developed countries tend to be clustered at the dusty latitudes highlighted in this study. It is likely that given a lack of equal voice and scientific resources, air quality guidelines targeting anthropogenic emissions represent a legacy of latitudinal bias that will continue to burden naturally dusty regions until more detailed understanding can level the playing field.

References

Adam, O., Grise, K. M., Staten, P., Simpson, I. R., Davis, S. M., Davis, N. A., ... Ming, A. (2018). The tropd software package (v1): standardized methods for calculating tropical-width diagnostics. *Geoscientific Model Development*, *11*(10), 4339–4357.

Adebiyi, A. A., & Kok, J. F. (2020). Climate models miss most of the coarse dust in the atmosphere. *Science advances*, *6*(15), eaaz9507.

Akritidis, D., Katragkou, E., Georgoulias, A. K., Zanis, P., Kartsios, S., Flemming, J., ... Eskes, H. (2020). A complex aerosol transport event over Europe during the 2017 Storm Ophelia in CAMS forecast systems: analysis and evaluation. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–31.

Allan, R., & Ansell, T. (2006). A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *Journal of Climate*, *19*(22), 5816-5842. Retrieved from `https://doi.org/10.1175/JCLI3937.1` doi: 10.1175/JCLI3937.1

Allen, R. J., & Ajoku, O. (2016). Future aerosol reductions and widening of the northern tropical belt. *Journal of Geophysical Research: Atmospheres*, *121*(12), 6765–6786.

Allen, R. J., Amiri-Farahani, A., Lamarque, J.-F., Smith, C., Shindell, D., Hassan, T., & Chung, C. E. (2019). Observationally constrained aerosol–cloud semi-direct effects. *npj Climate and Atmospheric Science*, *2*(1), 16, doi: 10.1038/s41612-019-0073-9.

Allen, R. J., Hassan, T., Randles, C. A., & Su, H. (2019). Enhanced land–sea warming contrast elevates aerosol pollution in a warmer world. *Nature Climate Change*, *9*(4), 300–305.

Allen, R. J., & Kovilakam, M. (2017a). The role of natural climate variability in recent tropical expansion. *Journal of Climate*, *30*(16), 6329–6350.

Allen, R. J., & Kovilakam, M. (2017b). The role of natural climate variability in recent tropical expansion. *Journal of Climate*, *30*(16), 6329–6350.

Allen, R. J., Landuyt, W., & Rumbold, S. T. (2016). An increase in aerosol burden and radiative effects in a warmer world. *Nature Climate Change*, *6*(3), 269.

Allen, R. J., & Luptowitz, R. (2017). El Niño-like teleconnection increases California precipitation in response to warming. *Nature communications*, *8*, 16055.

Allen, R. J., Norris, J. R., & Kovilakam, M. (2014). Influence of anthropogenic aerosols and the pacific decadal oscillation on tropical belt width. *Nature Geoscience*, *7*(4), 270.

Allen, R. J., & Sherwood, S. C. (2011). The impact of natural versus anthropogenic aerosols on atmospheric circulation in the community atmosphere model. *Climate dynamics*, *36*(9-10), 1959–1978.

Allen, R. J., Sherwood, S. C., Norris, J. R., & Zender, C. S. (2012a). The equilibrium response to idealized thermal forcings in a comprehensive gcm: Implications for recent tropical expansion. *Atmospheric Chemistry and Physics*, *12*(10), 4795–4816.

Allen, R. J., Sherwood, S. C., Norris, J. R., & Zender, C. S. (2012b). Recent northern hemisphere tropical expansion primarily driven by black carbon and tropospheric ozone. *Nature*, *485*(7398), 350.

Allen, R. J., Turnock, S., Nabat, P., Neubauer, D., Lohmann, U., Olivié, D., ... others (2020). Climate and air quality impacts due to mitigation of non-methane near-term climate forcers. *Atmospheric Chemistry and Physics*, *20*(16), 9641–9663.

Amaya, D. J., Siler, N., Xie, S.-P., & Miller, A. J. (2018a). The interplay of internal and forced modes of hadley cell expansion: lessons from the global warming hiatus. *Climate dynamics*, *51*(1-2), 305–319.

Amaya, D. J., Siler, N., Xie, S.-P., & Miller, A. J. (2018b). The interplay of internal and forced modes of Hadley Cell expansion: lessons from the global warming hiatus. *Climate Dynamics*, *51*(1-2), 305–319.

Ambient Air Quality Standards, GB 3095-2012. (2012). Ministry of Environmental Protection(MEP) and AQSIQ. (issued by the on Feb. 29, 2012, effective Jan. 1, 2016)

Apte, J. S., Marshall, J. D., Cohen, A. J., & Brauer, M. (2015). Addressing global mortality from ambient $PM_{2.5}$. *Environmental science & technology*, *49*(13), 8057–8066.

Archibald, A. T., O'Connor, F. M., Abraham, N. L., Archer-Nicholls, S., Chipperfield, M. P., Dalvi, M., ... Zeng, G. (2020). Description and evaluation of the UKCA stratosphere–troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1. *Geoscientific Model Development*, *13*(3), 1223–1266.

Arora, V., Scinocca, J., Boer, G., Christian, J., Denman, K., Flato, G., ... Merryfield, W. (2011). Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters*, *38*(5).

Bauer, S. E., Im, U., Mezuman, K., & Gao, C. Y. (2019). Desert Dust, Industrialization, and Agricultural Fires: Health Impacts of Outdoor Air Pollution in Africa. *Journal of Geophysical Research: Atmospheres*, *124*(7), 4104-4120.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., ... Stevens, B. (2019). Bounding global aerosol radiative forcing of climate change. *Reviews of Geophysics*, *57*, doi: 10.1029/2019RG000660.

Birner, T., Davis, S., & Seidel, D. (2014). Earth's tropical belt. *Phys. Today*, *67*(12), 38–44.

Bond, T. C., Doherty, S. J., Hahey, D. W., & et al. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *J. Geophys. Res.*, *118*, 5380-5552, doi:10.1002/jgrd.50171.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... Zhang, X. Y. (2013). Clouds and aerosols. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (pp. 571–657). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781107415324.016

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., ... Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*, *121*(21), 2331–2378.

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *The lancet*, *360*(9341), 1233–1242.

Buchard, V., da Silva, A., Randles, C., Colarco, P., Ferrare, R., Hair, J., ... Winker, D. (2016). Evaluation of the surface $PM_{2.5}$ in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States. *Atmospheric environment*, *125*, 100–111.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., ... Spadaro, J. V. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences*, *115*(38), 9592–9597. doi: 10.1073/pnas.1803222115

Cai, W., Cowan, T., & Thatcher, M. (2012). Rainfall reductions over southern hemisphere semi-arid regions: the role of subtropical dry zone expansion. *Scientific reports*, *2*, 702.

Cai, W., Santoso, A., Wang, G., Yeh, S.-W., An, S.-I., Cobb, K. M., ... others (2015). ENSO and greenhouse warming. *Nature Climate Change*, *5*(9), 849.

Caldeira, K., & Myhrvold, N. P. (2013). Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration. *Environmental Research Letters*, *8*(3), 034039, doi: 10.1088/1748-9326/8/3/034039.

Cane, M. A. (2005). The evolution of El Niño, past and future. *Earth and Planetary Science Letters*, *230*(3-4), 227–240.

Carslaw, K. S., Gordon, H., Hamilton, D. S., Johnson, J. S., Regayre, L. A., Yoshioka, M., & Pringle, K. J. (2017). Aerosols in the pre-industrial atmosphere. *Current Climate Change Reports*, *3*(1), 1–15.

Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., ... Pierce, J. R. (2013). Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, *503*(7474), 67.

Center for International Earth Science Information Network-CIESIN-Columbia University. (2018). Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11.

Chung, C. E., Jung-Eun, C., Lee, Y., van Noije, T., Jeoung, H., Ha, K.-J., & Marks, M. (2016). Global fine-mode aerosol radiative effect, as constrained by comprehensive observations. *Atmospheric Chemistry and Physics*, *16*(13), 8071.

Ciarelli, G., Aksoyoglu, S., Crippa, M., Jimenez, J.-L., Nemitz, E., Sellegri, K., ... Prévôt, A. S. H. (2016). Evaluation of European air quality modelled by CAMx including the volatility basis set scheme. *Atmospheric Chemistry and Physics*, *16*(16), 10313–10332.

Coats, S., & Karnauskas, K. (2017). Are simulated and observed twentieth century tropical pacific sea surface temperature trends significant relative to internal variability? *Geophysical Research Letters*, *44*(19), 9928–9937.

Colarco, P., da Silva, A., Chin, M., & Diehl, T. (2010). Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research: Atmospheres*, *115*(D14).

Collins, M., An, S.-I., Cai, W., Ganachaud, A., Guilyardi, E., Jin, F.-F., ... others (2010). The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geoscience*, *3*(6), 391.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., ... others (2017). Aerchemmip: quantifying the effects of chemistry and aerosols in cmip6.

Dai, A. (2011). Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(1), 45–65.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916.

Davis, N., & Birner, T. (2017). On the discrepancies in tropical belt expansion between reanalyses and climate models and among tropical belt width metrics. *Journal of Climate*, *30*(4), 1211–1231.

Davis, N. A., Seidel, D. J., Birner, T., Davis, S. M., & Tilmes, S. (2016a). Changes in the width of the tropical belt due to simple radiative forcing changes in the GeoMIP simulations. *Atmospheric Chemistry and Physics*, *16*(15), 10083–10095, doi: 10.5194/acp-16-10083-2016.

Davis, N. A., Seidel, D. J., Birner, T., Davis, S. M., & Tilmes, S. (2016b). Changes in the width of the tropical belt due to simple radiative forcing changes in the geomip simulations. *Atmospheric Chemistry and Physics*, *16*(15), 10083–10095.

Davis, S. M., & Rosenlof, K. H. (2012). A multidiagnostic intercomparison of tropical-width time series using reanalyses and satellite observations. *Journal of Climate*, *25*(4), 1061–1078.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., ... others (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553–597.

De Longueville, F., Hountondji, Y.-C., Henry, S., & Ozer, P. (2010). What do we know about effects of desert dust on air quality and human health in West Africa compared to other regions? *Science of the total environment*, *409*(1), 1–8.

DiNezio, P., Clement, A., & Vecchi, G. (2010). Reconciling differing views of tropical Pacific climate change. *Eos, Transactions American Geophysical Union*, *91*(16), 141–142.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., ... Zeng, F. (2011). The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, *24*(13), 3484–3519.

Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., ... Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate dynamics*, *40*(9-10), 2123–2165.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., ... Zadeh, N. (2012). GFDL ESM2 global coupled climate–carbon earth system models. Part I: Physical formulation and baseline simulation characteristics. *Journal of climate*, *25*(19), 6646–6665.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., ... PŽtron, G. (2020). The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(4), e2019MS001882.

England, M. H., McGregor, S., Spence, P., Meehl, G. A., Timmermann, A., Cai, W., ... Santoso, A. (2014). Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change*, *4*(3), 222.

European Council Directive. (2008). Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. *Official Journal of the European Union*, *L*(152), 1-44.

Evan, A. T., Flamant, C., Fiedler, S., & Doherty, O. (2014). An analysis of aeolian dust in climate models. *Geophysical Research Letters*, *41*(16), 5996–6001.

Evans, J., van Donkelaar, A., Martin, R. V., Burnett, R., Rainham, D. G., Birkett, N. J., & Krewski, D. (2013). Estimates of global mortality attributable to particulate air pollution using satellite imagery. *Environmental research*, *120*, 33–42.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development (Online)*, *9*(LLNL-JRNL-736881).

Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., ... Zeng, G. (2012). Global air quality and climate. *Chemical Society Reviews*, *41*(19), 6663–6683.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., ... Schulz, M. (2016). Recommendations for diagnosing effective radiative forcing from climate models for cmip6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12–460.

Fyfe, J. C., Gillett, N. P., & Zwiers, F. W. (2013). Overestimated global warming over the past 20 years. *Nature Climate Change*, *3*(9), 767.

Garfinkel, C. I., Waugh, D. W., & Polvani, L. M. (2015). Recent hadley cell expansion: The role of internal atmospheric variability in reconciling modeled and observed trends. *Geophysical Research Letters*, *42*(24), 10–824.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... others (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... Zhao, B. (2017, 06). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419-5454.

Giannadaki, D., Pozzer, A., & Lelieveld, J. (2014). Modeled global effects of airborne desert dust on air quality and premature mortality. *Atmospheric Chemistry & Physics*, *14*(2).

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C., & Zhao, M. (2012). Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. *Reviews of Geophysics*, *50*(3).

Glotfelty, T., He, J., & Zhang, Y. (2017). Impact of future climate policy scenarios on air quality and aerosol-cloud interactions using an advanced version of CESM/CAM5: Part I. Model evaluation for the current decadal simulations. *Atmospheric environment*, *152*, 222–239.

Goudie, A. S. (2014). Desert dust and human health disorders. *Environment international*, *63*, 101–113.

Grassi, B., Redaelli, G., Canziani, P. O., & Visconti, G. (2012). Effects of the pdo phase on the tropical belt width. *Journal of Climate*, *25*(9), 3282–3290.

Grise, K. M., Davis, S. M., Simpson, I. R., Waugh, D. W., Fu, Q., Allen, R. J., ... others (2019). Recent tropical expansion: natural variability or forced response? *Journal of Climate*, *32*(5), 1551–1571.

Grise, K. M., Davis, S. M., Staten, P. W., & Adam, O. (2018). Regional and seasonal characteristics of the recent expansion of the tropics. *Journal of Climate*, *31*(17), 6839–6856.

Grise, K. M., & Polvani, L. M. (2016). Is climate sensitivity related to dynamical sensitivity? *Journal of Geophysical Research: Atmospheres*, *121*(10), 5159–5176.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., ... Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*(5), 2197–2244.

Hamilton, D. S., Lee, L. A., Pringle, K. J., Reddington, C. L., Spracklen, D. V., & Carslaw, K. S. (2014). Occurrence of pristine aerosol environments on a polluted planet. *Proceedings of the National Academy of Sciences*, *111*(52), 18466–18471.

Heft-Neal, S., Burney, J., Bendavid, E., Voss, K. K., & Burke, M. (2020). Dust pollution from the sahara and african infant mortality. *Nature Sustainability*, 1–9.

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and Performance of GFDL's CM4.0 Climate Model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691-3727.

Held, I. M., & Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of climate*, *19*(21), 5686–5699.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., ... others (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the community emissions data system (ceds). *Geoscientific Model Development (Online)*, *11*(PNNL-SA-123932).

Horinouchi, T., Matsumura, S., Ose, T., & Takayabu, Y. N. (2019). Jet–precipitation relation and future change of the mei-yu–baiu rainband and subtropical jet in cmip5 coupled gcm simulations. *Journal of Climate*, *32*(8), 2247–2259.

Hu, Y., & Fu, Q. (2007). Observed poleward expansion of the hadley circulation since 1979. *Atmospheric Chemistry and Physics*, *7*(19), 5229–5236.

Hu, Y., Tao, L., & Liu, J. (2013). Poleward expansion of the hadley circulation in cmip5 simulations. *Advances in Atmospheric Sciences*, *30*(3), 790–795.

Im, U., Christensen, J. H., Geels, C., Hansen, K. M., Brandt, J., Solazzo, E., ... Galmarini, S. (2018). Influence of anthropogenic emissions and boundary conditions on multi-model simulations of major air pollutants over Europe and North America in the framework of AQMEII3. *Atmospheric chemistry and physics*, *18*(12), 8929.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., ... Suttie, M. (2019). The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, *19*(6), 3515–3556.

Johnson, B., Haywood, J., & Hawcroft, M. (2019). Are changes in atmospheric circulation important for black carbon aerosol impacts on clouds, precipitation and radiation? *Journal of Geophysical Research: Atmospheres*.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... Joseph, D. (1996, MAR). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437-471. doi: {10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2}

Karnauskas, K. B., Seager, R., Kaplan, A., Kushnir, Y., & Cane, M. A. (2009). Observed strengthening of the zonal sea surface temperature gradient across the equatorial Pacific Ocean. *Journal of Climate*, *22*(16), 4316–4321.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., ... Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333-1349. Retrieved from `https://doi.org/10.1175/BAMS-D-13-00255.1` doi: 10.1175/BAMS-D-13-00255.1

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., ... others (2020). GISS-E2. 1: Configurations and climatology. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS002025.

Kennedy, J., Rayner, N., Smith, R., Parker, D., & Saunby, M. (2011a). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research: Atmospheres*, *116*(D14).

Kennedy, J., Rayner, N., Smith, R., Parker, D., & Saunby, M. (2011b). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *Journal of Geophysical Research: Atmospheres*, *116*(D14).

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., ... others (2006). An AeroCom initial assessment–optical properties in aerosol component modules of global models. *Atmospheric Chemistry and Physics*, *6*(7), 1815–1834.

Kirkevåg, A., Grini, A., Olivié, D., Seland, Ø., Alterskjær, K., Hummel, M., ... Iversen, T. (2018). A production-tagged aerosol module for Earth system models, OsloAero5. 3-extensions and updates for CAM5. 3-Oslo. *Geoscientific Model Development*, *11*(10), 3945–3982.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... others (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., ... Zhao, Y. (2009). Evaluation of black carbon estimations in global aerosol models. *Atmos. Chem. Phys.*, *9*(22), 9001–9026.

Kociuba, G., & Power, S. B. (2015). Inability of CMIP5 models to simulate recent strengthening of the Walker circulation: Implications for projections. *Journal of Climate*, *28*(1), 20–35.

Kok, J. F. (2011). A scaling theory for the size distribution of emitted dust aerosols suggests climate models underestimate the size of the global dust cycle. *Proceedings of the National Academy of Sciences*, *108*(3), 1016–1021.

Kok, J. F., Ridley, D. A., Zhou, Q., Miller, R. L., Zhao, C., Heald, C. L., ... Haustein, K. (2017). Smaller desert dust cooling effect estimated from analysis of dust size and abundance. *Nature Geoscience*, *10*(4), 274.

Kok, J. F., Ward, D. S., Mahowald, N. M., & Evan, A. T. (2018). Global and regional importance of the direct dust-climate feedback. *Nature communications*, *9*(1), 241.

Kosaka, Y., & Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, *501*(7467), 403.

Kovilakam, M., & Mahajan, S. (2015). Black carbon aerosol-induced northern hemisphere tropical expansion. *Geophysical Research Letters*, *42*(12), 4964–4972.

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhao, M. (2018). *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.8597`

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., ... van Vuuren, D. P. (2010). Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, *10*(15), 7017-7039. (hal-00458149f) doi: ff10.5194/acp-10-7017-2010

Landrigan, P. J., Fuller, R., Acosta, N. J., Adeyi, O., Arnold, R., Baldé, A. B., ... others (2018). The lancet commission on pollution and health. *The lancet*, *391*(10119), 462–512.

L'Heureux, M. L., Lee, S., & Lyon, B. (2013). Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nature Climate Change*, *3*(6), 571.

Liu, J., Wang, B., Cane, M. A., Yim, S.-Y., & Lee, J.-Y. (2013). Divergent global precipitation changes induced by natural versus anthropogenic forcing. *Nature*, *493*(7434), 656.

Liu, L., Shawki, D., Voulgarakis, A., Kasoar, M., Samset, B., Myhre, G., ... others (2018). A pdrmip multimodel study on the impacts of regional aerosol forcings on global and regional precipitation. *Journal of Climate*, *31*(11), 4429–4447.

Lu, J., Chen, G., & Frierson, D. M. (2008). Response of the zonal mean atmospheric circulation to el niño versus global warming. *Journal of Climate*, *21*(22), 5835–5851.

Lu, J., Vecchi, G. A., & Reichler, T. (2007). Expansion of the hadley cell under global warming. *Geophysical Research Letters*, *34*(6).

Lucas, C., Timbal, B., & Nguyen, H. (2014). The expanding tropics: a critical assessment of the observational and modeling studies. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(1), 89–112.

Ma, J., Xu, J., & Qu, Y. (2020). Evaluation on the surface PM2. 5 concentration over China mainland from NASA's MERRA-2. *Atmospheric Environment*, 117666.

Ma, S., & Zhou, T. (2016). Robust strengthening and westward shift of the tropical Pacific Walker circulation during 1979–2012: a comparison of 7 sets of reanalysis data and 26 CMIP5 models. *Journal of Climate*, *29*(9), 3097–3118.

Mantsis, D. F., Sherwood, S., Allen, R., & Shi, L. (2017). Natural variations of tropical width and recent trends. *Geophysical Research Letters*, *44*(8), 3825–3832.

McNider, R. T., & Pour-Biazar, A. (2020). Meteorological modeling relevant to mesoscale and regional air quality applications: a review. *Journal of the Air & Waste Management Association*, *70*(1), 2–43.

Meng, Q., Latif, M., Park, W., Keenlyside, N. S., Semenov, V. A., & Martin, T. (2012). Twentieth century Walker Circulation change: data analysis and model experiments. *Climate Dynamics*, *38*(9), 1757–1773. Retrieved from `https://doi.org/10.1007/s00382-011-1047-8` doi: 10.1007/s00382-011-1047-8

Michou, M., Nabat, P., Saint-Martin, D., Bock, J., Decharme, B., Mallet, M., ... Voldoire, A. (2020). Present-day and historical aerosol and ozone characteristics in CNRM CMIP6 simulations. *Journal of Advances in Modeling Earth Systems*, *12*(1).

Min, S.-K., & Son, S.-W. (2013). Multimodel attribution of the southern hemisphere hadley cell widening: Major role of ozone depletion. *Journal of Geophysical Research: Atmospheres*, *118*(7), 3007–3015.

Molod, A., Takacs, L., Suarez, M., & Bacmeister, J. (2015). Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geoscientific Model Development*, *8*(5), 1339.

Mulcahy, J. P., Johnson, C., Jones, C. G., Povey, A. C., Scott, C. E., Sellar, A., ... Yool, A. (2020). Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3. 1 CMIP6 historical simulations. *Geoscientific Model Development Discussions*, 1–59.

Myhre, G., Forster, P., Samset, B., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S., ... others (2017). Pdrmip: A precipitation driver and response model intercomparison projectñprotocol and preliminary results. *Bulletin of the American Meteorological Society*, *98*(6), 1185–1198.

Myhre, G., Shindell, D., Bréon, F., Collins, W., Fuglestvedt, J., Huang, J., ... others (2013). *Anthropogenic and natural radiative forcing, climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change, 659–740.* Cambridge: Cambridge University Press.

Navinya, C. D., Vinoj, V., & Pandey, S. K. (2020). Evaluation of PM2. 5 Surface Concentrations Simulated by NASA's MERRA Version 2 Aerosol Reanalysis over India and its Relation to the Air Quality Index. *Aerosol and Air Quality Research*, *20*, 1329–1339.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., ... others (2010). Description of the NCAR community atmosphere model (CAM 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, *1*(1), 1-12.

Nguyen, H., Evans, A., Lucas, C., Smith, I., & Timbal, B. (2013). The hadley circulation in reanalyses: Climatology, variability, and change. *Journal of Climate*, *26*(10), 3357–3376.

Pan, X., Chin, M., Gautam, R., Bian, H., Kim, D., Colarco, P. R., ... Bellouin, N. (2015). A multi-model evaluation of aerosols over South Asia: common problems and possible causes. *Atmospheric Chemistry and Physics*, *15*(10), 5903–5928.

Petkova, E. P., Jack, D. W., Volavka-Close, N. H., & Kinney, P. L. (2013). Particulate matter pollution in African cities. *Air Quality, Atmosphere & Health*, *6*(3), 603–614.

Plesca, E., Grützun, V., & Buehler, S. A. (2018). How robust is the weakening of the Pacific Walker circulation in CMIP5 idealized transient climate simulations? *Journal of Climate*, *31*(1), 81–97.

Polvani, L. M., Waugh, D. W., Correa, G. J., & Son, S.-W. (2011). Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the southern hemisphere. *Journal of Climate*, *24*(3), 795–812.

Pope III, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: lines that connect. *Journal of the air & waste management association*, *56*(6), 709–742.

Power, S. B., & Kociuba, G. (2011a). The impact of global warming on the Southern Oscillation Index. *Climate dynamics*, *37*(9-10), 1745–1754.

Power, S. B., & Kociuba, G. (2011b). What caused the observed twentieth-century weakening of the Walker circulation? *Journal of Climate*, *24*(24), 6501–6514.

Power, S. B., & Smith, I. N. (2007). Weakening of the Walker Circulation and apparent dominance of El Niño both reach record levels, but has ENSO really changed? *Geophysical Research Letters*, *34*(18).

Provençal, S., Buchard, V., da Silva, A. M., Leduc, R., & Barrette, N. (2017). Evaluation of PM surface concentrations simulated by Version 1 of NASA's MERRA Aerosol Reanalysis over Europe. *Atmospheric pollution research*, *8*(2), 374–382.

Ramanathan, V., & Carmichael, G. (2008). Global and regional climate changes due to black carbon. *Nature Geosci.*, *1*, 221-227.

Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., ... Flynn, C. J. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *Journal of Climate*, *30*(17), 6823–6850.

Rao, S., Klimont, Z., Smith, S. J., Van Dingenen, R., Dentener, F., Bouwman, L., ... others (2017). Future air pollution in the shared socio-economic pathways. *Global Environmental Change*, *42*, 346–358.

Reddington, C. L., Carslaw, K. S., Stier, P., Schutgens, N., Coe, H., Liu, D., ... Zhang, Q. (2017). The Global Aerosol Synthesis and Science Project (GASSP): Measurements and Modeling to Reduce Uncertainty. *Bulletin of the American Meteorological Society*, *98*(9).

Rienecker, M. M., Suarez, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H., ... Nielsen, J. (2008). The GEOS-5 Data Assimilation System: Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0.

Rind, D., Orbe, C., Jonas, J., Nazarenko, L., Zhou, T., Kelley, M., ... Tausnev, N. (2020). GISS Model E2.2: A climate model optimized for the middle atmosphere. Part 1: Model structure, climatology, variability and climate sensitivity. *Journal of Geophysical Research: Atmospheres*, *125*(10), e2019JD032204.

Romm, J. (2011). The next dust bowl. *Nature*, *478*(7370), 450–451.

Rotstayn, L., Jeffrey, S. J., Collier, M. A., Dravitzki, S., Hirst, A., Syktus, J., & Wong, K. (2012). Aerosol-and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmospheric Chemistry and Physics*, *12*(14), 6377–6404.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., ... others (2010). The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058.

Samset, B., Myhre, G., Forster, P., Hodnebrog, Ø., Andrews, T., Faluvegi, G., ... others (2016). Fast and slow precipitation responses to individual climate forcers: A pdrmip multimodel study. *Geophysical Research Letters*, *43*(6), 2782–2791.

Sandeep, S., Stordal, F., Sardeshmukh, P. D., & Compo, G. P. (2014). Pacific Walker Circulation variability in coupled and uncoupled climate models. *Climate dynamics*, *43*(1-2), 103–117.

Scheff, J., & Frierson, D. M. (2012). Robust future precipitation declines in cmip5 largely reflect the poleward expansion of model subtropical dry zones. *Geophysical Research Letters*, *39*(18).

Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., ... Yao, M.-S. (2006). Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *Journal of Climate*, *19*(2), 153–192.

Schneider, T. (2006). The general circulation of the atmosphere. *Annual Review of Earth and Planetary Sciences*, *34*(1), 655–688, doi: 10.1146/annurev.earth.34.031405.125144.

Schwendike, J., Govekar, P., Reeder, M. J., Wardle, R., Berry, G. J., & Jakob, C. (2014). Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations. *Journal of Geophysical Research: Atmospheres*, *119*(3), 1322–1339, doi: 10.1002/2013JD020742.

Seager, R., Cane, M., Henderson, N., Lee, D.-E., Abernathey, R., & Zhang, H. (2019). Strengthening tropical pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nature Climate Change*, *9*(7), 517.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., ... others (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182–4227.

Seidel, D. J., Fu, Q., Randel, W. J., & Reichler, T. J. (2008). Widening of the tropical belt in a changing climate. *Nature geoscience*, *1*(1), 21.

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., ... Schulz, M. (2020). The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations. *Geoscientific Model Development Discussions*, *2020*, 1–68.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., ... Zerroukat, M. (2019). UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513-4558.

Shaddick, G., Thomas, M., Mudu, P., Ruggeri, G., & Gumy, S. (2020). Half the worldÕs population are exposed to increasing air pollution. *npj Climate and Atmospheric Science*, *3*(1), 1–5.

Shaw, T. A. (2014). On the role of planetary-scale waves in the abrupt seasonal transition of the Northern Hemisphere general circulation. *Journal of the Atmospheric Sciences*, *71*(5), 1724-1746, doi: 10.1175/JAS-D-13-0137.1.

Shaw, T. A., & Voigt, A. (2015). Tug of war on summertime circulation between radiative forcing and sea surface warming. *Nature Geoscience*, *8*(7), 560–566, doi: 10.1038/ngeo2449.

Shen, Z., & Ming, Y. (2018). The influence of aerosol absorption on the extratropical circulation. *Journal of Climate*, *31*(15), 5961–5975.

Silva, R. A., West, J. J., Zhang, Y., Anenberg, S. C., Lamarque, J.-F., Shindell, D. T., ... others (2013). Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change. *Environmental Research Letters*, *8*(3), 034005.

Smith, D. M., Booth, B. B., Dunstone, N. J., Eade, R., Hermanson, L., Jones, G. S., ... Thompson, V. (2016). Role of volcanic and anthropogenic aerosols in the recent global surface warming slowdown. *Nature Climate Change*, *6*(10), 936.

Smithson, P. A. (2002). IPCC, 2001: climate change 2001: the scientific basis. Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by JT Houghton, Y. Ding, DJ Griggs, M. Noguer, PJ van der Linden, X. Dai, K. Maskell and CA Johnson (eds). Cambridge University Press, Cambridge, UK, and New York, USA, 2001. No. of pages: 881. Price£ 34.95, US 49.95,ISBN0-521-01495-6(paperback).£90.00,US 130.00, ISBN 0-521-80767-0 (hardback). *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *22*(9), 1144–1144.

Soden, B. J., Jackson, D. L., Ramaswamy, V., Schwarzkopf, M., & Huang, X. (2005). The radiative signature of upper tropospheric moistening. *Science*, *310*(5749), 841–844.

Sohn, B., Yeh, S.-W., Schmetz, J., & Song, H.-J. (2013). Observational evidences of Walker circulation change over the last 30 years contrasting with GCM results. *Climate Dynamics*, *40*(7-8), 1721–1732.

Solazzo, E., Bianconi, R., Hogrefe, C., Curci, G., Tuccella, P., Alyuz, U., ... Galmarini, S. (2017). Evaluation and error apportionment of an ensemble of atmospheric chemistry transport modeling systems: multivariable temporal and spatial breakdown. *Atmospheric chemistry and physics*, *17*(4), 3001.

Solomon, A., & Newman, M. (2012). Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record. *Nature Climate Change*, *2*(9), 691.

Solomon, S., Qin, D., Manning, M., Alley, R., Berntsen, T., Bindoff, N., ... others (2007). Technical Summary, Climate Change 2007: The Physical Science Basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*.

Son, S.-W., Tandon, N. F., Polvani, L. M., & Waugh, D. W. (2009). Ozone hole and southern hemisphere climate change. *Geophysical Research Letters*, *36*(15).

Sousa, P., Trigo, R., Aizpurua, P., Nieto, R., Gimeno, L., & Garcia-Herrera, R. (2011). Trends and extremes of drought indices throughout the 20th century in the mediterranean. *Natural Hazards and Earth System Sciences*, *11*(1), 33–51.

Spracklen, D., Pringle, K., Carslaw, K., Chipperfield, M., & Mann, G. (2005). A global off-line model of size-resolved aerosol microphysics: I. Model development and prediction of aerosol properties. *Atmospheric Chemistry and Physics*, *5*(8), 2227–2252.

Staten, P. W., Grise, K. M., Davis, S. M., Karnauskas, K., & Davis, N. (2019). Regional widening of tropical overturning: Forced change, natural variability, and recent trends. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6104–6119, doi: 10.1029/2018JD030100.

Staten, P. W., Lu, J., Grise, K. M., Davis, S. M., & Birner, T. (2018). Re-examining tropical expansion. *Nature Climate Change*, *8*(9), 768–775.

Stjern, C. W., Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø., Andrews, T., ... others (2017). Rapid adjustments cause weak surface temperature response to increased black carbon concentrations. *Journal of Geophysical Research: Atmospheres*, *122*(21), 11–462.

Stocker, T. (2014). Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change.

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., & Nakajima, T. (2000). Global three-dimensional simulation of aerosol optical thickness distribution of various origins. *Journal of Geophysical Research: Atmospheres*, *105*(D14), 17853–17873.

Tandon, N. F., Gerber, E. P., Sobel, A. H., & Polvani, L. M. (2013). Understanding hadley cell expansion versus contraction: Insights from simplified models and implications for recent observations. *Journal of Climate*, *26*(12), 4304–4321.

Tang, T., Shindell, D., Samset, B. H., Boucher, O., Forster, P. M., Hodnebrog, Ø., ... Takemura, T. (2018). Dynamical response of Mediterranean precipitation to greenhouse gases and aerosols. *Atmospheric Chemistry and Physics*, *18*(11), 8439–8452, doi: 10.5194/acp-18-8439-2018.

Tao, L., Hu, Y., & Liu, J. (2016). Anthropogenic forcing on the hadley circulation in cmip5 simulations. *Climate dynamics*, *46*(9-10), 3337–3350.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012a). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93*(4), 485-498. Retrieved from `https://doi.org/10.1175/BAMS-D-11-00094.1` doi: 10.1175/BAMS-D-11-00094.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012b). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

Tilmes, S., Hodzic, A., Emmons, L. K., Mills, M. J., Gettelman, A., Kinnison, D. E., ... Liu, X. (2019). Climate forcing and trends of organic aerosols in the Community Earth System Model (CESM2). *Journal of Advances in Modeling Earth Systems*, *11*(12), 4323–4351.

Tokinaga, H., Xie, S.-P., Deser, C., Kosaka, Y., & Okumura, Y. M. (2012). Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature*, *491*(7424), 439.

Trenberth, K. E., Dai, A., Van Der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*(1), 17–22.

Trenberth, K. E., & Fasullo, J. T. (2013). An apparent hiatus in global warming? *Earth's Future*, *1*(1), 19–32.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., Artaxo, P., Bahadur, R., ... Zhang, X. (2014). The AeroCom evaluation and intercomparison of organic aerosol in global models. *Atmospheric Chemistry and Physics*, *14*(19), 10845–10895.

Tsigaridis, K., Koch, D., & Menon, S. (2013). Uncertainties and importance of sea spray composition on aerosol direct and indirect effects. *Journal of Geophysical Research: Atmospheres*, *118*(1), 220–235.

Turnock, S. T., Allen, R. J., Andrews, M., Bauer, S. E., Emmons, L., Good, P., ... Zhang, J. (2020). Historical and future changes in air pollutants from CMIP6 models. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–40.

Ukhov, A., Mostamandi, S., da Silva, A., Flemming, J., Alshehri, Y., Shevchenko, I., & Stenchikov, G. (2020). Assessment of natural and anthropogenic aerosol air pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem model simulations. *Atmospheric Chemistry and Physics*, *20*(15), 9281–9310.

United Nations. (2015). Sustainable Development Goals: 17 Goals to transform our world.

United Nations. (2020). Supplementary Information for Progress towards the Sustainable Development Goals 2020.

US EPA. (2016). *NAAQS Table.* US EPA.

van Donkelaar, A., Martin, R. V., Brauer, M., & Boys, B. L. (2015). Use of Satellite Observations for Long-Term Exposure Assessment of Global Concentrations of Fine Particulate Matter. *Environmental Health Perspectives*, *123*(2), 135-143.

Van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., ... Winker, D. M. (2016). Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environmental science & technology*, *50*(7), 3762–3772.

Vecchi, G. A., Soden, B. J., Wittenberg, A. T., Held, I. M., Leetmaa, A., & Harrison, M. J. (2006). Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, *441*(7089), 73.

Wang, B., Liu, J., Kim, H.-J., Webster, P. J., & Yim, S.-Y. (2012). Recent change of the global monsoon precipitation (1979–2008). *Climate Dynamics*, *39*(5), 1123–1135.

Watanabe, M., Shiogama, H., Tatebe, H., Hayashi, M., Ishii, M., & Kimoto, M. (2014). Contribution of natural decadal variability to global warming acceleration and hiatus. *Nature Climate Change*, *4*(10), 893.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., ... Kawamiya, M. (2011). MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development*, *4*(4), 845.

Watson-Parris, D., Bellouin, N., Deaconu, L. T., Schutgens, N. A. J., Yoshioka, M., Regayre, L. A., ... Stier, P. (2020). Constraining Uncertainty in Aerosol Direct Forcing. *Geophysical Research Letters*, *47*(9), e2020GL087141.

Watt-Meyer, O., Frierson, D., & Fu, Q. (2019). Hemispheric asymmetry of tropical expansion under co2 forcing. *Geophysical Research Letters*, *46*(15), 9231–9240.

Waugh, D. W., Garfinkel, C. I., & Polvani, L. M. (2015). Drivers of the recent tropical expansion in the southern hemisphere: Changing ssts or ozone depletion? *Journal of Climate*, *28*(16), 6581–6586.

Waugh, D. W., Grise, K. M., Seviour, W. J., Davis, S. M., Davis, N., Adam, O., ... others (2018). Revisiting the relationship among metrics of tropical expansion. *Journal of Climate*, *31*(18), 7565–7581.

Williams, A. P., & Funk, C. (2011). A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dynamics*, *37*(11-12), 2417–2435.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., ... Xavier, P. K. (2018). The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357-380.

World Health Organization. (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*.

World Health Organization. (2016). *Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease*.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., ... Liu, X. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. *Geoscientific Model Development*, *12*, 1573–1600.

Wu, T., Zhang, F., Zhang, J., Jie, W., Zhang, Y., Wu, F., ... Hu, A. (2020). Beijing Climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations. *Geoscientific Model Development*, *13*(3).

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., ... others (2019). The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan. Ser. II*.

Zeng, N. (2003). Drought in the Sahel. *Science*, *302*(5647), 999–1000.

Zeng, Y., Maxwell, S., Runting, R. K., Venter, O., Watson, J. E., & Carrasco, L. R. (2020). Environmental destruction not avoided with the Sustainable Development Goals. *Nature Sustainability*, 1–4.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B. (2018a). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 1. Simulation Characteristics With Prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*(3), 691-734.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B. (2018b). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 2. Model Description, Sensitivity Studies, and Tuning Strategies. *Journal of Advances in Modeling Earth Systems*, *10*(3), 735-769.