Examining the human influence on global climate using an empirical model

Austin Patrick Hope[1], Laura Anne McBride[1], Timothy P. Canty[1], Brian F Bennett[1], Walter R Tribett[1], and Ross J. Salawitch[1]

[1]University of Maryland, College Park

November 26, 2022

Abstract

We use the Empirical Model of Global Climate (EM-GC) to show that human activity has been responsible for ˜0.14 °C/decade (range: 0.08 to 0.20) of warming from 1979 to 2010. This EM-GC based quantification of Attributable Anthropogenic Warming Rate (AAWR) is constrained by the observed global mean surface temperature and ocean heat content records; the largest contribution to the uncertainty in our estimate of AAWR is imprecise knowledge of the radiative forcing due to tropospheric aerosols (AER RF). Our value of AAWR is noticeably lower than the mean value from the IPCC 2013 models, 0.22 °C/decade (range: 0.08 to 0.32) with no overlap of interquartile ranges. We also compute probabilistic forecasts of the rise in GMST where again the largest source of uncertainty is AER RF, and cast results in terms of the likelihood of achieving either 1.5 °C or 2.0 °C warmings relative to pre-industrial. We show that the likelihoods of limiting global warming to 2°C are 92%, 50%, and 20% if greenhouse gases follow the RCP 2.6, 4.5, and 6.0 scenarios; the likelihoods of limiting warming to 1.5°C drop to 67%, 10%, and 0.1% for these same three RCPs. Warming forecasts based upon our EM-GC are more optimistic than found by CMIP5 GCMs, following how many GCMs exhibit faster warming than inferred from the recent climate record. Our EM-GC forecasts show that aggressive controls on emissions of both CO_2 and CH_4 starting this decade are needed to limit global warming to 1.5°C with high probability.

Hosted file

`hope_etal20_v3a_supp.docx` available at https://authorea.com/users/555273/articles/605950-examining-the-human-influence-on-global-climate-using-an-empirical-model

Examining the human influence on global climate using an empirical model

Austin P. Hope[1], Laura A. McBride[2], Timothy P. Canty[1], Brian F. Bennett[1], Walter R. Tribett[1], Ross J. Salawitch[1,2,3]

[1]Department of Atmospheric and Oceanic Sciences, University of Maryland- College Park, College Park, 20740, USA

[2]Department of Chemistry and Biochemistry, University of Maryland -College Park, College Park, 20740, USA

[3]Earth System Science Interdisciplinary Center, University of Maryland -College Park, College Park, 20740, USA

Correspondence to: Austin P Hope (ahope12@umd.edu)

Key Points

- We use an observationally-driven simple climate model to argue with multiple lines of evidence that global climate models warm too quickly
- We unequivocally agree that greenhouse gas emissions must be limited soon to avoid extreme global warming
- Achieving the goals of the Paris Agreement is more likely than global climate models suggest but still requires significant global effort

Abstract

We use the Empirical Model of Global Climate (EM-GC) to show that human activity has been responsible for ~0.14 °C/decade (range: 0.08 to 0.20) of warming from 1979 to 2010. This EM-GC based quantification of Attributable Anthropogenic Warming Rate (AAWR) is constrained by the observed global mean surface temperature and ocean heat content records; the largest contribution to the uncertainty in our estimate of AAWR is imprecise knowledge of the radiative forcing due to tropospheric aerosols (AER RF). Our value of AAWR is noticeably lower than the mean value from the IPCC 2013 models, 0.22 °C/decade (range: 0.08 to 0.32) with no overlap of interquartile ranges. We also compute probabilistic forecasts of the rise in GMST where again the largest source of uncertainty is AER RF, and cast results in terms of the likelihood of achieving either 1.5 °C or 2.0 °C warmings relative to pre-industrial. We show that the likelihoods of limiting global warming to 2°C are 92%, 50%, and 20% if greenhouse gases follow the RCP 2.6, 4.5, and 6.0 scenarios; the likelihoods of limiting warming to 1.5°C drop to 67%, 10%, and 0.1% for these same three RCPs. Warming forecasts based upon our EM-GC are more optimistic than found by CMIP5 GCMs, following how many GCMs exhibit faster warming than inferred from the recent climate record. Our EM-GC forecasts show that aggressive controls on emissions of both CO_2 and CH_4 starting this decade are needed to limit global warming to 1.5°C with high probability.

Plain Language Summary

We use a simple climate model based on observations, the Empirical Model of Global Climate, to quantify past and future global warming due to human activity. Warming is driven by the release of greenhouse gases; the largest source of uncertainty in the human contribution to warming is due aerosols, small suspended particles that cool the planet by an amount that is not well known. We compare our results to more complex global climate models that are shown to warm faster than observed over the prior few decades. This pattern persists with projections of future temperature: while our model produces a wide range of future temperatures due to the relatively unknown strength of aerosol cooling, our ensemble generally produces lower temperatures by year 2100 than given by these more complex climate models. As such, our global warming forecasts provide a more optimistic view that society can keep global warming

beneath the thresholds set by the Paris Climate Agreement. Nonetheless, to achieve the target of the Paris Climate Agreement, society must greatly reduce its dependence on fossil fuels within this decade whereas to achieve the upper limit of this agreement, society must greatly limit fossil fuel emissions within the next two decades.

1. Introduction

Changes in Earth's climate on the decadal to century timescales are influenced by both anthropogenic and natural factors. Anthropogenic factors include rising concentrations of greenhouse gases (GHGs) that cause global warming (Lean & Rind, 2008; Santer et al., 2013) and increased burdens of tropospheric aerosols (hereafter, aerosols) that offset a portion of the GHG-induced warming (Bond et al., 2013; Kiehl, 2007; Smith & Bond, 2014; Stocker et al., 2013). Natural factors often cited as having a significant influence on global climate include the El Niño-Southern Oscillation (ENSO), the approximately 11-year solar cycle (total solar irradiance, TSI), and increases in the stratospheric aerosol optical depth (SAOD) that are the result of powerful volcanic eruptions (Chylek et al., 2016; Foster & Rahmstorf, 2011; Lean & Rind, 2008; Santer et al., 2013). Variations in total ocean heat content (OHC), the strength of the Atlantic Meridional Overturning Circulation (AMOC), and regional oceanic patterns like the Pacific Decadal Oscillation (PDO) and the Indian Ocean Dipole (IOD) also can influence global climate, though the extent each of these effects has on climate lacks consensus (Andronova & Schlesinger, 2000; Chylek et al., 2014; England et al., 2014; Rahmstorf et al., 2015; Rose et al., 2014; Saji et al., 1999; Steinman et al., 2015; Tokarska et al., 2019; Tung & Zhou, 2013). Feedbacks within the climate system driven by changes in atmospheric water vapor, lapse rate, clouds, and the surface albedo in response to radiative forcing (RF) induced by GHGs and aerosols also play a large role in the climate system (Andrews et al., 2012; Bony et al., 2006; Lin et al., 2019; Sherwood et al., 2020; Zelinka et al., 2013; Zhou et al., 2015).

Our focus is on quantification of the human influence on global climate. We examine the global monthly mean near surface air temperature anomaly relative to preindustrial (ΔT) from four data centers, collected over the past century and a half; for the purposes of this paper, we use a baseline of 1850-1900 as "preindustrial". We quantify the human influence on ΔT, termed the Attributable Anthropogenic Warming Rate (AAWR), using an Empirical Model of Global Climate (EM-GC) (Canty et al., 2013; Hope et al., 2017) that represents all of the factors described above. Our determination of AAWR is motivated by Box 10.1 of IPCC's Fifth Assessment Report, Working Group I (Stocker et al., 2013), except we divide their Attributable Anthropogenic Warming (AAW, units of °C) by time to arrive at warming rate (units of °C/decade). We primarily examine AAWR from the start of 1979 to the end of 2010 (hereafter 1979 to 2010) because AAW is nearly linear over this 32-year interval and this time period has also been studied by several other papers (Foster & Rahmstorf, 2011; Stocker et al., 2013; Zhou & Tung, 2013). We also quantify AAWR from archived output of the General Circulation Models (GCMs) used throughout Stocker et al. (2013) – hereafter, AR5 – as part of Phase 5 of the Climate Model Intercomparison Project (CMIP5) (Taylor et al., 2012).

Another key aspect of this study is projection of the rise in ΔT to year 2100 (ΔT_{2100}). Here, we use values of key model parameters (i.e. ocean heat export efficiency and the sum of climate feedback mechanisms, defined in section 2) obtained from fitting the historical climate record to forecast how ΔT and total ocean heat content will rise based on prescribed anthropogenic GHGs and aerosols. The projections focus solely on the anthropogenic component of ΔT so that our model results can be related to the Paris Agreement (UNFCCC, 2015). The agreement seeks to reduce future emissions of GHGs such that the increase in ΔT is "well below 2°C" and to "pursue efforts to limit the temperature increase to 1.5°C above preindustrial" (UNFCCC, 2015). Of course, future ΔT will also be influenced by natural variability, including TSI, ENSO, AMOC, and major volcanic eruptions (Chylek et al., 2016; Kavvada et al., 2013; Lean & Rind, 2009). Although variations in TSI have been forecast and could therefore be used in our projections, TSI exerts a relatively minor influence on ΔT (Lean & Rind, 2009; Zharkova et al., 2015). Since the other natural factors cannot be reliably predicted over the coming decades, we limit our projections of ΔT to the policy-relevant human component. Finally, the projections of ΔT are also framed in terms of the cumulative amount of carbon that can be emitted to achieve either the goal (1.5°C) or upper limit (2°C) of the Paris Agreement.

1.1 Previous estimates of AAWR

Multiple previous studies have examined AAWR, often focusing on 1979 to 2010 and using multiple linear regression (MLR) to quantify natural and anthropogenic influences on ΔT. Foster & Rahmstorf (2011) (FR11) suggested an AAWR of 0.170 ± 0.012 °C/decade based on analysis of version 3 of the ΔT record provided by the Climate Research Unit (HadCRUT3, hereafter CRU3) of East Anglia (Jones et al., 2012). They used MLR to remove the influence of ENSO, SAOD, and TSI on observed ΔT, and then fit the residual to quantify AAWR. Similar numerical values were reported for AAWR using ΔT from the Goddard Institute of Space Sciences (GISS, version 4) (Hansen et al., 2010) and the National Centers for Environmental Information (NCEI, blend of the Global Historical Climate Network-Monthly version 4 and the International Comprehensive Ocean-Atmosphere Data Set release 3) (Smith et al., 2008). Zhou & Tung (2013) (ZT13) examined version 4 of the CRU record (HadCRUT4, hereafter CRU4) and also used an MLR/residual method and concluded AAWR was 0.169 ± 0.019 °C/decade if temporal variations in the strength of the Atlantic Multidecadal Oscillation (AMO) are ignored. Most importantly, ZT13 stated that AAWR was 0.070 ± 0.019 °C/decade upon consideration of variations in the strength of the AMO. We highlight what we believe are shortcomings in the approaches of the FR11 and ZT13 studies in section §3.2.

Recently, Christy & McNider (2017), hereafter CM17, examined lower-tropospheric temperatures measured from satellite and radiosondes collected from the start of 1979 to the end of 2017. They concluded AAWR is 0.096 ± 0.023 °C/decade over this time period. This estimate covers a range of AAWR that includes the lower value of ZT13 but also suggests the value could be much higher, between the two possibilities for AAWR given by ZT13. Similar to CM17, we suggest the actual value of AAWR over 1979 to 2010 lies between the various estimates of FR11 and ZT13, though our value lies closer to the upper end of the range spanned between FR11 and ZT13.

1.2 Prior projections of future temperature

While it is certain that continued emissions of GHGs will cause a rise in ΔT, future warming is also subject to a wide range of uncertainties. One class of uncertainty, termed scenario uncertainty, is dependent on future atmospheric abundances of GHGs and aerosols. The CMIP5 community and AR5 adopted the Representative Concentration Pathways (RCPs) (Van Vuuren, Edmonds, et al., 2011) of GHGs and aerosols as part of an effort to address scenario uncertainty. Table SPM.2 of AR5 (IPCC, 2013) states that RCP 2.6 (Van Vuuren, Stehfest, et al., 2011) would result in 1.6 ± 0.7 °C warming (5-95% of model range of projections) relative to preindustrial temperature by the end of the 21st century, while RCP 8.5 (Riahi et al., 2011) would result in a warming of 4.3 ± 1.1 °C.

In addition to the two extreme RCP scenarios of RCP 2.6 and RCP 8.5, our study also focuses on the RCP 4.5 scenario (Thomson et al., 2011). While we have also examined RCP 6.0 (Masui et al., 2011) we choose to not focus on RCP 6.0 for several reasons. First, there is significant overlap between the projections for RCP 4.5 and RCP 6.0 among the CMIP5 GCMs used by AR5: ΔT_{2100} values of 2.4 °C warming (1.7 °C to 3.2 °C, 5-95% of model range of projections) for RCP 4.5 compared to 2.8 °C warming (2.0 °C to 3.7 °C) for RCP 6.0. Second, the CMIP5 GCMs collectively ran many more simulations for RCP 4.5 than for RCP 6.0. Third, there is much more literature concerning RCP 4.5 for comparison to our results. The time evolution of atmospheric CO_2, CH_4, and N_2O for the three RCPs we consider, together with recent observed globally averaged mixing ratios, are shown in **Figure 1**.

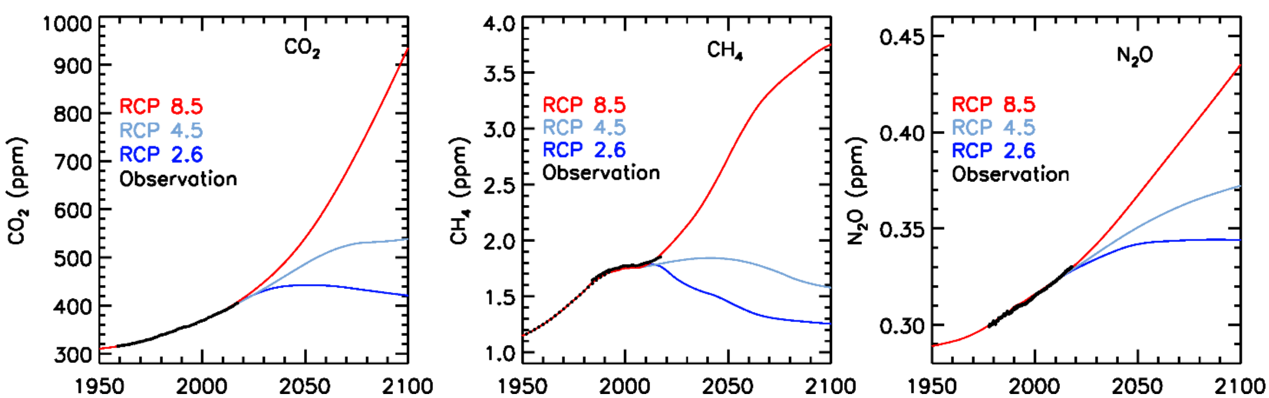

Figure 1. Greenhouse gas abundances for CO_2, CH_4, and N_2O, 1950 to 2100, taken from the RCP scenarios. Each scenario's GHG abundances are portrayed in a different color: red for RCP 8.5, light blue for RCP 4.5, and dark blue for RCP 2.6 (Meinshausen et al., 2011). These mixing ratio time series are combined with those of other minor GHGs to create the GHG RF times series used in this study. With the exception of CH_4, the RCP scenarios are visually hard to distinguish from each other between their divergence in 2005 and present, though their divergence becomes clear within the next ten years. Also shown for comparison are observations (black) for each GHG (data from https://www.esrl.noaa.gov/gmd/ccgg/trends/).

For a specific GHG scenario, such as RCP 4.5, there is a considerable range in end-of-century warming among various CMIP5 GCMs (e.g., figure SPM.7 of AR5 (IPCC, 2013)), i.e. model uncertainty. Primary drivers of these differences are uncertainties in climate feedback, the radiative forcing of climate due to aerosols, and the uptake of heat by the oceans (Forster et al., 2013; Kiehl, 2007; Knutti & Hegerl, 2008). Such model uncertainty can cause a large range of ΔT_{2100} found by different GCMs even if they use the same prescribed evolution of GHGs. Policy decisions geared towards meeting the Paris Climate Agreement must be made considering scenario and model uncertainty.

Several other approaches have been developed to forecast ΔT. Some approaches use similar regression analyses of historical records of ΔT (Chylek et al., 2016; Folland et al., 2018; Lean & Rind, 2009; Suckling et al., 2017). Other simple models use a small number of boxes to represent the atmosphere, ocean, and/or global carbon cycle (Meinshausen et al., 2009; Schwartz, 2012). Projections of future ΔT have also been constructed from simple calculations using emissions or mixing ratios of carbon dioxide that are from prescribed scenarios such as the RCPs or are based on forecasts of population, economic growth, and other factors (Raftery et al., 2017; Zeng & Geil, 2016).

Many of these studies reach conclusions concerning future global warming generally in agreement with the CMIP5 GCMs. Often this consensus is due to their models or analyses being driven by CMIP5 inputs and/or results. Fawcett et al. (2015) used a reduced complexity climate model constrained by the climate sensitivities from CMIP5 GCMs to conclude GHG emissions reductions based on the Paris Climate Agreement fall well short of the reductions needed to limit global warming to 2 °C, and suggest emissions scenarios similar to that associated with RCP 2.6 are needed. Similarly, Raftery et al. (2017) examined projections of population, global economic output, and the carbon intensity of the world's economies to conclude that median warming in 2100 would be 3.2 °C (likely range 2 °C to 4.9 °C) with only a 5% chance of remaining beneath 2 °C. Raftery et al. (2017) relied on the relationship between carbon emissions and global warming from the CMIP5 GCMs.

In slight contrast, Chylek et al. (2016) used a standalone regression model to project a rise in future ΔT of slightly less than 2 °C by end-of-century for RCP 4.5. This warming is somewhat less than the projected 2.5 °C multi-model mean of 42 CMIP5 GCMs. Another empirical analysis of ΔT using an energy balance model (Mauritsen & Pincus, 2017) examines the relationship between transient, equilibrium, and committed warming for scenarios that either omit or include the uptake of CO_2 by the world's oceans. They determined that there is a 50% chance of global warming remaining below 1.5 °C if additional future radiative forcing of climate (RF) does not exceed 1.2 W m^{-2}. This limit would be realized in year 2053 if the current rate of RF increase (+0.033 W m^{-2} yr^{-1}) was kept constant. We reach a broadly similar conclusion based on our modeling effort with the EM-GC.

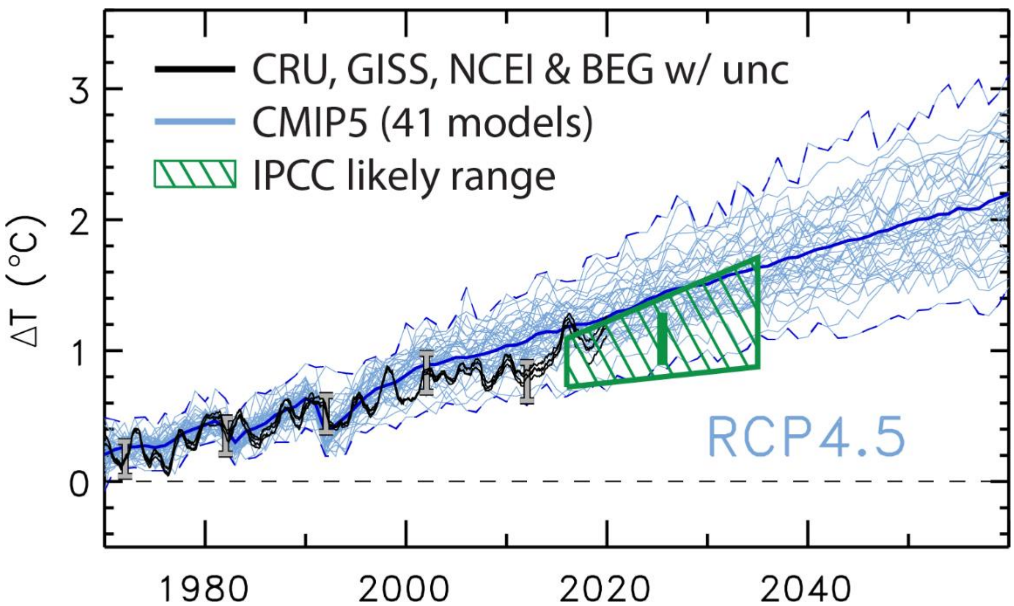

Figure 2. Observed and GCM-simulated global warming, 1970-2060. The observed ΔT time series are taken from four data centers (CRU4, GISS, NCEI, and a fourth option from the Berkeley Earth group notated as BEG) and are shown in black, with grey error bars representing the uncertainty from the CRU4 record every ten years. The modeled ΔT time series are taken from the output of 41 GCMs that participated in CMIP5 over the RCP historical and RCP 4.5 future experiment time periods and are shown individually in light blue. The maximum, mean, and minimum from the GCM ensemble are shown in dark blue. The green trapezoid represents the indicative likely range for annual average ΔT for the years 2016 to 2035, and the green bar represents the likely range for the mean value of ΔT for this two-decade time period, both given in Chapter 11 of AR5.

Chapter 11 of AR5 (Kirtman et al., 2013) showed that the CMIP5 GCMs tend to overestimate ΔT for the early part of the 21st century, as shown in **Figure 2**. This figure compares time series of ΔT from 41 CMIP5 GCMs (light blue) with the observed temperature record from the four data centers shown in figure 11.25 of AR5. Due to the tendency for observed ΔT to be overestimated by the climate models, the authors of Chapter 11 of AR5 prepared an expert judgement of the expected rise in ΔT over the next few decades (green trapezoid in Figure 2). Notably, these likely ranges of global warming lie below the GCM ensemble mean. This expert judgement of global warming covers a time period for which all four RCPs have similar values of RF. As will be shown in section §3, global warming forecasts by the EM-GC are in close quantitative agreement with this green trapezoid.

Crafting environmental policy based on such a wide range of possible futures is difficult, even if the physical link between rising GHGs and increasing temperature is well established. The Transient Climate Response to cumulative carbon Emissions (TCRE) is a metric that was developed to link global warming to future anthropogenic emissions of CO_2 (Gregory et al., 2009). As such, TCRE provides a means for policy makers to exert direct control of global warming through regulation of CO_2 emissions. Chapter 12 of AR5 (Collins et al., 2013) defines TCRE as the modeled transient increase in ΔT per 1000 GtC of CO_2 released to the atmosphere. Most climate models show that future ΔT increases in a nearly linear fashion with respect to cumulative emissions of CO_2, but this relationship is dependent on the physics and structure of the climate model, as well as assumptions regarding emissions of other GHGs and the time rate of change of emitted CO_2 (Figure 12.45 of AR5). For example, model experiments that have CO_2 concentrations increasing at the rapid rate of 1% per year find a relatively low value for TCRE compared to simulations with a slower rate of CO_2 increase (e.g. Figure SPM.10 (IPCC, 2013)) because the inertia of the climate system limits the transient temperature response relative to faster emissions. It is commonly accepted that the transient response of ΔT to rising GHGs is less than the equilibrium response because certain aspects of the climate system such as the cryosphere and the transfer of heat into the ocean occur on multi-year timescales. Chapter 12 of AR5 (Collins et al., 2013) states that TCRE likely lies between 0.8°C to 2.5°C per 1000 GtC. The CMIP5 GCMs tend to lie toward the high end of this range for TCRE, whereas projections of global warming found using our approach (section §3) lie toward the lower end.

1.3 Overview of this work

The Empirical Model of Global Climate (EM-GC) used in this study builds upon the framework first described by Canty et al. (2013). This work also builds on Hope et al. (2017), who used an earlier version of the model to conduct similar analysis. The EM-GC uses MLR combined with a two-module ocean-atmosphere approach to simulate observed monthly variations in ΔT. The MLR component uses an equation that represents numerous anthropogenic and natural factors that drive variations in global climate. The version of the EM-GC used here considers several factors not present in most other MLR-based analyses, and includes several important updates since Hope et al. (2017), especially in the ocean module, as described in section 2.

Section 3 presents results from the EM-GC concerning AAWR and projection of ΔT out to year 2100. Differences in these quantities between the EM-GC and other works are also described in section 3, as are analyses of these differences and how the EM-GC works. We present our conclusions and how the EM-GC results fit into the climate modeling community's knowledge of the climate system in section 4.

2. Model construction

The EM-GC provides a mathematical representation of the factors that govern the global mean surface temperature anomaly (ΔT). We compute numerical values of climate amplification (γ) and the efficiency of heat transfer from the atmosphere to the ocean (κ) based on the observed climate record. We then use these values of γ and κ to project future ΔT. While MLRs have been used to conduct similar calculations by other groups, our EM-GC includes several components not included in these other models. These differences include the long-term export of heat from the atmosphere to the ocean, a comprehensive treatment of tropospheric aerosols, and the influence on global climate from variations in the strength of Atlantic Multidecadal Variability (AMV, which we use as a proxy for AMOC), and a new probability weighting method for the large ensemble of ΔT projections that incorporates the expert judgement of aerosol RF given in Chapter 8 of AR5 (Myhre et al., 2013).

The EM-GC is an ensemble-based model whose parameter space spans a large range of possible values for both the strength of the climate feedback and the historical strength of anthropogenic aerosol forcing. The overall ensemble is filtered based on a set of model parameters that quantify a statistically acceptable fit ($\chi^2 \leq 2$, described below) between observed and modeled historical ΔT and ocean heat content (OHC). These simulations of historical ΔT and OHC are then used to create a corresponding ensemble forecasts of ΔT, which allows for a detailed statistical analysis of the impact of uncertainty in aerosol RF and climate feedback on future global warming.

In the four following sections, we describe the model equations and input data, representation of the ocean component, climate sensitivity and feedbacks, and the different modes of the EM-GC.

2.1 EM-GC core equations

The EM-GC simulation of observed ΔT uses a MLR-based analysis of the flow of energy between major components of Earth's climate system. The main equations of the EM-GC are:

$$\Delta T_{MDL\,i} = \frac{1+\gamma}{\lambda_p}\{GHG\ RF_i + AER\ RF_i + LUC\ RF_i - 0.671 Q_{OCEAN\,i}\} + C_0 + C_1 \times SAOD_{i-6} + C_2 \times TSI_{i-1} + C_3 \times ENSO_{i-2} + C_4 \times AMV_i + C_5 \times PDO_i + C_6 \times IOD_i \quad \text{1a)}$$

$$Cost\ Function = \sum_{i=1}^{N_{MONTHS}} \frac{(\Delta T_{MDL\,i} - \Delta T_{OBS\,i})^2}{\sigma_{OBS\,i}^2} \quad \text{1b)}$$

$$\chi^2 = \frac{1}{(N_{YRS} - N_{DOF} - 1)} \sum_{j=1}^{N_{YRS}} \frac{\left(<\Delta T_{MDL}>_j - <\Delta T_{OBS}>_j\right)^2}{<\sigma_{OBS}>_j^2} \quad \text{(1c)}$$

where λ_P =3.2 W m^{-2} °C^{-1} (Planck response), γ is the dimensionless climate amplification term, C_0 to C_6 are regression coefficients, and *i* is an index for month. This model representation of $\Delta T_{MDL\,i}$ considers four anthropogenic factors (GHGs, net anthropogenic tropospheric aerosols [AER], land use change [LUC], and ocean heat export [Q_{OCEAN}]) as well as six natural factors (SAOD, TSI, ENSO, AMV, PDO, and IOD). The data inputs for all factors aside from Q_{OCEAN} are either taken directly or modified slightly from outside sources, while Q_{OCEAN} is calculated within the EM-GC. (Section §2.2 below defines Q_{OCEAN} and its governing equations; the multiplication of Q_{OCEAN} by 0.671 in Equation 1a represents an area correction between Q_{OCEAN} as computed in the ocean module and its effect on the atmosphere.) The sources for all of our input data and the small changes we apply to them are fully documented in sections §2.1.1 and §2.1.2 below.

Each member of an EM-GC ensemble uses the same natural factors, ΔT_{OBS} record, OHC record, GHG forcing time series, and LUC forcing time series. The ensemble members vary as each is constrained to different assumed values for AER radiative forcing and γ. The anthropogenic components for each run are fed into the OHC submodule of the EM-GC to calculate Q_{OCEAN}. The submodule produces estimates for global average sea surface temperatures (SSTs) and the human component of ΔT_{MDL}. These two temperature time series are used to recalculate Q_{OCEAN}, and the submodule iterates this process until those three quantities (i.e. SSTs, $T_{MDL\text{-}HUMAN}$, and Q_{OCEAN}) remains stable between iterations.

We then solve for the seven regression coefficients (C_i) by minimizing the cost function (Equation 1b), accounting for the 1σ uncertainty in each value of monthly global mean observed temperature ($\Delta T_{OBS\ i}$). The temperature records used to prescribe $\Delta T_{OBS\ i}$ are based on one of the previously mentioned CRU4, GISS, and NCEI data sets, as well as a fourth option from the Berkeley Earth group (BEG) (Rohde et al., 2013). The terms in the cost function are indexed over the total number of months (N_{MONTHS}) for which $\Delta T_{OBS\ i}$ are available, either 2040 months (Jan.1850 to Dec.2019) for CRU4 and BEG or 1680 months (Jan.1880 to Dec.2019) for GISS and NCEI.

Our modeling approach also makes use of reduced chi-squared (χ^2) that defines the goodness-of-fit between observed and modeled ΔT (Equation 1c). In equation 1c, N_{YRS} represents the total number of years for which $\Delta T_{OBS\ i}$ are available (170 years for CRU4 and BEG; 140 years for GISS and NCEI) and < > denotes annual average. Unless otherwise stated, the number of degrees of freedom (N_{DOF}) in this study is nine: the climate amplification term, an ocean heat surface diffusivity parameter, and the seven regression coefficients. The EM-GC can run with any selection of our natural and anthropogenic variables included or excluded, potentially reducing N_{DOF}. The equation for χ^2 is based on annual averages of observed and modeled ΔT because the autocorrelations of ΔT_{OBS} and ΔT_{MDL} exhibit similar shapes when examined as annual averages, but do not match on the monthly time grid (Canty et al., 2013). Therefore, the use of annual averages reduces the effect of high-frequency variations of ΔT_{OBS} that are not captured by the model. Nonetheless, the model framework is expressed in terms of monthly time series for all quantities to properly quantify the effect of factors such as major volcanic eruptions and ENSO on global climate.

We compute two other versions of χ^2 as well. While equation 1c as described above is defined for ΔT over the full time period of available observations, we additionally compute reduced chi-squared using the same framework both for OHC over the full time period of OHC and for ΔT over the most recent 80 years. As the EM-GC is designed to fit both atmospheric and oceanic observations, we do not consider simulations with acceptable fits to ΔT_{OBS} if they do not also provide an acceptable fit between OHC_{MDL} and OHC_{OBS}. We then consider χ^2 for ΔT over the past 80 years because a combination of factors makes it possible to achieve an acceptable fit (i.e., $\chi^2 \leq 2$) for the full time series of ΔT_{OBS} while significantly over- or under-estimating warming during the last 30 to 50 years. Since the most recent several decades are a focus of our study, a lack of fit of ΔT_{OBS} over this period would confound meaningful analysis. We choose a length of 80 years to assure that all semi-oscillatory natural forcing factors experience at least one full cycle within the years of consideration (PDO and AMV vary on time scales of up to 60 or 70 years).

Figure 3 provides a visual representation of our model (i.e. Equations 1a-c) by showing a single run from an EM-GC ensemble with the best fits to ΔT_{OBS} when using a time series for aerosol radiative forcing (AER RF) that matches the IPCC most-likely value of AER RF in 2011. We refer to this depiction of the EM-GC's components as a "ladder plot". The top rung of Figure 3a shows model input ($\Delta T_{OBS\ i}$, black) and output ($\Delta T_{MDL\ i}$, red). The second and third rungs of Figure 3a show the anthropogenic components of ΔT_{MDL}, both separated (second rung) and combined (third rung). GHG RF in this run is based on the RCP 4.5 time series for CO_2, CH_4, N_2O, and other GHGs (Meinshausen et al., 2011; Thomson et al., 2011). Nearly identical results are found upon use of RCP 8.5 or RCP 2.6, because the RCP scenarios use the same historical data for all species until 2005 (Figure 1). The time series for AER RF used in Figure 3a is based upon our analysis of direct RF due to six aerosol types provided by RCP (Lamarque et al., 2011), combined and expanded to include the indirect aerosol effect as well, as described in section §2.1.2. LUC RF is taken from Table AII.1.2 of AR5. Our Q_{OCEAN}, the export of heat from the atmosphere to the ocean, is found by simulating the long-term observed rise in OHC, as described in section §2.2. The net human time series in the third rung (gold line) serves as the basis for calculating AAWR and is further discussed in section §2.3.

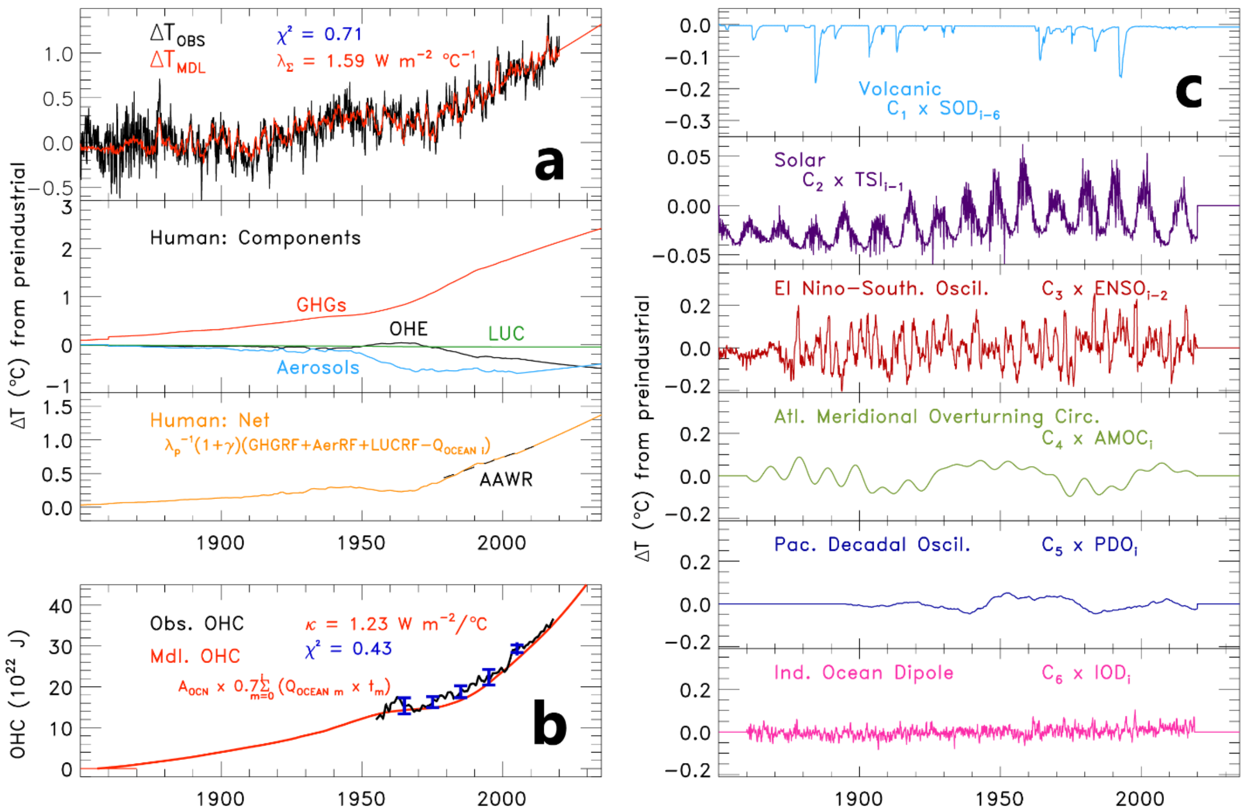

Figure 3. Observed and modeled ΔT, 1850 to 2019, relative to a preindustrial baseline (1850 to 1900). Ladder plots showing observed ΔT from CRU4 (top rung of Figure 3a, black line), the EM-GC simulation of ΔT from CRU4 (top rung, red line), ΔT_{HUMAN} (third rung of Figure 3a, gold line) and its components (second rung of 3a), AAWR (third rung of 3a, dashed black line), and contributions to ΔT from various natural factors (TSI, second rung of 3c, purple line; SAOD, first rung of 3c, light blue line; ENSO as measured by the Multivariate ENSO Index version 2, third rung of 3c, crimson line; AMV with 9-year Fourier filtering as our AMOC proxy, fourth rung on 3c, green line; PDO with 10-year smoothing, fifth rung of 3c, indigo line; IOD, sixth rung of 3c, pink line). Figure 3b shows the model-simulated rise in OHC (red line) from the top 700 m compared to the average of five OHC data sets (black line; blue error bars show the 1σ standard deviation of these data about mean OHC).

The lower-left rung of the ladder plot (Figure 3b) depicts the modeled increase in OHC (red curve, proportional to the summation of Q_{OCEAN} via equation 2) and the observed rise in OHC (black line). In Figure 3b, the observed OHC is based upon the average of five data sets (Balmaseda et al., 2013; Carton et al., 2018; Cheng et al., 2016; Ishii et al., 2017; Levitus et al., 2012) such that the average is taken for each year when at least three of the five data sets provide an annual value (see **Figure S1**). The modeled increase in OHC is related to Q_{OCEAN} as follows:

$$OHC_{MDL\,i} = A_{OCEAN} \times 0.7 \sum_{m=0}^{i} [Q_{OCEAN\,m} \times t_m] \tag{2}$$

In equation 2, A_{OCEAN} is 3.3×10^{14} m^2 (Domingues et al., 2008), and t_m is the time length for month m in seconds, with m=0 representing the first month of a model run (e.g. January 1850 for runs using CRU4). The factor of 0.7 is used to account for the fact that the OHC data sets we use represent only the top 700 m of the oceans, which hold roughly 70% of the total heat content of the ocean (Sect. 5.2.2.1 of Solomon (2007)). We verified this value of roughly 70% by comparing the OHC time series for the upper 700 m to the time series for the upper 2000 m or full ocean for the three data sources that provided time series for multiple depths (Balmaseda et al., 2013; Carton et al., 2018; Levitus et al., 2012).

The influences on ΔT of solar irradiance, volcanoes, ENSO, AMOC, PDO, and IOD are shown on the six rungs on Figure 3c. The sum of the variations in ΔT due to the six natural factors (Figure 3c) and the four anthropogenic factors (third rung of Figure 3a) plus the regression constant (C_0) equal ΔT_{MDL}. The overall agreement between the black and red lines in Figure 3a demonstrates the ability of the EM-GC to capture much of the variability and rise in global mean surface temperature over the past century and a half.

2.1.1 Model input, natural factors

In this section, we describe the data used to define model inputs for stratospheric optical depth (SAOD), total solar irradiance (TSI), the El Niño-Southern Oscillation (ENSO), the Atlantic Multidecadal Variability (AMV), Pacific Decadal Oscillation (PDO), and the Indian Ocean Dipole (IOD). Table A1 provides URLs of the websites that host these data records.

The time series for SAOD is a combination of two data sources. The first source is the recommended input for Phase 6 of the Climate Model Intercomparison Project (CMIP6). This CMIP6 SAOD time series is a combination of the Volcanic Model Intercomparison Project

(Zanchettin et al., 2016) for enhanced values of SAOD before the satellite era, Arfeuille et al. (2014) for background values before the satellite era, and the GloSSAC data set (Thomason et al., 2016) for all values during the satellite era. GloSSAC has since been extended past the last year of CMIP6 (Thomason et al., 2018). We then use the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data (Vaughan et al., 2004) for year 2019 as our second source since the GloSSAC record available at time of paper submission ends in Dec 2018. The GloSSAC and CALIPSO data are available as SAOD values by latitude; we create SAOD values by latitude for years before 1979 by taking extinction coefficients and integrating them from the tropopause to the top of the atmosphere. We then weight each latitude band by area to reach a near-global SAOD time series (80°S to 80°N). The three data sets, as presently available, match very closely during their respective time periods of overlap, though we do apply a small adjustment to bring CALIPSO in line with GloSSAC. We treat SAOD after 2019 in the same way that we treat all of our natural factors from 2020 onward: we flatline the data at a value near zero representative of the current, non-volcanic background. As SAOD has a small background value, it is the only natural time series in the EM-GC that uses a nonzero value into the future (specifically, the December 2019 value of SAOD from CALIPSO). The final SAOD time series is then lagged by six months in equation 1a to match the delay between surface forcing and the thermodynamic response to major volcanic eruptions found by Thompson et al. (2009) which is the same time lag that has been used in other MLR studies (Foster & Rahmstorf, 2011; Lean & Rind, 2008).

The time series of TSI used in the EM-GC is constructed from two data sets. TSI data up to 2014 are an average of two solar models, one empirical and one semi-empirical (Matthes et al., 2017); TSI data after 2014 are from satellite-based solar radiance measurements (Dudok de Wit et al., 2017). These two data sets agree well at the point of merging. The EM-GC can propagate the underlying 11-year solar cycle past 2019, but in this analysis we flatline TSI in the future. We make this choice so that future projections of ΔT are based solely on anthropogenic forcing. A one-month lag for TSI is used in equation 1a, because this lag yields the largest value of C_2. Lagging TSI is a common approach for quantifying the slight temporal offset of ΔT in response to variations in total solar output (Lean & Rind, 2008).

The proxy for variations in the strength of the El Niño-Southern Oscillation used here is built around the Multivariate ENSO Index version 2 (MEIv2) (Wolter & Timlin, 1993; Zhang et al., 2019). The MEI time series, regularly updated, consists of a principle component analysis of five physical quantities (sea level pressure, sea surface temperature, surface zonal winds, surface meridional winds, and outgoing longwave radiation) that represent the state of the tropical ocean-atmosphere system. The MEIv2 record begins in 1979. To provide data back to 1870, we prepend the MEI-extended record (Wolter & Timlin, 2011), which uses a weighted combination of two components (SSTs and the Southern Oscillation time series); the MEI-extended was created as an extension of the original MEI record (Wolter & Timlin, 1993) that uses six physical quantities instead of five. To extend the MEI-based data record from 1869 back to 1850, we manually compute the SST average over the ocean surface area considered in the MEI records. To prevent data shock, we increase the MEIv2 values by a constant offset so that its average from 1979 to 2005 matches the average of the MEI-extended over the same time period (which is the extent of their overlap). A two-month lag has been applied to the ENSO index in equation 1a, because this lag provides the highest correlation with the simulated response of ΔT to ENSO found using a thermodynamic approach (Thompson et al., 2009). The process used to determine this lag from ENSO is described in Canty et al. (2013); this two-month lag is the same as used in other MLR studies (e.g. Lean & Rind (2008) and FR11). The EM-GC is capable of using five other ENSO data sets: the original MEI-based record (i.e. not MEIv2-based) as used in previous iterations of the EM-GC (Canty et al., 2013; Hope et al., 2017), the Tropical Pacific Index (Zhang et al., 1997), the Niño 3.4 index (Trenberth, 1997), the Cold Tongue Index (Deser & Wallace, 1990), and the thermodynamic index of Thompson et al. (2009). Our scientific conclusions are entirely insensitive to which ENSO index is used. Here we choose to focus on the MEIv2 as its multiple-component construction leads to a robust time series with less noise than other ENSO time series, particularly those that are based only on SST averages.

Our AMV index is based on the area weighted, monthly mean SST in the Atlantic Ocean, between the equator and 60°N (Schlesinger & Ramankutty, 1994). We detrend the AMV index using the RF anomaly due to human activity over the century-and-a-half time period of analysis, as described in section §3.2.3 of Canty et al. (2013). This detrending process removes the influence of long-term global warming on the AMV time series; without this external influence, the detrended index can serve as a proxy for variations in the strength of the Atlantic Meridional Overturning Circulation (AMOC) (Knight et al., 2005; Medhaug & Furevik, 2011; Stouffer et al., 2006). Since AMOC is slowly varying, if it affects the climate with AMV as a proxy, then high-frequency components of the AMV would be indicative of influence from non-AMOC factors (such as ENSO) or noise. As such, our AMV index is also Fourier-filtered to remove all components with temporal variations shorter than nine years, as described in Canty et al. (2013). The resulting index represents anomalies in the north Atlantic SST that vary on time scales of a decade or longer and are decoupled from human influence.

For the Atlantic signal, we have also tested the LOWESS filtering of ZT13, the “Atlantic Water Variability index” of Pausata et al. (2015), and the “Atlantic gyre index” of Rahmstorf et al. (2015), and two other levels of Fourier filtering of the AMV. Our main results concerning AAWR are insensitive to the proxy used for AMOC, though the water variability index and gyre index both have little to no expression in the climate record (i.e. the EM-GC returns near-zero values for regression coefficient C_4). As such, we favor using the AMV with nine-year filtering and anthropogenic detrending as the AMOC proxy in our regressions. Including AMV produces relatively low values for χ^2, allowing our ensembles to include more members without biasing either AAWR or γ (which determines the trend of future temperatures). We have not yet tested the North Atlantic Variability Index (NAVI), an alternative to AMV (Haustein et al., 2019). However, near-zero values for C_4 are found for several other proxies for AMOC and the resulting values of AAWR and ΔT_{2100} are similar to those shown when AMV is used in the regression. We expect the NAVI (Haustein et al., 2019) to be inconsequential for our general scientific conclusions (section §3.1) because values of AAWR and ΔT_{2100} found in our model framework are insensitive to various other proposed proxies for AMOC, as well as the omission of a proxy for AMOC from our regression.

Directly measuring the AMOC, namely its overall rate and volume of flow, is inherently difficult and has not been done over a long enough time period to be used in the EM-GC. However, observations of AMOC have been made in recent years. An analysis of a 14 year (April 2004 to September 2018) time series of data from the RAPID-AMOC program (Duchez et al., 2014) reveals a decline in the strength of AMOC over this time period (Smeed et al., 2017; Srokosz & Bryden, 2015), similar to that shown by the AMV over these same years (Figure 3c).

The PDO represents the temporal evolution of temperature and sea level pressure of the Pacific Ocean poleward of 20°N (Zhang et al., 1997). The PDO index, which begins in 1900 and extends to present, represents the response of circulation in the Pacific Ocean to atmospheric forcing (Saravanan & McWilliams, 1998; Wu & Liu, 2003). This index is regularly updated by the University of Washington. The EM-GC also has the capability to use the Interdecadal Pacific Oscillation (IPO) index to represent the influence of the Pacific Ocean on global climate, rather than the PDO. For comparison to our decadal-filtered Atlantic signal, we use a 10-year running mean of PDO (or IPO) in our analysis, a method reflected in other studies that attempt to link the Pacific signal to global temperature patterns (England et al., 2014). In this paper, we use the 10-year average of the PDO; using the IPO or using different smoothing times does not change our results. We have also attempted to take the integral of the PDO to address the idea that the sign of PDO affects the trend of ΔT, but taking the integral of PDO produces a time series with an uncharacteristically large peak in the middle of the time series and few other features relating to ΔT_{OBS}, leading to very small values for C_5. As detailed in Canty et al. (2013), variations in ΔT are more strongly influenced by the Atlantic Ocean than the Pacific, regardless of the treatment of PDO. That said, the Pacific signal is stronger than the Atlantic signal in some select ensemble members with strong aerosol forcing or with Atlantic proxies that exhibit inherently weak fits to the ΔT_{OBS} record.

A proxy for variations in the circulation of the Indian Ocean is also used, so that all three major ocean basins are represented. We use the Indian Ocean Dipole (IOD) index as defined Saji et al. (1999), which represents the temperature gradient between the Western and Southeastern portions of the equatorial Indian Ocean. The IOD time series we use is made with SSTs from the Centennial in situ Observation-Based Estimate record (Ishii et al., 2005). We find that influence of the Indian Ocean on ΔT is small, likely due to the size of this ocean basin relative to those of the Atlantic and Pacific.

2.1.2 Model inputs, anthropogenic factors

The anthropogenic inputs included in the EM-GC are greenhouse gases, tropospheric aerosols, land use change, and the long-term export of heat into the ocean in response to anthropogenic atmospheric warming. All data mentioned below can be found online, with source websites and other comments listed in Table A2.

Atmospheric abundances of the main drivers of the anthropogenic RF of climate – the well-mixed greenhouse gases CO_2, CH_4, and N_2O – are taken directly from the RCP scenarios (Meinshausen et al., 2011). For these and other GHGs referenced to Meinshausen et al. (2011), we specifically use data from files provided by the Potsdam Institute for Climate Research

(PICR, URL in Table A2). Annual mixing ratios for each gas are converted to radiative forcing, using the equations from Myhre (1998) and initial values from table AII.1.1a in AR5. The RF values are interpolated onto the EM-GC's monthly time grid, with annual RF treated as midyear conditions. The projected mixing ratio of CH_4 in year 2100 is dramatically higher in RCP 8.5 (3.48 ppm) compared to RCP 4.5 (1.54 ppm) (Figure 1). To quantify the sensitivity of global warming to CH_4, we have defined four new hybrid scenarios for CH_4 that are linear combinations of RCP 4.5 and RCP 8.5 (**Figure S2**). The RF of CH_4 for the RCPs used here differs slightly from values of RF for CH_4 archived at PICR, due to the update for the preindustrial value for CH_4 given in table AII.1.1a of AR5.

Other greenhouse gases include tropospheric ozone (O_3), stratospheric-ozone-depleting substances (CFCs, HCFCs, CCl_4, CH_3Cl, CH_3Br, etc.), and other F-gases (HFCs, PFCs, and SF_6). Prior and future RF of climate due to tropospheric O_3 is taken directly from Meinshausen et al. (2011) for each RCP scenario. The increase in RF of climate due to tropospheric O_3 between 1750 and 2011 is nearly equal to that of CH_4, albeit with much larger uncertainty. The various RCPs project different future RFs due to tropospheric O_3, with RCP 8.5 being the largest. The RF of climate due to ozone-depleting substances (ODS), HFCs, PFCs, and SF_6 are all also taken directly from the RCPs, via PICR.

We consider numerous anthropogenic aerosol scenarios that represent a wide range of total (direct and indirect) RF of climate due to the aerosols. This wide range is essential for consideration because the historical effect of aerosols on climate is not well-known (Myhre et al., 2013), whereas future AER RF is projected to decline due to air quality regulations (Smith & Bond, 2014). The climate record can be well-simulated by an aerosol scenario for which the RF of climate due to GHGs has been considerably offset by aerosol cooling: in this case, large values for the sum of climate feedback mechanisms (section §2.3) are needed to match the observed rise in ΔT. The climate record can be fit just as well by a scenario for which RF due to GHGs has barely been offset by aerosol cooling, in which case small values for the sum of climate feedback mechanisms are required to match ΔT. If we assume that the feedback inferred from the prior climate record will persist into the future, the strong aerosol cooling case will lead to much larger future warming than the weaker cooling case (Knutti & Hegerl, 2008). The need to consider this relation between AER RF and feedback drives the wide range of scenarios for AER RF described below. We will often refer to a given AER RF time series by the value in year

2011 (AER RF_{2011}) in order to relate our results to estimates of AER RF_{2011} given in chapter 8 of AR5 (Myhre et al., 2013).

We construct our AER RF scenarios based on forcing data from the RCP database. First, the direct RF for six types of aerosols (sulfate, black carbon, nitrate, dust, organic carbon, and biomass burning products) are obtained from PICR for each RCP scenario (Lamarque et al., 2011). These direct RF estimates were tied to the state of knowledge that guided the fourth IPCC assessment report (Solomon, 2007). As was done in Canty et al. (2013), we use direct RF as given by PICR for five of the six aerosol species; for sulfate, Smith et al. (2011) is used instead because the PICR sulfate data do not reflect sulfate emissions well. In our study, the direct RF time series for each component has been scaled to match values of direct RF in 2011 given by chapter 8 of AR5 (Myhre et al., 2013), as noted in the caption of **Figure S3**. This matching process includes eliminating the effect of biomass burning on RF of climate, as AR5 chapter 8 estimated that the RF due to biomass burning in 2011 was zero. Physically, biomass burning can conceivably provide no RF impact as a result of cancellation between the warming due to black carbon and the cooling due to organic aerosols products (sections 7.5.2.2 and 8.3.4.2 of AR5). Our scientific conclusions would be unaffected had we used the RCP AER RF time series directly, as archived by PICR. Nonetheless, we scale to AR5 values of direct RF in 2011 so that our study is consistent with the consensus of the scientific community at the time of paper submission.

We perform a second scaling on the aerosol direct RF time series to mathematically simulate the aerosol indirect effect, e.g. cloud-aerosol interactions, with the goal of reaching the AR5 best estimate for total aerosol forcing of $-0.9\ W/m^2$ in 2011. First we separate the direct RF time series into a cooling group – sulfates, mineral dust, primary and secondary organic aerosols, and nitrates – and a warming "group" of black carbon; as biomass burning products were zeroed earlier, they do not factor into the remainder of this analysis. Second, we take the ratio of the sulfate total RF to the sulfate direct RF based on Smith et al. (2011) and Stern (2006) as described in Canty et al. (2013), a scaling of 2.432. We apply this ratio to the cooling group overall as the scaling factor to change the direct RF time series of the cooling group to the respective total RF time series. We term the value used to scale direct-to-total RF for the cooling group as α_{COOL}. Next, we find the respective direct-to-total ratio for the heating group – α_{HEAT}, in this case 2.188 – needed to make the total cooling RF times series and total heating RF time series add together such that the value in 2011 is –0.9 W/m^2, which is AR5's best estimate of effective AER RF_{2011}.

To simulate the uncertainty in historical AER RF, we create other AER RF time series by varying α_{COOL} and α_{HEAT}. In doing so, we can create a potentially infinite number of AER RF times series of different shapes and strengths. For each ensemble run of the EM-GC, we constrain our AER RF times series to those constructed using a finite length of a cross-section through this α-space, which we term "roads", shown visually in **Figure S4a**. The "middle road" of aerosol scenarios is anchored by the α-space point obtained by attaining the best-estimate AER RF_{2011} (–0.9 W/m^2) as described in the previous paragraph. The slope of the middle road is found by determining four other statistical combinations of the six aerosol direct RF time series that produce values of total AER RF_{2011} corresponding to the AR5 confidence intervals (–0.4 W/m^2 to –1.5 W/m^2, likely range, and –0.1 W/m^2 to –1.9 W/m^2, 5 to 95% confidence range). The "low road" and "high road" are then anchored by points in α-space that also have an AER RF_{2011} value of –0.9 W/m^2. **Figure S4b** then shows the fifteen resulting AER RF time series that correspond to the intersections of the three roads and five forcing isopleths.

Variations in the composition of Earth's surface due to deforestation and other human activities can also exert a change in the radiative forcing of climate. The ΔRF effect of anthropogenic land use change (LUC) is taken directly from table AII.1.2 of AR5, with the annual values linearly interpolated to the EM-GC's monthly time grid. We assume the annual values are centered at midyear in the interpolation. The release of GHGs from land-use-change activities such as deforestation or concrete laying are factored into the GHG term itself, as the LUC term only represents surface reflectivity.

We consider the rise in OHC as an anthropogenic signal because increases in the RF of climate due to human activity cause a rise in temperature of both the atmosphere and the oceans (Raper et al. 2002; Hansen et al. 2011; Schwartz 2012). The focus of many OHC studies has been the top 700 m or top 2 km of the world's oceans, and our work considers data from five such studies (Balmaseda et al., 2013; Carton et al., 2018; Cheng et al., 2016; Ishii et al., 2017; Levitus et al., 2012); three of these five studies consider both depths. For proper comparison, the five data sets are normalized to a common value of 0 in 1986 (the midpoint year for the range of time where three or more of the five OHC records are provided) before being averaged together. The magnitude of an input OHC data set at any given point in time is not important in our model framework, because we rely upon change in OHC over time. In this study, we focus on EM-GC runs that use 700 m OHC data, in which case we multiply the OHC values by 1.429 (1/0.7) before the model computes κ, the ocean heat diffusivity term, so as to scale the OHC from the upper 700 m to a value that approximates OHC for the full ocean.

2.2 EM-GC ocean components

The formulation for Q_{OCEAN} is based on finding the value of κ that best fits observed OHC data. Raper, Gregory, & Stouffer (2002) define κ as the ratio between the atmosphere-to-ocean temperature difference and the heat lost to the ocean. We assume Q_{OCEAN} is anthropogenically driven and we define the monthly values of $Q_{OCEAN\ i}$ as:

$$Q_{OCEAN\ i} = \kappa\left(\Delta T_{ATM,HUMAN\ i} - \Delta T_{OCN,HUMAN\ i}\right)$$
$$= \kappa\left(\left[\frac{1+\gamma}{\lambda_p}\{GHG\ RF_i + AER\ RF_i + LUC\ RF_i\}\right] - \Delta T_{OCN,HUMAN\ i}\right) \qquad 3)$$

$$\kappa = \frac{\Delta OHC \div A_{OCEAN}}{\int_{t_{START}}^{t_{END}}\left(\Delta T_{ATM,HUMAN} - \Delta T_{OCN,HUMAN}\right)_{i-72} dt} \qquad (4)$$

$$= \frac{OHE\ \times\ \Delta t}{\int_{t_{START}}^{t_{END}}\left(\left[\frac{1+\gamma}{\lambda_p}\{GHG\ RF_{i-72} + AER\ RF_{i-72} + LUC\ RF_{i-72}\}\right] - \left[f_o \sum_0^{i-72} Q_{OCEAN}\ \right]\right) dt}$$

The main improvement from Hope et al. (2017) is the inclusion of $\Delta T_{OCN,HUMAN}$, which represents the temperature response of the well-mixed upper 100 m of the ocean due to the total rise in OHC. In the previous version of the EM-GC, heat exited the atmosphere without any modulation from an ocean response (i.e. $T_{OCN,HUMAN\ i} = 0$ for all *i*) allowing the ocean to function as an infinite sink. By allowing the model ocean to warm in response to the increase in atmospheric temperature, the amount of heat lost to the ocean per month is reduced over time due to the smaller difference in temperatures between the ocean surface and overlying air, providing a more realistic description of the climate system than earlier versions of our model. The new interactive ocean module provides a reduction in Q_{OCEAN} over time compared to the earlier version of the model, resulting in a slight rise in total anthropogenic RF and in computed future global mean surface temperature relative to that found using a static ocean. This new model formulation thus also introduces a mechanism for the climate system to continue warming even after total anthropogenic RF plateaus, as it does in both RCP 4.5 and RCP 2.6.

The integral in the denominator and the delta time in the numerator of equation 4 are both taken over the entire time extent of the OHC record being considered, i.e. t_{START} and t_{END} are the first and last months corresponding to the years of the OHC record being used. Ocean Heat Export (OHE) is an average over area and time of the export of heat from the atmosphere to the ocean, obtained by estimating the total rise in OHC over time with a linear fit (Canty et al., 2013). We apply a six-year (72 month) lag to account for the time needed for a given amount of heat leaving the atmosphere to penetrate to depth (Schwartz, 2012). Other studies (Lean & Rind, 2008; Suckling et al., 2017) infer or apply a ten-year lag; key model outputs such as AAWR are insensitive to choices for the time delay between atmospheric perturbation and mean oceanic response (equation 4) for any timescales ranging from annual to multidecadal. The new formulation for Q_{OCEAN} allows the model parameter κ to be directly compared to literature values derived from GCMs (Raper et al., 2002).

The term f_o in the denominator of the last part of equation 4 represents a combination of the heat capacity of ocean water, the fraction of ocean volume in the surface layer of interest, and the fraction of total Q_{OCEAN} that warms the surface layer. To calculate f_o, decadal ocean warming as a function of depth was extracted from a selection of CMIP5 models' output, smoothed, and then normalized to the warming in the layer from 0-100 m. A simplified warming profile was then selected for the remaining depth of the ocean down to 4 km, (green segments of **Figure S5**) favoring the group of warming profiles from RCP 4.5 and RCP 8.5. This stratified warming profile was used in combination with the ocean depth profile to determine the percentage of ocean heat export that warmed the 0-100 m layer, producing $\Delta T_{OCN,HUMAN}$. This 100 m top layer is used as the section of ocean that interacts directly with the atmosphere, because it is well-mixed. We represent the ocean as being 1 km deep for 10% of the ocean area (representing the continental shelves) and 4 km deep for the remaining 90%. This simplified depth profile approximates the average depth of the real ocean to within 3%, 3.7 km compared to 3.682-3.814 km (Charette & Smith, 2010); using the ocean surface area estimate of 3.3×10^8 km^2 from (Domingues et al., 2008), our simplified ocean also approximates the total volume of the real ocean to within 10%, 1.221×10^9 km^3 compared to $1.33\text{-}1.37\times10^9$ km^3. Taken together, this CMIP5-based warming profile with depth implies that 13.7% of the rise in total OHC occurs in the well mixed, upper 100 m of the ocean, resulting in the $\Delta T_{OCN,HUMAN}$ term in equations 3 and 4. As a result, the value of f_o in equation 4 is 8.76×10^{-5} °Cm^2/W.

Output from the ocean module, Q_{OCEAN}, is area corrected to scale the average forcing applied to the atmosphere by the ocean before this quantity is used in the MLR. The ocean module is based upon the total surface area of the world's oceans, but the inputs to the atmospheric module are applied to the entire surface area of the Earth. As such, we scale Q_{OCEAN} in the model atmosphere by the ratio of ocean surface area to Earth's total surface area, (i.e. the multiplier 0.671 in equation 1a,) to ensure that the total amount of energy leaving the atmosphere is the same as the total amount of energy entering the oceans.

Four alternate values for the fraction of OHC in the upper 100 m were also considered to test the sensitivity of future atmospheric temperatures to the ocean's response to global warming. At one extreme, warming due to the rise in OHC is distributed linearly in just the upper 1 km of the ocean with no warming deeper, putting 18.2% of the rise in OHC into the top 100 m of the global ocean. At the other extreme, a warming profile that assumes a constant warming rate throughout the entire ocean has only 2.7% of the rise in OHC going into the upper 100 m. Both scenarios are physically unrealistic but provide bounds for the range of how much $\Delta T_{OCN,HUMAN}$ can change for a given OHC record. All five warming profiles with depth and their associated top 100 m fractions are summarized in **Table S1**. Choice of ocean warming profile does not affect our results significantly, as the interplay between $\Delta T_{OCN,HUMAN}$ and κ means that Q_{OCEAN} is largely driven by the choice of OHC_{OBS}.

2.3 Climate feedback and sensitivity

Climate feedback processes and overall climate sensitivity define how changes in RF, particularly the rise in anthropogenic RF, drive ΔT. In the EM-GC, the sum of RF due to GHGs, aerosols, LUC, and OHE is multiplied by $(1+\gamma)/\lambda_P$, where λ_P is the Planck response parameter (3.2 W m^{-2} °C^{-1}) and γ is the dimensionless climate amplification term, to determine ΔT (equation 1a). If the net response of changes in humidity, lapse rate, clouds, and surface albedo that occur in response to anthropogenic RF of climate is positive, as is the case for the vast majority of simulations conducted for this study, then the numerical value of γ is positive. This model framework is based on that described in Bony et al. (2006) and section 8.6 of the previous

IPCC report (Solomon, 2007). The EM-GC's variable for the sum of climate feedback mechanisms, λ_Σ, has units of W m^{-2} °C^{-1} and is related to γ and λ_P via:

$$1 + \gamma = \frac{1}{1 - \frac{\lambda_\Sigma}{\lambda_P}} = \frac{\lambda_P}{\lambda_p - \lambda_\Sigma} \qquad 5)$$

This relation between γ and λ_Σ is commonly used in the climate modeling community (section 8.6 of Solomon (2007)). We can also relate λ_Σ to the traditional climate feedback parameter λ (Bony et al., 2006; Gregory, 2000; Schwartz et al., 2014; Sherwood et al., 2020) by reducing equation 1a to just the anthropogenic terms to produce the relations:

$$\Delta T_{Human\ i} = \frac{1+\gamma}{\lambda_P}\{GHG\ RF_i + AER\ RF_i + LUC\ RF_i - Q_{OCEAN\ i}\} \qquad 6a)$$

$$\Delta T_{Human} = \frac{1+\gamma}{\lambda_P}\Delta F_{Human} \qquad 6b)$$

$$\lambda^{-1} = \frac{1+\gamma}{\lambda_P} \rightarrow \lambda = \lambda_P - \lambda_\Sigma \qquad 6c)$$

We choose to focus on λ_Σ instead of λ for the majority of this paper. This choice allows for an intuitive comparison between λ_Σ, γ, and ΔT – as one quantity rises, so do the others. This intuitive relationship highlights how uncertainty in various climate feedback mechanisms (i.e. aside from the blackbody response) can be the driving force in future ΔT uncertainty. We assume a constant value for λ_Σ (and λ) for each ensemble member in most of the results shown below, as this assumption provides a multitude of simulations of ΔT with χ^2 values less than 1, well below our acceptable fitting limit of 2. We view this as a reasonable approximation because section 12.5.3 of AR5 and references therein suggest λ changes slowly over millennia; any changes in λ over a few centuries should be unnoticeable unless gradual changes force the climate system past a significant tipping point. For completeness, we also examine the effect on EM-GC ΔT of a slowly- or moderately-varying λ in section §3.4 of this paper (also see **Table S2** and **Figure S6**).

3. Results and Analysis

Our results focus mainly on Attributable Anthropogenic Warming Rate (AAWR) and projections of the global mean surface temperature anomaly relative to preindustrial (ΔT). We first present here a summary of the probabilistic distribution of these two quantities for our best representative ensemble of EM-GC simulations, and next describe how these distributions compare to CMIP5 and other studies. Then, in each subsection to follow, we delve further into our results for AAWR and ΔT_{2100}, providing a detailed description of their sensitivities as well as comparisons to other published results. The first three subsections present discussion of AAWR, in which we describe our approach, possible shortcomings in prior efforts used to evaluate AAWR, and the uncertainties involved in proper quantification of AAWR. The last five subsections present results for ΔT, including quantification of the sensitivity to uncertainty in future emissions of CH_4 and relating our projections of future warming to cumulative, anthropogenic emissions of CO_2.

Overall, our numerical estimates of AAWR for 1979 to 2100 fall between prior estimates. Our best estimate of AAWR is ~0.14°C/decade, which is noticeably lower than the value for AAWR from CMIP5 GCMs (~0.22°C/decade). Our value for AAWR falls between estimates of AAWR from FR11 (0.170±0.012) and the AMOC-based AAWR from ZT13 (0.070±0.019). Below, we describe the sensitivity of AAWR to various estimates of radiative forcing by aerosols, the sum of climate feedback mechanisms, and multiple records for $\Delta T_{OBS.}$ Notably, AAWR in our model is largely insensitive to whether AMOC is included (see §2.1.1 and **Figure S7a and S7b**).

Our ensemble median value for global warming at the end of this century, ΔT_{2100}, is consistently cooler than the CMIP5 ensemble median value for ΔT_{2100}. Indeed, our ensemble median of ΔT_{2100} often lies close to the CMIP5 ensemble minimum warming. The EM-CG framework, with its tendency for cooler results, assigns each RCP scenario a higher probability of fulfilling the Paris Agreement warming limitations, compared to the CMIP5 GCMs. The near-term warming found by our EM-GC also closely matches the expert assessment of CMIP5 results shown in chapter 11 of AR5, represented by the green trapezoid in Figure 2.

Figure 4 shows AAWR and ΔT_{2100} for the same ensemble run of the EM-GC and depicts the weighting function we use to create probabilistic summaries of our results. Each simulation in this ensemble has the same set of inputs except for varying λ_Σ and varying the shape and strength of anthropogenic aerosol forcing, pinned to the value of AER RF in 2011. Computed values of AAWR are sensitive to AER RF and λ_Σ because of differences in the shape of the aerosol term (blue line Figure 2a) that is subtracted from the GHG term (red line Fig 2a), for various members of the ensemble. Values of ΔT_{2100} are particularly sensitive to climate feedback, because by end of century the RF due to aerosols is expected to be considerably smaller than contemporary values (Smith and Bond, 2014). The EM-GC ensemble shown in Figure 4 is based on RCP 4.5 GHGs and constrained by the CRU4 record for ΔT_{OBS} from 1850 to 2019. The values of AAWR and ΔT_{2100} shown in Figure 4 are for those members of the ensemble for which all three χ^2 filters yield a value less than or equal to 2: i.e., those sets of model results able to provide a "good fit" to ΔT_{OBS} from 1850 to 2019, from 1940 to 2019, and to OHC averaged among five data centers from 1955 to 2018. We also eliminate any simulations for which AER RF_{2011} does not lie between –1.9 and –0.1 W m^{-2}, the 5% and 95% confidence intervals for RF due to anthropogenic aerosols given in AR5, which is why model results shown on the left side of Figure 4b and 4c are shown in grey. **Figure 5** then aggregates these results into probability density functions (PDFs) of AAWR and ΔT_{2100} from the EM-GC (blue) using a weighting method, described in section §3.1, that is based on the AR5 likelihoods for the values of AER RF_{2011} (Figure 4a). Similar PDFs based on results from 41 CMIP5 GCMs (without any weighting, red) are also shown in Figure 5.

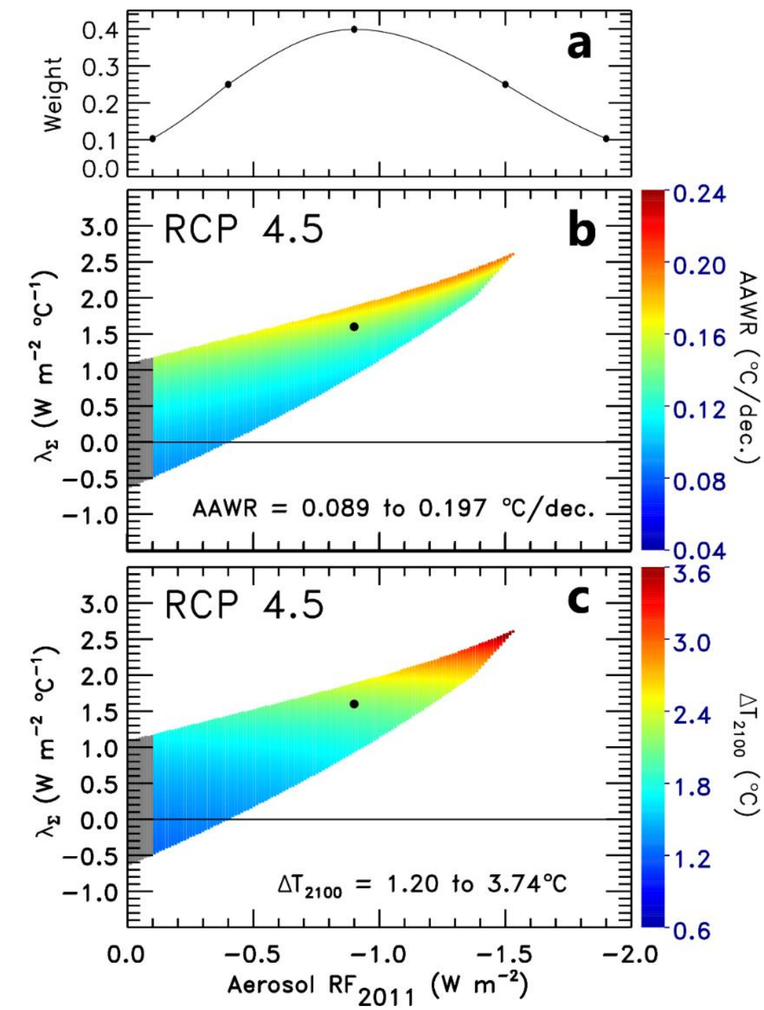

Figure 4. AAWR and ΔT_{2100} as a function of AER RF$_{2011}$ and λ_Σ over a single EM-GC ensemble (panels b and c) and the AER RF$_{2011}$-based weighting function used to aggregate our ensemble statistics (panel a). Values of AAWR are for 1979 to 2010 and values of ΔT_{2100} are relative to the preindustrial baseline. We show all simulations for which $\chi^2 \leq 2$ for all three fitting comparisons, (i.e. fitting ΔT for the full time period, fitting ΔT for the most recent 80 years, and fitting OHC over its time period,) though any model results that for AER RF$_{2011}$ outside of the range –0.1 to –1.9 W m^{-2}, the 5% and 95% confidence intervals given in AR5, are covered using the color grey. All runs in this ensemble use RCP 4.5 GHG RF and RCP 4.5-based AER RF scenarios along the middle road of Figure S4 to simulate the CRU4 ΔT_{OBS} record. The black dot in panels b and c represents the single run from the ensemble with the lowest χ^2 over the full ΔT record among simulations forced with the AER RF times series that gives −0.9 W/m^2 in 2011, the best estimate for AER RF$_{2011}$ stated in AR5. Results from this single simulation, broken into component time series, are shown in Figure 3.

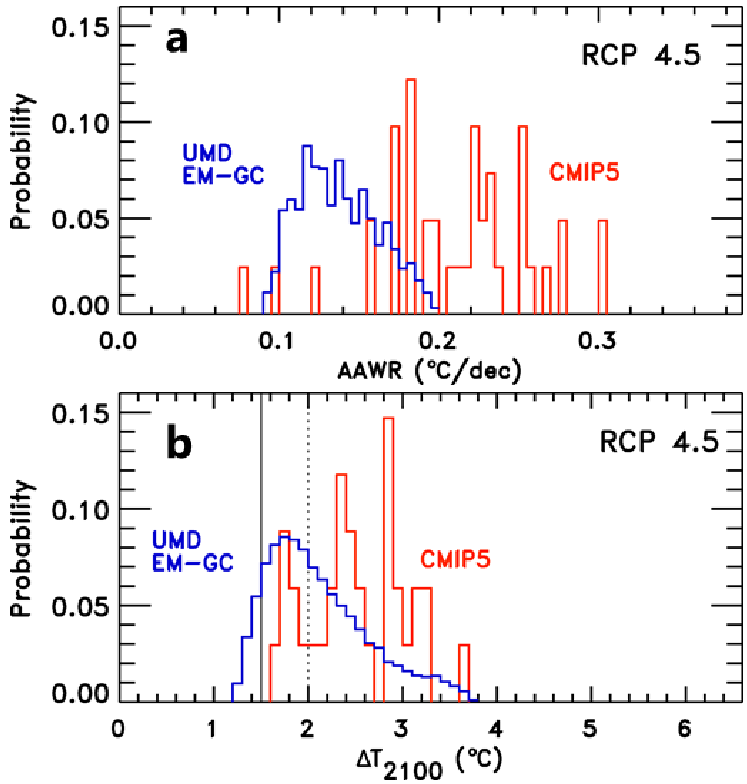

Figure 5. Probability density functions (PDFs) of the EM-GC computations of AAWR and ΔT_{2100} for RCP 4.5 shown in Figs. 4b and 4c, weighted by the associated value of AER RF_{2011} using the weighting function shown in Figure 4a (blue lines). PDFs of AAWR and ΔT_{2100} from CMIP5 GCMs, also for RCP 4.5, are also shown (red lines). The height for each bin (0.1°C width for ΔT_{2100}, 0.05°C/decade width for AAWR) for the UMD EM-GC PDFs represents the probability of a run with that value being randomly selected from the respective model output shown by non-grey colors in Figure 4b and 4c, when each model run is weighted by the AER RF_{2011}-based weighting function shown in Figure 4a. Similar probabilities can be taken from the CMIP5 ensemble giving results from each GCM equal weighting (red). For panel b, the Paris Agreement goal of 1.5°C warming and upper limit of 2°C warming are shown by the vertical solid and dotted lines, respectively.

3.1 AAWR from the EM-GC

We base our estimate of AAWR on the slope of ΔT_{HUMAN} over the years 1979 to 2010. To calculate AAWR, our empirical model is first run over a chosen time period to produce the ΔT_{HUMAN} series; then a linear fit is calculated from this series over the years 1979 to 2010.

Except when otherwise stated, the chosen time period for the model run is the entire available ΔT record, which for the CRU4 record used in Figure 3 is January 1850 to December 2019. To choose which ensemble member from Figure 4 would be shown in Figure 3, we selected the run with the best estimate of AER RF_{2011} from AR5 (−0.9 W/m^2) that had the lowest value of χ^2 for fitting the full ΔT_{OBS} record. This selected run gives a value for AAWR of 0.146 °C/decade.

To aggregate the EM-GC ensemble results from Figure 4b, we assign probabilities to each ~150 long time series for the RF due to aerosols, tied to the value of AER RF_{2011} for each time series. We create an approximate Gaussian distribution of AER RF_{2011} based upon AR5 estimates of this quantity (Figure 4a). This weighting function peaks at –0.9 W/m^2 (AR5 best estimate of AER RF_{2011}) and the cumulative probability of AER RF_{2011} values between –0.4 and –1.5 W/m^2 is set at 66.7%. Similarly, the cumulative probability of the weighting function between –0.1 and –1.9 W/m^2 is 90%, which corresponds to the AR5 specification of –0.1 and –1.9 W/m^2 being the 5% and 95% confidence intervals for this quantity (Myhre et al., 2013). We then take all runs shown as non-grey colors for AAWR in Figure 4b (i.e. all runs for which a good fit to ΔT for the full time period, ΔT for the most recent 80 years, and OHC from 1955 to 2018 can be obtained), bin by AER RF_{2011}, and find the probability distribution for AAWR within each of these bins. The PDFs for AAWR within each bin are then aggregated using the IPCC-based weightings (Figure 4a) for each value of AER RF_{2011} (Figure 4a) to create the final PDF shown as a blue line in Figure 5a. We use this superposition of PDFs weighting method to account for the fact that the EM-GC finds many more acceptable fits to the climate record (i.e. $\chi^2 \leq 2$) for combinations of λ_Σ and AER RF_{2011} associated with less-negative values of AER RF_{2011}, whereas AR5 suggests that –1.5 W/m^2 is as likely as –0.4 W/m^2 for the RF of climate in 2011 due to aerosols. This weighting method gives model runs with stronger aerosol cooling the same weight as runs with weaker aerosol cooling. For the AAWR ensemble shown in Figure 4b, this weighting process produces a median of 0.135 °C/decade, with a full range of 0.089 °C/decade to 0.197 °C/decade. Through most of the ensemble, the resulting time series for ΔT_{HUMAN} and the resulting values of AAWR agree well with another recent estimate found using a similar approach (Chylek et al., 2014).

Our estimate of AAWR is sensitive to which aerosol forcing time series is used, especially in relation to λ_Σ, and is partially sensitive to ΔT_{OBS}, but insensitive to the inclusion of terms for AMV, the PDO, and the IOD in the model framework. The insensitivity of AAWR to

AMV extends to all of the other proxies for variations in the strength of the AMOC we have considered. The modeled strength of the PDO varies noticeably depending on the proxies and filtering methods chosen for both this climate signal and as well as AMV and RF due to aerosols. Specifically, the contribution of the PDO to ΔT_{MDL} increases in magnitude with stronger-cooling aerosol scenarios – but these model results do not show any strong effect on AAWR. We can also run the EM-GC with specific single records for OHC instead of using the average of five OHC data records; varying the input OHC time series does not produce any noticeable variation in AAWR.

One EM-GC simplification that deserves mention is the lack of spatial variability in the effect of the oceans. Rose et al. (2014) showed that the climatic effect of ocean heat uptake is weaker if heat export from the atmosphere is concentrated in the tropics and stronger if heat export is concentrated in high latitudes. While the EM-GC cannot directly separate the locality of ocean heat export, it corroborates the Rose et al. (2014) result in the sense that almost all runs show a stronger climatic signal from the AMOC (driven by high-latitude deep water formation) and a weaker signal from the PDO (an expression of comparatively shallow-water heat storage in the tropics (England et al., 2014); see §2.1.1 for a summary of the various AMOC and PDO proxies tested). While various other MLR studies (Chylek et al., 2016; Zhou & Tung, 2013) focus on the AMOC as the main oceanic driver of the climate system, other literature suggests the PDO has a stronger influence on global temperature, either overall or specifically for the last few decades (England et al., 2014; Steinman et al., 2015; Tokarska et al., 2019). Due to the structure of MLR models, finding regression coefficients for time spans less than the multidecadal characteristic time of known natural variability is not practical, nor is attempting to define climate drivers using fewer total years than this characteristic time scale. Some research suggests that the sign of the PDO is what drives trends in ΔT (England et al., 2014), meaning an integral of the original PDO time series might be a stronger regressor. However, using a time series calculated as such did not produce lower values of χ^2 or higher values of the PDO regression coefficient than found using the raw PDO signal, further suggesting that AMOC is likely the stronger driver of variations in ΔT.

Our method of determining AAWR is also relatively insensitive to the choice of beginning and end years (**Table S3**). For example, using the EM-GC simulation shown in Figure 3, AAWR only varies from 0.130 to 0.156 °C/decade when we vary both the initial year (1975 to

1985) and final year (2006 to 2016) around the default AAWR time range of 1979 to 2010. This insensitivity derives from the fact that ΔT_{HUMAN} follows from the known rise in CO_2, CH_4, and N_2O that leads to a RF of climate due to GHGs that rises in a roughly linear manner over the past four to five decades. Our calculation of AAWR is thus robust and, as detailed in the following section, does a better job of isolating the underlying anthropogenic climate trend than methods that rely on analysis of ΔT using the residual method.

3.2 Comparison to previous AAWR estimates

We assert that the slope of ΔT_{HUMAN} provides a more accurate method for quantifying AAWR than the use of a residual method. As described in section §3.1, our median estimate of AAWR with RCP 4.5 GHGs and "middle road" aerosols is 0.135 °C/decade, with range of possible values extending from 0.089 °C/decade to 0.197 °C/decade based on uncertainty in RF due to aerosols and climate feedback. The best estimate of AAWR given by FR11 for 1979 to 2010, upon analysis of ΔT from CRU3, is 0.170 °C/decade. The residual method used by FR11 involves finding the slope of observed ΔT after the contributions from solar irradiance, volcanoes, and ENSO have been removed. By not including AMV (green curve in Figure 3c) and by focusing their analysis solely on a 31 year period of time, FR11 do not account for the significant warming trend that occurred from 1979 to 2010 that our analysis suggests is due to natural variability and instead attributed this component of the rise in ΔT to anthropogenic warming. Although the precise magnitude of the AMV influence is sensitive to how North Atlantic SST is detrended (Canty et al., 2013) and smoothed, an independent analysis of SST using spectral methods (DelSole et al., 2011) supports our suggestion that internal climate variability contributed significantly to the relative warming over our default time period for AAWR.

We note that Haustein et al. (2019) emphatically state "we argue that AMV must not be used as a regressor" because "AMV is found to be primarily controlled by external forcing". Booth et al. (2012) implicate tropospheric aerosols due to pollution and stratospheric sulfate aerosols due to major volcanic eruptions of the primary driver of SST variability in the North Atlantic. There are numerous other studies making similar claims (Knight et al., 2005; Medhaug & Furevik, 2011; Meehl et al., 2011; Stouffer et al., 2006), including a study of paloecurrent speed that extends over a time period of 230 years (Boessenkool et al., 2007), that suggests our

AMV proxy does represent interval variability of the climate record. Regardless, we find nearly identical values of AAWR based on the slope of ΔT_{HUMAN}, with or without the use of AMV as a term in the regression model (Figure S7a vs S7b). We are therefore confident FR11 have overestimated the true value AAWR, either because there is a component of natural variability present in the residual they have computed, or because their analysis is restricted to such a short period of time. In contrast, our computation of AAWR for 1979 to 2010 is found using a physical model that provides consistent treatment of RF due to GHGs, aerosols, and natural factors such as ENSO, TSI, and SAOD, over a century and a half period, which mitigates a host of potential complications present when one examines a residual (Silver, 2012).

Other studies suggest the PDO or changes in SAOD from minor volcanic eruptions could have also played a role in driving variations of ΔT over this time period (e.g. Tokarska (2019) and the references therein). Our estimates of AAWR include all of these factors and based on analysis of data collected over a ~150 year time period; in our model framework the most important natural drivers of ΔT over the past four to five decades are ENSO, major volcanic eruptions, and AMV. If temperature is affected by a natural process not represented by the exogenous factors used to compute the residual, then quantification of AAWR will be unduly influenced (Supporting Information Text S2 and Figure S7). As shown in our SI, the difference between our best estimate of AAWR and that given by FR11 is nearly completely explained by the proper attribution of the signal from variations in the strength of AMOC.

Conversely, the estimate of AAWR over 1979 to 2010 provided by ZT13 upon consideration of the variations in the strength of AMOC is likely biased low. They suggest AAWR drops from 0.170 °C/decade in a regression without AMV to 0.070 °C/decade with AMV included. Even though they considered AMV as a proxy for variations in the strength of AMOC, they used a linear function to describe ΔT_{HUMAN} over the entire 1860 to 2010 time period as an input to their MLR. We are able to closely reproduce their estimate of AAWR if we replace our formulation of ΔT_{HUMAN} with a linear function spanning 1860 to 2010 (Supporting Information Figure S7c+d). However, it is well known that anthropogenic RF of climate, which drives ΔT_{HUMAN}, has varied in a non-linear manner that generally follows human population over the past century and a half. While ZT13 state that the use of an RCP-shaped anthropogenic forcing causes trends in their computed residual between ΔT_{OBS} and ΔT_{MDL}, we cannot reproduce this result. With our use of RCP-based anthropogenic forcing that underlies the

CMIP5 GCMs, for the entire historical record (1850-present) and our method of calculating AAWR, we find that both ΔT_{HUMAN} and AAWR are insensitive to the inclusion or exclusion of a proxy for AMOC in the regression analysis (Supporting Information Text S2).

The uncertainty in AAWR is likely much higher than the small values suggested by FR11 and ZT13. As detailed in section §3.3, our estimate of AAWR based on the full uncertainty in AER RF and analysis of ΔT_{OBS} from multiple data centers spans the range 0.08 °C/decade to 0.20 °C/decade. FR11 state that the computation of AAWR upon use of ΔT_{OBS} from various data centers provides a range of 0.158 °C/decade to 0.187 °C/decade (these values are the 1σ lower and upper uncertainties of the standard error of their regression). The final estimate of AAWR given by ZT13 is 0.05 °C/decade to 0.09 °C/decade based solely on the mathematical uncertainty from calculating a linear fit to their ΔT_{HUMAN}. In our model framework, uncertainties in the strength and temporal shape of AER RF over the past four decades cause ΔT_{HUMAN} to vary much more than allowed by uncertainties from any linear fit to ΔT_{HUMAN}. The variation of AER RF used in our study results in a range for AAWR of 0.089 to 0.197 °C/decade for a single ensemble (Figure 4); this range extends slightly further to 0.084 to 0.202 °C/decade when considering all ensembles. Figure 4 and **Figure S8** show that EM-GC runs with small amounts of aerosol cooling tend to have both lower values of χ^2 (i.e good fits to the climate record span a wider range of values for λ_Σ) and lower values of AAWR than model runs constrained by larger aerosol cooling.

Individual runs demonstrating the effect AER RF has on ΔT_{HUMAN} and χ^2 are highlighted in **Figure 6**. As AER RF cooling was largest in the 1970s and decreased in past decade (Smith and Bond, 2014), larger aerosol cooling implies higher values of AAWR due to the nature of the definition of ΔT_{HUMAN} (equation 6a). This relationship explains why Figure 6a, with a relatively weak AER RF_{2011} value of –0.4 W/m^2, has a relatively low AAWR value of 0.128 °C/decade; conversely, the simulation in Figure 7c with strong aerosol cooling (AER RF_{2011} of –1.4 W/m^2) results in a relatively high AAWR value of 0.169 °C/decade. If the global warming due to black carbon aerosols and co-emitted species over the industrial era were as large as the best-estimate of Bond et al. (2013), this term would place the actual value of AER RF_{2011} close to –0.4 W/m^2 rather than the AR5 best estimate of –0.9 W/m^2, rendering AAWR well below the best estimate of 0.170 °C/decade given by FR11. On the other hand, if the cooling of climate due to anthropogenic aerosols was as large as suggested by the recent study of Shen et al. (2020), this finding would likely place AER RF_{2011} close to –1.4 W/m^2, leading to a value of AAWR similar to the 0.170 °C/decade estimate of FR11. FR11 do not address the quite large uncertainty in AAWR due to imprecise knowledge of the RF of climate by tropospheric aerosols.

There remains an important distinction to be made when comparing values of AAWR based on specific estimates of AER RF_{2011}. While Figure 6b shows that using an AER RF time series with the AR5 best estimate for AER RF_{2011} (–0.9 W/m^2) gives an AAWR of 0.146 °C/decade, whereas our weighted ensemble median value of AAWR is 0.135 °C/decade. The lower value for the ensemble median follows from how our χ^2 strength-of-fit filters eliminate more runs with stronger aerosol cooling than runs with weaker aerosol cooling. Runs that use weaker aerosol cooling have lower resulting values for AAWR, so even though our ensemble weighting method theoretically assigns equal weighting to runs with –0.1 W/m^2 and –1.9 W/m^2, the relative lack of high values of AAWR corresponding to stronger aerosol cooling that pass this filter causes the weighted median AAWR to be slightly lower than 0.146 °C/decade.

We state our estimate of AAWR for 1979 to 2010 as 0.14 ± 0.06 °C/decade, where the uncertainty covers the full range of model runs that yield a good fit to ΔT_{OBS} from CRU4. Our estimate of AAWR is larger than the trend in lower tropospheric temperature of 0.096 ± 0.012 °C/decade reported by CM17. Their estimate is based upon analysis of satellite and radiosonde measurements of temperature throughout the global lower troposphere (GLT, the atmospheric layer from the surface to approximately 300 hPa) over the time period Jan. 1979 to June 2017. Similar to FR11, CM17 do not address the contribution of imprecise knowledge of AER RF to their estimate of AAWR, which leads to their small uncertainty for AAWR compared to our uncertainty. We reach similar results, though, when comparing the drop seen between the trend in observed ΔT and the trend after removing natural components of ΔT. They report a significant difference between the temperature trend for raw data (0.155 °C/decade) and the trend after the data have been adjusted to account for natural influences due to major volcanoes and ENSO (0.095 °C/decade). Our best estimate of AAWR for the Jan. 1979 to June 2017 time period, based on a linear fit to ΔT_{HUMAN} shown in Figure 3a, remains ~0.14 °C/decade, compared to the trend in ΔT_{OBS} of ~0.18 °C/decade for this time period. As such, we compute the drop from the trend in ΔT_{OBS} to trend in ΔT_{HUMAN} to be about two-thirds of the corresponding drop reported by CM17. Given the presence of two major volcanic eruptions in the first half of this time period and a major ENSO event in 2015-16, the drop in value from the observed trend to the anthropogenic tread should be expected. We caution that precise determination of the effect of major volcanic eruptions for analysis of GLT data collected during the satellite era is affected by whether or not one includes the effect of AMV in the analysis (section 4.5 of Canty et al., 2013), which may explain the difference between our drop and the drop from CM17. Finally, and most importantly, CM17 emphasize (i.e. their figure 2) that CMIP5 GCM simulations of GLT result in considerably more rapid warming than is discerned from their adjusted observations. We reach a similar conclusion based on our analysis of ΔT, as described in the following section.

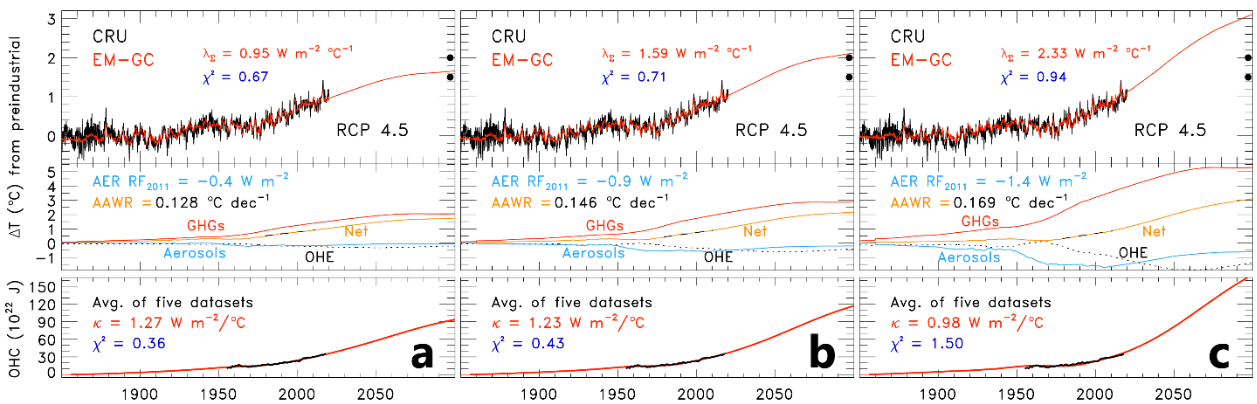

Figure 6. Observed and modeled ΔT, 1850 to 2019, as well as projected global warming to 2100. The model runs pictured are identical to the run in Figure 3, except 7a and 7c use alternate middle road constructions of AER RF. We show a run with AER RF_{2011} of −1.4 W m^{-2} instead of −1.5 W m^{-2} because the simulation with −1.5 W m^{-2} exists at the far edge of acceptable χ^2 values, producing unrealistic individual anthropogenic components and OHC_{2100}. (AER RF_{2011} of −0.4 W m^{-2} to −1.5 W m^{-2} would match the upper and lower limits respectively of AR5's likely range of anthropogenic, tropospheric forcing values in 2011 relative to preindustrial values). The upper rung of each abbreviated ladder plot here is the same format as those in figure 3a. The second rungs show the anthropogenic effect on the climate in gold as well as three of its four components: the temperature rise from GHG forcing (red), the temperature fall from aerosol cooling (light blue), and the temperature fall from OHE (dashed black). For clarity, the LUC component is not shown as its value is consistently near-zero compared to the other components.

3.3 Comparison to AAWR from GCMs

In this section, we conduct a comparison of estimates of AAWR found using our EM-GC to AAWR inferred from CMIP5 GCMs. First, we further characterize uncertainties in AAWR. **Figure 7** shows the sensitivity of AAWR to AER RF and the choice of data record for ΔT_{OBS}. The middle of box and whicker (BW) plot under the label CRU4 summarizes the weighted median (0.135 °C/decade), weighted interquartile range (IQR), and extrema of AAWR for the EM-GC determined PDF shown in Figure 5a.

The EM-GC ensembles shown thus far relied only on time series of AER RF found using scaling factors for aerosol cooling (α_{COOL}) and heating (α_{HEAT}) along the "middle road" of Figure S4 (section §2.1.2). The shape of the AER RF time series varies by choosing values for α_{COOL} and α_{HEAT} along either the "low" or "high" road of Figure S4. The dashed BW plots surrounding the solid BW labeled CRU4 show the AAWR ensemble changes by a small amount, upon modification of the shape of the input AER RF time series. Scatter plots of AAWR versus AER RF_{2011} and λ_Σ as well as ladder plots documenting the computation of ΔT_{HUMAN} and AAWR for the AR5 best estimate of AER RF_{2011}, for these "high" and "low" road simulations, are shown in **Figures S8** and **S9**. The value of AAWR exhibits only a small sensitivity to variation of the shape of AER RF. We include this comparison for the CRU4 record of ΔT_{OBS} because the various time series for AER RF that underlie the model inputs for these simulations cover a large range of possibilities, similar to that shown in figure 4 of Smith & Bond (2014).

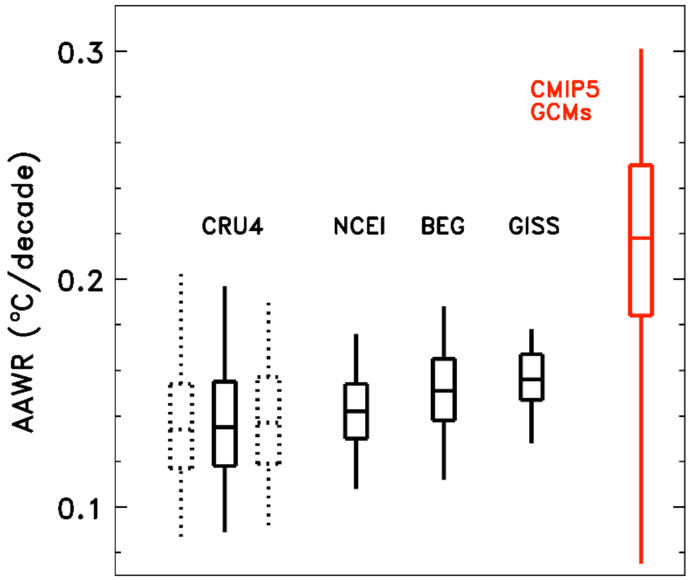

Figure 7. Comparison of EM-GC values of AAWR to CMIP5 GCM values of AAWR, for 1979 to 2010. Each black box-and-whisker plot shows the weighted median, weighted IQR, and extrema from PDFs for AAWR found using the EM-GC, as shown in Figure 5. The solid black box-and-whisker plots represent EM-GC ensembles fit to records of ΔT_{OBS} from various data centers (as indicated), found using RF from RCP 4.5 with the middle road construction of AER RF as shown in Figure S4. The dashed box-and-whisker plots show AAWR for ΔT_{OBS} from CRU4, found using the low (left) and high (right) road constructions for AER RF (section §2.1.2); for simplicity, this comparison is not shown for the other temperature records. Among all of the EM-GC results, the maximum value of AAWR is 0.202 °C/decade and minimum is 0.084 °C/decade. The red box and whisker plot at the right shows AAWR found using the regression method described in section §2.4, using archived output of ΔT from 112 individual CMIP5 GCM runs, all constrained by RCP 4.5.

The choice of data center for the ΔT_{OBS} record contributes another small uncertainty to AAWR. Figure 7 shows BW plots for ΔT_{OBS} from other data centers. The median values of AAWR for AER RF computed with α_{COOL} and α_{HEAT} along the "middle road" of Figure S4 are 0.142, 0.151, and 0.156 °C/decade for the use of temperature data from NCEI, BEG, and GISS, respectively. The choice of data center exerts a difference of 0.021 °C/decade between the largest and smallest median values, which is quite a bit larger than the value of 0.005 °C/decade difference reported by FR11 for selection of data between CRU3, NCDC (this dataset is now termed NCEI), and GISS. We find the largest value of AAWR upon use of ΔT_{OBS} from GISS and the smallest value upon use of data from CRU4. We have featured data from CRU4 throughout our paper, as well as in our earlier studies (Canty et al., 2013; Hope et al., 2017) because so many other published papers over the prior decade have used CRU temperature records as their baseline dataset. The 0.021 °C/decade difference in AAWR that arises between our use of ΔT from CRU4 compared to GISS is much smaller than the difference between any AAWR from the EM-GC and the Coupled Model Intercomparison Project (CMIP5) (Taylor et al., 2012) GCM multi-model mean value of AAWR.

The far-right red BW plot in Figure 7 shows AAWR from the CMIP5 GCMs used by AR5. Archived output from the CMIP5 (Taylor et al., 2012) for 112 runs of 41 GCMs driven by RCP 4.5 (Thomson et al., 2011) has been used to estimate AAWR from these GCM results (Supporting Information Text S1). We find good agreement between values of AAWR from these 112 runs found using two analysis methods, termed linear fit (LIN) and regression (REG) (Supporting Information **Figure S10, Figure S11, and Table S4**). For the AAWR found using LIN, we perform a linear least squares regression to archived output of global mean two-meter air temperature (TAS) or years 1979 to 2010, ignoring years with obvious of major volcanic eruptions (1982, 1983, 1991, and 1992). The rationale behind this method is natural variability in TAS due to internally model generated ENSO events will be randomly distributed in time; the influence of variations in the strength of AMOC on TAS tend to be small within these GCMs (Kavvada et al., 2013). For AAWR found using REG, we perform a multiple linear regression of TAS versus TSI and SAOD in a two-step process, as described in section 2.6 of Hope et al. (2017). Values of AAWR found using both methods are also given in Table S4 for each GCM. Comparing AAWR from LIN versus REG shows the two methods result in values of AAWR with a high correlation coefficient ($r^2 \geq 0.95$) and a mean ratio close to 1, providing confidence that AAWR has been computed accurately from the CMIP5 GCMs.

All of the CMIP5 GCM output used here are global, two-meter air temperature (TAS) data. According to Cowtan et al. (2015), the blending of TAS (over land) with GCM output of sea surface temperature (SST, termed TOS in the CMIP5 archive) provides a more appropriate manner for sampling GCM output than use of global TAS, since datasets such as CRU4-based ΔT are a combination of near surface air temperature over land and SST over water. Cowtan et al. (2015) state the use of blended temperature rather than air temperature accounts for 25% of the difference between the GCM-based and observed variations in global temperature over 2009–2013 and 38% of this difference over 1975-2014. Our own analysis using TAS-TOS blended temperature from CMIP5 GCMs results in a reduction of AAWR by roughly 2-5%, depending on which GCM is considered. This 2-5% reduction in GCM-AAWR translates to explaining 6-14% of the difference between median GCM-based and median EMGC-based values of AAWR, as well as 11-28% of the difference between GCM-based AAWR and the observed CRU4 slope of 0.18 °C/decade over the AAWR time period. While our use of blended temperature from a handful of CMIP5 GCMs rather than TAS does lead to a reduction in GCM-based values of AAWR, we find this effect is small (i.e. 2 to 5 %). A similar conclusion was reached by Tokarska et al. (2020). Therefore, other than this paragraph, our paper focuses entirely on TAS from the GCMs because the use of blended temperature introduces a modest effect that does not alter any of our major conclusions, plus the information needed to produce blended temperature is no longer available on the CMIP5 archive for enough GCMs to complete an ensemble similar in size to our initial the GCM ensemble shown in Figures 4 and 5.

The median value of the CMIP5 GCM-based AAWR found with the regression method 0.22 °C/decade. This value for AAWR is slightly more than 50% larger than our best empirical estimate of 0.14 °C/decade. The IQR for AAWR inferred from CMIP5 GCMs is 0.184 to 0.250 °C/decade, and the extrema are 0.075 and 0.301 °C/decade. More than two-thirds of the 112 archived CMIP5 GCM runs (Table S4) exhibit a value for AAWR larger than our upper limit of 0.202 °C/decade, and there is no overlap between the CMIP5 IQR and any IQR from the EM-GC – the 25^{th} percentile of the CMIP5 ensemble is 0.184 °C/decade, while the highest 75^{th} percentile from an EM-GC ensemble is 0.167 °C/decade (Figure 7, GISS). Also, only 3 of the 41 CMIP5 GCMs exhibit a value of AAWR less than 0.14 °C/decade: INM-CM4 (Volodin et al., 2010), IPSL-CM5B-LR (Hourdin et al., 2013), and MRI-CGCM3 (Yukimoto et al., 2012) (Table S4). We conclude therefore that the large majority of the CMIP5 GCMs exhibit anthropogenically induced warming that is considerably more rapid than what has actually occurred over the time period 1979 to 2010. This finding is not closely tied to the chosen time period for AAWR: whereas AAWR does exhibit some dependence on start and end year (Table S3), the median value of CMIP5 GCM-based AAWR exceeds these empirical values by about the same amount for any similar time period. The tendency of most GCMs to overestimate empirical AAWR is evident in plots of time series of archived ΔT shown in AR5 (i.e. figure 11.25a) and persists whether the GCM output is examined in terms of individual runs, various GCMs, or specific modeling centers (Supporting Information Figure S11, Table S4). A similar tendency of GCM-based warming rates to lie considerably above empirical estimates has been noted by multiple recent, complementary studies (Christy & McNider, 2017; Chylek et al., 2014; Fyfe et al., 2013; Millar et al., 2017; Tokarska et al., 2020).

A GCM retrospective paper by Hausfather et al. (2020) shows that many past GCMs have predicted the observed rise in ΔT quite well. By comparing older GCM predictions of ΔT to the rise in ΔT_{OBS} since those predictions were made, Hausfather et al. (2020) show that the skill in predicting $\Delta T/\Delta t$ up to 2017 increased through the first three IPCC assessment reports. This ability to predictively match $\Delta T/\Delta t$ with the realized ΔT_{OBS} is quantified with unitless skill values, which increase from 0.63 to 0.73 to 0.81 for the first three assessment reports. However, the predictions of ΔT_{OBS} for GCMs used in the fourth IPCC assessment report overestimated observed $\Delta T/\Delta t$ from 2007-2017, resulting in an overall skill value of 0.56 despite being a short-term prediction using the most advanced GCMs available at the time. While this excess warming in the GCMs that underlie the fourth report is argued to be due to overestimated scenario RF, excess warming continued into the AR5 GCMs with CMIP5, as documented here and in figure 11.25b of AR5. Early indications are that the tendency for GCMs to warm more quickly than the actual climate system extends into the CMIP6 GCMs being used as the backbone of the sixth assessment report (Belcher et al., 2019; Tokarska et al., 2020; Voosen, 2019). While the increase in GCM complexity over the years certainly provides many benefits, it appears that such complexity has had the unintended consequence of providing a noticeable warming bias compared to older GCMs.

It is possible that some of the differences between AAWR found using our EM-GC and that inferred from CMIP5 GCMs are due to unaccounted internal variability in the observed temperature record. In particular, common explanations for the inability of models to match the lack of warming from 1998-2012 include shifts in the PDO and the strength of SAOD from minor volcanic eruptions (England et al., 2014; Tokarska et al., 2019), as well as variations in transport of heat to the deep ocean (Meehl et al, 2011) that we have attempted to simulate using AMV as a proxy for the strength of AMOC. Some members of the EM-GC ensemble produce results consistent with a climatically important role for the PDO: as mentioned in sections §2.1.1 and §3.1, simulations with high values of AER RF_{2011} show relatively stronger influence of the

PDO and relatively weaker influence of AMOC compared to the results shown in Figure 3. These ensemble members also result in an increase in the climatic importance of SAOD following major volcanic eruptions (Canty et al., 2013). However, consistent with the conclusion of Chylek et al. (2020), we find enhancements of SAOD due to recent minor eruptions to have a negligible effect on ΔT for all members of our EM-GC ensemble. The very low values of globally averaged SAOD following minor eruptions in the past decade will not noticeably affect ΔT, unless the climate response to SAOD is highly nonlinear.

3.4 The effects of aerosols and climate feedback on future ΔT

We turn our attention to the effects of the radiative forcing due to AER RF and λ_{Σ} on future projections of ΔT. Figure 6, in addition to showing the effect on AAWR of varying the input model time series for AER RF, also shows how these three simulations differ in ΔT out to year 2100 (ΔT_{2100}). The projection of ΔT is produced by applying the solutions for λ_{Σ} and κ that best fit the historical record to the prescribed GHG forcing pathways of RCP 4.5 out to the end of this century. Natural variations of all climatically important factors are assumed to be zero in the future to highlight only the rise in human driven rise in the global mean surface temperature anomaly (ΔT_{HUMAN}). We focus on projecting the underlying trend of future warming due to anthropogenic GHGs, as opposed to attempting to predict year-to-year variations in temperatures.

The full historical time series for ΔT can be fit reasonably well ($\chi^2 \leq 2$) for many combinations of time series of AER RF (indexed by their value in 2011) and value of λ_{Σ}. As a result, there exist a wide range of possible future temperature projections assuming the value for the sum of climate feedback mechanisms needed to simulate prior warming will persist into the future. This ability to fit the historical global temperature record with a wide range of possible climate feedback values is the main source of the resulting uncertainty in our estimate of ΔT_{2100}. If aerosol cooling to date has been low, then aerosols have counteracted only a small amount of the GHG forcing that warms the atmosphere, necessitating a low value of λ_{Σ}, resulting in modest future warming. By the same logic, a high amount of aerosol cooling to date leads to a large amount of future warming (Goodwin et al., 2018; Kiehl, 2007; Knutti & Hegerl, 2008).

Figure 6a shows a value for ΔT_{2100} of just 1.7°C for weak aerosol cooling (AER RF_{2011} of –0.4 W/m^2) offsetting the warming from RCP 4.5 GHGs. Conversely, if aerosol cooling has been large (AER RF_{2011} of –1.4 W/m^2), global warming will be much more intense: Figure 6c shows a ΔT_{2100} slightly above 3.1°C. Warming can reach over 4°C by 2100, still with RCP 4.5 GHGs, for the upper range of aerosol cooling scenarios stated in AR5 (i.e. AER RF_{2011} from –1.6 W/m^2 to –1.9 W/m^2). However, under these strong aerosol cooling scenarios, it is not possible in our modeling framework to obtain values of χ^2 below 2 for the full historical ΔT fit and particularly for the OHC fit, which is why the largest aerosol cooling case shown in Figure 6c is for AER RF_{2011} of –1.4 W/m^2. The individual simulations pictured in Figure 6 were chosen by finding the value of λ_Σ that minimizes χ^2 over the entire CRU4 record of ΔT (equation 1c), for each value of AER RF_{2011}. This difference in ΔT_{2100} results from the fact that aerosol concentrations, and thus aerosol forcings, are set to return to near-zero values in the future as an effect of air quality regulations that arise from human health concerns (Smith & Bond, 2014). As such, all simulations approach the same net human RF by 2100 but have different values of λ_Σ based on the amount of GHG RF that was offset by anthropogenic aerosols over the historical record.

Figure 4c shows ΔT_{2100} as a function of λ_Σ and AER RF_{2011} for RCP 4.5. Figure 5b shows a PDF of ΔT_{2100} for the ensemble members shown in Figure 4c, computed using the weighting method based on the AR5 likelihoods for the values of AER RF_{2011} (Figure 4a) described in section §3.1. Figures 4c and 5b illustrate a vitally important aspect of the climate system: the present uncertainty in the amount of GHG warming offset by aerosols causes a large spread in future warming, for a single future GHG abundance scenario, in this case RCP 4.5. If warming due to black carbon aerosols and co-emitted species were as large as the best-estimate of Bond et al. (2013), this term would place the actual value of AER RF_{2011} close to –0.4 W/m^2, resulting in values of ΔT_{2100} in our model framework close to the low end (i.e. 1.20 °C) of this forecast. On the other hand, if the climate cooling due to aerosols was as large as suggested by Shen et al. (2020), values of ΔT_{2100} would lie towards the high close on end (i.e. 3.74 °C) of our forecast. A reduction in the uncertainty of the amount of warming offset by tropospheric aerosols for the contemporary atmosphere, which requires obtaining consensus on the role of black carbon (Bond et al., 2013) as well as various aerosol indirect effects (Chen & Penner, 2005; Gryspeerdt et al., 2020), would enable more accurate forecasts of end of century warming.

A PDF of ΔT_{2100} for the output from 41 GCMs is also shown in Figure 5b. For GCMs that have submitted multiple runs using RCP 4.5 to the CMIP5 archive, ΔT in year 2100 is first averaged for these runs, such that the PDF consists of the distribution of ΔT_{2100} for the 41 GCMs shown in Table S4. The median warming of T_{2100} from our EM-GC simulations for RCP 4.5 is 2.00 °C, with lower and upper limits of 1.20 and 3.74 °C, respectively. The median ΔT_{2100} from the 41 GCMs is 2.52 °C, with lower and upper limits of 1.69 and 3.64 °C. Only 7 of the 41 CMIP5 GCMs exhibit a value for ΔT_{2100} less than the EM-GC median of 2 °C.

Figure 8 shows a probabilistic forecast of the future rise in ΔT from our EM-GC for RCP 4.5. Colors denote the probability of reaching at least that temperature by each year. The figure also contains the CMIP5 GCM ensemble minimum, multi-model mean, and maximum values of ΔT (gray lines) as well as the likely range of warming (green trapezoid) from figure 11.25b of AR5 (Kirtman et al., 2013). Temperature projections from our EM-GC agree well with the expert judgement of the near-future rise ΔT provided by Chapter 11 of AR5. The white color in Figure 8 for EM-GC probability is the median warming projection in our model framework. Similar to the comparison shown in section §3.3 for AAWR, projections of warming from the CMIP5 GCMs tend, on average, to be larger than the warming projection from our empirical model of global climate. Notably from a policy perspective, our most likely outcome for ΔT lies slightly above the CMIP5 GCM multi-model minimum, with both the EM-GC-based median and the GCM-based minimum lying below the Paris Climate upper limit of 2 °C. As explored further in sectjon §3.7, carbon emissions consistent with the CO_2 trajectory of RCP 4.5 provide a more likely chance of limiting global warming to either the Paris goal (1.5 °C) or upper limit (2 °C) than is projected by CMIP5 GCMs constrained by RCP 4.5. Most notably, observed ΔT over the years ~2005 to 2020 lies between the CMIP5 GCM multi-model minimum and mean, which of course was the driving factor behind the formulation of the green trapezoid in Figure 8 by the authors of Chapter 11 of AR5 (Kirtman et al., 2013).

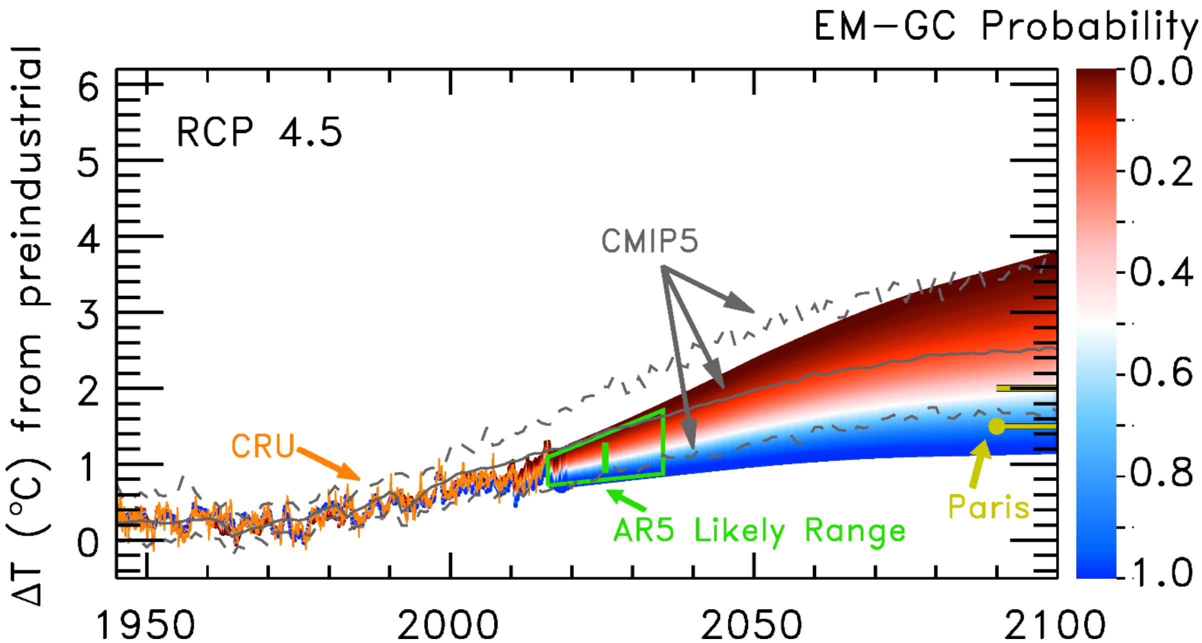

Figure 8. Global warming projections for RCP 4.5 relative to the preindustrial baseline. The EM-GC ensemble is shown with red-to-blue envelope and the CMIP5 GCM ensemble is shown with grey lines. Color at any given point within the EM-GC envelope represents the chance of ΔT reaching at least that temperature at that time. The three CMIP5 lines represent the minimum, multi-model mean, and maximum of ΔT from the GCMs that submitted projections of each RCP scenario respectively to the CMIP5 archive (grey lines). One set of observed temperatures to date (CRU4, orange line), the expert judgement from Figure 11.25 of AR5 (green trapezoid and vertical bar), and the targets of the Paris Agreement (gold spikes at right) are also shown for comparison.

The values of λ_Σ for the EM-GC-based projections shown above suggest less future warming than similar values provided by most CMIP5 GCMs. Table 9.5 of AR5 (Flato et al., 2013) suggests model mean values for λ_Σ of 1.6, 2.04, or 2.15 W/m^2/°C depending on which quantities are used to infer λ_Σ; specifically, the first value is the sum of the average value of the four individual feedbacks, the second value is based upon their estimate of ECS and the RF associated with 2×CO_2, and the third value is based upon the values for the climate sensitivity parameter and climate feedback parameter given. Table 1 of Sherwood et al. (Sherwood et al., 2020) gives a value for λ_Σ of 1.9 W/m^2/°C, with a 66% confidence range of 1.46 to 2.34

$W/m^2/°C$. Although tabulations of λ_Σ from CMIP5 models exist (Andrews et al., 2012; Forster et al., 2013; Vial et al., 2013), particularly Table 9.5 of AR5 (Flato et al., 2013) and Table 1 and Figure 4 of Sherwood et al. (Sherwood et al., 2020), comparison to our values is complicated by the sensitivity of λ_Σ to AER RF for good fits to the climate record (Figure 4). In general, our λ_Σ (and thus λ) suggest less future warming than those from the CMIP5 GCMs.

An important assumption for our quantification of both AAWR and ΔT_{2100} using the EM-GC is that λ_Σ (and thus λ) has remained constant over time. However, we can also simulate time dependent climate feedback to address the possibility that λ_Σ may change over time (section 12.5.3 of AR5 and references therein; also Rose et al. (2014), Shindell (2014), and Marvel et al. (2018)). Recall that $1/\lambda = (1+\gamma)/\lambda_p = 1/(\lambda_p - \lambda_\Sigma)$ from equations 5 and 6; we have used λ_Σ up to this point as simulations with higher values of λ_Σ have higher future ΔT. For this reason, we prefer to examine a time dependent λ in terms of its inverse, $1/\lambda$, also known as the climate sensitivity parameter, because this quantity also has a positive correlation with ΔT. Figure S6 and Table S2 summarize how ΔT_{2100} changes if we allow $1/\lambda$ to vary over time while still keeping the strength of fit between ΔT_{OBS} and ΔT_{MDL} at acceptable levels ($\chi^2 \leq 2$) for either the full historical time period or the most recent 80 years.

For four different cases of aerosol forcing, we find that allowing $1/\lambda$ to scale with anthropogenic forcing while still keeping $\chi^2 \leq 2$ over the full historical time period results in roughly doubling of ΔT_{2100} compared to the constant feedback case (Table S2). This scenario with time-varying feedback, which results in our maximum warming, implies an increase in $1/\lambda$ by nearly a factor of three over two and a half centuries (Figure S6e). This rise in $1/\lambda$ is much more rapid than expected. Even one of the most extreme estimates from recent work (Marvel et al., 2018) suggests an increase in median estimated equilibrium climate sensitivity (ECS) from 1.8°C (for simulations constrained to match data acquired over 1979-2005) to a long-term (end of century) value of 3.1°C, which corresponds to a 72% increase. This rise in ECS postulated by Marvel et al. (2018) is predicated on the assumption that current atmospheric and oceanic conditions are truly exceptional. The validity of preliminary results for a handful of CMIP6 models suggesting even higher ECS (Belcher et al., 2019; Gettelman et al., 2019; Zelinka et al., 2020) has been questioned by numerous recent papers based upon analysis of paleoclimate data as well as climatic conditions over the past several decades (Forster et al., 2020; Nijsse et al., 2020; Sherwood et al., 2020; Voosen, 2019; Zhu et al., 2020). Such scenarios that greatly increase $1/\lambda$ by the end of the century also produces a time dependent drift in the residual between observed and modeled over the historical record in our model framework (Figure S6f). Coincidentally, the comparison between modeled and measured ΔT in Figure S6f looks similar to the comparison of the CMIP5 GCM multi-model mean and ΔT_{OBS} shown in Figure 2. This time dependent drift between ΔT_{OBS} and our modeled ΔT, combined with the large temporal change in λ that underlies this simulation, suggests this might be an unreasonable scenario for use in CO_2 emission mitigation strategies.

We also calculate a medium-varying feedback case by considering a $\chi^2 \leq 2$ strength-of-fit restriction that focuses only on the most recent 80 years of the ΔT_{OBS} record (Figure S6c,d) instead of $\chi^2 \leq 2$ over the full ΔT_{OBS} time series. This scenario results in a simulation of ΔT that appears more reasonable upon inspection of the residuals and the smaller rise in $1/\lambda$. However, depending on the strength of AER RF, the increase in $1/\lambda$ can still range from roughly 50% to more than a factor of two (Table S2). Changes in $1/\lambda$ of this magnitude over two and a half centuries are faster than the millennia-order timescale changes usually referenced (e.g. section 12.5.3 of AR5 and references therein) when discussing noticeable changes in λ, ECS, and other related quantities such as transient climate sensitivity (TCS). While a factor of two or more rise in $1/\lambda$ does not match literature, an increase of roughly 50% falls in line with a 50% increase in TCS (Shindell, 2014) and neatly between the 28% increase (1.8 °C to 2.3 °C) and the 72% increase (1.8 °C to 3.1 °C) (Marvel et al., 2018) seen in other re-analyses of historical forcing results from GCMs. As there is no strong evidence from the climate record for a noticeable rise in $1/\lambda$ on the multidecadal time scale consistent with the simulations shown in Fig S6, we assert that the assumption of constant feedback within the EM-GC framework seems to be a reasonable assumption for the next few decades. There also certainly exists the possibility that by end of century, the rise in ΔT could be a few tenths of a degree warmer than our current best estimates assuming constant λ_{Σ} due to a slow rise in $1/\lambda$.

We also assume for our computation of Q_{OCEAN} that κ is constant over time. This assumption follows from the fact that, like λ, the rate of change of κ is most likely small enough to not have a significant effect on the time scale of our calculations of ΔT (Raper et al., 2002). Our application of κ requires a monotonic increase in the magnitude of this term, we solve for Q_{OCEAN} based on $\Delta T_{ATM,HUMAN}$ instead of total ΔT_{MDL} because the latter displays strong natural variability and thus is not monotonically increasing. The anthropogenically-forced temperature itself is not strictly monotonic either, especially for AER RF time series corresponding to more negative values of AER RF_{2011}, but the short, small, instances of cooling in those scenarios are relatively insignificant. Also, those cooling instances are pre-1950, and the OHC record we fit does not extend earlier than 1950. As such, the few instances when ΔT_{HUMAN} includes short, small cooling periods should not affect the overall approximation of κ as a constant. As with λ and fitting ΔT_{OBS}, a constant value of κ results in modeled OHC that fits the observed OHC quite well (Figure 3b and Figure 6).

Model treatment of aerosols and clouds are two possible explanations for why λ from the EM-GC differs from CMIP5 models. About half of the CMIP5 GCMs do not include aerosol indirect effects (Schmidt et al., 2014). A lack of the indirect effect in our EM-GC would result in cooler projections of future ΔT, as less total AER RF over the historical record would favor lower values of λ_{Σ} and thus lower ΔT_{2100}. Such a relationship between ΔT_{2100} and the presence of the indirect effect does seem to appear in CMIP5 results as well: models without the indirect aerosol effect consistently warm less from 2014 to 2100 than models that do include it (Chylek et al., 2016). Considering that the CMIP5 GCMs tend to warm more than the EM-GC in terms of both AAWR (§3.3) and ΔT_{2100} (§3.5), a lack of the indirect effect in some GCMs does not explain the excess warming in GCMs, meaning it seems more likely that the difference between the EM-GC and CMIP5 GCMs lies in cloud feedback. (This should not eliminate considerations of aerosols and their complex interactions, however, especially given that aerosol indirect effects and cloud feedback processes are related.) It is widely known that uncertainty in the cloud feedback is much larger than that of other major feedbacks and this uncertainty is a main driver for the spread between CMIP5 models (Dolinar et al., 2015; Stocker et al., 2013; Zhou et al., 2015). While the fourth IPCC report (Solomon, 2007) suggested a cloud response spread centered around zero feedback within the CMIP3 GCMs, AR5 suggests a largely positive cloud feedback, in line with some recent observations (Dolinar et al., 2015; Zhou et al., 2015). However, there is considerable spread in the determination of cloud feedback from observations, including the possibility of a neutral or even negative feedback (Ceppi et al., 2017; Vial et al., 2013; Zelinka et al., 2016). While some recent studies suggest that cloud feedback and overall ECS interpreted from observation are higher than those from modeling studies based solely on observations since 2000 (Dessler, 2013; Sherwood et al., 2020), other observational studies offer lower ECS values than those found in modeling studies (Lewis & Curry, 2018; Masters, 2014;

Otto et al., 2013; Schwartz, 2012). If the actual cloud feedback is less positive than the models currently suggest, that could also be a factor in the high bias of GCMs for AAWR, ECS, and ΔT_{2100} (Hope et al., 2017; Tokarska et al., 2020; Weaver et al., 2020; Zelinka et al., 2020).

3.5 Other RCPs and comparisons to projections from GCMs

One advantage of simple models such as the EM-GC is the ability to perform sensitivity testing by completing many more runs of the model with less computing power. For example, each ensemble represented in Figure 7, originally consisting of 160,000 simulations, takes roughly two hours to complete. All those ensembles focus on ΔT_{HUMAN} driven by RF from GHG abundances from RCP 4.5, and a full treatment of the effects of uncertainty in RF due to aerosols. **Figures 9a and 10a**, driven by RCP 2.6, and **Figures 9b and 10b**, driven by RCP 8.5, show the results of EM-GC ensembles constrained by low and high ends of RF tested in CMIP5, respectively. The panels of Figure 9 and Figure 10 are the same as Figure 4b and Figure 5b, respectively, except for the different RCP scenario driving the ensembles. As the RCP scenarios are identical up to 2005 and do not differ greatly until after 2020, the shape of the model output shown in Figure 4b and Figure 9 are nearly identical, because the three χ^2 calculations (fits to total ΔT, recent ΔT, and OHC) used to select good fits to the climate record consider only historical data. The difference in end-of-century RF drives the differences in ΔT_{2100}, shown both in the colors of the model output and the positions of the PDFs in Figure 5b, Figure 10a, and Figure 10b. The probabilities of the rise in ΔT_{2100} staying beneath 2°C are 92%, 50%, and 0% for RCP 2.6, 4.5, and 8.5 respectively; probabilities for ΔT_{2100} remaining below 1.5°C fall to 67%, 10%, and 0%. For RCP 6.0, (not pictured,) the probability of staying beneath 2°C is 20%, which falls to 0.1% for 1.5°C.

Figure 10c and 10d also compare ΔT_{2100} from the EM-GC to temperatures at the end of the century presented in Sherwood et al. (2020). The probabilistic estimate of end of century warming given in Sherwood et al. (2020) is based upon their expert evaluation of climate sensitivity combined with the assumption of a linear relation between the transient climate response and radiative forcing. They estimate that by end of century warming will be less than 2°C relative to pre-industrial are 83%, 17%, and 0% for RCP 2.6, 4.5, and 8.5, respectively. The probabilistic estimate of the upper end of warming given by Sherwood et al. (2020) is considerably less than indicated by the CMIP5 GCMs (green versus red lines in Fig. 10). They also compute a lower probability for the low end of the distribution of ΔT_{2100} than we find using our EM-GC, which is traceable to their judgement that the most likely value of total cloud feedback is positive (Klein et al., 2017).

There are three clear takeaways from Figures 5b and 10. First, earth's climate allows for a wide range of future temperatures, even when a model such as our EM-GC is sufficiently trained with historical data due to uncertainty in quantities such as AER RF_{2011}. Second, consistent with the expert assessment of temperature projections from CMIP5 GCMs given in Chapter 11 of AR5 (Kirtman et al., 2013), our EM-GC projects smaller future increases in ΔT than provided by most of the CMIP5 GCMs. Third, while temperature projections from our EM-GC agree with CMIP5 GCM results in that society must avoid a GHG pathway consistent with RCP 8.5 to achieve the goals of the Paris Climate Agreement, our model simulations show that RCP 4.5 and particularly RCP 2.6 are GHG pathways more likely to achieve limited warming than indicated by GCM results within the CMIP5 archive. Our model projections suggest that adhering to the RCP 4.5 pathway is as likely as not (=50%) to give Earth a future that limits global warming to 2°C above preindustrial; placing GHGs on the RCP 2.6 pathway is highly likely (>90%) to limit global warming to 2°C and likely (>66.7%) to stay beneath 1.5°C.

Our probabilistic temperature projections disagree somewhat with AR5. According to CMIP5 GCM results as presented in AR5, RCP 2.6 is likely but not highly likely (that is, >66.7% but not >90%) to keep global temperatures beneath 2°C. AR5 also says RCP 4.5 is more likely than not (>50%) to *exceed* 2°C. The power of RCP 2.6 to keep us beneath 2°C of warming appeared in another recent study (Goodwin et al. 2018) that determined staying beneath 2°C to be highly likely, much closer to our result than to those of CMIP5.

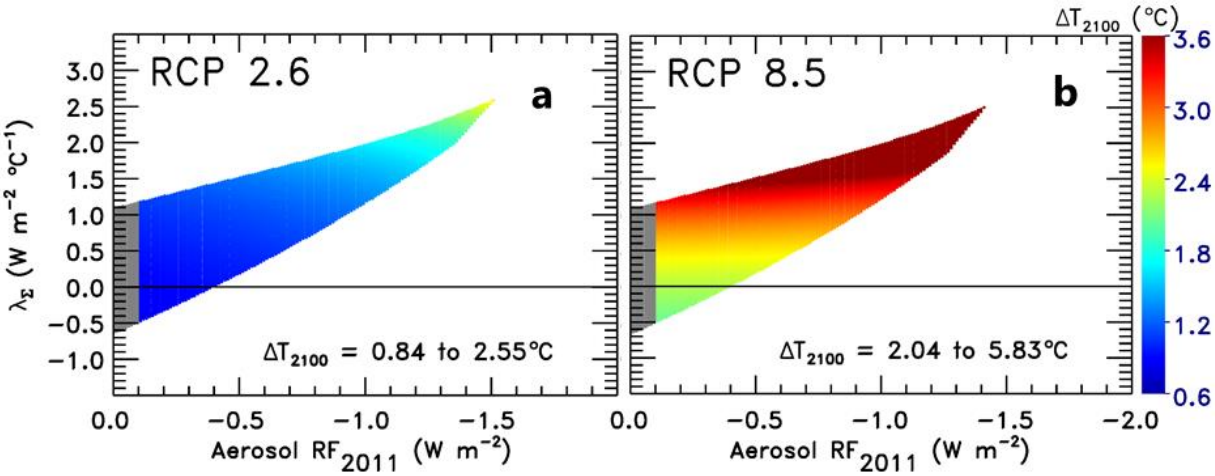

Figure 9. Same as Figure 4a except for ensembles using RCP 2.6 and RCP 8.5 anthropogenic forcing instead of RCP 4.5.

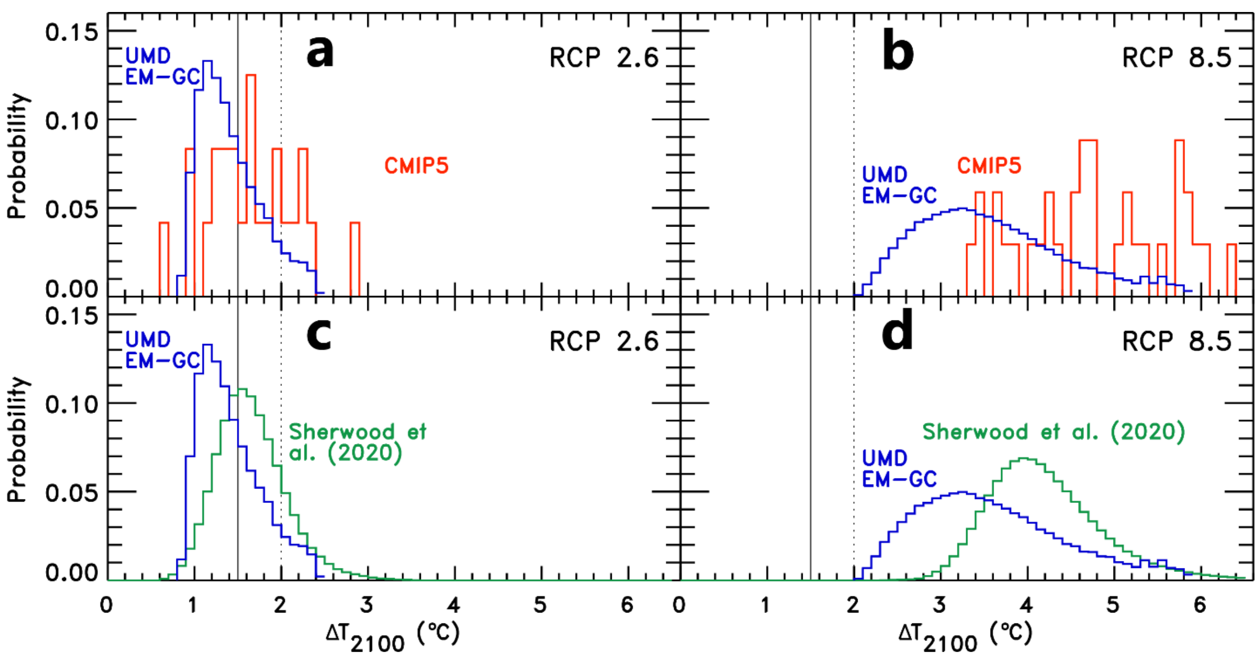

Figure 10. Panels a and b are the same as Figure 5a except for ensembles using RCP 2.6 and RCP 8.5 anthropogenic forcing instead of RCP 4.5. Panels c and d then exchange CMIP5 data for data taken from figure 23 of Sherwood et al. (2020), binned to match the structure presented in the previous PDFs of this study.

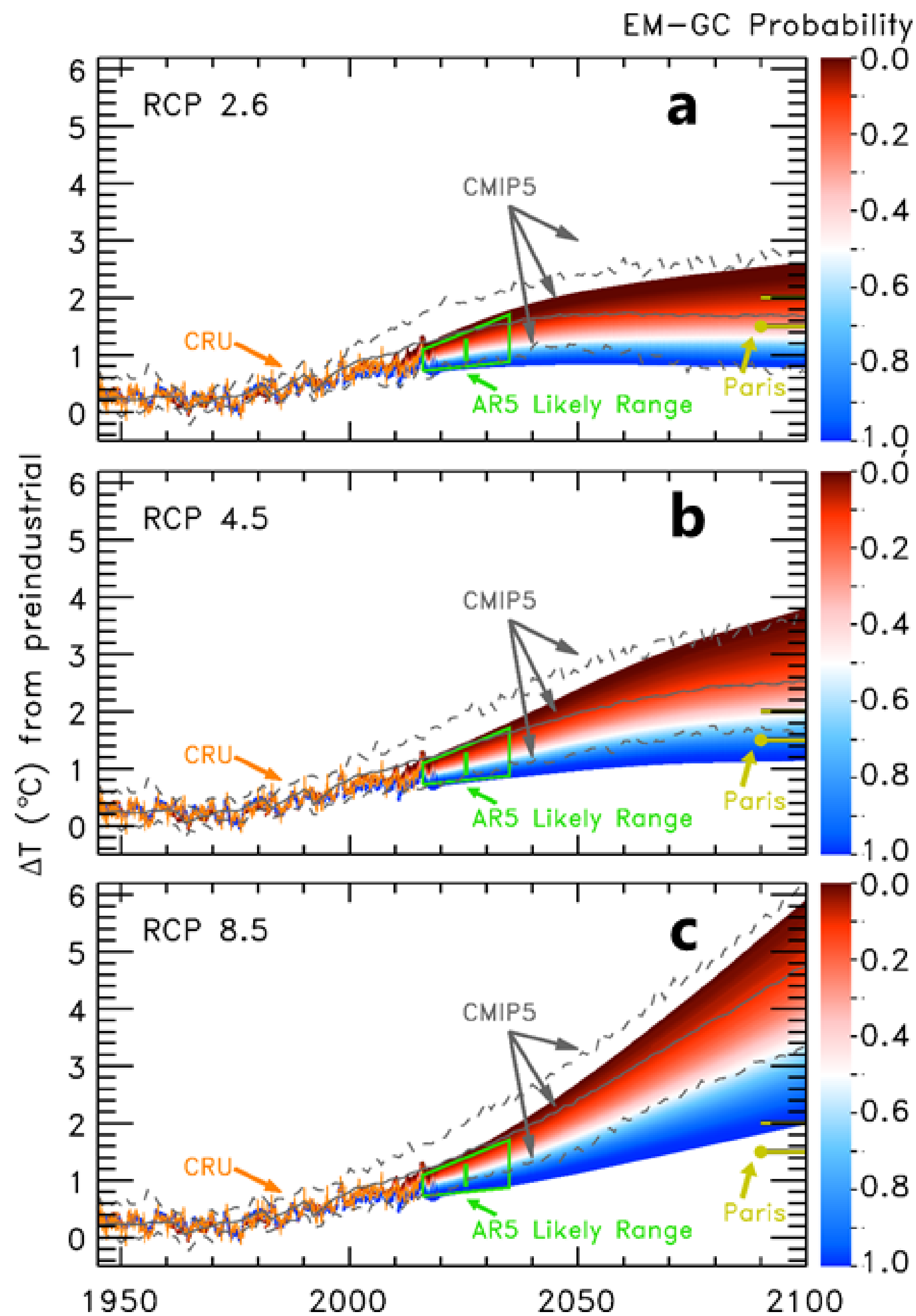

Figure 11. Global warming projections for RCP 2.6, RCP 4.5, and RCP 8.5 relative to the preindustrial baseline. Figure 11b is the same as Figure 8; all three panels are of similar construction. Color at any given point within the EM-GC envelope represents the chance of ΔT reaching at least that temperature at that time.

Figure 11 shows probabilistic projections of global warming for RCP 2.6, 4.5, and 8.5. This figure is the same as Figure 8 (that showed results for RCP 4.5), using the same vertical axis for all three ensembles. This figure demonstrates a fourth key takeaway from our modeled projections of future temperature. Projections of global warming computed using our EM-GC agree well with the expert judgement of near-future ΔT from AR5 of Chapter 11 (Kirtman et al., 2013), shown as a trapezoid on each panel of Figure 11. The colored envelope for each panel of Figure 11 is based upon a representative sample of the runs from each respective RCP ensemble (Figure 4b and Figure 9) and displays the rise in ΔT out to the end of the century. At each time along this envelope, the color represents the probability within the ensemble of reaching at least that temperature. Whatever RCP scenario we examine, the EM-GC results match the near-future projections (trapezoid) based on the expert judgement of AR5's Chapter 11 authors. Our projections of warming using a physically based model tied to observations of ocean heat content, natural as well as anthropogenic drivers of variations in ΔT, and the consideration of uncertainty in AER RF are thus remarkably similar to the expert assessment of the CMIP5-GCM-based future rise in ΔT sketched out in figure 11.25b of AR5 (Kirtman et al., 2013).

3.6 The effect of increased future emissions of CH_4

As many countries around the world transition from coal to natural gas as a primary fossil fuel for electricity production, assumptions about future methane scenarios must shift to account for faster growth (Molnár, 2018; Saunois et al., 2020). The RCP 4.5 scenario has future CH_4 leveling off by midcentury at near-current atmospheric mixing ratios, then decreasing until the end of the 21st century (Figure 1b). The RCP 2.6 scenario has even more drastic and immediate reductions in atmospheric CH_4. This leveling off may be difficult for society to achieve due to natural gas becoming a primary source of energy for the foreseeable future (Jackson et al., 2018). This energy shift would imply more leakage of CH_4 from utility infrastructure as demand increases (Saunois et al., 2020). The atmospheric mixing ratio might also rise through anthropogenically-induced releases of natural CH_4 reservoirs, such as permafrost melting, or increased biogenic activity (Comyn-Platt et al., 2018; Voigt et al., 2017). Finally, the observed abundance of atmospheric CH_4 has already been rising faster in the past few years than in the previous two decades (Nisbet et al., 2019; Saunois et al., 2020), exceeding the RCP 4.5

projection and mapping closer to the most intense RCP 8.5 scenario. These factors taken together suggest that we should expect that the abundance of CH_4 in the atmosphere may increase over time, and not level off as suggested in the RCP 4.5 scenario (Saunois et al., 2016). Therefore, we have created blended CH_4 scenarios, noted in section §2.1.2 and shown in Figure S2, to test the sensitivity of warming computed using our EM-GC to various future for atmospheric CH_4.

Figure 12 shows the probability of achieving the Paris Climate Agreement goals as a function of the atmospheric abundance of CH_4 in 2100. Each symbol in Figure 12 shows ΔT_{2100} for an ensemble of EM-GC runs where the only change between the ensembles is the input time series of CH_4. Starting from either the RCP 2.6 (squares) and RCP 4.5 (diamonds) as scenario for all GHGs other than CH_4, the time series for each ensemble calculation is based upon either RCP 2.6, RCP 4.5, RCP 8.5, or one of four linear combinations of CH4 versus time between RCP4.5 and RCP 8.5 shown in Figure S2. Otherwise the ensembles are identical to those in Figs. 10a and 10b for RCP 2.6 and RCP 4.5 respectively. In RCP 4.5, the abundance of CH_4 in 2100 is 1578 ppb; in the ensemble driven by RCP 4.5 with no changes to CH_4, we compute a 50% probability of ΔT_{2100} remaining beneath 2°C and a 10% chance of remaining beneath 1.5°C. These probabilities correspond to the leftmost diamonds in each panel of Figure 12. As the CH_4 time series approaches the RCP 8.5 pathway, which has a CH_4 mixing ratio of 3748 ppb in 2100, the probabilities of future warming remaining beneath 2°C and 1.5°C fall to 30% and 2%, respectively (rightmost diamonds). Similarly, switching from RCP 2.6 for all GHGs to a combined scenario that uses CH_4 from RCP 8.5 causes the probability of ΔT_{2100} staying below 2°C to decrease from 92% to 73% and the likelihood of staying below 1.5°C warming to decrease from 66% to 33% (left- and rightmost squares). This analysis indicates that failure to limit methane to the RCP 2.6 trajectory will have a large impact on the achievement of the 1.5°C goal of the Paris Climate Agreement.

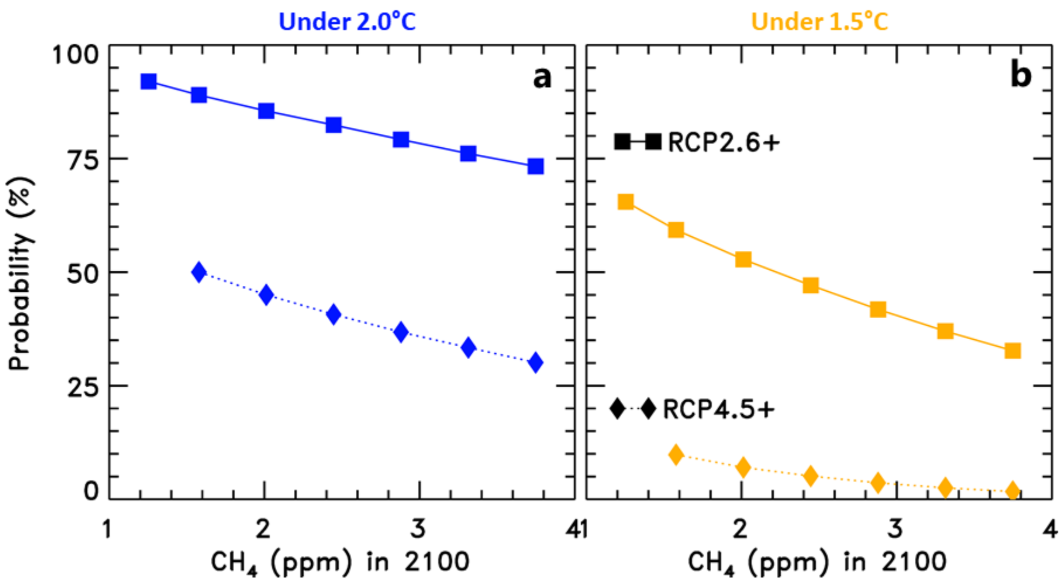

Figure 12. Impact of CH_4 on EM-GC projections. Shown are the probabilities that ΔT_{2100} remains below 2°C (11a, blue) or below 1.5°C (11b, gold) for both the RCP 2.6-based ensembles (squares, solid lines) and RCP 4.5-based ensembles (diamonds, dashed lines) relative to preindustrial. Each ensemble based on RCP 4.5 uses all GHG and aerosol forcing inputs from RCP 4.5 except replacing the RCP 4.5 CH_4 time series with one of six linear combinations between the RCP 4.5 CH_4 scenario and the RCP 8.5 CH_4 scenario, inclusive. Likewise, each ensemble based on RCP 2.6 uses all forcing inputs from RCP 2.6 except substituting the RCP 2.6 CH_4 with one of the six linear combinations (save for the seventh ensemble, far left, which is purely RCP 2.6). Ensembles are placed in this figure based on the CH_4 mixing ratio in 2100 (i.e. end values of Figure S2).

3.7 Response to cumulative emissions

The Transient Climate Response to cumulative carbon Emissions (TCRE), which relates ΔT to accumulated anthropogenic emissions of CO_2, is an important policy metric (Gregory et al., 2009; Millar et al., 2017). As the EM-GC projects generally less future warming than most of the CMIP5 GCMs, it follows that TCRE derived from the EM-GC falls on the low side of the range for TCRE given in AR5 (0.8 to 2.5°C per 1,000 GtC). **Figure 13a** shows the results of the ensemble median EM-GC simulations for the four RCP pathways. All four EM-GC ensemble medians suggest a TCRE of roughly 1.4°C per 1,000 GtC. The other four lines on Figure 13a show the multi-model mean projections from the CMIP5 GCMs for the four RCP scenarios, taken from figure SPM.10 of AR5. TCRE from the CMIP5 GCMs in this figure has a value of 2.3°C per 1,000 GtC, which lies well above the EM-GC estimate and in the high end of the assessed range given by AR5. Following AR5 and Millar et al. (2017), future cumulative emissions of CO_2 in figure 13a are based on the rise since 1870, with ΔT shown relative to the two-decade average for 1861-1880. The emissions along the horizontal axis represent global, atmospheric release of CO_2 due to combustion of fossil fuels, flaring, cement production, and LUC from the RCP database.

Figure 13b shows an adjustment of the RCP 8.5 lines from Figure 13a that are set to zero for the most recent decade (2010 to 2019), as done in Millar et al. (2017). This adjustment clarifies the allowable remaining carbon budget for limiting future warming to remain beneath a given amount. Millar et al. (2017) presented this adjustment as one manner of accounting for the overestimate of ΔT_{MDL} provided by CMIP5 GCMs compared to ΔT_{OBS}; setting the ΔT baseline to recent years is also the process behind the expert judgement of Kirtman et al. (2013) that produced the AR5-based trapezoid shown in previous figures. The central finding of Millar et al. (2017) is that the best estimate of the remaining carbon budget needed to limit future warming to 0.6°C relative to 2015 (which translates to 1.5°C relative to preindustrial) is higher than the best estimate suggested by AR5. They later issued a public clarification saying their estimate for future temperature (and thus their carbon budget) still lies within the AR5 uncertainty (Allen & Millar, 2017), albeit with their carbon budget on the high extreme of the AR5 range. As shown in Figure 13 and further discussed below, our EM-GC projection of ΔT_{MDL} indicates the remaining carbon budget is even larger than the values given by AR5 and Millar et al. (2017).

We can also use full ensemble simulations within the EM-GC to compute probabilistic forecasts of emissions thresholds for the Paris Agreement targets. **Figure 14** displays ΔT with respect to cumulative emissions of CO_2, using the same color scheme adopted for Figures 8 and 11. The colors represent the probability that a particular future value of ΔT will reach at least that temperature for the specified cumulative emission of CO_2. Figure 14 is based on the RCP 8.5 scenario for GHGs, to cover the widest range of future anthropogenic emissions of CO_2.

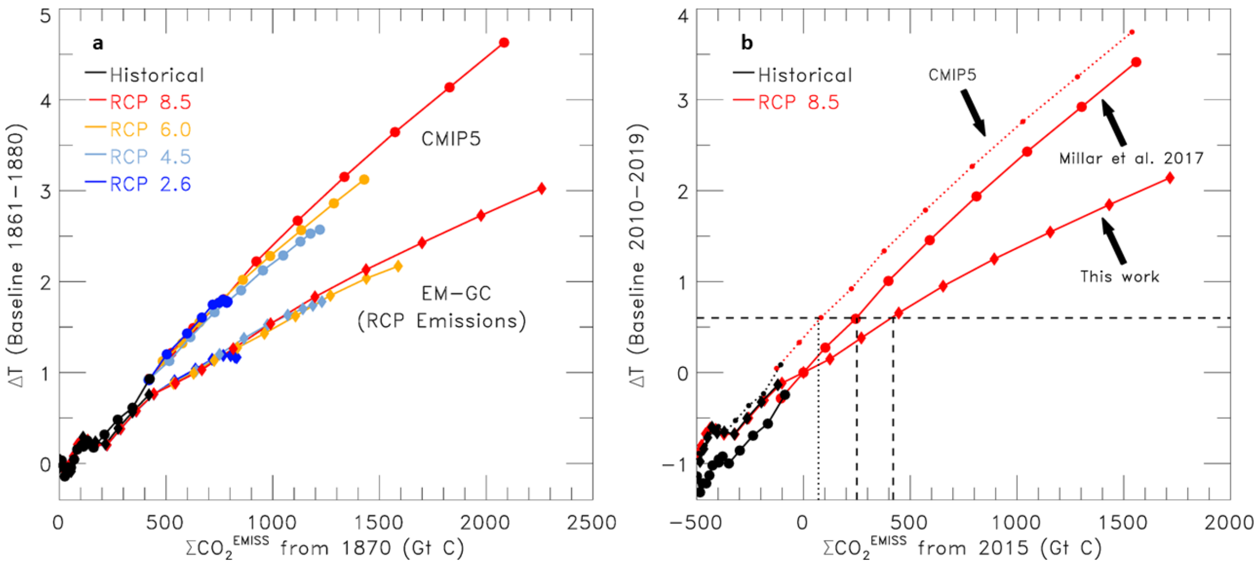

Figure 13. (**a**) Transient climate response to cumulative CO_2 emissions, in units of GtC. Average ΔT from CMIP5 GCMs, as taken from figure SPM.10 of AR5, is plotted against the average cumulative emissions since 1870 modeled to meet RCP prescribed concentrations (circles). EM-GC results show ΔT from a single EM-GC simulation for each RCP scenario representing the median of the ensemble; CO_2 emissions for each RCP ensemble are taken directly from the RCP database. (**b**) TCRE for different studies of RCP 8.5, illustrated using the same axes as Millar et al. (2017). Both the EM-GC projection and the Millar et al. (2017) projection from CMIP5 are plotted such that the point representing the decade of the 2010s (centered on 2015) is set to the origin, based on current ΔT and estimated cumulative emission to date. For comparison, the CMIP5 projection is also shown such that it matches the EM-GC results for roughly the first century of the simulation, as done in panel a, i.e. instead of matching in the 2010s, so as to demonstrate the effective shift that Millar et al. (2017) applied to the CMIP5 projection. Vertical dotted and dashed lines indicate the remaining amount of CO_2 that can be released prior to having ΔT rise 1.5°C above preindustrial (i.e., 0.6°C above the observed 0.9°C rise in ΔT for the 2010s) according to the three studies.

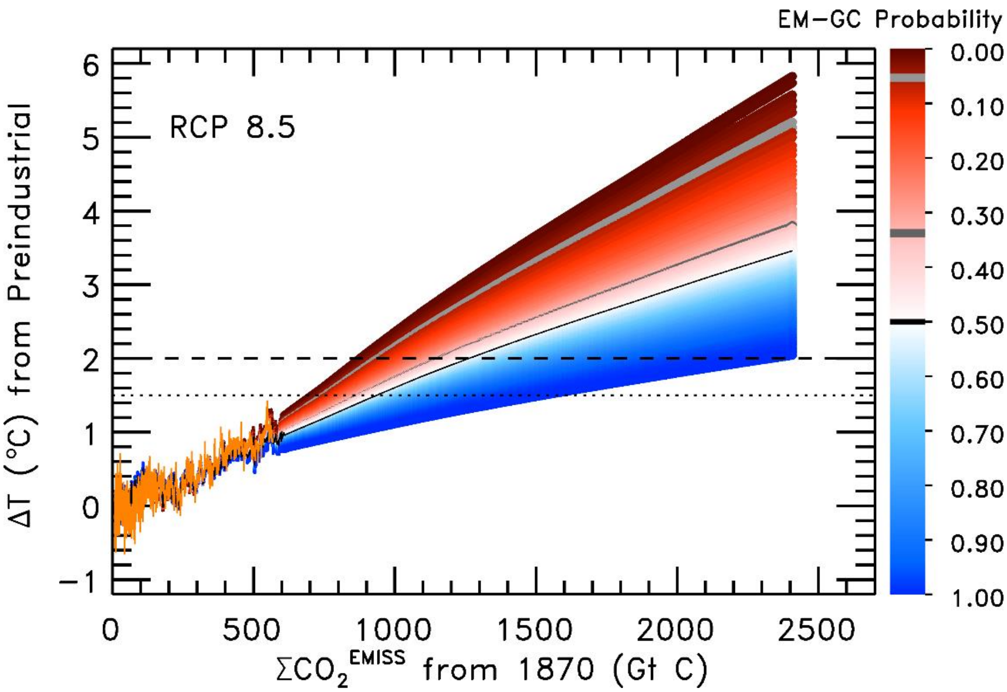

Figure 14. Full EM-GC projection of transient climate response to cumulative CO_2 emissions for RCP 8.5. Colors and structure mirror those of figure 11c. The three grey lines of varying darkness provide a visual guide to highlight the lines of 50%, 66%, and 95% probability for keeping temperatures cooler than those temperatures, while the black horizontal lines are a visual guide to the two Paris Climate Agreement target ΔT values. As such, the intersection of a grey line with a black line determines the maximum amount of cumulative emissions that would be allowed while remaining cooler than one of the Paris Climate Agreement targets with the given probability.

Figure 14 shows that according to calculations conducted in our model framework, if society can keep cumulative carbon emissions below 900 GtC, then we will have a 95% chance of preventing global warming from exceeding 2°C relative to preindustrial. Cumulative emissions of 1140 GtC or 1250 GtC would lower the chance of ΔT remaining below 2°C to 66% or 50%, respectively. Likewise, the 95%, 66%, or 50% probabilities for ΔT remaining below 1.5°C correspond to cumulative CO_2 emissions of 710 GtC, 850 GtC, or 930 GtC, respectively. For comparison, these emission numbers are listed in **Table 1**, alongside comparable numbers from other studies (Goodwin et al., 2018; Millar et al., 2017; Tokarska & Gillett, 2018). Similar quantities from Chapter 2 of the IPCC 1.5 C Special Report are also included, and show lower allowable carbon emission budgets driven by two intermediate-complexity climate models, FAIR and MAGICC, that are constrained to approximate climate sensitivity from the CMIP5 GCMs (Rogelj et al., 2018).

For reference, while Millar et al. (2017) suggest that human activity has emitted roughly 545 GtC from 1870 to 2014, leading to a rise in ΔT of 0.9 °C for that time (Millar et al., 2017). Improvements in the understanding of LUC emission and five more years of emitting CO_2 suggest these values increase to 635 GtC (Friedlingstein et al., 2019; Le Quéré et al., 2018) and 1.0 °C respectively for 1870 to 2019. Using only emissions prescribed by the RCP scenarios, the carbon budget values are 532 GtC to 2014 and 590 GtC to 2019, but similar to the 545 GtC value from Millar et al. (2017), these numbers do not reflect updated understanding of past emissions from LUC. Society has been emitting 11.0 ± 0.8 GtC per year over the past decade, a rate that has also been increasing over previous decades, with a current annual value of nearly 12 GtC per year (Friedlingstein et al., 2019).

Table 2 shows the remaining, post end of 2019 carbon budget for limiting warming to either 1.5°C or 2.0°C in our model framework. For all entries in Table 2 we assume at the end of 2019 that cumulative carbon emissions are 590 GtC. We tie this cumulative carbon emission estimate to the time series used to drive the global carbon cycle model that underlies the RCP 8.5 specification of atmosphere CO_2. The current annual rate of global carbon emission of nearly 12 GtC would imply surpassing 710 GtC, our RCP-based threshold for staying below 1.5 °C with 95% confidence, in 2029 and surpassing the 2.0 °C, 95% confidence threshold of 900 GtC in 2045. These dates of passing the 95% confidence intervals for the target and upper limit of the Paris Agreement fall to 2028 and 2040, respectively, if we allow total carbon emissions to rise at the rate that underlies RCP 8.5. Similarly, year ranges for passing the 66% confidence interval for 1.5°C and 2.0°C warming are 2037-2041 and 2052-2065, respectively (Table 2). The 50% confidence interval estimates for these thresholds rise to 2042-2048 and 2057-2074. We emphasize that even though our EM-GC exhibits slower warning than the CMIP5 GCMs, the Paris Climate Agreement target of limiting warming to 1.5°C cannot be achieved unless society begins a near immediate transition to a low carbon future, and at the same time slows or eliminates the rise in atmospheric CH_4 (section §3.6).

Table 1

Carbon Budget Comparisons

Source & reference emissions	Threshold Budget for 1.5°C			Threshold Budget for 2.0°C		
	50% prob.	65% prob.	90/95% prob.	50% prob.	65% prob.	90/95% prob.
Millar et al. (2017) - 545 GtC, 2014	768 GtC	749 GtC	709 GtC (90%)	961 GtC	940 GtC	875 GtC (90%)
Tokarsak & Gillet (2018) - 555 GtC, 2015	763 GtC	685 GtC	*	*	*	*
Rogelj et al. (2018) – 605 GtC 2017	158 GtC	114 GtC	*	*	*	*
Goodwin et al. (2018) - 572 GtC, 2016	*	767-777 GtC	*	*	967-1027 GtC	*
This study - 532 GtC, 2014 ; 590 GtC, 2019	930 GtC	850 GtC	710 GtC (95%)	1250 GtC	1140 GtC	900 GtC (95%)

Table 2

Metrics for Future Carbon Budgets Leading to Crossing the Pairs Thresholds

Warming Targets		Future Budget to Threshold Budget [a]	Future Budget as % of Past Budget [a]	Range of Years to Meet Threshold Budget [b]
1.5 °C	95%	120 GtC	20%	2028-2029
	66%	260 GtC	44%	2037-2041
	50%	340 GtC	58%	2042-2048
2.0 °C	95%	310 GtC	53%	2040-2045
	66%	550 GtC	93%	2052-2065
	50%	660 GtC	112%	2057-2074

[a]Future Budget = Threshold Budget – Past Budget, with threshold and past budgets taken from the final row of Table 1 (2019 reference year)

[b]First year of range is taken from RCP 8.5 prescribed emissions; last year of range assumes the current rate of emissions of roughly 12GtC continues into the future

4. Conclusions

The value of the anthropogenic contribution to global warming over the past three three decades, termed attributable anthropogenic warming rate (AAWR), has been analyzed in detail using both an empirical model of global climate (EM-GC) and output from CMIP5 GCMs. We find AAWR to be as 0.14 ± 0.06 °C/decade (full range of possible values) for 1979 to 2010 using our EM-GC, where the uncertainty covers the full range of model runs that yield a good fit to ΔT_{OBS} from CRU4. The CMIP5 GCMs exhibit values of AAWR of 0.22 ± 0.10 °C/decade (standard deviation among AAWR values from the CMIP5 GCMs), considerably larger than inferred from the climate record using our EM-GC. More than two-thirds of the 112 archived CMIP5 GCM runs exhibit a value for AAWR larger than our upper limit of 0.20 °C/decade. The uncertainty in the EM-GC based derivation of AAWR is driven by imprecise knowledge of the radiative forcing of climate due to tropospheric aerosols, whereas the largest source of spread in the GCM simulation of GMST is due to uncertainty in cloud feedback (Ceppi et al., 2017; Vial et al., 2013; Zelinka et al., 2016). Our finding that the CMIP5 GCMS exhibit a considerably faster rise in GMST than observed is consistent with finding of Chapter 11 of AR5 (Kirtman et al., 2013). Attempts to improve the understanding of aerosol species is an active area of current research, but large uncertainties persist (Bond et al., 2013; Collins et al., 2017; Pincus et al., 2016; Shen et al., 2020; Smith et al., 2011; Smith & Bond, 2014; Thornhill et al., 2020). Similarly, while some recent studies suggest total cloud feedback is positive (Klein et al., 2017; Sherwood et al., 2020), a recent analysis of a 40 year satellite record shows no trend in cloud reflectivity (Weaver et al., 2020), which is thought to be the largest driver of this positive trend.

While it is beyond the scope of this study to thoroughly assess the possible shortcomings of the CMIP5 GCM simulations over the past three decades, we suggest that high values for the sum of climate feedback mechanisms (λ_{Σ}), in particular the various cloud feedback processes, could be responsible for the apparent warm bias of the CMIP5 GCMs. Indeed, a recent analysis of CMIP6 shows that the next generation of GCMs displays high correlation between high ECS (i.e. high λ_{Σ} in the EM-GC) and a poor fit to observed AAWR from 1981 to 2014/17 (Tokarska et al., 2020), and these high-ECS models have strongly positive cloud feedbacks (Zelinka et al., 2020). If the actual cloud feedback is less positive than currently exhibited by CMIP5 and CMIP6 GCMs, this could explain the apparent warm bias of these models.

Accurate projections of ΔT are critical to the successful implementation of the Paris Agreement. However, the wide span of possible futures even in our EM-GC framework confounds policy-making efforts and confidence in achieving the desired warming limits. While our model projections show a wide range of possible warming by end of century for the same GHG scenario, our forecasts produce a more optimistic likelihood for achieving the goals of the Paris Agreement than is provided by the CMIP5 GCMs. The temperature forecasts given by our EM-GC tend to lie among the lower half of the projections provided by the CMIP5 GCMs, with our maximum forecast warming tending to lie near the CMIP5 multi-model mean. Most importantly, the projections of ΔT from our EM-GC agree extremely well with the "indicative likely range for annual mean ΔT" from Chapter 11 of AR5 (Kirtman et al., 2013) – lending important computational support for this expert assessment of the CMIP5 GCMs driven by the fact many of the GCM-based values of ΔT have exceeded ΔT_{OBS} over the past few decades (Figure 2 here and figure 11.25b of Kirtman et al. (2013)).

Projections of ΔT versus cumulative emissions provides a policy relevant framework for achieving the goals of the Paris Climate Agreement. From 1870 to date, humans have emitted roughly 600 GtC, and are currently emitting nearly 12 GtC per year. If society is to achieve the Paris Agreement by keeping the rise in ΔT below 2^{o}C with 95% probability by 2100, then only 20 to 25 years of cumulative carbon emissions remain in the allowable budget (Table 2). However, in reality society has less than 20 to 25 years when considering practical complications. Since it is unreasonable to assume that annual emissions can drop from 12 GtC to zero instantaneously, the reduction in emission rate must begin earlier to reach the same cumulative emissions total. While the 20 to 25 year limit found in our model framework suggests society has more time to act than indicated by the CMIP5 GCMs, we emphasize that the goal of the Paris Climate Agreement can only be achieved by near immediate reductions in global carbon emissions.

Methane, a potent GHG, also needs to be a large part of policy considerations when considering how cumulative emissions compare to global warming projections. As the United States and other major coal-burning nations switch to natural gas, the risk of significant CH_4 leakage into the atmosphere increases, potentially negating the climate benefit of switching to the less carbon intensive fossil fuel source (Jackson et al., 2018; Saunois et al., 2020). Carbon cycle feedbacks, such as higher activity in natural or agricultural wetland methane production or leakage from previously locked natural reservoirs, could further increase the atmospheric abundance of CH_4 (Comyn-Platt et al., 2018; Voigt et al., 2017). It is probably not reasonable to expect CH_4 to follow the peak-and-decline component pattern of RCP 4.5; on the other hand, the aggressive methane growth of RCP 8.5 also seems unreasonable (Figure 1). A projection of methane mid-range between RCP 4.5 and RCP 8.5 is perhaps a more likely scenario, implying a value for atmospheric methane between 2 and 3 ppm by 2100. Placing all GHGs other than methane along the RCP 2.6 trajectory would place us on a trajectory for having a reasonably favorable probability of liming warming to 2°C, irrespective of the future methane scenario (Figure 12). However, even for RCP 2.6, achievement of the ambitious Paris Climate Agreement target of 1.5°C drops noticeably as future atmospheric methane intensifies. Quite simply, limiting warming to 1.5°C will require aggressive future controls on atmospheric release of both CO_2 and CH_4.

Appendix A: Input data sources

Table A1.

Natural Factor Input Sources

Variable	Data Name	Years	Location
SAOD	CMIP6	1850-2014	https://esgf-node.llnl.gov/search/input4mips/
	GloSSACv2	1979-2018	https://opendap.larc.nasa.gov/opendap/GloSSAC/contents.html
	CALIPSO	2019	https://opendap.larc.nasa.gov/opendap/CALIPSO/contents.html
TSI	CMIP6	1850-2014	https://esgf-node.llnl.gov/search/input4mips/
	SORCE	2003-2019	https://lasp.colorado.edu/home/sorce/data/tsi-data/
ENSO[a]	MEIv2	1979-2019	https://psl.noaa.gov/enso/mei/data/meiv2.data
	MEI-ext	1871-2005	https://psl.noaa.gov/enso/mei.ext/table.ext.html
AMOC[b]	AMV	1850-2019	https://crudata.uea.ac.uk/cru/data/temperature/HadSST.3.1.1.0.median.nc
PDO		1900-2018	http://research.jisao.washington.edu/pdo/PDO.latest.txt
IOD		1870-2019	http://www.jamstec.go.jp/frcgc/research/d1/iod/iod/dipole_mode_index.html
		1850-1870	http://www.jamstec.go.jp/frcgc/research/d1/iod/kaplan_sst_dmi_new.txt

[a]1850-1870 ENSO constructed as an area SST average over the Nino3.4 region using https://crudata.uea.ac.uk/cru/data/temperature/HadSST.3.1.1.0.median.nc

[b]AMV calculated as an area average of Atlantic SSTs; multiple detrending and Fourier filtering options can be applied

Table A2.

Anthropogenic Factor Input Sources

Variable	Data Name	Years	Location
GHG RF	RCPs	1850-2099	http://www.pik-potsdam.de/~mmalte/rcps/
AER RF[c]			
	Sulfates	1850-2005	http://www.sterndavidi.com/datasite.html http://ciera-air.org/sites/default/files/Total SO2.xls
OHC	Levitus	1955-2019	http://data.nodc.noaa.gov/woa/DATA_ANALYSIS/3M_HEAT_CONTENT/DATA/basin/yearly/h22-w0-700m.dat
	Balmaseda	1958-2017	http://www.cgd.ucar.edu/cas/catalog/ocean/OHC700m.tar.gz
	Cheng	1955-2019	http://159.226.119.60/cheng/
	Ishii	1955-2017	http://159.226.119.60/cheng/
	Carton	1982-2017	https://www.atmos.umd.edu/~ocean/index_files/soda3_readme.htm

[c]The six time series of AER RF for each of the six types of aerosol species considered were combined as described in section §2.1.2

Appendix B: Model Input & Output Databases

All EM-GC input files (for natural and anthropogenic forcings) are available at http://dsrs.atmos.umd.edu/DATA/EMGCv4/INPUTS/ . Included are fifteen sample total AER RF time series.

All ΔT_{OBS} and OHC_{OBS} fitting files are available at http://dsrs.atmos.umd.edu/DATA/EMGCv4/OBS/ .

Original ΔT_{OBS} data was downloaded from the following sites:

CRU – https://www.metoffice.gov.uk/hadobs/hadcrut4/data/current/download.html

NCEI – https://www.ncdc.noaa.gov/cag/global/time-series

GISS – https://data.giss.nasa.gov/gistemp/

BEG – http://berkeleyearth.org/data-new/

CMIP5 model output is available at https://esgf-node.llnl.gov/search/esgf-llnl/ . All CMIP5 atmospheric temperature data were downloaded over several months in 2013 and 2014, and all CMIP5 ocean data were downloaded in early 2017.

Author contributions and other acknowledgements

TPC, APH, and LAM developed the model code used in this analysis; APH, LAM, and BFB collected data; RJS and TPC supervised, administrated, and developed the project; APH wrote the original draft; LAM, BFB, WRT, TPC, and RJS participated in the review and editing of the manuscript.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table S4 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

We appreciate financial support for this work from the NASA Climate Indicators and Data Products for Future National Climate Assessments (INCA) program (award NNX16AG34G) as well as the NOAA Cooperative Institute for Climate and Satellites (award WAIVER5278050)

References

Allen, M., & Millar, R. (2017, September 20). *Clarification on recent press coverage of our '1.5 degrees' paper in Nature Geoscience*. Oxford Martin School. Retrieved December 1 from https://www.oxfordmartin.ox.ac.uk/opinion/view/379

Andrews, T., Gregory, J., Webb, M., & Taylor, K. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models [Article]. *Geophysical Research Letters*, *39*, Article ARTN L09712. https://doi.org/10.1029/2012GL051607

Andronova, N., & Schlesinger, M. (2000). Causes of global temperature changes during the 19th and 20th centuries [Article]. *Geophysical Research Letters*, *27*(14), 2137-2140. https://doi.org/10.1029/2000GL006109

Arfeuille, F., Weisenstein, D., Mack, H., Rozanov, E., Peter, T., & Brönnimann, S. (2014). Volcanic forcing for climate modeling: a new microphysics-based data set covering years 1600–present. *Climate of the Past*, *10*(1), 359-375.

Balmaseda, M., Trenberth, K., & Kallen, E. (2013). Distinctive climate signals in reanalysis of global ocean heat content [Article]. *Geophysical Research Letters*, *40*(9), 1754-1759. https://doi.org/10.1002/grl.50382

Belcher, S., Boucher, O., & Sutton, R. (2019). Why results from the next generation of climate models matter. *Carbon Brief*.

Boessenkool, K. P., Hall, I. R., Elderfield, H., & Yashayaev, I. (2007). North Atlantic climate and deep-ocean flow speed changes during the last 230 years. *Geophysical Research Letters*, *34*(13).

Bond, T., Doherty, S., Fahey, D., Forster, P., Berntsen, T., DeAngelo, B., Flanner, M., Ghan, S., Karcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P., Sarofim, M., Schultz, M., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S., Hopke, P., Jacobson, M., Kaiser, J., Klimont, Z., Lohmann, U., Schwarz, J., Shindell, D., Storelvmo, T., Warren, S., & Zender, C. (2013). Bounding the role of black carbon in the climate system: A scientific assessment [Review]. *Journal of Geophysical Research-Atmospheres*, *118*(11), 5380-5552. https://doi.org/10.1002/jgrd.50171

Bony, S., Colman, R., Kattsov, V., Allan, R., Bretherton, C., Dufresne, J., Hall, A., Hallegatte, S., Holland, M., Ingram, W., Randall, D., Soden, B., Tselioudis, G., & Webb, M. (2006). How well do we understand and evaluate climate change feedback processes? [Review]. *Journal of Climate*, *19*(15), 3445-3482. https://doi.org/10.1175/JCLI3819.1

Booth, B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., & Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, *484*(7393), 228-232.

Canty, T., Mascioli, N., Smarte, M., & Salawitch, R. (2013). An empirical model of global climate - Part 1: A critical evaluation of volcanic cooling [Article]. *Atmospheric Chemistry and Physics*, *13*(8), 3997-4031. https://doi.org/10.5194/acp-13-3997-2013

Carton, J., Chepurin, G., & Chen, L. (2018). SODA3: A new ocean climate reanalysis. *Journal of Climate*, *31*(17), 6967-6983.

Ceppi, P., Brient, F., Zelinka, M. D., & Hartmann, D. L. (2017). Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdisciplinary Reviews: Climate Change*, *8*(4), e465.

Charette, M. A., & Smith, W. H. (2010). The volume of Earth's ocean. *Oceanography*, *23*(2), 112-114.

Chen, Y., & Penner, J. E. (2005). Uncertainty analysis for estimates of the first indirect aerosol effect.

Cheng, L., Trenberth, K., Palmer, M., Zhu, J., & Abraham, J. (2016). Observed and simulated full-depth ocean heat-content changes for 1970–2005. *Ocean Science*, *12*(4), 925-935.

Christy, J. R., & McNider, R. T. (2017). Satellite bulk tropospheric temperatures as a metric for climate sensitivity. *Asia-Pacific Journal of Atmospheric Sciences*, *53*(4), 511-518.

Chylek, P., Folland, C., Klett, J. D., & Dubey, M. K. (2020). CMIP5 climate models overestimate cooling by volcanic aerosols. *Geophysical Research Letters*, *47*(3), e2020GL087047.

Chylek, P., Klett, J. D., Dubey, M. K., & Hengartner, N. (2016). The role of Atlantic Multi-decadal Oscillation in the global mean temperature variability. *Climate Dynamics*, *47*(9-10), 3271-3279.

Chylek, P., Klett, J. D., Lesins, G., Dubey, M. K., & Hengartner, N. (2014). The Atlantic Multidecadal Oscillation as a dominant factor of oceanic influence on climate [Article]. *Geophysical Research Letters*, *41*(5), 1689-1697. https://doi.org/10.1002/2014GL059274

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., & Krinner, G. (2013). Long-term climate change: projections, commitments and irreversibility. In T. Stocker, D. Qin, G. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, & B. Midgley (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. 1029-1136). Cambridge University Press.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., & Shindell, D. (2017). AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6. *Geoscientific Model Development*, *10*(2), 585-607.

Comyn-Platt, E., Hayman, G., Huntingford, C., Chadburn, S. E., Burke, E. J., Harper, A. B., Collins, W. J., Webber, C. P., Powell, T., & Cox, P. M. (2018). Carbon budgets for 1.5 and 2 C targets lowered by natural wetland and permafrost feedbacks. *Nature geoscience*, *11*(8), 568.

Cowtan, K., Hausfather, Z., Hawkins, E., Jacobs, P., Mann, M., Miller, S., Steinman, B., Stolpe, M., & Way, R. (2015). Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophysical Research Letters*, *42*(15), 6526-6534.

DelSole, T., Tippett, M., & Shukla, J. (2011). A Significant Component of Unforced Multidecadal Variability in the Recent Acceleration of Global Warming [Article]. *Journal of Climate*, *24*(3), 909-926. https://doi.org/10.1175/2010JCLI3659.1

Deser, C., & Wallace, J. M. (1990). Large-scale atmospheric circulation features of warm and cold episodes in the tropical Pacific. *Journal of Climate*, *3*(11), 1254-1281.

Dessler, A. (2013). Observations of climate feedbacks over 2000–10 and comparisons to climate models. *Journal of Climate*, *26*(1), 333-342.

Dolinar, E. K., Dong, X., Xi, B., Jiang, J. H., & Su, H. (2015). Evaluation of CMIP5 simulated clouds and TOA radiation budgets using NASA satellite observations. *Climate Dynamics*, *44*(7-8), 2229-2247.

Domingues, C., Church, J., White, N., Gleckler, P., Wijffels, S., Barker, P., & Dunn, J. (2008). Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature*, *453*(7198), 1090-1093.

Duchez, A., Hirschi, J.-M., Cunningham, S., Blaker, A., Bryden, H., de Cuevas, B., Atkinson, C., McCarthy, G., Frajka-Williams, E., & Rayner, D. (2014). A new index for the Atlantic Meridional Overturning Circulation at 26° N. *Journal of Climate*, *27*(17), 6439-6455.

Dudok de Wit, T., Kopp, G., Fröhlich, C., & Schöll, M. (2017). Methodology to create a new total solar irradiance record: Making a composite out of multiple data records. *Geophysical Research Letters*, *44*(3), 1196-1203.

England, M., McGregor, S., Spence, P., Meehl, G., Timmermann, A., Cai, W., Sen Gupta, A., McPhaden, M., Purich, A., & Santoso, A. (2014). Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus [Article]. *Nature Climate Change*, *4*(3), 222-227. https://doi.org/10.1038/NCLIMATE2106

Fawcett, A. A., Iyer, G. C., Clarke, L. E., Edmonds, J. A., Hultman, N. E., McJeon, H. C., Rogelj, J., Schuler, R., Alsalam, J., & Asrar, G. R. (2015). Can Paris pledges avert severe climate change? *Science*, *350*(6265), 1168-1169.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., & Eyring, V. (2013). Evaluation of climate models. In T. Stocker, D. Qin, G. Plattner, M.

Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, & B. Midgley (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. 741-866). Cambridge University Press.

Folland, C. K., Boucher, O., Colman, A., & Parker, D. E. (2018). Causes of irregularities in trends of global mean surface temperature since the late 19th century. *Science advances*, *4*(6), eaao5297.

Forster, P., Andrews, T., Good, P., Gregory, J., Jackson, L., & Zelinka, M. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models [Article]. *Journal of Geophysical Research-Atmospheres*, *118*(3), 1139-1150. https://doi.org/10.1002/jgrd.50174

Forster, P. M., Maycock, A. C., McKenna, C. M., & Smith, C. J. (2020). Latest climate models confirm need for urgent mitigation. *Nature Climate Change*, *10*(1), 7-10.

Foster, G., & Rahmstorf, S. (2011). Global temperature evolution 1979-2010 [Article]. *Environmental Research Letters*, *6*(4), Article ARTN 044022. https://doi.org/10.1088/1748-9326/6/4/044022

Friedlingstein, P., Jones, M., O'sullivan, M., Andrew, R., Hauck, J., Peters, G., Peters, W., Pongratz, J., Sitch, S., & Le Quéré, C. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*(4), 1783-1838.

Fyfe, J., Gillett, N., & Zwiers, F. (2013). Overestimated global warming over the past 20 years [Editorial Material]. *Nature Climate Change*, *3*(9), 767-769.

Gettelman, A., Hannay, C., Bacmeister, J., Neale, R., Pendergrass, A., Danabasoglu, G., Lamarque, J. F., Fasullo, J., Bailey, D., & Lawrence, D. (2019). High climate sensitivity in the Community Earth System Model version 2 (CESM2). *Geophysical Research Letters*, *46*(14), 8329-8337.

Goodwin, P., Katavouta, A., Roussenov, V. M., Foster, G. L., Rohling, E. J., & Williams, R. G. (2018). Pathways to 1.5 C and 2 C warming based on observational and geological constraints. *Nature Geoscience*, *11*(2), 102-107.

Gregory, J. (2000). Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, *16*(7), 501-515.

Gregory, J., Jones, C., Cadule, P., & Friedlingstein, P. (2009). Quantifying carbon cycle feedbacks. *Journal of Climate*, *22*(19), 5232-5250.

Gryspeerdt, E., Muelmenstadt, J., Gettelman, A., Malavelle, F. F., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., & Wang, H. (2020). Surprising similarities in model and observational aerosol radiative forcing estimates. *Atmospheric Chemistry and Physics*, *20*(1), 613-623.

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). GLOBAL SURFACE TEMPERATURE CHANGE [Review]. *Reviews of Geophysics*, *48*, Article ARTN RG4004. https://doi.org/10.1029/2010RG000345

Hausfather, Z., Drake, H., Abbott, T., & Schmidt, G. (2020). Evaluating the performance of past climate model projections. *Geophysical Research Letters*, *47*(1), e2019GL085378.

Haustein, K., Otto, F., Venema, V., Jacobs, P., Cowtan, K., Hausfather, Z., Way, R., White, B., Subramanian, A., & Schurer, A. (2019). A limited role for unforced internal variability in 20 th century warming. *Journal of Climate*(2019).

Hope, A. P., Canty, T. P., Salawitch, R. J., Tribett, W. R., & Bennett, B. F. (2017). Forecasting Global Warming. In *Paris Climate Agreement: Beacon of Hope* (pp. 51-113). Springer.

Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., Rochetin, N., Fairhead, L., Idelkadi, A., & Musat, I. (2013). LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, *40*(9-10), 2193-2222.

IPCC. (2013). Summary for Policymakers. In T. Stocker, D. Qin, G. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, & B. Midgley (Eds.), *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3-29). Cambridge University Press.

Ishii, M., Fukuda, Y., Hirahara, S., Yasui, S., Suzuki, T., & Sato, K. (2017). Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, *13*, 163-167.

Ishii, M., Shouji, A., Sugimoto, S., & Matsumoto, T. (2005). Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *25*(7), 865-879.

Jackson, R. B., Le Quéré, C., Andrew, R., Canadell, J. G., Korsbakken, J. I., Liu, Z., Peters, G. P., & Zheng, B. (2018). Global energy growth is outpacing decarbonization. *Environmental Research Letters*, *13*(12), 120401.

Jones, P., Lister, D., Osborn, T., Harpham, C., Salmon, M., & Morice, C. (2012). Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010 [Article]. *Journal of Geophysical Research-Atmospheres*, *117*, Article ARTN D05127. https://doi.org/10.1029/2011JD017139

Kavvada, A., Ruiz-Barradas, A., & Nigam, S. (2013). AMO's structure and climate footprint in observations and IPCC AR5 climate simulations [Article]. *Climate Dynamics*, *41*(5-6), 1345-1364. https://doi.org/10.1007/s00382-013-1712-1

Kiehl, J. (2007). Twentieth century climate model response and climate sensitivity [Article]. *Geophysical Research Letters*, *34*(22), Article ARTN L22710. https://doi.org/10.1029/2007GL031383

Kirtman, B., Power, S., Adedoyin, A., Boer, G., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore, A., Kimoto, M., & Meehl, G. (2013). Near-term climate change: Projections and predictability. In T. Stocker, D. Qin, G. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, & B. Midgley (Eds.), *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 953-1028). Cambridge University Press.

Klein, S. A., Hall, A., Norris, J. R., & Pincus, R. (2017). Low-cloud feedbacks from cloud-controlling factors: a review. In *Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity* (pp. 135-157). Springer.

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., & Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, *32*(20).

Knutti, R., & Hegerl, G. C. (2008). The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geoscience*, *1*(11), 735.

Lamarque, J., Kyle, G., Meinshausen, M., Riahi, K., Smith, S., van Vuuren, D., Conley, A., & Vitt, F. (2011). Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways [Article]. *Climatic Change*, *109*(1-2), 191-212. https://doi.org/10.1007/s10584-011-0155-0

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., & Canadell, J. G. (2018). Global carbon budget 2018. *Earth System Science Data*, *10*(4), 2141-2194.

Lean, J. L., & Rind, D. H. (2008). How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006 [Article]. *Geophysical Research Letters*, *35*(18), Article ARTN L18701. https://doi.org/10.1029/2008GL034864

Lean, J. L., & Rind, D. H. (2009). How will Earth's surface temperature change in future decades? *Geophysical Research Letters*, *36*(15).

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J., Seidov, D., & Yarosh, E. S. (2012). World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophysical Research Letters*, *39*(10).

Lewis, N., & Curry, J. (2018). The impact of recent forcing and ocean heat uptake data on estimates of climate sensitivity. *Journal of Climate*, *31*(15), 6051-6071.

Lin, Y. J., Hwang, Y. T., Ceppi, P., & Gregory, J. M. (2019). Uncertainty in the evolution of climate feedback traced to the strength of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*, *46*(21), 12331-12339.

Marvel, K., Pincus, R., Schmidt, G. A., & Miller, R. L. (2018). Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, *45*(3), 1595-1601.

Masters, T. (2014). Observational estimate of climate sensitivity from changes in the rate of ocean heat uptake and comparison to CMIP5 models. *Climate dynamics*, *42*(7-8), 2173-2181.

Masui, T., Matsumoto, K., Hijioka, Y., Kinoshita, T., Nozawa, T., Ishiwatari, S., Kato, E., Shukla, P., Yamagata, Y., & Kainuma, M. (2011). An emission pathway for stabilization at 6 Wm− 2 radiative forcing. *Climatic change*, *109*(1-2), 59.

Matthes, K., Funke, B., Anderson, M., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M., Dudok de Wit, T., Haberreiter, M., & Hendry, A. (2017). Solar forcing for CMIP6 (v3. 2). *Geoscientific Model Development*, *10*(6), 2247-2302.

Mauritsen, T., & Pincus, R. (2017). Committed warming inferred from observations. *Nature Climate Change*, *7*(9), 652-655.

Medhaug, I., & Furevik, T. (2011). North Atlantic 20th century multidecadal variability in coupled climate models: sea surface temperature and ocean overturning circulation.

Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A., & Trenberth, K. E. (2011). Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change*, *1*(7), 360-364.

Meinshausen, M., Meinshausen, N., Hare, W., Raper, S. C., Frieler, K., Knutti, R., Frame, D. J., & Allen, M. R. (2009). Greenhouse-gas emission targets for limiting global warming to 2 C. *Nature*, *458*(7242), 1158.

Meinshausen, M., Smith, S., Calvin, K., Daniel, J., Kainuma, M., Lamarque, J., Matsumoto, K., Montzka, S., Raper, S., Riahi, K., Thomson, A., Velders, G., & van Vuuren, D. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300 [Article]. *Climatic Change*, *109*(1-2), 213-241. https://doi.org/10.1007/s10584-011-0156-z

Millar, R. J., Fuglestvedt, J. S., Friedlingstein, P., Rogelj, J., Grubb, M. J., Matthews, H. D., Skeie, R. B., Forster, P. M., Frame, D. J., & Allen, M. R. (2017). Emission budgets and pathways consistent with limiting warming to 1.5C. *Nature Geoscience*, *10*(10), 741-747.

Molnár, A. (2018). Leaking away the future: The role of methane emission and natural gas supply chains in global warming. *Energy Policy Studies*.

Myhre, G., Highwood, E. J., Shine, K. P., & Stordal, F. (1998). New estimates of radiative forcing due to well mixed greenhouse gases. *Geophysical research letters*, *25*(14), 2715-2718.

Myhre, G., Shindell, D., Bréon, F., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J., Lee, D., & Mendoza, B. (2013). Anthropogenic and Natural Radiative Forcing. In T. Stocker, D. Qin, G. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, & B. Midgley (Eds.), *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 659-740). Cambridge University Press.

Nijsse, F. J., Cox, P. M., & Williamson, M. S. (2020). Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models. *Earth System Dynamics*, *11*(3), 737-750.

Nisbet, E., Manning, M., Dlugokencky, E., Fisher, R., Lowry, D., Michel, S., Myhre, C. L., Platt, S. M., Allen, G., & Bousquet, P. (2019). Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris Agreement. *Global Biogeochemical Cycles*, *33*(3), 318-342.

Otto, A., Otto, F. E., Boucher, O., Church, J., Hegerl, G., Forster, P. M., Gillett, N. P., Gregory, J., Johnson, G. C., & Knutti, R. (2013). Energy budget constraints on climate response. *Nature Geoscience*, *6*(6), 415-416.

Pausata, F. S., Chafik, L., Caballero, R., & Battisti, D. S. (2015). Impacts of high-latitude volcanic eruptions on ENSO and AMOC. *Proceedings of the National Academy of Sciences*, *112*(45), 13784-13788.

Pincus, R., Forster, P. M., & Stevens, B. (2016). The radiative forcing model intercomparison project (RFMIP): Experimental protocol for CMIP6. *Geoscientific Model Development (Online)*, *9*(9).

Raftery, A. E., Zimmer, A., Frierson, D. M., Startz, R., & Liu, P. (2017). Less than 2 ^{o}C warming by 2100 unlikely. *Nature Climate Change*, *7*(9), 637-641.

Rahmstorf, S., Box, J., Feulner, G., Mann, M., Robinson, A., Rutherford, S., & Schaffernicht, E. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation [Article]. *Nature Climate Change*, *5*(5), 475-480. https://doi.org/10.1038/NCLIMATE2554

Raper, S., Gregory, J., & Stouffer, R. (2002). The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response [Article]. *Journal of Climate*, *15*(1), 124-130. https://doi.org/10.1175/1520-0442(2002)015<0124:TROCSA>2.0.CO;2

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., & Rafaj, P. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic change*, *109*(1-2), 33.

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forester, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., & Kriegler, E. (2018). Mitigation pathways compatible with 1.5° C in the context of sustainable development. IPCC Special Report on global warming of 1.5° C.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., Curry, J., Wickhams, C., & Mosher, S. (2013). Berkeley Earth Temperature Averaging Process. Geoinfor Geostat: An Overview 1: 2. *of*, *13*, 20-100.

Rose, B., Armour, K., Battisti, D., Feldl, N., & Koll, D. (2014). The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake [Article]. *Geophysical Research Letters*, *41*(3), 1071-1078. https://doi.org/10.1002/2013GL058955

Saji, N., Goswami, B., Vinayachandran, P., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean [Article]. *Nature*, *401*(6751), 360-363. https://doi.org/10.1038/43855

Santer, B., Painter, J., Bonfils, C., Mears, C., Solomon, S., Wigley, T., Gleckler, P., Schmidt, G., Doutriaux, C., Gillett, N., Taylor, K., Thorne, P., & Wentz, F. (2013). Human and natural influences on the changing thermal structure of the atmosphere [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(43), 17235-17240. https://doi.org/10.1073/pnas.1305332110

Saravanan, R., & McWilliams, J. C. (1998). Advective ocean–atmosphere interaction: An analytical stochastic model with implications for decadal variability. *Journal of Climate*, *11*(2), 165-188.

Saunois, M., Jackson, R., Bousquet, P., Poulter, B., & Canadell, J. (2016). The growing role of methane in anthropogenic climate change. *Environmental Research Letters*, *11*(12), 120207.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., & Patra, P. K. (2020). The global methane budget 2000–2017. *Earth system science data*, *12*(3), 1561-1623.

Schlesinger, M. E., & Ramankutty, N. (1994). An oscillation in the global climate system of period 65–70 years. *Nature*, *367*(6465), 723.

Schmidt, G., Shindell, D., & Tsigaridis, K. (2014). Reconciling warming trends [Editorial Material]. *Nature Geoscience*, *7*(3), 158-160. https://doi.org/10.1038/ngeo2105

Schwartz, S. E. (2012). Determination of Earth's Transient and Equilibrium Climate Sensitivities from Observations Over the Twentieth Century: Strong Dependence on Assumed Forcing [Review]. *Surveys in Geophysics*, *33*(3-4), 745-777. https://doi.org/10.1007/s10712-012-9180-4

Schwartz, S. E., Charlson, R. J., Kahn, R., & Rodhe, H. (2014). Earth's climate sensitivity: apparent inconsistencies in recent assessments. *Earth's Future*, *2*(12), 601-605.

Shen, Z., Ming, Y., & Held, I. M. (2020). Using the fast impact of anthropogenic aerosols on regional land temperature to constrain aerosol forcing. *Science Advances*, *6*(32), eabb5297. https://doi.org/10.1126/sciadv.abb5297

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., Heydt, A. S. v. d., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., & Zelinka, M. D. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *n/a*(n/a), e2019RG000678. https://doi.org/10.1029/2019rg000678

Shindell, D. T. (2014). Inhomogeneous forcing and transient climate sensitivity. *Nature Climate Change*, *4*(4), 274-277.

Silver, N. (2012). *The signal and the noise: why so many predictions fail--but some don't*. Penguin.

Smeed, D., Rayner, D., McCarthy, G., Moat, B., Johns, W. E., Baringer, M. O., Meinen, C. S., & Collins, J. (2017). RAPID-MOC time series April 2004 to February 2017.

Smith, S., Aardenne, J. v., Klimont, Z., Andres, R., Volke, A., & Delgado Arias, S. (2011). Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmospheric Chemistry and Physics*, *11*(3), 1101-1116.

Smith, S., & Bond, T. (2014). Two hundred fifty years of aerosols and climate: the end of the age of aerosols [Article]. *Atmospheric Chemistry and Physics*, *14*(2), 537-549. https://doi.org/10.5194/acp-14-537-2014

Smith, T., Reynolds, R., Peterson, T., & Lawrimore, J. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006) [Article]. *Journal of Climate*, *21*(10), 2283-2296. https://doi.org/10.1175/2007JCLI2100.1

Solomon, S. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge University Press.

Srokosz, M., & Bryden, H. (2015). Observing the Atlantic Meridional Overturning Circulation yields a decade of inevitable surprises. *Science*, *348*(6241), 1255575.

Steinman, B. A., Mann, M. E., & Miller, S. K. (2015). Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science*, *347*(6225), 988-991.

Stern, D. I. (2006). Reversal of the trend in global anthropogenic sulfur emissions. *Global Environmental Change*, *16*(2), 207-220.

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, B., & Midgley, B. (2013). IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.

Stouffer, R. J., Yin, J., Gregory, J., Dixon, K., Spelman, M., Hurlin, W., Weaver, A., Eby, M., Flato, G., & Hasumi, H. (2006). Investigating the causes of the response of the thermohaline circulation to past and future climate changes. *Journal of Climate*, *19*(8), 1365-1387.

Suckling, E. B., van Oldenborgh, G. J., Eden, J. M., & Hawkins, E. (2017). An empirical model for probabilistic decadal prediction: global attribution and regional hindcasts. *Climate Dynamics*, *48*(9-10), 3115-3138.

Taylor, K., Stouffer, R., & Meehl, G. (2012). AN OVERVIEW OF CMIP5 AND THE EXPERIMENT DESIGN [Article]. *Bulletin of the American Meteorological Society*, *93*(4), 485-498. https://doi.org/10.1175/BAMS-D-11-00094.1

Thomason, L., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., & Peter, T. (2018). A global space-based stratospheric aerosol climatology: 1979-2016. *Earth System Science Data*, *10*(1), 469-492.

Thomason, L., Vernier, J.-P., Bourassa, A., Millan, L., & Manney, G. (2016). A global, space-based stratospheric aerosol climatology: 1979 to 2014. AGU Fall Meeting Abstracts,

Thompson, D. W., Wallace, J. M., Jones, P. D., & Kennedy, J. J. (2009). Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights. *Journal of Climate*, *22*(22), 6120-6141.

Thomson, A. M., Calvin, K. V., Smith, S. J., Kyle, G. P., Volke, A., Patel, P., Delgado-Arias, S., Bond-Lamberty, B., Wise, M. A., & Clarke, L. E. (2011). RCP4. 5: a pathway for stabilization of radiative forcing by 2100. *Climatic change*, *109*(1-2), 77-94.

Thornhill, G., Collins, W., Olivié, D., Archibald, A., Bauer, S., Checa-Garcia, R., Fiedler, S., Folberth, G., Gjermundsen, A., Horowitz, L., Lamarque, J.-F., Michou, M., Mulcahy, J., Nabat, P., Naik, V., O'Connor, F. M., Paulot, F., Schulz, M., Scott, C. E., Seferian, R., Smith, C., Takemura, T., Tilmes, S., & Weber, J. (2020). Climate-driven chemistry and aerosol feedbacks in CMIP6 Earth system models. *Atmospheric Chemistry and Physics Discussion (in review)*. https://doi.org/https://doi.org/10.5194/acp-2019-1207

Tokarska, K. B., & Gillett, N. P. (2018). Cumulative carbon emissions budgets consistent with 1.5 C global warming. *Nature Climate Change*, *8*(4), 296.

Tokarska, K. B., Schleussner, C.-F., Rogelj, J., Stolpe, M. B., Matthews, H. D., Pfleiderer, P., & Gillett, N. P. (2019). Recommended temperature metrics for carbon budget estimates, model evaluation and climate policy. *Nature Geoscience*, *12*(12), 964-971.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., & Knutti, R. (2020). Past warming trend constrains future warming in CMIP6 models. *Science advances*, *6*(12), eaaz9549.

Trenberth, K. E. (1997). The definition of el nino. *Bulletin of the American Meteorological Society*, *78*(12), 2771-2777.

Tung, K., & Zhou, J. (2013). Using data to attribute episodes of warming and cooling in instrumental records [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(6), 2058-2063. https://doi.org/10.1073/pnas.1212471110

[Record #124 is using a reference type undefined in this output style.]

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., & Lamarque, J.-F. (2011). The representative concentration pathways: an overview. *Climatic change*, *109*(1-2), 5.

Van Vuuren, D. P., Stehfest, E., den Elzen, M. G., Kram, T., van Vliet, J., Deetman, S., Isaac, M., Goldewijk, K. K., Hof, A., & Beltran, A. M. (2011). RCP2. 6: exploring the possibility to keep global mean temperature increase below 2 C. *Climatic change*, *109*(1-2), 95.

Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., Hu, Y., & Hostetler, C. A. (2004). Fully automated analysis of space-based lidar data: An overview of the CALIPSO retrieval algorithms and data products. Laser radar techniques for atmospheric sensing,

Vial, J., Dufresne, J., & Bony, S. (2013). On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates [Article]. *Climate Dynamics*, *41*(11-12), 3339-3362. https://doi.org/10.1007/s00382-013-1725-9

Voigt, C., Lamprecht, R. E., Marushchak, M. E., Lind, S. E., Novakovskiy, A., Aurela, M., Martikainen, P. J., & Biasi, C. (2017). Warming of subarctic tundra increases emissions of all three important greenhouse gases–carbon dioxide, methane, and nitrous oxide. *Global Change Biology*, *23*(8), 3121-3138.

Volodin, E., Dianskii, N., & Gusev, A. (2010). Simulating present-day climate with the INMCM4. 0 coupled model of the atmospheric and oceanic general circulations. *Izvestiya, Atmospheric and Oceanic Physics*, *46*(4), 414-431.

Voosen, P. (2019). New climate models forecast a warming surge. *Science (New York, NY)*, *364*(6437), 222-223.

Weaver, C. J., Wu, D. L., Bhartia, P. K., Labow, G. J., & Haffner, D. P. (2020). A Long-Term Cloud Albedo Data Record Since 1980 from UV Satellite Sensors. *Remote Sensing*, *12*(12), 1982.

Wolter, K., & Timlin, M. S. (1993). Monitoring ENSO in COADS with a Seasonally Adjusted Principal. Proc. of the 17th Climate Diagnostics Workshop, Norman, OK, NOAA/NMC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, 52, Wolter, K., & Timlin, M. S. (2011). El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext). *International Journal of Climatology*, *31*(7), 1074-1087.

Wu, L., & Liu, Z. (2003). Decadal variability in the North Pacific: the eastern North Pacific mode. *Journal of Climate*, *16*(19), 3111-3131.

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T. Y., Shindo, E., Tsujino, H., & Deushi, M. (2012). A new global climate model of the Meteorological Research Institute: MRI-CGCM3—Model description and basic performance—. *Journal of the Meteorological Society of Japan. Ser. II*, *90*, 23-64.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E., Hegerl, G., Robock, A., Pausata, F., & Ball, W. (2016). The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, *9*(8), 2701-2719.

Zelinka, M., Klein, S., Taylor, K., Andrews, T., Webb, M., Gregory, J., & Forster, P. (2013). Contributions of Different Cloud Types to Feedbacks and Rapid Adjustments in CMIP5 [Article]. *Journal of Climate*, *26*(14), 5007-5027. https://doi.org/10.1175/JCLI-D-12-00555.1

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., & Taylor, K. E. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782.

Zelinka, M. D., Zhou, C., & Klein, S. A. (2016). Insights from a refined decomposition of cloud feedbacks. *Geophysical Research Letters*, *43*(17), 9259-9269.

Zeng, X., & Geil, K. (2016). Global warming projection in the 21st century based on an observational data-driven model. *Geophysical Research Letters*, *43*(20).

Zhang, H., Lawrimore, J., Huang, B., Menne, M., Yin, X., Sánchez-Lugo, A., Gleason, B., Vose, R., Arndt, D., & Rennie, J. (2019). Updated temperature data give a sharper view of climate trends. *Eos*, *100*.

Zhang, Y., Wallace, J. M., & Battisti, D. S. (1997). ENSO-like interdecadal variability: 1900-93. *Journal of climate*, *10*(5), 1004-1020.

Zharkova, V. V., Shepherd, S. J., Popova, E., & Zharkov, S. I. (2015). Heartbeat of the Sun from Principal Component Analysis and prediction of solar activity on a millenium timescale. *Scientific reports*, *5*, 15689.

Zhou, C., Zelinka, M. D., Dessler, A. E., & Klein, S. A. (2015). The relationship between interannual and long-term cloud feedbacks. *Geophysical Research Letters*, *42*(23).

Zhou, J., & Tung, K. (2013). Deducing Multidecadal Anthropogenic Global Warming Trends Using Multiple Regression Analysis [Article]. *Journal of the Atmospheric Sciences*, *70*(1), 3-8. https://doi.org/10.1175/JAS-D-12-0208.1

Zhu, J., Poulsen, C. J., & Otto-Bliesner, B. L. (2020). High climate sensitivity in CMIP6 model not supported by paleoclimate. *Nature Climate Change*, *10*(5), 378-379.