

AR6 updates to RF by GHGs and aerosols lowers the probability of accomplishing the Paris Agreement compared to AR5 formulations

Endre Z. Farago¹, Laura A. McBride², Austin P. Hope², Timothy P. Canty³, Brian F. Bennett³, Ross J. Salawitch^{1,3,4}

¹Department of Chemistry and Biochemistry, University of Maryland at College Park, College Park, MD, 20740, USA
²Science and Technology Corporation, Columbia, MD, 21046, USA
³Department of Atmospheric and Oceanic Science, University of Maryland at College Park, College Park, MD, 20740, USA
⁴Earth System Science Interdisciplinary Center, University of Maryland at College Park, College Park, MD, 20740, USA

Correspondence to: Endre Z. Farago (efarago@umd.edu) and Ross Salawitch (rsalawit@umd.edu)

Abstract. We provide a reduced complexity climate model (RCM) evaluation of how the IPCC WG1 Sixth Assessment Report (AR6) updates to the time series of the future atmospheric concentrations of GHGs, the effective radiative forcing (ERF) of GHGs, and the ERF of tropospheric aerosols (ERF_{AER}) affect attributable anthropogenic warming rate, climate sensitivity, and the likelihood of achieving either the goal (1.5 °C) or upper limit (2 °C) global warming thresholds of the Paris Agreement. This evaluation is conducted for four selected Shared Socioeconomic Pathway (SSP) scenarios: SSP1–1.9, SSP1–2.6, SSP4–3.4, and SSP2–4.5. Throughout, we compare and contrast these AR6 updates to the state of knowledge that existed prior to the publication of AR6, and provide probabilistic model simulations based on an evaluation of the impact in the uncertainty of ERF_{AER} and climate feedback. Our most important findings are that the rate of human-induced warming between 1975 and 2014 is 0.18 [0.13 to 0.21] °C decade^{−1} within the AR6 framework (range reflects the 5th and 95th percentiles), which is considerably lower than values found by many Earth System Models (ESMs) that participated in Phase 6 of the Coupled Model Intercomparison Project (CMIP6). Effective Climate Sensitivity (EffCS) inferred from the historical Global Mean Surface Temperature (GSMT) record was found to be 2.29 [1.54 to 3.11] °C using the ERF datasets from AR6 as model inputs. Upon adoption of the AR6 best estimate for the pattern effect (that is, 0.5 W m^{−2} °C^{−1}), we find values for Equilibrium Climate Sensitivity (ECS) of 3.24 [1.92 to 5.15] °C, which is quite similar to the AR6 assessment of 3.0 [2.0 to 5.0] °C for ECS. The hallmark of our RCM is the ability to conduct large (here, 160,000 member) ensemble forecasts of global warming. These calculations show that AR6 updates to the ERF of GHGs and aerosols result in a considerable decline in the likelihood of limiting warming to either 1.5 or 2 °C, compared to prior knowledge, for the same future emissions scenarios of GHGs. The likelihood of limiting global warming to 2.0 °C by end-of-century is found to be 100%, 85%, 40%, and 8%, for the SSP1–1.9, SSP1–2.6, SSP4–3.4, and SSP2–4.5 scenarios, respectively, based on the AR6 ERF datasets. Similarly, the ensembles run using the AR6 updates yield likelihoods of 70%, 32%, 3%, and 0% of limiting warming to 1.5 °C by end-of-century, for the same four SSPs. Many Reduced-complexity climate models (RCMs) and Earth System Models (ESMs) use prescribed concentrations or Effective Radiative Forcing (ERF) of Greenhouse Gases (GHGs) and tropospheric aerosols as inputs for

projections. Revisions to these datasets, made in Chapter 7 and Annex III of the Sixth IPCC Assessment Report: The Physical Science Basis (AR6, 2021) are vital to ensure the accuracy of climate model forecasts. AR6 provided updates to the formulation of ERF for most GHGs and tropospheric aerosols, relative to values in AR5 (2013). In this work, we provide a comprehensive assessment of how the changes to the ERF datasets impact projections of future warming, using our multiple linear regression energy balance RCM, the Empirical Model of Global Climate (EM-GC). We provide an analysis of the rate of human-induced warming (AAWR) between 1974 and 2014, and Effective Climate Sensitivity (EffCS) from the regression to the observation-based historical climate record with ERF datasets predating the AR6 report (which we term Baseline Framework) and AR6 ERF data (AR6 Framework). Probabilistic projections on future warming that consider the uncertainty in the magnitude of climate feedback and ERF from tropospheric aerosols are provided for four policy-relevant Shared Socioeconomic Pathway (SSP) scenarios. We find AAWR within the AR6 Framework to be 0.18 [0.13 to 0.21 °C decade⁻¹, 5–95% range], a slight increase to the values of 0.16 [0.12 to 0.20 °C decade⁻¹] within the Baseline Framework. The central estimate of EffCS is found to be nearly identical between the two Frameworks, but a narrower range is found for the AR6 Framework at 2.29 [1.54 to 3.11 °C, 5–95% range] relative to 2.26 [1.45 to 4.37 °C] within the Baseline Framework. We find Equilibrium Climate Sensitivity (ECS) to be 3.24 [1.92 to 5.15 °C] for the AR6 best estimate of the pattern effect. Our estimates of AAWR, EffCS and ECS are highly consistent with recent studies and observationally constrained CMIP6 model output. Projections of future warming for the AR6 Framework compared to the Baseline Framework show an increase of 0.2 to 0.4 °C in the end-of-century median warming for the SSP scenarios studied. This increase corresponds to a significantly lowered possibility of accomplishing the goals of the Paris Agreement (PA). In particular, the SSP2-4.5 scenario, that is widely considered to be consistent with current climate policies, only offers an 8% chance of accomplishing the PA upper limit of 2.0 °C warming by the end of the century within the AR6 Framework.

1. Introduction

Reduced Complexity Models (RCMs) that compute the response of the global mean surface temperature (GMST) to a prescribed radiative forcing (RF) due to anthropogenic greenhouse gases (GHGs) and tropospheric aerosols are becoming increasingly important for evaluating important climate metrics, such as the likelihood of limiting global warming to either the goal (1.5 °C) or upper limit (2.0 °C) of the Paris Agreement (Hope et al., 2017; McBride et al., 2021; Smith et al., 2018a; Nicholls et al., 2020; Nicholls et al., 2021). There are various types of RCMs, including some with interactive carbon cycles capable of computing atmospheric concentrations of CO₂ and CH₄ from emissions (Meinshausen et al., 2020). Typically,

Formatted: Normal, Tab stops: Not at 0.25"

RCMs represent the global mean energy balance between the atmosphere and the world's oceans on either a monthly or annual timescale, using various types of parameterizations that provide significant computational efficiency compared to three-dimensional Earth System Models (ESMs) (Nicholls et al., 2020; Nicholls et al., 2021). The computational efficiency of RCMs allows for the impact on GMST to be quantified for ensembles with hundreds of thousands of combinations for the RF due to GHGs, aerosols, and climate feedback. Of course, ESMs are essential for providing comprehensive simulations of the changes in the climate system in response to rising anthropogenic RF, such as the spatial distribution of warming, which is a key indicator for the impacts of climate change on local communities (Eyring et al., 2016). Furthermore, ESMs provide a more sophisticated treatment of atmospheric and oceanic interactions than is possible to achieve with RCMs (Nicholls et al., 2020; Nicholls et al., 2021).

We use a multiple linear regression (MLR) energy balance model, termed the Empirical Model of Global Climate (EM-GC) (Canty et al., 2013; Mascioli et al., 2012; Hope et al., 2017; McBride et al., 2021), to provide data-driven, probabilistic forecasts of GMST. The model is constrained by the Hadley Centre Climatic Research Unit version 5 (HadCRUT5) record for GMST over 1850–2019 (Morice et al., 2021), as described by McBride et al. (2021). In the McBride et al. (2021) paper, our projections of GMST were based on the time series for the atmospheric concentration of GHGs and the RF of tropospheric aerosols from Shared Socioeconomic Pathway (SSP) scenarios published in between the time of the fifth (AR5) Working Group 1 (WG1) Intergovernmental Panel on Climate Change (IPCC) report (IPCC, 2013) and the sixth (AR6) WG1 IPCC report (IPCC, 2021b). We had used formulations for the RF of GHGs from AR5 in McBride et al. (2021). Here, we rely upon formulations for the RF of GHGs given in AR6, which as detailed in Sect. 2.1.2 differ from the earlier formulations due to several new considerations within AR6. Similarly, we rely on updated time series of the atmospheric concentrations of GHGs given in AR6 for four selected SSP scenarios (that is, SSP1–1.9, SSP1–2.6, SSP4–3.4, and SSP2–4.5), which also differ from the pre-AR6 time series of GHGs due to the use of an updated model to compute atmospheric concentrations from prescribed emissions (Sect. 2.1.1). Finally, we use the AR6 update for the RF due to tropospheric aerosols, which also differs considerably from the AR5-based values used by McBride et al. (2021) (Sect. 2.1.3).

Section 2 provides a brief overview of the EM-GC model. Section 3 describes the simulations of GMST, the attributable anthropogenic warming rate, climate sensitivity and global warming projections found using our reduced complexity model. Throughout, we focus on comparing results found using the model inputs of McBride et al. (2021) (that is, the pre-AR6 estimates termed “Baseline simulations”) with results found using the AR6 values of the model inputs. We show that considerable differences are found for the end-of-century GMST and the probability of achieving either the goal (1.5°C) or upper limit (2.0°C) of the Paris Agreement, between the Baseline and the AR6 values for model inputs, despite the fact that for any given SSP scenario, both formulations are based upon the same time series of GHG emissions (Sect. 3.2). We conclude by comparing our GMST projections to output from ESMs that participated in Phase 6 of the Coupled Model Intercomparison Project (CMIP6) (Eyring et al., 2016). A brief set of concluding remarks are given in Section 4.

95 **1-Introduction**

The Paris Agreement (PA), negotiated in 2015, established the goal of limiting global warming to 1.5 °C (PA target) relative to the pre-industrial baseline, with an upper limit of 2.0 °C (PA upper limit). Forecasting the rise in global mean surface temperature, for various estimates of the future atmospheric abundances of greenhouse gases and aerosols, is important for assessing the feasibility of accomplishing the goal of the PA.

100 — Numerous modeling efforts focus on assessing the feasibility of various climate policies, and the implications of these policies on various elements of Earth’s climate. Earth System Models (ESMs) are the most comprehensive climate models available and are the primary tool of global warming research. Although ESMs are able to compute the atmospheric concentration of CO₂ from emissions, most often these models use prescribed time series of atmospheric CO₂ (Jones et al., 2016; Lawrence et al., 2016; Meinshausen et al., 2020). For non-CO₂ GHGs, ESMs of Phase 6 of the Coupled Model Intercomparison Project (CMIP6), also use prescribed atmospheric concentration trajectories (Meinshausen et al., 2020). While 105 there is an effort to make ESMs emissions-based for more greenhouse gases (GHGs) (Meinshausen et al., 2024), many ESMs will need to continue to rely on concentration input for many non-CO₂ GHGs.

— Reduced Complexity Climate Models (RCMs) are used to compute “best estimate” GHG concentration trajectories from emissions for various scenarios, which can serve as an input to ESMs (Smith et al., 2018a; Meinshausen et al., 2020; 110 Meinshausen et al., 2024). RCMs require substantially lower computational resources than ESMs and allow the exploration of a wide variety of possible emissions scenarios. Outputs from RCMs serve as the basis for projections of atmospheric abundances of CO₂, CH₄, N₂O and other GHGs that constitute the Shared Socioeconomic Pathways (SSPs) (Meinshausen et al., 2020).

— Chapters 10 (Bindoff et al., 2013) and 11 (Kirtman et al., 2013) of the 2013 IPCC Report (IPCC, 2013b) raised the 115 possibility that “some CMIP5 models have a higher transient response to GHGs and a larger response to other anthropogenic forcings (dominated by the effects of aerosols) than the real world”. The tendency of ESMs to warm more quickly than the observed rise in global mean surface temperature (GMST) over the past three to four decades has also been noted for CMIP6 model output (Tokarska et al., 2020b; Zelinka et al., 2020; McBride et al., 2021). Hausfather et al. (2022) recently termed this tendency the “hot model problem”. Nicholls et al. (2021) examined future projections of GMST from a suite of RCMs and 120 concluded that “the most extreme CMIP6 model projections [of GMST] are outside the range of most RCMs’ 5–95th percentiles [of GMST]”.

— The Physical Science Basis document published in 2021 (IPCC, 2021a), commonly known as the Sixth Assessment Report (AR6), provided important updates to the projections of future abundances of GHGs and aerosols compared to previously published values in the SSP database (Riahi et al., 2017; Van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; 125 Calvin et al., 2017; Kriegler et al., 2017; Rogelj et al., 2018). Furthermore, AR6 also updated values of the effective radiative forcing (ERF) due to the major GHGs and aerosols relative to the Fifth Assessment Report (AR5) (IPCC, 2013b).

— In this paper, we review the changes to the projections of GHGs and aerosols, as well as the ERF formulations adapted by AR6, and compare to prior SSP-based projections of the radiative forcing due to GHGs and aerosols (Sect. 2.3). Then, we use our multiple linear regression energy balance RCM, the Empirical Model of Global Climate (EM-GC, Sect. 2.4), (Cantý et al., 2013; Mascioli et al., 2012; Hope et al., 2017; McBride et al., 2021), to assess the implications of the AR6 updates on Effective Climate Sensitivity (EffCS, Sect. 3.1), the rate of warming due to human activity (Attributable Anthropogenic Warming Rate, AAWR; Sect. 3.1), and projected future warming (Sect. 3.2) for a large ensemble of aerosol trajectories. A major strength of our EM-GC is the ability to provide probabilistic forecasts of the future rise in GMST for each SSP scenario, based upon the uncertainty in climate feedback and the radiative forcing due to tropospheric aerosols (Cantý et al., 2013; Mascioli et al., 2012; Hope et al., 2017; McBride et al., 2021). Therefore, we conclude the paper by comparing the likelihood of achieving the target (1.5 °C warming in 2100) and upper limit (2.0 °C) of the Paris Agreement within the two frameworks, for various SSP scenarios (Sect. 3.2).

— Previous projections of the GMST anomaly found using our EM-GC (McBride et al., 2021) show good agreement with results from other RCMs that participated in the reduced complexity model intercomparison project (RCMIP) exercise (Nicholls et al., 2020; Nicholls et al., 2021). In particular, projections of GMST from our model showed good agreement with results from the Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC7) RCM (Meinshausen et al., 2011a; Meinshausen et al., 2011b; Meinshausen et al., 2020), that was used for the computation of GHG trajectories adapted by AR6. Consequently, the EM-GC simulations presented here should serve as a good reference point for the impact of the new AR6 ERF time series on projections of the GMST anomaly found using other RCMs.

2 Data and methods

Our EM-GC is designed to quantify the influence of a variety of anthropogenic and natural influences on GMST, using a multiple linear regression (MLR), energy balance approach. The anthropogenic contribution to GMST is simulated by the energy balance component of the model from the radiative forcing due to GHGs, tropospheric aerosols, and land use change (LUC), while also accounting for the export of heat from the atmosphere to the world's oceans (ocean heat export or OHE). The MLR component of the model is responsible for quantifying the influence of various natural factors on the GMST in a manner similar to other MLR-based analyses of the climate system (Lean and Rind, 2008, 2009; Foster and Rahmstorf, 2011; Zhou and Tung, 2013). Natural factors include increases in stratospheric aerosols due to major volcanic eruptions, the approximate 11-year variation in total solar irradiance (TSI), as well as interactions between the ocean and the atmosphere due to the El Niño–Southern Oscillation (ENSO), the Atlantic Meridional Overturning Circulation (AMOC), the Pacific Decadal Oscillation (PDO), and the Indian Ocean Dipole (IOD). The reader is directed towards Sect. 2.1 of McBride et al. (2021) for a more complete description of the model, as well as the governing equations.

Here, we examine four policy-relevant SSP scenarios: SSP1–1.9, SSP1–2.6, SSP2–4.5, and SSP4–3.4 from Tier 1 and Tier 2 of the ScenarioMIP protocol (O'Neill et al., 2016). These were chosen because SSP2–4.5 is the SSP scenario most

160 consistent with recent trends in the anthropogenic emissions of GHGs and aerosols (Meinshausen et al., 2024), while the other
three SSPs we have chosen all offer more aggressive means for climate mitigation than the SSP2–4.5 scenario. The second
number in the name of the SSP scenario is the target RF at the end of the century in units of W m^{-2} , commonly referred to as
the “nameplate RF” (O'Neill et al., 2014). We show model results for two frameworks: a Baseline that represents the state of
knowledge prior to AR6, and one that is based upon data given in Chapter 7 and Annex III of AR6 (Forster et al., 2021; IPCC,
2021c; Smith et al., 2021a; Smith et al., 2021b).

165 **2.1 Model Inputs**

2.1.1 Atmospheric Concentrations of Greenhouse Gases

Our EM-GC uses prescribed abundances of GHGs, including CO_2 , CH_4 , N_2O , tropospheric O_3 , chlorinated and/or brominated
ozone depleting substances, as well as hydrofluorocarbons, perfluorocarbons, and sulfur hexafluoride, as described in Sect.
2.2.3 of McBride et al. (2021). Here, we focus on describing changes in the atmospheric abundances of CO_2 , CH_4 , and N_2O
170 that occurred within the four SSP scenarios of our study, before and after the publication of AR6, since this change drives the
differences central to this study.

Before and after publication of AR6, the atmospheric concentrations of GHGs were computed by other groups using the
Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC) RCM (Meinshausen et al., 2011a;
Meinshausen et al., 2011b; Meinshausen et al., 2020). These times series rely on emissions data from various Integrated
Assessment Models (IAMs), as described by Riahi et al. (2017). The Baseline (that is, pre-AR6) projections, which we obtained
175 from the SSP database (Riahi et al., 2017; van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; Calvin et al.,
2017; Kriegler et al., 2017; Rogelj et al., 2018), were found using version 6.8 of MAGICC. The AR6 projections, which we
obtain from Annex III of AR6 (IPCC, 2021c), are based upon model runs with version 7 of MAGICC. MAGICC7 includes
important updates relative to MAGICC6.8, such as a permafrost feedback module that results in additional emissions of CO_2
and CH_4 from the melting of permafrost (Meinshausen et al., 2020).

180 The top three panels of Fig. S1 compare time series of the concentrations of CO_2 , CH_4 , and N_2O for the Baseline and AR6
frameworks. All of the GHG projections were found for the same underlying emissions scenarios. The impact of updates to
MAGICC7 are modest but noticeable for CO_2 , quite small for N_2O , and substantial for CH_4 . The MAGICC7 updates result in
the atmospheric concentration of CH_4 projected in 2100 to be higher by 180 ppb and 400 ppb for SSP2–4.5 and SSP4–3.4,
185 respectively, compared to the projections provided in the SSP database (Riahi et al., 2017).

2.1.2 Radiative Forcing of Greenhouse Gases

The EM-GC relies on time series of the RF due to GHGs, computed from time series of the atmospheric abundance of each
gas. For the Baseline framework, each RF term is found using parameterizations from the AR5 report, as described in Section
2.2.3 of McBride et al. (2021). These AR5 formulations, termed effective radiative forcing (ERF), allow for stratospheric

temperature adjustments to the instantaneous RF for many GHGs. The AR5 ERF formulas are based on the analysis by Myhre et al. (1998) of output found using line-by-line models (Edwards, 1992; Myhre and Stordal, 1997), and are identical to those given in the third WG1 IPCC report (IPCC, 2001) as well as the fourth report (IPCC, 2007).

Chapter 7 of the AR6 report introduced two important changes to the parameterizations of ERF due to GHGs (Forster et al., 2021). First, the parameterizations were updated to reflect the spectral overlaps of CO₂ and N₂O, the shortwave RF due to CH₄, and a new representation of the H₂O continuum (Etminan et al., 2016; Meinshausen et al., 2020). Second, the AR6 values of ERF accounts for both stratospheric and tropospheric temperature adjustments (Smith et al., 2018b). Consequently, while AR5 considers stratospheric temperature adjusted RF (SARF) to be equal to ERF, these two quantities differ in the AR6 formulation of RF. The AR5 and AR6 formulas for the ERF due to CO₂, CH₄, and N₂O are given in the Supplement.

The middle row of Fig. S1 compares time series of ERF due to CO₂, CH₄, and N₂O, for the Baseline (dotted lines) and AR6 (solid lines) frameworks. The results shown in this middle row reflect the AR6 updates to both the ERF and the future atmospheric abundances of GHGs. Values of ERF are higher in the AR6 framework compared to Baseline, with particularly large increases found for the ERFs of CO₂ and CH₄ for the SSP4–3.4 and SSP2–4.5 scenarios. Finally, Fig. S1h compares ERF due to all GHGs, for the Baseline and AR6 frameworks. The largest increase in ERF, among the four SSP scenarios considered, is found for SSP4–3.4 and SSP2–4.5, with end-of-century increases of 0.6 and 1.0 W m⁻², respectively. A similar qualitative conclusion was reached by Fredriksen et al. (2023), who contrasted projections of ERF from CMIP5 models with those from CMIP6 models, and found that CMIP6 models project higher levels of ERF by the end of the century relative to CMIP5 models.

2.1.3 Radiative Forcing of Tropospheric Aerosols

Time series of the RF due to tropospheric aerosols (hereafter, aerosols) is another very important input to our EM-GC. Figure S1g compares the total ERF of aerosols (ERF_{AER}), for the Baseline and AR6 frameworks, which is the sum of direct cooling by aerosols and the effect of aerosols on clouds (that is, the indirect effect). A summary of the ERF_{AER} time series for the Baseline framework, which considers a range of aerosol types such as sulphate, dust, organic carbon, black carbon, and biomass burning products, is given in Sect. 2.2.4 of McBride et al. (2021). These time series, published prior to the AR6 report, originate from the Potsdam Institute of Climate Research (PICR) website (<https://www.pik-potsdam.de/~mmalte/rcps/>, last opened: May 21, 2024) and are based on future emissions of aerosol precursors in the SSP database (Riahi et al., 2017; van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; Calvin et al., 2017; Kriegler et al., 2017; Rogelj et al., 2018).

Considerably stronger aerosol cooling is evident in the AR6 time series of ERF_{AER}, compared to the Baseline time series (Fig. S1g). Further, the temporal evolution (that is, the shape) of the historical cooling differs substantially between the AR6 and Baseline. Finally, AR5 and AR6 each provide best estimates, and possible ranges for ERF_{AER}, over the time periods 1750–2011 and 1750–2019, respectively. These estimates were assessed to be −0.9 W m⁻² with a possible range of −0.1 to −1.9 W m⁻² in AR5 (Myhre et al., 2013), and to be −1.1 [−0.4 to −1.7] W m⁻² in AR6 (Forster et al., 2021). Formally, the possible range limits correspond to 5th and 95th percentiles.

It is beyond the scope of this paper to delve deeply into the cause of the differences between the AR5 and AR6 estimates of ERF_{AER} . It is somewhat surprising that the AR6 update to the best estimate of ERF_{AER} in the year 2019 exhibits more cooling than the AR5 best estimate that reflected conditions out to 2011, because individual time series of ERF_{AER} in both AR5 and AR6 (Fig. S1g) exhibit a considerable decline in the absolute value of ERF_{AER} over the 2011 to 2019 period of time. This decline was driven by successful efforts to reduce the emissions of aerosol precursors, by various entities throughout the world, due to the public health concerns of aerosols (Smith and Bond, 2014; Fu et al., 2021). The primary reason for larger aerosol cooling in the AR6 best estimate of ERF_{AER} , despite the 8-year extension in end year, is the nearly factor of two increase in the assessed value of cooling due to the aerosol indirect effect from AR5's best estimate of -0.45 [0.0 to -1.2] $W\ m^{-2}$ to the AR6 best estimate of -0.84 [-0.25 to -1.45] $W\ m^{-2}$. A significant decline in the best estimate of black carbon warming in AR6 (0.11 [-0.20 to 0.42] $W\ m^{-2}$) compared to AR5 (0.4 $W\ m^{-2}$ [0.05 to 0.80] $W\ m^{-2}$) also contributes to the decline in the absolute value of ERF_{AER} in AR6, compared to AR5. There are other updates in the AR6 approach for ERF_{AER} , as summarized in Sect. 7.3.3 of Forster et al. (2021).

Recently, Zelinka et al. (2023) pointed out two coding errors in the Smith et al. (2020) paper that influenced the AR6 evaluation of ERF_{AER} . These two errors largely cancel for the evaluation of ERF_{AER} . The Zelinka et al. (2023) best estimate and standard deviation of ERF_{AER} , over 1750 to 2014, is -1.09 ± 0.24 $W\ m^{-2}$, which is slightly less aerosol cooling than the AR6 estimate of -1.3 [-0.6 to -2.0] $W\ m^{-2}$ for the same time period. Given the “medium confidence” associated with the assessed value of ERF_{AER} noted in Chapter 7 of AR6 (Forster et al., 2021), the lack of evaluation of ERF_{AER} by Zelinka et al. (2023) for the 1750 to 2019 time period that is central to our study, and the focus within Zelinka et al. (2023) on the evaluation of the various components of ERF_{AER} for contemporary periods of time rather than the historical evolution of ERF_{AER} , we have decided to use the AR6 historical time series for aerosol cooling as presented in the assessment.

Time series of ERF_{AER} are vitally important inputs to our EM–GC. A hallmark of our approach is to span a wide range of possible time series of ERF_{AER} , as well as a model parameter λ_E that represents the sum of all climate feedbacks, retaining for further analysis the members of this ensemble that satisfy three goodness-of-fit constraints, to the: 1) 170-year GMST record; 2) GMST record over the past 8 decades (formally, 1940 to 2019); 3) the ocean heat content record that begins in 1955. Further details of this ensemble approach are given in Sect. 2.1 of McBride et al. (2021). Figure S2 illustrates our approach for generating an ensemble of ERF_{AER} time series for the SSP2–4.5 scenario, within the AR6 framework. The solid black line shows the AR6 assessed best value of the time series of ERF_{AER} . An ensemble is created by scaling this time series by various constant multiplicative factors, with the color scheme chosen to highlight the numerical value of ERF_{AER} in 2019. A similar approach is used for the Baseline framework, relying upon time series of ERF_{AER} obtained from the aforementioned PICR website, as detailed in Sect. 2.5 and Fig. S7 of McBride et al. (2021). While one can envision a more sophisticated approach that allows for the alteration of the shape of ERF_{AER} , in addition to the magnitude, the actual ERF_{AER} responds quickly to changes in precursor emissions due to the short lifetime of tropospheric aerosols. Generally, historical aerosol precursor emissions are fairly well known (e.g. Hoesly et al. (2018)). The more sophisticated approach of Smith and Bond (2014), which

relied upon a RF parametrization tied to the emission of sulfate, black carbon, and organic carbon aerosols, resulted in an ensemble of time series for ERF_{AER} that exhibit nearly the same shape, with quite different peak cooling.

Figure S1i shows time series of total anthropogenic ERF (ERF_{ANTH}). As noted above, the design of the SSPs was predicated on the end-of century RF due to all human activity being close to the last numerical value in the scenario (that is, 4.5 W m^{-2} for SSP2–4.5) (O'Neill et al., 2014; Tebaldi et al., 2021). Close agreement of end-of-century ERF_{ANTH} and the SSP nameplate is found using GHG concentrations together with ERF formulations for the Baseline framework. Conversely, within the AR6 framework, ERF_{ANTH} in 2100 exceeds the nameplate values, with the difference being particularly large for SSP4–3.4 (0.6 W m^{-2}) and SSP2–4.5 (0.9 W m^{-2}). One final, important difference between the two frameworks is the steeper rise in ERF_{ANTH} between about 1960 and present within AR6 compared to Baseline, which is attributable to an assessed best value of much stronger aerosol cooling over the latter part of the prior century in AR6 relative to the Baseline (Fig. S1g).

2.1.4 Other Model Inputs

The EM-GC output shown below relies entirely on simulations constrained to match the HadCRUT5 GMST anomaly (ΔT) record (Morice et al., 2021) over the years 1850–2019. Model simulations are also constrained by observations of Ocean Heat Content (OHC) that start in 1955. Here, we use the average OHC from data provided by five groups: Levitus et al. (2012),

Balmaseda et al. (2013), Cheng et al. (2017), Ishii et al. (2017), and Carton et al. (2018). Further details, including our evaluation of the uncertainties in observed ΔT and OHC, are given in Sect. 2.2 of McBride et al. (2021).

The EM-GC simulations consider a variety of natural factors alongside the anthropogenic component of warming. The ENSO time series used during the training period (1850–2019) is based on Version 2 of the Multivariate ENSO Index (MEI.v2) (Wolter and Timlin, 1993; Zhang et al., 2019). The MEI.v2 dataset provides data starting in 1979. For 1850 to 1978, a historical extension based on Wolter and Timlin (2011) and the HadSST3 dataset (Kennedy et al., 2011) is used, as detailed in Sect. 2.2.6 of McBride et al. (2021). The input time series that is used to reflect changes in the strength of the AMOC is based on sea surface temperature (SST) data from HadSST4 (Kennedy et al., 2019) between the Equator and 60°N in the Atlantic Ocean, detrended using the magnitude of global anthropogenic radiative forcing, then Fourier-filtered to remove frequencies above $1/9 \text{ yr}^{-1}$ as described in Sect. 3.2.3 and 4.1.2 of Canty et al. (2013), as well as Sect. 2.2.7 of McBride et al. (2021). Input time series for the PDO and the IOD, which are found to have little effect on our historical simulations of ΔT , are the same as described by McBride et al. (2021). Indices for all of the oceanic proxies after 2019, which marks the end of the training period, are set to zero.

Our model also considers the impact on ΔT of variations in total solar irradiance (TSI) and major volcanic eruptions. The input time series for TSI anomalies is constructed from CMIP6 model data between 1850 and 2014 (Matthes et al., 2017), while values for 2015–2019 are obtained from the Solar Radiation and Climate Experiment (SORCE) (Dudok De Wit et al., 2017). The input time series for stratospheric aerosol optical depth (SAOD), for 1850 to 1978, is based on extinction coefficients obtained from the Volcanic Forcing Dataset (Arfeuille et al., 2014) that had been prepared for CMIP6 GCM runs. For 1979 to 2018, we use a time series of SAOD at 550 nm from the Global Space-based Stratospheric Aerosol Climatology

(GloSSAC v2.0) (Thomason et al., 2018). For the earlier time period (1850 to 1978), the extinction coefficients from the Volcanic Forcing Dataset were integrated from the tropopause to 39.5 km, to obtain a globally averaged SAOD, weighted by the cosine of latitude from 80 °S to 80 °N. For the latter time period (1979 to 2018), we calculate globally averaged SAOD from the GloSSAC dataset using cosine-latitude weighting over the same range of latitudes. For the year 2019, level 3 gridded SAOD product from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) (Vaughan et al., 2004) is used to obtain a global average SAOD, which is then offset by the average difference between the GloSSAC and CALIPSO datasets for the period of overlap (2006–2018) between the two datasets, as described in Sect. 2.2.5 of McBride et al. (2021). Values of the TSI anomaly beyond 2019 are set to zero, while for SAOD we use the value from December 2019 for 2020 to 2100. The input time series for all natural and anthropogenic factors are archived on Zenodo (Farago et al., 2025).

2.2 Model Outputs

Here we provide a brief overview of various outputs of the EM-GC simulations that will be described in Sect. 3.

2.2.1 Attributable Anthropogenic Warming Rate and Effective Climate Sensitivity

Attributable Anthropogenic Warming Rate (AAWR) is defined as the rate of change of the GMST anomaly (ΔT) due to anthropogenic activity, between 1975 and 2014. This time interval spans a 40-year period in which ΔT rose in a near-linear manner due to human activity (McBride et al., 2021). AAWR is determined as the slope of a linear fit to the anthropogenic component of global warming, defined by Eq. (9) of McBride et al. (2021). Our method for the evaluation of AAWR is similar to earlier, MLR based studies (Lean and Rind, 2008, 2009; Foster and Rahmstorf, 2011; Zhou and Tung, 2013), except that we quantitatively account for the impact of the uncertainty in the RF of aerosols and the strength of climate feedback on the possible range of AAWR.

Equilibrium Climate Sensitivity (ECS) is defined as the rise in ΔT after climate has equilibrated to a theoretical doubling of the pre-industrial concentration of CO_2 (IPCC, 2001, 2021a; Forster et al., 2021). Since equilibrium can take centuries to reach due to the slow transfer of heat to the deep oceans (Hansen et al., 2011; Church et al., 2013; Tokarska et al., 2020a), often the more short-term Effective Climate Sensitivity (EffCS) is used (Gregory et al., 2020; Tokarska et al., 2020a; Spencer and Christy, 2023). We compute EffCS from the ERF due to the doubling of the pre-industrial CO_2 concentration ($\Delta\text{ERF}_{2\times\text{CO}_2}$) as shown in Eq. (1).

$$\text{EffCS} = \frac{1}{\lambda_p - \lambda_{\Sigma}} \times \Delta\text{ERF}_{2\times\text{CO}_2} \quad (1)$$

where λ_p is a constant equal to $3.2 \text{ W m}^{-2} \text{ }^\circ\text{C}^{-1}$, λ_{Σ} represents the sum of all feedbacks that varies for different members of the ensemble, and $\Delta\text{ERF}_{2\times\text{CO}_2}$ is the rise in RF due to a doubling of CO_2 . This methodology, based on the terminology of Bony et al. (2006) is consistent with Box 7.1 of AR6 (Forster et al., 2021). Within the Baseline framework, we use the RF formula of Myhre et al. (1998), which leads to $\Delta\text{ERF}_{2\times\text{CO}_2} = 5.35 \times \ln(2) = 3.71 \text{ W m}^{-2}$. For the AR6 Framework, we use a value for

320 $\Delta\text{ERF}_{2\times\text{CO}_2}$ of 3.93 W m^{-2} , which is the best estimate for $\Delta\text{ERF}_{2\times\text{CO}_2}$ given in Sect. 7.3.2.1 and 7.SM.1.2 of AR6 (Forster et al., 2021; Smith et al., 2021a). Consequently, EffCS computed using the AR6 formula is 6% larger than that found using the Myhre et al. (1998) formula, for a given value of λ_{Σ} . Finally, we note that climate sensitivity deduced from historical warming may be different from true ECS, as the historical climate feedback could differ from the climate feedback under an abrupt $4\times\text{CO}_2$ forcing scenario that is often used to evaluate ECS in ESMs (Andrews et al., 2018; Andrews et al., 2019; Winton et al., 2020; Forster et al., 2021).

325 **2.2.2 Future Temperature Projections**

Projections of ΔT , for various SSP scenarios, are central to assessing the likelihood of achieving either the Paris Agreement goal (1.5°C) or upper limit (2.0°C) for the rise of GMST relative to pre-industrial (Hope et al., 2017; McBride et al., 2021; Smith et al., 2018a; Nicholls et al., 2020; Nicholls et al., 2021). An important aspect of EM-GC simulations is the ability to compute probabilistic forecasts of the rise in ΔT , taking into account the uncertainty in the radiative forcing of climate due to tropospheric aerosols.

330 To consider the uncertainty in the magnitude of net climate feedback and the strength of aerosol cooling, we use a 160,000 member ensemble, comprised of 400 possible values for the parameter λ_{Σ} , each combined with 400 time series for ERF_{AER} . This ensemble serves as the basis for the probabilistic forecasts of ΔT , as well as the numerical evaluations of AAWR and EffCS. We consider only the members of the ensemble that satisfy three χ^2 based metrics given by Eqs. S1–S3 (that is, each χ^2 value must be ≤ 2), which serve as the observational constraints of the model. Two of these metrics quantify how well the modelled GMST anomaly represents the observed temperature anomaly of the atmosphere for the entire training period (1850 – 2019, χ^2_{ATM}) and over the last 80 years (1940 – 2019, χ^2_{RECENT}). The third metric (χ^2_{OCEAN}) is a goodness-of-fit value between the observed and modeled ocean heat content. The χ^2_{RECENT} metric is used because without this constraint, some solutions with values of $\chi^2_{\text{ATM}} \leq 2$ have a visually poor simulation of the rise in GMST over the past 4 to 5 decades, due to the large uncertainty associated with early measurements of ΔT (McBride et al., 2021).

340 For the results shown in Section 3, the observationally constrained ensemble is then weighted by an asymmetrical Gaussian function, shown in Fig. 1a–b, centered around the IPCC best estimate of ERF_{AER} in the reference year (-0.9 Wm^{-2} in 2011 for the Baseline framework and -1.1 Wm^{-2} in 2019 for the AR6 framework). The 1σ and 2σ boundaries of the Gaussians are derived from the possible and likely ranges for ERF_{AER} provided by the AR5 (Baseline framework) and AR6 (AR6 framework) reports (Table S3). The Gaussians are asymmetrical because the likely and possible ranges of ERF_{AER} specified in AR5 and AR6 are not symmetric around the respective best estimates. The weighted ensemble is then used to compute probabilistic estimates of AAWR and EffCS (Sect. 3.1), as well as probabilistic forecasts on the GMST anomaly (Sect. 3.2).

345 The projections of ΔT shown in Section 3 assume that the climate feedback parameter, λ_{Σ} , is constant over time. Support for this assumption is given by the temporal invariance of the residual between measured and modeled values of ΔT , over the past century and a half, as shown in Fig. 14 of McBride et al. (2021). If the true value of λ_{Σ} varies over time, as has been suggested based on analysis of CMIP5 (Marvel et al., 2018; Rugenstein et al., 2020) and CMIP6 (Dong et al., 2020; Salvi et

Formatted: Normal

al., 2023), then the analysis conducted by McBride et al. (2021) indicates that our end-of-century projections of global warming could be biased low by a few tenths of the degree Celsius. Regardless, the primary contributor to the uncertainty in end-of-century warming is the imprecise knowledge of ERF_{AER} .

2-Data and methods

2.1 Shared Socioeconomic Pathway scenarios

Shared Socioeconomic Pathway (SSP) scenarios are used to represent various future outcomes in the emissions of GHGs and tropospheric aerosols, as well as aerosol precursors. SSP scenarios are denoted with the nomenclature SSP_x-y , where x represents the identifier of the baseline SSP pathway (1–5) related to mitigation and adaptation, and y is the target radiative forcing (in $W\ m^{-2}$) at the end of the century (commonly referred to as headline or nameplate RF) (O'Neill et al., 2014). As an example, $SSP1-2.6$ aims to keep anthropogenic ERF close to $2.6\ W\ m^{-2}$ in the year 2100 under a sustainable socioeconomic pathway. Different baseline SSP pathways associated with the same nameplate RF represent the notion that various socioeconomic trajectories can lead to a similar end-of-century RF (O'Neill et al., 2016). AR6 provides a detailed description of nine Tier 1 and Tier 2 (O'Neill et al., 2016) SSP scenarios (Chen et al., 2021). Of these nine scenarios, we extensively analyze the following four: $SSP1-1.9$, $SSP1-2.6$, $SSP4-3.4$, $SSP2-4.5$.

2.2 Modeling framework

We assess the impact on global warming projections of changes to the ERF datasets to GHGs and aerosols introduced by AR6. We define two frameworks, a Baseline Framework that represents the state of knowledge prior to AR6, and an AR6 Framework based on ERF data from Chapter 7 and Annex III of AR6 (Forster et al., 2021; IPCC, 2021b; Smith et al., 2021b; Smith et al., 2021a).

For the Baseline Framework, ERF for GHGs is primarily based on GHG concentrations from the SSP database (Riahi et al., 2017; Van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; Calvin et al., 2017; Kriegler et al., 2017; Rogelj et al., 2018) converted to ERF using the formulae from AR5, as described in Sect. 2.5.3. The ERF due to tropospheric aerosols in the Baseline Framework considers six aerosol types: sulfate, mineral dust, ammonium nitrate, fossil fuel organic carbon, fossil fuel black carbon and biomass burning organic and black carbon based on data from the Potsdam Institute for Climate Research (Meinshausen et al., 2011c), (Sect. 2.5.3). ERF input data for the AR6 Framework is obtained from Annex III and Chapter 7 of the AR6 report (Forster et al., 2021; IPCC, 2021b; Smith et al., 2021b; Smith et al., 2021a) as described in Sect. 2.5.4.

2.3 Effective radiative forcing

Effective Radiative Forcing (ERF) is described as the top-of-the-atmosphere (TOA) energy flux difference (in $W\ m^{-2}$) due to an imposed perturbation (Myhre et al., 2013a; Boucher and Randall, 2013; Sherwood et al., 2015; Forster et al., 2021). ERF accounts for both tropospheric and stratospheric temperature adjustments (Smith et al., 2018b; Forster et al., 2021), and as such is more representative of the impacts of various forcing agents on the GMST anomaly than stratospheric temperature-adjusted RF (SARF) or instantaneous RF (IRF) (Forster et al., 2021). ERF describes Earth's energy imbalance due to anthropogenic factors such as GHGs and tropospheric aerosols; hence, ERF is connected quantitatively to the anomaly in GMST that is central to projections of global warming.

Formatted: Normal, Tab stops: Not at 0.25"

2.3.1 Greenhouse gases

ERF from GHGs can be computed from their atmospheric concentrations (Byrne and Goldblatt, 2014). Formulations of Myhre et al. (1998) based on results from a line-by-line model (Edwards, 1992; Myhre and Stordal, 1997) were the generally accepted method for converting atmospheric concentrations to radiative forcing for GHGs. These parametrizations were adapted by multiple IPCC reports (TAR, AR4, AR5) (IPCC, 2001, 2007, 2013b). Updates to these formulations, based on the Oslo line-by-line model (Myhre et al., 2006) were provided by Etminan et al. (2016) and Meinshausen et al. (2020), with the latter being adapted into Chapter 7 of the AR6 report (Forster et al., 2021). The formulations by Etminan et al. (2016) and Meinshausen et al. (2020) account for the band overlaps between CO_2 and N_2O , include shortwave forcing from CH_4 , and provide an update to the water vapor continuum. Both of these formulations have their own respective advantages and limitations, as discussed in Section 2.7 of Meinshausen et al. (2020) and Section 7.SM.1.1 of AR6 (Smith et al., 2021a).

— The AR5 and AR6 formulations are used to compute time series of SARF for each GHG, based on time series of historical and future concentrations of GHGs. In the AR5 RF formulations, SARF computed with the formulations of Myhre et al. (1998) was considered to be equal to ERF, based on an analysis of the RF due to CO_2 by Vial et al. (2013). Since no similar analyses were available for other GHGs, ERF was considered equal to SARF for all other well-mixed GHGs in AR5 (Forster et al., 2021). In AR6, tropospheric adjustments for CO_2 , CH_4 and N_2O are assessed to be non-zero (see Sections 7.3.2 and 7.3.5.1 of Forster et al. (2021) for further detail). Consequently, the formulations for ERF given by AR6 allow for tropospheric adjustments, whereas the formulations given by AR5 allow for only stratospheric adjustments.

— Annex III of the AR6 report (IPCC, 2021b) adapted not only the ERF formulae described in Meinshausen et al. (2020), but also projections for GHG concentrations for various SSP scenarios published in the same paper. Here and throughout, we consider concentration to be synonymous with volume mixing ratio. The Meinshausen et al. (2020) GHG projections are provided using version 7 of the Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC7) RCM (Meinshausen et al., 2011a; Meinshausen et al., 2011b; Meinshausen et al., 2020). Prior to the release of the AR6 report, one of the most up-to-date sources for GHG concentrations for various SSP scenarios was the SSP database (Riahi et al., 2017; Van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; Calvin et al., 2017; Kriegler et al., 2017; Rogelj et al., 2018). These concentration projections were computed with version 6.8 of the MAGICC model, based upon emissions data from various Integrated Assessment Models (IAMs) (Riahi et al., 2017), that were also published in the SSP database. Meinshausen et al. (2020) computed the GHG concentrations using the same GHG emissions database with a newer version of the MAGICC model. MAGICC7 includes several updates and changes to certain gas cycles, as well as the enablement of the permafrost feedback module of the model, which results in additional emissions of CO_2 and CH_4 from the melting of permafrost.

— Figure 1 shows the time evolution of the concentration (Fig. 1a–c) and ERF (Fig. 1d–f) for CO_2 , CH_4 and N_2O within the Baseline and AR6 Frameworks, as well as ERF from tropospheric aerosols (ERF_{AER}), the total ERF from all GHGs, and ERF from overall anthropogenic activity (Fig. 1g–i). There is a considerable difference between the Baseline and AR6 datasets, particularly for SSP4–3.4 and SSP2–4.5. Concentrations of CO_2 are larger within the AR6 Framework by as much as 30 ppm

(Fig. 1a) in the year 2100 compared to the Baseline. Projected concentrations of CH_4 within the AR6 Framework are much higher for SSP2-4.5 and SSP4-3.4, compared to the Baseline. For the three major GHGs, an increase for the ERF was found in the AR6 Framework for most SSP scenarios, with a particularly high increase for ERF_{CO_2} and ERF_{CH_4} of SSP2-4.5 and SSP4-3.4 (Fig. 1d-f). This increase can be attributed to a combined effect of the increase in projected GHG concentrations (Fig. 1a-c) and the updated ERF formulations of Meinshausen et al. (2020), as well as updates to the tropospheric adjustments of RF (Smith et al., 2018b; Hodnebrog et al., 2020; Forster et al., 2021).

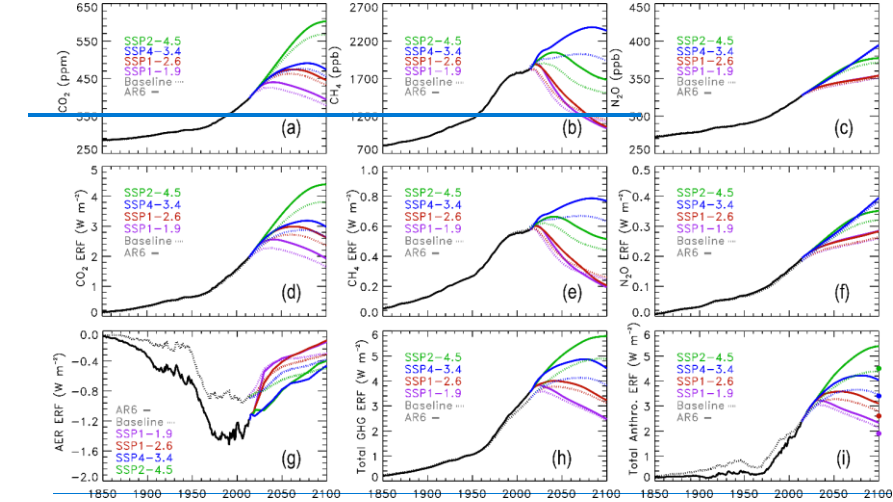


Figure 1: Concentrations and ERF due to various forcing agents for the Baseline (dashed) and AR6 (solid) frameworks. Colors represent SSP scenarios, as indicated on the individual panels. (a) Concentration of CO_2 . (b) Concentration of CH_4 . (c) Concentration of N_2O . (d) ERF from CO_2 . (e) ERF from CH_4 , including ERF from the enhancement of stratospheric water vapor due to methane. (f) ERF from N_2O . (g) ERF due to tropospheric aerosols. (h) ERF from all GHGs, including halogens and tropospheric O_3 , as well as land-use change. (i) Total anthropogenic ERF. Colored circles on the right-hand axis represent the target ERF of each SSP scenario as indicated in the name of the scenario.

AR6 adapted the SARF formulations of Meinshausen et al. (2020) for CO_2 , CH_4 and N_2O , replacing the formulae of Myhre et al. (1998) used in previous IPCC Assessment Reports (Table S1). To convert $\text{SARF}_{\text{CO}_2}$ to ERF_{CO_2} , an additional +5% tropospheric adjustment was also introduced in AR6 (Forster et al., 2021). This tropospheric adjustment increases ERF_{CO_2} by as much as 0.2 W m^{-2} in the year 2100 for SSP2-4.5. Concentration projections for CO_2 in Annex III of AR6 (IPCC, 2021b; Smith et al., 2021b) — which are largely based on Meinshausen et al. (2020) — are also higher than the concentration projections within the SSP Database (Riahi et al., 2017; Fricko et al., 2017; Calvin et al., 2017). Thus, the combination of increased projections on the concentration of CO_2 and updates to the conversion of CO_2 concentrations to ERF introduced in AR6 result

Formatted: Justified, Tab stops: Not at 0.25"

Formatted: Normal, Tab stops: Not at 0.25"

in increased projections of ERF_{CO_2} within the AR6 Framework relative to the Baseline, for all four SSP scenarios studied (Fig. 1d).

— The inclusion of permafrost emissions of methane result in as much as 50 ppb increase in the atmospheric concentration of methane in 2100, for the scenarios studied in our work (Fig. 3b of Meinshausen et al. (2020)). However, the inclusion of permafrost emissions does not fully account for the difference in the projected concentration of CH_4 between the Baseline and AR6 Frameworks (Fig. 1b).

— Updates to the atmospheric lifetime of CH_4 were introduced in MAGICC7 resulting in a shorter average lifetime for low radiative forcing scenarios, and longer lifetimes for high forcing scenarios (Sect. 2.4.1 of Meinshausen et al. (2020)). This update, alongside updates to the parametrization of the CH_4 cycle in MAGICC7 result in the difference between the projected CH_4 concentrations within the Baseline and AR6 frameworks for SSP4-3.4 and SSP2-4.5 depicted in Fig. 1b. In 2100, the atmospheric concentration of CH_4 is projected by Meinshausen et al. (2020) to be higher by 180 ppb and 400 ppb for SSP2-4.5 and SSP4-3.4, respectively, compared to the projections in the SSP database (Riahi et al., 2017; Fricko et al., 2017; Calvin et al., 2017). These findings are consistent with Meinshausen et al. (2020) who concluded “the comparison of mid-century CO_2 and CH_4 concentrations also reveals that the main reason for higher implied warming of SSP4-3.4 in comparison to SSP1-2.6 are elevated CH_4 concentrations”.

— While the difference in the projected concentration of CH_4 between our two frameworks is considerable for SSP4-3.4 and SSP2-4.5, the increase in ERF_{CH_4} is proportionally smaller in magnitude. The comparably smaller increase in ERF_{CH_4} is mostly due to the application of a multiplicative factor of 0.86 used to relate the new SARF formulation of CH_4 to ERF. This 14% reduction offsets a significant portion of the would-be increase in ERF due to CH_4 caused by higher projected concentrations and spectroscopy-based updates to the formulation of SARF (Forster et al., 2021). Section 7.3.2.2 of AR6 (Forster et al., 2021) assesses the uncertainty in the multiplicative factor used to relate SARF to ERF to be ± 0.15 , which introduces a sizeable new uncertainty in the radiative forcing of CH_4 for all of the SSPs. Overall, we find a major increase in the ERF_{CH_4} within the AR6 Framework compared to the Baseline for SSP4-3.4 and SSP2-4.5, driven by the elevated concentration projections of CH_4 . In contrast, differences in ERF_{CH_4} between the Baseline and AR6 Frameworks for SSP1-1.9 and SSP1-2.6 are quite small.

New parametrizations in MAGICC7 were also introduced for N_2O (Meinshausen et al., 2020). These parametrizations result in a minimal increase in the projected concentrations shown on Fig. 1e. A small increase in $\text{ERF}_{\text{N}_2\text{O}}$ is found between the two frameworks for all four SSP scenarios (Fig. 1f), mostly attributable to the $+7 \pm 13\%$ tropospheric adjustment described in Section 7.3.2.3 of AR6 (Forster et al., 2021).

2.3.2 Tropospheric aerosols and overall anthropogenic ERF

Tropospheric aerosols (AER) exhibit an overall cooling effect on the atmosphere, both by directly reflecting and absorbing incoming radiation (aerosol direct effect), and through interactions with clouds (aerosol indirect effect) (Forster et al., 2021). ERF from these two factors is usually denoted ERF_{aer} and $\text{ERF}_{\text{aer}^*}$, respectively. We use ERF_{AER} to denote the total ERF due to tropospheric aerosols ($\text{ERF}_{\text{aer}} + \text{ERF}_{\text{aer}^*}$).

— In AR5, the best estimate of the ERF_{AER} was assessed to be -0.9 Wm^{-2} , for the time period 1750–2011. In AR6, the best estimate of ERF_{AER} was determined to be -1.1 Wm^{-2} , between 1750–2019 (Forster et al., 2021). The corresponding 5th and 95th percentile ranges were -0.1 to -1.9 Wm^{-2} in 2011 and -0.4 to -1.7 Wm^{-2} in 2019, for the AR5 and AR6 assessments, respectively. AR6 also provided the assessment of -1.3 [range -0.6 to -2.0 Wm^{-2}] for the 1750–2014 period. We use the 1750–2019 time period below.

— Furthermore, AR5 assessed ERF_{net} and ERF_{aer} to be nearly equal in magnitude (see Fig. 8.15 of AR5 (Myhre et al., 2013a)). A major difference between these two assessments of ERF_{AER} is that, in AR6, it was stated that the cooling due to the aerosol indirect effect (ERF_{aci}) is much larger than the cooling due to the direct effect (ERF_{adi}). The change of ERF_{AER} by AR6 to a value that exhibits more cooling, for a time period that is 8 years more recent than the 2011 end-year used by AR5 implies a major shift in the shape of ERF_{AER} as a function of time, because tropospheric aerosol loading is believed to have declined over the past two decades due to successful air quality regulations implemented throughout the world (Smith and Bond, 2014; Fu et al., 2021).

— The methodology to develop the time evolution of RF due to aerosols in AR6 is partially based on the modeling study of Smith et al. (2021c). Similarly to GHGs, there is a direct relationship between ERF_{AER} and the atmospheric abundance of various types of aerosols, which is commonly expressed as a relation between the ERF and the emission of aerosols and aerosol precursors. The parametrization used in AR6 assumes that ERF_{net} scales linearly with the emissions of various types of aerosol precursors (Smith et al., 2018a; Smith et al., 2021c), while the relation between ERF_{aci} and emissions is logarithmic (Carslaw et al., 2013; Ghan et al., 2013; Stevens, 2015; Smith et al., 2018a; Smith et al., 2021c). Smith et al. (2021c) used aerosol ERF from various CMIP6 models, which were decomposed to the direct and indirect components using the Approximate-Partial Radiative Perturbation (APRP) method (Taylor et al., 2007; Zelinka et al., 2014), from which coefficients for the linear and logarithmic relations (for ERF_{net} and ERF_{aci} , respectively) for each CMIP6 model were determined.

— Next, Smith et al. (2021c) created a large ensemble of ERF_{AER} time series of various shapes and magnitudes by a Monte-Carlo sampling around the sets of coefficients obtained from the CMIP6 model-based fits. This aerosol ensemble accounts for the uncertainty in both the historical time evolution and the magnitude of ERF_{AER} . This methodology is adapted by Chapter 7 of AR6 for ERF_{aci} , as described in Sections 7.SM.1.3 and 7.SM.1.4 of AR6 (Smith et al., 2021a). The median of the ensemble is adapted as the best estimate for the historical time series of ERF_{aci} , scaled to match a value of -1.0 Wm^{-2} for 1750 to 2005–2014. For ERF_{adi} , AR6 uses the same assumption of a linear connection between aerosol emissions and forcing (Smith et al., 2018a; Smith et al., 2021c), but unlike Smith et al. (2021c), also considers emissions of NH_3 . Additionally, the AR6 report does not use a Monte-Carlo ensemble to derive a best estimate time series for ERF_{net} (and corresponding coefficients). Instead, AR6 determines coefficients such that the contributions from the individual aerosol precursor species match the values derived from Myhre et al. (2013b), as described in Section 7.SM.1.3.2 of the AR6 report (Smith et al., 2021a). The AR6 methodology provides a single set of “best estimate” coefficients that can be used to convert emissions data to a time series of ERF_{adi} and ERF_{aci} and consequently, ERF_{AER} .

— In our model framework, the radiative forcing of aerosols is treated in a probabilistic manner that expands upon the single best estimate ERF_{AER} time series that is associated with each SSP. Therefore, the comparison of single ERF_{AER} time series shown in Fig. 1g highlights only a portion of the difference between the Baseline and AR6 Frameworks.

— There are two major takeaways from Fig. 1g. First, for SSP2-4.5 and SSP4-3.4, aerosol cooling is much stronger between 2020 and 2060 in the AR6 Framework compared to the Baseline, whereas for SSP1-1.9 and SSP1-2.6, aerosol cooling is similar or weaker within the AR6 Framework than in the Baseline. In general, for all four SSP scenarios, the rate of decline in aerosol cooling is more rapid within the AR6 Framework relative to the Baseline. Second, if we use the aerosol ensemble member from our probabilistic treatment that corresponds to the best estimate of ERF_{AER} in the Baseline Framework (-0.9 Wm^{-2} in 2011 from AR5) as well as the best estimate in the AR6 Framework (-1.1 Wm^{-2} in 2019) to compute the overall anthropogenic ERF for the respective Frameworks (Fig. 1i), the total anthropogenic ERF between 1850 and 2019 is lower in the AR6 framework compared to the Baseline, whereas the rise in total anthropogenic ERF between 1950 and 2019 is more pronounced within the AR6 framework (black dotted and solid lines, Fig. 1i). This difference has important implications for our estimates of the rate of warming due to human activity, which we discuss in Sect. 2.4.

— Finally, we highlight how total anthropogenic ERF at the end of the century considerably exceeds, by as much as 0.7 Wm^{-2} for SSP2-4.5, the nameplate RF target for each of the four SSP scenarios (Fig. 1i). At the time the SSP scenarios were constructed, the value of the nameplate RF at the end of the century was designed to be interpreted as SARF (Tebaldi et al., 2021). The original purpose of the nameplate RF metric was to provide a numerical representation of a certain climate outcome (O'Neill et al., 2016). As highlighted by Sect. 7.3.1 of AR6 (Forster et al., 2021), “AR5 recommended ERF as a more useful measure of the climate effects of a physical driver than [SARF] adopted in earlier assessments”. Additionally, as described above, AR5 considered SARF to be equal to ERF, while this is no longer the case for AR6. Thus, in AR5 consistent settings, the nameplate RF of a given SSP scenario could be interpreted as either SARF or ERF. Lastly, RF for tropospheric aerosols is computed as ERF directly from emissions in AR6 (Smith et al., 2021a), which leads to additional differences with the end-of-century design of the original SSPs. As such, we conclude that the end-of-century RF is larger than the nameplate target for all four SSPs.

2.4 Empirical Model of Global Climate

To quantify the impacts of the updates to ERF datasets outlined above on AAWR, EffCS, and projected warming, we use the Empirical Model of Global Climate (EM-GC) model (Canty et al., 2013; Maseioli et al., 2012; Hope et al., 2017; McBride et al., 2021). EM-GC uses a multiple linear regression energy balance approach and computes the GMST anomaly ($\Delta T_{MDL,t}$) based on anthropogenic ERF and natural factors as shown in Eq. (1):

$$\Delta T_{MDL,t} = \frac{1+\gamma}{\lambda_F} (\Delta ERF_{GHG,t} + s \times \Delta ERF_{AER,t} + \Delta ERF_{LUC,t} - Q_{ULCAN,t}) + C_0 + C_1 \times SAOD_{t=6} + C_2 \times TSI_{t=1} + C_3 \times ENSO_{t=2} + C_4 \times AMOC_t + C_5 \times PDO_t + C_6 \times IOD_t \quad (1)$$

In Eq. (1), $\Delta ERF_{x,t}$ represent the effective radiative forcing from greenhouse gases (GHG), tropospheric aerosols (AER) and land use change (LUC) compared to a 1750 baseline. The model uses a monthly time grid, with i being the index of a specific

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

month. The dimensionless scaling parameter, s , is used to account for the uncertainty in the ERF due to aerosols as described in Sect. 2.4.1.

The natural factors considered when modeling the GMST anomaly are Stratospheric Aerosol Optical Depth (SAOD) from volcanic eruptions, Total Solar Irradiance (TSI), El Niño Southern Oscillation (ENSO), Atlantic Meridional Overturning Circulation (AMOC), Pacific Decadal Oscillation (PDO) and Indian Ocean Dipole (IOD), as outlined in Sect. 2.5.5. Similarly to previous analyses with EM-GC (McBride et al., 2021), SAOD, TSI and ENSO are lagged by 6, 1 and 2 months, respectively. Additionally, EM-GC quantitatively accounts for the export of heat to Earth's oceans ($Q_{OCEAN,i}$ in Eq. 1), as described below.

The dimensionless parameter γ in Eq. (1) is defined as the sensitivity of the global climate to feedbacks due to changes in the anthropogenic ERF due GHGs, AER and LUC. γ is related to the climate feedback parameter of EM-GC (λ_{Σ}) as shown in Eq. (2):

$$1 + \gamma = \frac{1}{1 - (\frac{\lambda_{\Sigma}}{\lambda_p})} \quad (2)$$

Here, λ_{Σ} is the sum of all climate feedbacks from factors such as water vapor, lapse rate, surface albedo and clouds. λ_p in Eqs. (1) and (2) corresponds to the response of a black body to a perturbation with no climate feedback present, with a value of $3.2 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ (Bony et al., 2006). The above mathematical representation is used in lieu of incorporating the Planck feedback into λ_{Σ} . For the EM-GC simulations shown below, we assume λ_{Σ} to be constant over time, which was found to be a very good approximation within EM-GC, as described by Sect. 3.3.6 of McBride et al. (2021).

Values of the regression coefficients C_0 – C_6 in Eq. (1) are computed by minimizing the cost function shown in Eq. (3):

$$\text{Cost function} = \sum_{i=1}^{N_{\text{MONTHS}}} \frac{1}{\sigma_{OBS,i}^2} (\Delta T_{OBS,i} - \Delta T_{MDL,i})^2 \quad (3)$$

Here, $\Delta T_{OBS,i}$ and $\Delta T_{MDL,i}$ represent the observed and EM-GC modeled GMST anomaly for a given month (i), respectively, while σ_{OBS} is the 1σ uncertainty associated with each temperature observation.

EM-GC simulations quantitatively account for heat being absorbed by Earth's oceans (ocean heat export, OHE). Time-dependent OHE is computed as shown by Eq. (4), where the constant κ is the ocean heat uptake efficiency (in $\text{W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$) as expressed by Eq. (S1). Further details of our model treatment of OHE is given in the Supplement as well as Sect. 2.1 of McBride et al. (2021).

$$Q_{OCEAN,i} = \kappa (\Delta T_{ATM,HUMAN,i} - \Delta T_{OCEAN,HUMAN,i}) \quad (4)$$

Time-series of the natural and anthropogenic factors between 1850 and 2019 are used as inputs to compute the coefficients (C_0 – C_6) in Eq. (1). We refer to this step as the training of the model. We define the historical training period as 1850–2019, and the future SSP-based projection period as 2020–2100, to be consistent with the definition of historical and SSP-based timeframes outlined in Section 7.SM.1.3 (Smith et al., 2021a) and Annex III (IPCC, 2021b) of AR6. Each fit to the historical GMST record (a set of coefficients) can be extrapolated to the future using Eq. (1), with effects of natural forcings zeroed out, allowing EM-GC to provide a time-dependent GMST forecast for each set of regression coefficients. The model also uses three χ^2 goodness-of-fit constraints, as summarized in Sect. 2.7.

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

Figure 2 shows the EM–GC best fit to the Hadley Centre Climatic Research Unit version 5 (HadCRUT5) (Morice et al., 2021) historical GMST anomaly record for the Baseline and AR6 Frameworks, assuming the IPCC best estimate RF trajectories for tropospheric aerosols shown on Fig. 1g. The top panels of Fig. 2 show the EM–GC modelled best fit (red) to the HadCRUT5 GMST record (black). Other panels represent the contribution to the modelled GMST from anthropogenic activity (Fig. 2e,d), as well as the natural factors of SAOD, TSI, ENSO, AMOC, PDO and IOD included in Eq. (1) (Fig. 2e–l). The bottom panels of both columns of Fig. 2 show the modeled (red) and measured (black) Ocean Heat Content (OHC), as well as the corresponding measurement uncertainty (blue). Description of the GMST and OHC datasets used to train the model are provided in Sect. 2.5. EM–GC runs simulate the amount of heat exported to Earth’s oceans such that the change in modeled OHC matches the change in observed OHC, over the period where observed OHC data are available. For panel (m) of Fig. 2, the modeled OHC time series is initialized to zero in 1850, and the observed OHC (black line, which is an anomaly) is adjusted to match the model mean over 1955 to 2019. For panel (n), the observed OHC is identical to that shown in panel (m), and the modeled time series is again adjusted to match the observed mean over 1955 to 2019. This final adjustment leads to a non-zero value for OHC in 1850 for the AR6 Framework run.

We provide the value of the λ_{Δ} and κ from Eqs. (1) and (4) on Fig. 2a,b and Fig. 2m,n, respectively. The χ^2 goodness-of-fit parameters (Sect. 2.7) for the modelled GMST anomaly (χ^2_{ATM}) and OHC (χ^2_{OCEAN}) are also provided on the same panels. The rate of increase in GMST between 1975 and 2014 due to human activity (AAWR, Sect. 2.6) is provided as the slope of a linear fit to the anthropogenic contribution to the GMST anomaly in this period on Fig. 2e,d. Similar values of AAWR are found for modest (that is, shifts within a decade) changes in the start and end year used to compute AAWR, due to the linearity of the anthropogenic contribution to GMST (orange lines in Fig. 2e–d).

Formatted: Tab stops: Not at 0.25"

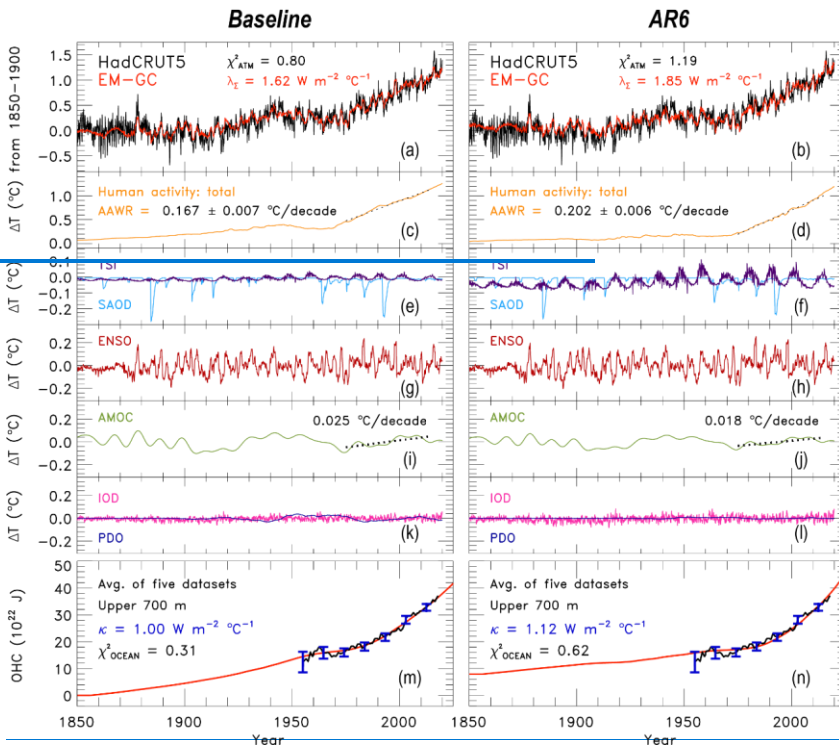


Figure 2: EM-GC assessed contribution of natural and anthropogenic factors to warming for the best fits to the HadCRUT5 GMST record within the Baseline (left) and AR6 (right) Frameworks. Fits are shown for the IPCC best estimate of ERF_{AER} (see text) for each Framework. Both simulations use a training period of 1850 to 2019. (a,b) Observed (black) and modelled (red) GMST anomaly (ΔT) relative to a pre-industrial (1850–1900) baseline. Values of λ_T and χ^2_{ATM} for the best-fit simulation are shown at the top. (c,d) Contribution of anthropogenic activity to the GMST anomaly (orange). AAWR, computed as the slope of a linear fit (dashed black line) as described in Section 2.6 is shown at the top. The 2σ uncertainty for the slope of the linear fit is also provided (see text). (e–h): Influence of TSI, SAOD and ENSO on the GMST anomaly. (i,j): Contribution of AMOC to the GMST anomaly. A linear fit between 1975 and 2014 is provided in a similar manner to the AAWR on panels (e,d). (k,l): Influences of PDO and IOD on the GMST anomaly. (m,n) Observed (black) and modelled (red) ocean heat content (OHC), with blue bars corresponding to the observational uncertainty. Observed OHC shown in this figure is based on the

Formatted: Justified, Tab stops: Not at 0.25"

Formatted: Normal, Tab stops: Not at 0.25"

average of five OHC datasets as described in Sect. 2.5.6. The χ^2_{OCEAN} goodness-of-fit parameter, as well as the ocean heat uptake efficiency (κ) are displayed at the top.

For the AR6 Framework, an AAWR of 0.202 ± 0.006 °C/decade is obtained from the best fit to the HadCRUT5 GMST record. A considerably smaller value for AAWR of 0.167 ± 0.006 °C/decade is found for the Baseline framework. As noted above, the rise in total anthropogenic ERF since about 1950 is more pronounced within the AR6 Framework compared to the Baseline (Fig. 1i), due to differences in the best estimate time series of ERF_{AER} between the two frameworks (Fig. 1g). Since AAWR is based on the slope of the linear fit to the anthropogenic contribution to warming, a higher value of AAWR is found for the AR6 framework. Here, both uncertainties for AAWR are based on the goodness of the linear fit for a single model run and are noted only for completeness. A more realistic estimate of the true uncertainty in AAWR is found using the ensemble runs of our EM-GC, as described in Section 3.1.

The value of 0.202 °C/decade found within the AR6 Framework for AAWR is consistent with the assessment of 0.19 °C/decade between 1980–2020 described in Table 2.4 of AR6 Chapter 2 (Gulev et al., 2021; Forster et al., 2023). Furthermore, AAWR computed within the AR6 framework shows closer agreement—compared to the Baseline framework—with AAWR derived from a CMIP6 multi-model ensemble, the median of which McBride et al. (2021) found to be 0.221 °C/decade, with the 5th, 25th, 75th and 95th percentiles being 0.151 , 0.192 , 0.245 and 0.299 °C/decade, respectively. However, the median value of AAWR for the CMIP6 ensemble, as well as AAWR obtained from numerous CMIP6 ensemble members, are higher than our regression-based estimate of AAWR using EM-GC. This finding is consistent with several recent assessments finding that some CMIP6 GCMs overestimate the extent of warming in response to GHG emissions (Tokarska et al., 2020b; McBride et al., 2021; Hausfather et al., 2022).

Below, we provide a description of the probabilistic approach utilized by EM-GC to account for the uncertainty in ERF_{AER} . We quantify AAWR and EffCS for the whole range of uncertainty in ERF_{AER} and climate feedback in Section 3.1.

2.4.1 Uncertainty in ERF_{AER} within EM-GC

Uncertainty in the magnitude of ERF_{AER} is accounted for by performing the above regression for a large ensemble of aerosol scenarios. We scale a best estimate time series of aerosol ERF ($\Delta\text{ERF}_{\text{AER}}$ in Eq. (1)) by a series of constant multiplicative factors (s in Eq. (1), Fig. S1). The best estimate time series is derived from data from the Potsdam Institute for Climate Research and Annex III of AR6 for the Baseline and AR6 frameworks, respectively; as outlined in Sections 2.5.3 and 2.5.4. We use ERF_{AER} in a given reference year (hereinafter $\text{ERF}_{\text{AER},t}$ where t denotes the reference year) as the identifier of each of the aerosol ensemble members. The reference year was chosen as 2011 and 2019 for the Baseline and AR6 frameworks, respectively, to be consistent with the years for which AR5 and AR6 provided a best estimate and likely range for ERF_{AER} .

The regression described in Eqs. 1–3 is performed with each member of the aerosol ensemble being combined with a series of various values of the climate feedback parameter (λ_{E}). In total, 400 values of λ_{E} are used together with 400 values of the scaling factor s (hence $\text{ERF}_{\text{AER},t}$), to create a $\lambda_{\text{E}} - \text{ERF}_{\text{AER},t}$ grid with 160,000 elements (hereinafter EM-GC grid). Each element of the grid corresponds to a fit of the historical GMST record, and thus, a unique set of $C_u - C_6$ regression coefficients. These sets of coefficients are then used to forecast the future GMST anomaly for each grid member, using ERF projections

for GHGs and aerosols from SSPs. These GMST projections are then filtered and weighed as described in Section 2.7 based on both the goodness of the fit to the historical GMST record, as well as the magnitude of ERF_{AER} compared to the best estimate and likely range of ERF_{AER} prescribed by AR5 and AR6. Lastly, we use the weighted projections to compute probabilistic forecasts on the future rise of GMST.

2.5 Model inputs

2.5.1 Temperature data

In this study, we use the Hadley Centre Climatic Research Unit version 5 (HadCRUT5) (Morice et al., 2021) GMST anomaly records for the training of the EM-GC. Temperature anomalies are with respect to an 1850–1900 pre-industrial baseline. Similarly to previous works using EM-GC (McBride et al., 2021), the uncertainty time series given with the HadCRUT4 temperature record are used for the EM-GC simulations in this paper. The input GMST anomaly is shown by the black line in Fig. 2a–b.

2.5.2 Shared Socioeconomic Pathways

In this work, we provide EM-GC output for four different SSP scenarios, two from Tier 1 (SSP1–2.6, SSP2–4.5) and two from Tier 2 (SSP1–1.9, SSP4–3.4) of the ScenarioMIP experiment (O'Neill et al., 2016). Three of these SSP scenarios (SSP1–1.9, SSP1–2.6 and SSP2–4.5) are highlighted as significant scenarios both in Annex III of AR6 (IPCC, 2021b) and Table 1 of Meinshausen et al. (2024). SSP4–3.4 is considered as a scenario which was closest to a 50% probability of limiting warming below the PA upper limit of 2.0 °C in 2100, based on the previous work of McBride et al. (2021) using the EM-GC model. Higher forcing scenarios like SSP3–7.0 and SSP5–8.5 are not considered here, as our intention is to focus on the SSP scenarios that are consistent with current or potential future climate policies. Scenarios like SSP3–7.0 and SSP5–8.5 fall into the “The World We Avoided” category of Meinshausen et al. (2024), who considers this category to be of a lower priority (see Table 1 of Meinshausen et al. (2024)).

2.5.3 Baseline Framework

The Baseline framework is constructed based on datasets available shortly before the AR6 report was published. Historical concentrations of GHGs between 1850 and 2014 are from Meinshausen et al. (2017), while GHG concentrations between 2015 and 2100—with the exception of ODSs—are from the SSP database (Riahi et al., 2017; Van Vuuren et al., 2017; Fricko et al., 2017; Fujimori et al., 2017; Calvin et al., 2017; Kriegler et al., 2017; Rogelj et al., 2018) available at <https://tntcat.hiasa.ac.at/SspDb/> (last opened: January 8, 2024). Data in these sources are provided on a yearly and decadal grid, respectively, which are interpolated to obtain a monthly time series of GHG concentrations between 1850 and 2100. Concentrations of CO_2 , CH_4 and N_2O are converted to SARF using the formulae of Myhre et al. (1998). SARF attributable to stratospheric water vapor is computed as 15% of $SARF_{CH_4}$ based on Myhre et al., (2007). Concentrations of ODSs beyond 2015 are from table 6.4 of the 2018 Ozone Assessment Report (Carpenter et al., 2018), and are converted to RF using radiative efficiencies for these obtained from the WMO (2018). The SSP database does not provide SARF time series for tropospheric O_3 , hence our input of $SARF_{O_3}$ is developed based on Meinshausen et al. (2011c) with the same methodology used by McBride et al. (2021): $SARF_{O_3}$ between 1850 and 2005 are obtained from Meinshausen et al. (2011c) through the Potsdam Institute of

Climate Research (PICR) website (<https://www.pik-potsdam.de/~mmalte/reps/>, last opened: May 21, 2024), while from 2006 onwards, for each SSP scenario, we use SARF_{O3} from a corresponding RCP pathway in Meinshausen et al. (2011e). SARF_{O3} from RCP2.6 and RCP4.5 are used for SSP1–2.6 and SSP2–4.5, respectively. We use SARF_{O3} from RCP2.6 for SSP1–1.9, while for SSP4–3.4, a new time series was created, using the linear combination of the RCP2.6 and RCP8.5 time series, weighed by the total GHG SARF in 2100. We note that RCP2.6 is sometimes referred to as RCP3–PD (Meinshausen et al., 2011e), including on the PICR website. Consistent with the pre-AR6 formulations, within the Baseline framework, we consider ERF to be equal to SARF for all GHGs (Myhre et al., 2013a; Forster et al., 2021). Finally, ERF due to land-use change between 1850 and 2011 is obtained from Table AII.1.2 of Annex II of the AR5 (IPCC, 2013a), which is interpolated onto a monthly time grid. Between 2012 and 2100, we used the 2011 value for ERF_{LUC}.

The initial time series for the RF due to tropospheric aerosols is developed in a manner that is similar to that described by McBride et al. (2021). Between 1850 and 2005, we sum RF from six different aerosol types (sulfate, mineral dust, ammonium-nitrate, fossil fuel black carbon and biomass burning organic and black carbon) to obtain a time series of the total aerosol direct effect. RF data for each of these aerosol types are from the PICR (Meinshausen et al., 2011e) for RCP4.5, except for sulfate, where we use the estimate of Smith et al. (2011). The PICR datasets for aerosol direct effects are nearly identical until 2005, so the specific choice of RCP4.5 between 1850 and 2004 carries no significance. Time series for aerosol indirect effects between 1850 and 2005 are developed by scaling the time series of the direct effect using methodology described in Section 3.2.2 of Canty et al. (2013), as well as Hope et al. (2017). The time series for direct and indirect effects are added to obtain the time series of total RF due to tropospheric aerosols between 1850 and 2005. From 2005 onwards, values of total aerosol RF published in the SSP database are used. In order to obtain a continuous time series, values of total aerosol RF between 1850 and 2005 are scaled such that the value in 2005 matches the total aerosol RF value in 2005 obtained from the SSP database. The time series of RF due to aerosols described above are treated as ERF for the Baseline Framework. The initial ERF_{AER} time series is then scaled to obtain an ensemble of aerosol forcing scenarios, as described in Sect. 2.4.1.

2.5.4 AR6 Framework

EM–GC input data for the AR6 Framework are based on the datasets published in Annex III of AR6 (IPCC, 2021b), and the corresponding data repository (<https://doi.org/10.5281/zenodo.5705391>, last opening: 9th January, 2024) (Smith et al., 2021b). Yearly concentrations for CO₂, CH₄ and N₂O are obtained from this repository, which are based on the time series by Meinshausen et al. (2017) and Meinshausen et al. (2020) for the 1850–2014 and 2019–2100 periods, respectively, with a linear transition between the two datasets between 2015 and 2020, as described by Sections 7.SM1.3 and 7.SM.1.4 of AR6 (Smith et al., 2021a). Similarly to AR6, the atmospheric concentrations are converted to SARF with the formulae of Meinshausen et al. (2020), using pre-industrial concentrations of 278.3 ppm, 270.1 ppb and 729 ppb for CO₂, N₂O and CH₄ respectively. These values are ice-core-based best assessments of pre-industrial concentrations for these GHGs as described in Chapter 2 of AR6 (Gulev et al., 2021). Tropospheric adjustments of +5%, +7% and –14% are used (Smith et al., 2018b; Hodnebrog et al., 2020) to convert SARF to ERF for CO₂, N₂O and CH₄, respectively, which are based on Section 7.3.2 and 7.SM.1.3.1 of AR6 (Forster et al., 2021; Smith et al., 2021a). ERF from stratospheric water vapor is computed as 9.19% of methane ERF; our use of this

sealing value was found to have the highest agreement with the values provided in Annex III of AR6. ERF time series for halogenated compounds and tropospheric ozone are directly adapted from the AR6 Annex III datasets. This dataset provides total ERF from both stratospheric and tropospheric O_3 , which is dominated by the contribution from tropospheric O_3 (see note below Table AIII.3 in Annex III of AR6). All of the above input time series are interpolated onto a monthly time grid, and then summed to provide the ΔERF_{CHG} time series in Eq. (1) for the AR6 Framework.

———— The ERF time series for land-use change is adapted directly from the AR6 Annex III dataset and interpolated onto a monthly time grid to obtain ΔERF_{LUC} to be used in Eq. (1). Finally, the initial ERF time series for tropospheric aerosols is also created using the AR6 Annex III datasets by summing ERF_{aer} (direct effect) and ERF_{aer} (indirect effect) for the entire timeframe for each SSP scenario. The sum is then scaled to create an ensemble for ERF_{AER} as described in Section 2.4.1. The resulting time series are then interpolated to obtain monthly time series for ERF_{AER} .

2.5.5 Natural factors

EM-GC simulations consider a variety of natural factors alongside the anthropogenic component of warming, as shown in Eq. (1). The input time series for TSI anomalies used in Eq. (1) is constructed from CMIP6 model data between 1850 and 2014 (Matthes et al., 2017), while values for 2015–2019 are obtained from the Solar Radiation and Climate Experiment (SORCE) (Dudok De Wit et al., 2017). The SAOD input time series from 1850 to 1978 is based on extinction coefficients at 550 nm from 80°S to 80°N obtained from the Volcanic Forcing Dataset (Arfeuille et al., 2014) that was prepared for CMIP6 GCM runs. For 1979 to 2018, we use a time series of SAOD at 550 nm from the Global Space-based Stratospheric Aerosol Climatology (GloSSAC v2.0) (Thomason et al., 2018). For the earlier time period (1850 to 1978), the extinction coefficients from the Volcanic Forcing Dataset were integrated from the tropopause to 39.5 km, to obtain a globally averaged SAOD, weighted by the cosine of latitude from 80°S to 80°N. For the latter time period (1979 to 2018), we calculate globally averaged SAOD from the GloSSAC dataset using cosine-latitude weighting over the same range of latitudes. For the year 2019, level 3 gridded SAOD product from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) (Vaughan et al., 2004) is used to obtain a global average SAOD, which is then offset by the average difference between the GloSSAC and CALIPSO datasets for the period of overlap (2006–2018) between the two datasets, as described in Section 2.2.5 of McBride et al. (2021). The contribution of TSI and SAOD to the GMST anomaly for the best-fit runs of the Baseline and AR6 frameworks are shown in Fig. 2e-f. Values of TSI beyond 2019, which marks the end of the training period, are set to zero, while for SAOD, we use the value from December 2019 for the 2020 to 2100 period.

———— The ENSO time series used during the training period (1850–2019) is based on Version 2 of the Multivariate ENSO Index (MEI.v2) (Wolter and Timlin, 1993; Zhang et al., 2019). The MEI.v2 dataset provides data starting in 1979. For 1850 to 1978, a historical extension based on Wolter and Timlin (2011) and the HadSST3 dataset (Kennedy et al., 2011) is used, as detailed in Section 2.2.6 of McBride et al. (2021). Input values for ENSO are also set to zero beyond 2019. The contribution of ENSO to the GMST anomaly for the best-fit runs are shown in Fig. 2g-h.

———— PDO input data are obtained from the University of Washington PDO index for 1900–2018, with values of zero used outside this time period. IOD input for 1850–2019 was created from the $1^\circ \times 1^\circ$ Sea Surface Temperatures (SSTs) from the

Centennial in situ Observation-Based Estimate (COBE) (Ishii et al., 2005), with values of zero assigned beyond 2019. PDO and IOD were found to have little effect on GMST in EM–GC simulations (McBride et al., 2021), but for completeness, we include these factors. The input AMOC time series is based on SST data from HadSST4 (Kennedy et al., 2019) between the Equator and 60 °N in the Atlantic Ocean, detrended using the magnitude of anthropogenic radiative forcing, then Fourier-filtered to remove frequencies above $1/9 \text{ yr}^{-1}$ following Sect. 3.2.3 and 4.1.2 of Canty et al. (2013) and Sect. 2.2.7 of McBride et al. (2021). Further detail regarding model inputs for ENSO, AMOC, IOD, PDO and SAOD is given in Sections 2.2.5–2.2.7 of McBride et al. (2021). The computed contribution of variations in the strength of AMOC to the GMST anomaly for the best-fit runs is shown in Fig. 2i–j.

2.5.6 Ocean heat content

For both Frameworks in this paper, the average of five ocean heat content datasets (Levitus et al., 2012; Balmaseda et al., 2013; Cheng et al., 2017; Ishii et al., 2017; Carton et al., 2018) is used. The uncertainty for the average OHC record in each month is obtained by computing the 1σ standard deviation between the five datasets for a given month, or by using the uncertainty from Cheng et al. (2017) for that month, whichever is greater. The uncertainties from Cheng et al. (2017) were chosen for this purpose as these are the largest uncertainties of the five datasets. EM–GC simulations normalize modelled OHC (Fig. 2m–n, red line) to 0 in 1850, which is shown along the average observed OHC time series (black line) and the uncertainty series described above (blue markers) on Fig. 2m–n for the upper 700m of the ocean. EM–GC simulations assume that the upper 700m of the ocean holds 70% of the total OHC (McBride et al., 2021). Section 2.2.8 and Fig. S9 of McBride et al. (2021) provide additional detail on the average OHC dataset and the corresponding uncertainty time series.

2.6 Attributable anthropogenic warming rate and effective climate sensitivity

We define Attributable Anthropogenic Warming Rate (AAWR) as the rate of change in GMST due to anthropogenic activity between 1975 and 2014. The period 1975–2014, chosen in accordance with previous work using our model, spans a 40-year period in which GMST rose in a near-linear fashion due to human activity (McBride et al., 2021). AAWR is determined as the slope of a linear fit to the anthropogenic component ($\Delta T_{\text{MDL,anth}}$) of global warming, defined below in Eq. (5), between 1975 and 2014. The time series of $\Delta T_{\text{MDL,anth}}$ for the best fits of the Baseline and AR6 Framework simulations are shown by the orange lines in Fig. 2c–d.

$$\Delta T_{\text{MDL,anth}} = \frac{1+\gamma}{\lambda_p} (\Delta RF_{\text{GHG},t} + \Delta RF_{\text{AER},t} + \Delta RF_{\text{EUC},t} - Q_{\text{OCEAN},t}) \quad (5)$$

Equilibrium Climate Sensitivity (ECS) is defined as the warming after climate equilibrated to a theoretical doubling of the pre-industrial concentration CO_2 (IPCC, 2001, 2021e; Forster et al., 2021). Since equilibrium can take centuries to reach due to the slow heat transfer to deep oceans (Hansen et al., 2011; Church et al., 2013; Tokarska et al., 2020a), often the more short-term Effective Climate Sensitivity (EffCS) (Gregory et al., 2020; Tokarska et al., 2020a; Spencer and Christy, 2023) is used. EM–GC output is used to provide an estimate of EffCS which is based upon the regression to the historical GMST record. We infer EffCS from the ERF due to the doubling of the pre-industrial CO_2 concentration ($\Delta \text{ERF}_{2\times\text{CO}_2}$) as shown in Eq. (6), which is consistent with the methodology described in Box 7.1 of AR6 (Forster et al., 2021).

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

$$EffCS = \frac{1}{\lambda_p - \lambda_E} \times \Delta ERF_{2 \times CO_2} \quad (6)$$

This method of computing EffCS depends on the formulation of ERF. Within the Baseline Framework, we use the RF formula of Myhre et al. (1998), which leads to $\Delta ERF_{2 \times CO_2} = 5.35 \times \ln(2) = 3.71 \text{ Wm}^{-2}$. For the AR6 Framework, we use a $\Delta ERF_{2 \times CO_2}$ value of 3.93 Wm^{-2} based on Section 7.3.2.1 and 7.SM.1.2 of AR6 (Forster et al., 2021; Smith et al., 2021a). Thus, EffCS computed using the AR6 ERF_{CO_2} formula is 6% larger relative to EffCS calculated with the Myhre et al. (1998) formula for a given value of λ_E . Finally, we note that climate sensitivity deduced from historical warming may be different from true ECS, as the historical climate feedback could differ from the climate feedback under an abrupt $4 \times CO_2$ forcing scenario that is often used to evaluate ECS in ESMs (Andrews et al., 2018; Andrews et al., 2019; Winton et al., 2020; Forster et al., 2021).

2.7 Observational constraining and ensemble weighting

The 160,000 member λ_E -ERF_{AER} grid described in Sect. 2.4.1 is constrained by observational data of the GMST anomaly and OHC, using three χ^2 -based metrics (McBride et al., 2021) given by Eqns. S2–S4. Two of these χ^2 metrics quantify how well the modelled GMST anomaly represents the observed temperature anomaly of the atmosphere for the entire training period (1850–2019, χ^2_{ATM}) and the GMST anomaly of the last 80 years (1940–2019, χ^2_{RECENT}). The third χ^2 metric (χ^2_{OCEAN}) is a goodness-of-fit value between the observed and modeled ocean heat content. Numerical values of χ^2_{ATM} and χ^2_{OCEAN} for the best fit run of the Baseline and AR6 Frameworks are given in Fig. 2a–b and 2m–n, respectively. The χ^2_{RECENT} metric is used because without this constraint, some solutions with values of χ^2_{ATM} less than or equal to 2 have a visually poor simulation of the rise in GMST over the past 4 to 5 decades (McBride et al., 2021). Only members of the λ_E -ERF_{AER} grid that yield a good fit to the historical GMST and OHC data are accepted; that is, grid members for which all three χ^2 metrics yield values less than or equal to 2.

For the results shown in Section 3, the accepted grid members that are considered to be good fits to the observational data are then weighted by the assessed likelihood that their specific value of ERF_{AER} in the reference year 2011 (for the Baseline Framework) or 2019 (for the AR6 Framework) is “true”. To derive these weightings, we use an asymmetrical Gaussian function that is centered around the IPCC best estimate of the ERF of aerosols in the reference year (-0.9 Wm^{-2} in 2011 within the Baseline Framework, and -1.1 Wm^{-2} in 2019 for the AR6 Framework). These weighting functions are shown in Fig. 3a–b. The 2σ boundaries of the Gaussians are based on the boundaries of the likely range of ERF_{AER} in the same year provided by the corresponding IPCC reports (see Table S3). The Gaussians are asymmetrical, as the likely range of ERF_{AER} specified in both the AR5 and AR6 reports is not symmetrical around the respective best estimate. The weighted grid is then used to provide probabilistic estimates of AAWR and EffCS (Sect. 3.1), as well as probabilistic forecasts on the GMST anomaly (Sect. 3.2), which serve as the primary output of EM-GC.

Formatted: Space Before: 0 pt, After: 0 pt, Tab stops: Not at 0.25"

Formatted: Tab stops: Not at 0.25"

3. Results

3.1 Attributable Anthropogenic Warming Rate and Effective Climate Sensitivity

We begin by analyzing values of AAWR found from the EM-GC ensemble simulation of ΔT over the 1850 to 2019 time period. As noted above, our estimates of AAWR quantify the human contribution to the rate of global warming from 1975 to 2014. Figure 1 shows the values of AAWR (panels c–d) and EffCS (panels e–f), as the function of climate feedback (vertical axis) and the strength of aerosol cooling (horizontal axis) for the Baseline (left column) and AR6 frameworks (right column). Colors corresponding to values of AAWR and EffCS are only shown for combinations of λ_S and ERF_{AER} for which a good fit to the historical GMST and OHC record was obtained, defined by the values of all three reduced χ^2 indicators being ≤ 2 . Panels a–b show the asymmetrical Gaussian functions used to weight the EM-GC output, described in Sect. 2.2.2. Notably, the highest values of λ_S for which the model can achieve a good fit to the historical climate record is lower for the AR6 framework relative to Baseline, which allows for a tighter constraint on the upper limit of EffCS. This difference is driven by the considerable variation in the shape of the best estimate of the ERF_{AER} associated with each framework (Fig. S1g), that drives considerable variations in the value of total anthropogenic ERF between about 1960 and 2000 (Fig. S1i).

The weighted median estimate and 5–95% range for AAWR are 0.16 [0.12 to 0.20] $^{\circ}\text{C decade}^{-1}$ and 0.18 [0.13 to 0.21] $^{\circ}\text{C decade}^{-1}$ for the Baseline and AR6 frameworks, respectively. The probability distribution function (PDF) of AAWR, for both frameworks, is shown in Fig. S3a. Our estimates of AAWR are quite similar for the two frameworks, which is of course to be expected since both estimates result from observations of the GMST anomaly that are identical for both sets of ensembles. The slightly higher median value of AAWR in the AR6 framework is due to a larger slope of ERF_{ANTH} over 1975 to 2014, 0.47 W m^{-2} per decade, compared to the 0.36 W m^{-2} per decade slope of ERF_{ANTH} in the Baseline framework. Our values of AAWR within both frameworks are considerably lower than the median and 5–95% range of 0.221 [0.151 to 0.299] $^{\circ}\text{C decade}^{-1}$ from the CMIP6 multi-model ensemble derived by McBride et al. (2021). This finding is consistent with Samset et al. (2023), who found that “virtually all CMIP6 simulations have higher 50-year warming rates than the observations”, using an ensemble of 119 ESM simulations from CMIP6. Our empirical estimates of AAWR align well with several other recent empirical estimates, suggesting that our quantification of the natural and anthropogenic drivers of the variations in ΔT is consistent with other studies. Recent estimates of AAWR are between 0.17 to 0.20 $^{\circ}\text{C decade}^{-1}$ based upon the 1973–2022 period (Samset et al., 2023) and the 1980–2020 period (Table 2.4 of AR6 Chapter 2 (Gulev et al., 2021; Forster et al., 2023)). Finally, the rate of warming has likely accelerated since 1990 at a rate of 0.008 to 0.025 $^{\circ}\text{C decade}^{-1}$ per decade (Samset et al., 2023), with Forster et al. (2023) finding a rate of 0.2 $^{\circ}\text{C decade}^{-1}$ for human-induced warming for the 2013–2022 period, while Ribes et al. (2021) found this rate to be 0.23 $^{\circ}\text{C decade}^{-1}$ over the 2010–19 period.

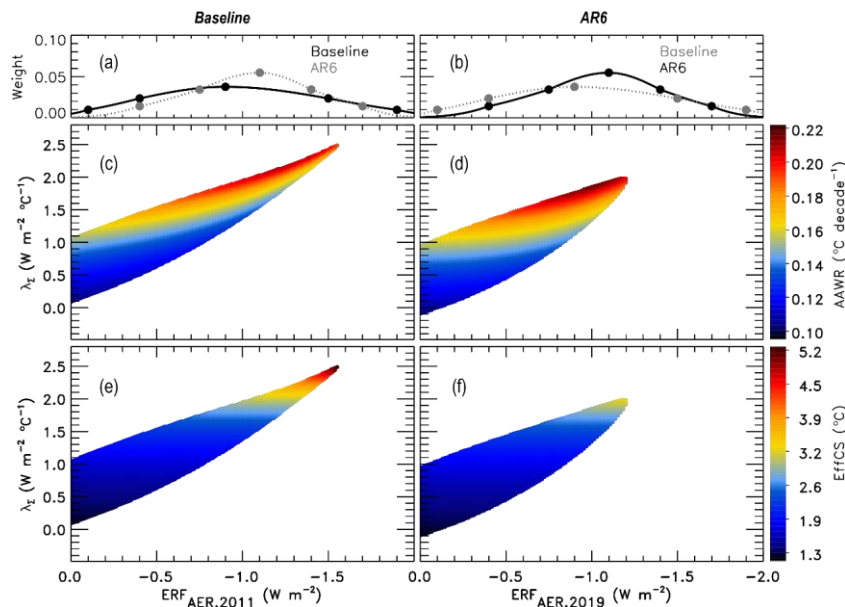


Figure 1: Aerosol weighting method and EM-GC computed values of AAWR and EffCS for combinations of ERF_{AER} and λ_{Σ} . (Left): Simulations using the Baseline framework. (Right): Simulations using the AR6 framework. (a–b) Asymmetrical Gaussians used to weight aerosol scenarios for probabilistic forecasts, as described in Sect. 2.2.2. Points marked on the Gaussians represent specific ERF values used as the central values, as well as 1σ and 2σ boundaries of each Gaussian (Table S3). The Gaussians are overlaid for visual comparison. The Gaussians shown with the solid black line are used to weight the EM-GC output in each column. (c–d) EM-GC computed values of AAWR for the $\lambda_{\Sigma} - ERF_{AER}$ ensemble. Colors denote the specific values of AAWR as indicated by the color bar on the right, and are only shown for the combinations of ERF_{AER} and λ_{Σ} for which a good fit to the HadCRUT5 historical climate record was found. (e–f) EM-GC computed values of EffCS for the $\lambda_{\Sigma} - ERF_{AER}$ ensemble.

Figure 1e and 1f show values of EffCS obtained for the Baseline and AR6 frameworks. The median EffCS is nearly identical for the Baseline and AR6 frameworks, at 2.26 and 2.29 °C, respectively. However, the corresponding 5–95% ranges of EffCS differ considerably, with the Baseline range of [1.45 to 4.37] °C being considerably larger than that found using the AR6 framework [1.54 to 3.11] °C. This difference is readily apparent in the PDFs of EffCS for both frameworks, as shown in Fig. S3b. Simulations based on the AR6 framework allow for a tighter constraining of the upper estimate of EffCS, due to the fact that within this ensemble, we are not able to obtain good fits to the historical GMST and OHC records with values of λ_{Σ} greater than about $2.0 \text{ W m}^{-2} \text{ °C}^{-1}$. In contrast, simulations in the Baseline framework can achieve good fits for stronger levels of climate feedback, upwards to about $2.5 \text{ W m}^{-2} \text{ °C}^{-1}$, that also correspond to higher amounts of aerosol cooling (Fig. 1e–f). Our estimate for EffCS of 2.29 [1.54 to 3.11] °C within the AR6 Framework exhibits close agreement with the results of Skeie et al. (2024), who found the best estimate and 90% uncertainty range for EffCS to be 2.2 [1.6 to 3.0] °C using a Bayesian

estimation model and ERF datasets from AR6, while assuming climate feedback to be constant over the historical period, similar to our approach.

We now discuss the relationship between the estimates of EffCS from our model simulations and ECS. Values of ECS can be approximated from observationally constrained estimates of EffCS by applying a correction factor to the strength of the climate feedback term, termed α' , as shown in Eq. (2). Here, α' represents the difference between climate feedback inferred from historical warming and the climate feedback consistent with an equilibrated climate following an abrupt doubling of the pre-industrial concentrations of CO₂ (hereafter *abrupt2xCO₂*), as described in Sects. 7.4.4.3 and 7.5.2 of Forster et al. (2021). This difference in climate feedback is associated with differences in the spatial pattern of warming over the historical period and the equilibrium warming pattern, termed the pattern effect.

$$ECS = \frac{1}{\lambda_p - \lambda_E - \alpha'} \times \Delta ERF_{2\times CO_2} \tag{2}$$

Section 7.4.4.3 of AR6 assessed α' to have a value of $0.5 \pm 0.5 \text{ W m}^{-2} \text{ }^\circ\text{C}^{-1}$, at a low confidence level. The large uncertainty is due to the fact that estimates for the value of α' vary greatly between individual studies and specific CMIP models (Armour, 2017; Proistosescu and Huybers, 2017; Andrews et al., 2018; Andrews et al., 2019; Dong et al., 2020; Winton et al., 2020). The black symbols in Fig. 2 show ECS found using Eq. (2) as a function of the value of α' for the AR6 ensemble EM-GC simulation. The dots and error bars represent the median and 5–95% range of ECS, respectively. The green dashed line and shaded area correspond to the central value and very likely range of ECS given in Table 7.13 of the AR6 report (Forster et al., 2021).

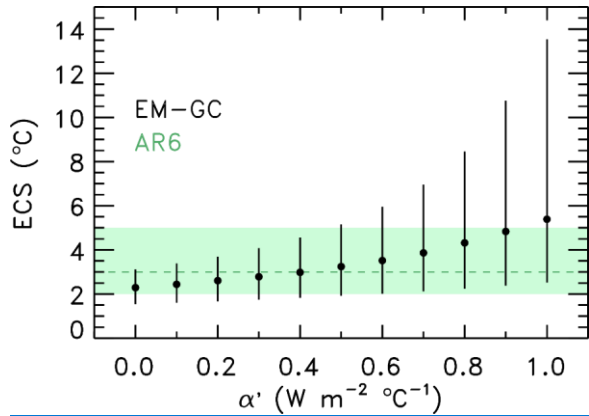


Figure 2: Equilibrium Climate Sensitivity (ECS) as the function of the pattern effect (α' , see text). Black vertical bars and circles correspond to the EM-GC 5–95% range, and 50% probability, respectively. The green shaded area and horizontal dashed line represent the AR6 very likely range, and central estimate of ECS respectively, from Table 7.13 of Forster et al. (2021). All results shown in this figure are based on simulations that use inputs from the AR6 framework.

For $\alpha' = 0 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ (corresponding to no pattern effect), EffCS is equal to ECS and we obtain the median and range for ECS of 2.29 [1.54 to 3.11] $^{\circ}\text{C}$, which had been given above. For the AR6 best estimate of $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$, we find ECS to be 3.24 [1.92 to 5.15] $^{\circ}\text{C}$, which is in very good agreement with the AR6 assessment of ECS (green shaded region). The AR6-based upper limit of $\alpha' = 1.0 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ yields a median and 5–95% range of 5.39 [2.52 to 13.54] $^{\circ}\text{C}$, which tends to exceed the assessed value of ECS from AR6. These results highlight the sensitivity of ECS to the pattern effect, a concept first introduced in the latest WG1 IPCC report. Our value of ECS found for $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ is consistent with Skeie et al. (2024), who performed a similar analysis and found their estimate of ECS to be “almost identical” to the AR6 central value and very likely range of 3.0 [2.0 $^{\circ}\text{C}$ to 5.0 $^{\circ}\text{C}$], upon using the AR6 best estimate for the value of α' . Finally, our estimates of ECS for $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ also show good consistency with the range for ECS of 2.42 to 5.83 $^{\circ}\text{C}$ obtained from millennial-long CMIP model simulations (Rugenstein et al., 2020) performed under the auspices of the LongRunMIP protocol (Rugenstein et al., 2019).

Figure 2 shows that the 5th percentile values of ECS vary between about 1.5 to 2.5 $^{\circ}\text{C}$ depending on the magnitude of α' , consistent with the AR6 assessment of ECS being greater than 1.5 $^{\circ}\text{C}$ at a virtually certain level of confidence (Forster et al., 2021). Further, the 50th percentile estimates for ECS (black dots) fall into the range of ECS assessed by AR6 (green shading), except for $\alpha' = 1.0 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$. Conversely, our 95th percentile estimates for ECS vary greatly with α' , and exhibit a substantially larger level of variation than the 5th and 50th percentile estimates. These findings are consistent with Chapter 7 of AR6, which stated that “warming over the instrumental record provides robust constraints on the lower end of the ECS range (high confidence), but owing to the possibility of future feedback changes it does not, on its own, constrain the upper end of the range, in contrast to what was reported in AR5” (Forster et al., 2021). Finally, numerous studies have linked strong, positive feedback between clouds and anthropogenic RF as being a causal factor in the tendency of some ESMs to exhibit values of ECS that are considerably higher than our median value of 3.24 $^{\circ}\text{C}$ found for $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ (Gettelman et al., 2019; Zelinka et al., 2020; Wang et al., 2021).

We now discuss how the rate of warming in recent decades relates to EffCS/ECS and the spatial pattern of global warming. Armour et al. (2024) found that the rate of warming between 1981 and 2014 of 0.18 [0.15 to 0.21] $^{\circ}\text{C decade}^{-1}$ inferred from the HadCRUT5 record corresponds to ECS of 2.7 [1.5 to 3.9] $^{\circ}\text{C}$ and EffCS of 2.3 [1.9 to 2.7] $^{\circ}\text{C}$ within a set of CMIP5/6 models. The rate of warming obtained by Armour et al. (2024) is in close agreement with our AAWR estimate of 0.18 [0.13 to 0.21] $^{\circ}\text{C decade}^{-1}$ between 1974 and 2014. Furthermore, our EffCS estimate of 2.29 [1.54 to 3.11] $^{\circ}\text{C}$ closely matches that of Armour et al. (2024), albeit with a wider 5–95% range. Finally, our ECS estimate of 3.24 [1.92 to 5.15] $^{\circ}\text{C}$ for $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ is broadly consistent with the Armour et al. (2024) value. The ECS in Armour et al. (2024) was computed from the first 150 years of the CMIP model simulation, which on average, is about 17% lower than ECS in full equilibrium (Rugenstein et al., 2020); a 17% increase to the ECS values of Armour et al. (2024) brings their estimates closer to our values of ECS for $\alpha' = 0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$. Consequently, our estimates of EffCS and ECS are consistent with values obtained from CMIP models that accurately capture the observed rate of rise in GMST over the recent decades.

3 Results

3.1 Ensemble-based assessment of AAWR and climate sensitivity

In Sect. 2.4, we briefly analyzed the AAWR obtained from EM–GC best fits to the GMST record, assuming the AR5 and AR6 prescribed radiative forcing for aerosols, for the Baseline and AR6 Frameworks, respectively. We found a significant increase in the value of AAWR from 0.167 °C/decade within the Baseline Framework, to 0.202 °C/decade for the AR6 Framework. In this section, we evaluate AAWR and EffCS for the entire 160,000 member EM–GC grid, which accounts for the uncertainty in the magnitude of climate feedback and ERF of aerosols (Sect. 2.4.1).

Figure 3 shows AAWR (Fig. 3e–d) and EffCS (Fig. 3e–f) for the EM–GC λ_{Σ} –ERF_{AER,t} grid, within the Baseline (left) and AR6 Frameworks (right). Colors denote the values of AAWR and EffCS found for simulations that satisfy all three χ^2 metrics, as described in Sect. 2.7. Fig. 3a–b shows the Gaussians used to convert the λ_{Σ} –ERF_{AER,t} grid output into a probabilistic quantification of AAWR and EffCS, which are presented in the form of Probability Distribution Functions (PDFs) in Fig. 4. The height of bars on Fig. 4 corresponds to the probability of AAWR (Fig. 4a) and EffCS (Fig. 4b) obtained from the Gaussian weighting of the EM–GC grid—being in the range defined by the width of each column. We use a bin size of 0.005 °C decade^{−1} and 0.1 °C in Fig. 4 for AAWR and EffCS, respectively. Figure 4 also shows the EM–GC median for AAWR and EffCS, which we define as the values of AAWR and EffCS corresponding to 50% probability, alongside the 5th to 95th percentile range for AAWR and EffCS.

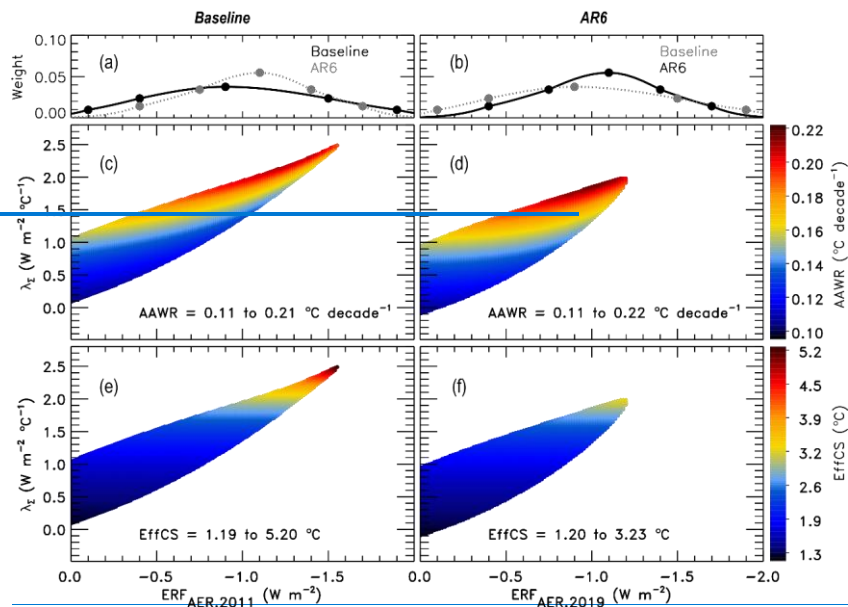


Figure 3: Aerosol weighting method and EM-GC computed values of AAWR and EffCS for combinations of ERF_{AER} and λ_{Σ} . (Left): Simulations using the Baseline Framework. (Right): Simulations using the AR6 Framework. (a–b) Asymmetrical Gaussians used to weight aerosol scenarios for probabilistic forecasts, as described in Sect. 2.7. Points marked on the Gaussians represent specific ERF values used as the central values, as well as 1 σ and 2 σ boundaries of each Gaussian (Table S3). The Gaussians are overlaid for visual comparison. The Gaussians shown with the solid black line are used to weight the EM-GC output in each column. (c–d) EM-GC computed values of AAWR for the λ_{Σ} — ERF_{AER} grid. Colors denote the specific values of AAWR as indicated by the color bar on the right, and are only shown for the combinations of ERF_{AER} and λ_{Σ} for which a good fit to the HadCRUT5 historical climate record was found. (e–f) EM-GC computed values of EffCS for the λ_{Σ} — ERF_{AER} grid.

Within the Baseline Framework (Fig. 3c), the historical GMST record can be fit with combinations of high λ_{Σ} and strong aerosol cooling. For example, good fits to the historical GMST record were found with aerosol cooling as negative as -1.6 W m^{-2} in 2011, which corresponds to much stronger aerosol cooling than the AR5 best estimate of -0.9 W m^{-2} within the same year. Conversely, within the AR6 Framework, the strongest aerosol cooling for which a good fit was found (about -1.2 W m^{-2} in 2019) is very close to the AR6 best estimate of -1.1 W m^{-2} in 2019. Thus, within the AR6 Framework, it is not possible to fit the historical GMST record with aerosol cooling much stronger than the corresponding IPCC best estimate. Within the AR6 Framework, the $ERF_{AER,2019}$ range for which a good fit to the historical GMST record was found is 0.0 to -1.2 W m^{-2} , which is broadly consistent with the AR6 likely range for $ERF_{AER,2019}$ of -0.4 to -1.7 W m^{-2} (Forster et al., 2021). While good fits to the historical GMST record were found with weak aerosol cooling of 0 to -0.4 W m^{-2} , these simulations are given very low weights (Fig. 3b) because the value of ERF_{AER} lies outside of the AR6-assessed likely range. These weak aerosol cooling

simulations are also associated with values of climate feedback ($\lambda_{\text{a}}^{\text{aer}}$) that fall outside the “likely” range of total feedback derived from Table 7.10 of AR6 (see Supplement for details). Finally, even though EM–GC simulations can obtain good fits to the HadCRUT5 record with ERF_{AER} being greater than 0 W m^{-2} and with $\lambda_{\text{a}}^{\text{aer}}$ being negative, these simulations are discarded (and not shown in Fig. 3) because Section 7.3.3.4 of AR6 states that it is “virtually certain” that total aerosol ERF is negative. The ERF_{AER} cutoff in Fig. 3e–f is thus set at 0 W m^{-2} .

The range of AAWR found using the entire EM–GC grid is near identical between the Baseline and AR6 Frameworks at 0.11 to $0.21 \text{ }^{\circ}\text{C decade}^{-1}$ and 0.11 to $0.22 \text{ }^{\circ}\text{C decade}^{-1}$, respectively (Fig. 3e–d). However, the median AAWR increases from 0.16 to $0.18 \text{ }^{\circ}\text{C decade}^{-1}$ from the Baseline to the AR6 Framework (Fig. 4a). Similarly to the range of AAWR derived from the entire $\lambda_{\text{a}}^{\text{aer}}$ – ERF_{AER} grid, the 5^{th} to 95^{th} percentile ranges are found to be consistent between the Baseline and AR6 Frameworks, at 0.12 to $0.20 \text{ }^{\circ}\text{C decade}^{-1}$ and 0.13 to $0.21 \text{ }^{\circ}\text{C decade}^{-1}$, respectively. As described above, AAWR found for the best fit to the GMST record using the IPCC prescribed best estimate for the RF of aerosols is $0.167 \text{ }^{\circ}\text{C decade}^{-1}$ and $0.202 \text{ }^{\circ}\text{C decade}^{-1}$ for the Baseline and AR6 Frameworks, respectively (Fig. 2e–d). For the Baseline Framework, the best-fit AAWR of $0.167 \text{ }^{\circ}\text{C decade}^{-1}$ found using the IPCC best estimate aerosol cooling is close to the EM–GC median of $0.16 \text{ }^{\circ}\text{C decade}^{-1}$. Conversely, for the AR6 Framework, the median AAWR of $0.18 \text{ }^{\circ}\text{C decade}^{-1}$ is about 10% smaller than the AAWR of $0.202 \text{ }^{\circ}\text{C decade}^{-1}$ shown in Fig. 2d. This difference is caused by the fact that the IPCC AR6 best estimate of -1.1 W m^{-2} for $\text{ERF}_{\text{AER},2010}$ is very close to the upper end of the range of ERF_{AER} for which a good fit was found (Fig. 3d), whereas for the Baseline Framework, the AR5 best estimate for aerosol cooling of -0.9 W m^{-2} is approximately in the middle of the range of ERF_{AER} for which good fits to the historical GMST record are found.

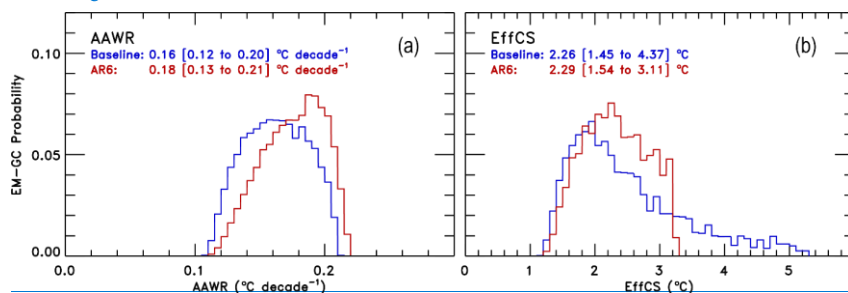


Figure 4: (a) Probability Distribution Functions (PDFs) for AAWR obtained from EM–GC simulations trained with the HadCRUT5 temperature dataset. AAWR from the Baseline Framework is shown in blue, while AAWR from the AR6 Framework is shown in red. The EM–GC median (defined as the 50^{th} percentile probability), and 5^{th} to 95^{th} percentile range is displayed at the top. (b) PDFs of EffCS within the Baseline (blue) and AR6 Framework (red), with the corresponding median and range.

The EM–GC computed median and 5 – 95% range of AAWR are 0.18 [0.13 to $0.21 \text{ }^{\circ}\text{C decade}^{-1}$] between 1974 and 2014 for the AR6 Framework. This range and median are consistent with several other studies, such as Samset et al. (2023), who found a rate of increase in warming to be $0.19 \pm 0.01 \text{ }^{\circ}\text{C decade}^{-1}$ between 1973 and 2022 for the HadCRUT5 GMST record. Samset et al. (2023) also found rates for other temperature records such as GISTEMP v4, NOAA v5.1 and Berkeley Earth to

be 0.19, 0.18 and 0.17 °C decade⁻¹, respectively. Our HadCRUT5-based median of 0.18 °C decade⁻¹ is consistent with the analysis of Samset *et al.* (2023) who employed a variety of historical temperature records, as well as the assessment of 0.19 °C decade⁻¹ between 1980–2020 described in Table 2.4 of AR6 Chapter 2 (Gulev *et al.*, 2021; Forster *et al.*, 2023). As highlighted by Samset *et al.* (2023) however, warming has accelerated since 1990 at a rate of 0.008 to 0.025 °C decade⁻¹ per decade depending on the temperature record studied, with Forster *et al.* (2023) finding a rate of 0.2 °C decade⁻¹ for human-induced warming for the 2013–2022 period, while Ribes *et al.* (2021) found this rate to be 0.23 °C decade⁻¹ for the 2010–19 period.

—Our median and range of AAWR of 0.18 [0.13 to 0.21 °C decade⁻¹] within the AR6 Framework is considerably lower than the median and 5–95% range of 0.221 [0.151 to 0.299 °C decade⁻¹] derived from a CMIP6 multi-model ensemble by McBride *et al.* (2021). The CMIP6 median obtained by McBride *et al.* (2021) is greater than the upper limit of the range found in our study. Samset *et al.* (2023) also found that “virtually all CMIP6 simulations have higher 50-year warming rates than the observations”, using an ensemble of 119 ESM simulations from CMIP6. In particular, Samset *et al.* (2023) found that models with ECS being greater than 3.0 °C tend to quantify the rate of warming to be above 0.2 °C decade⁻¹ up to about 0.43 °C decade⁻¹ for certain models with ECS around 5 °C (Fig. 2d of Samset *et al.* (2023)). Armour *et al.* (2024) showed that CMIP5/6 models that tended to over-estimate the observed warming trend over the 1981 to 2014 time period tended to have larger values of EffCS, which they attribute to gross differences in the observed versus modeled warming patterns, particularly in the tropical Pacific Ocean. Overall, AAWR found using our AR6 Framework is consistent with other lines of empirical evidence in the recent literature, and tends to be quite a bit lower than the CMIP6 multi-model mean.

—EffCS found using the three χ^2 -based metrics has a range of 1.20 to 3.23 °C in the AR6 Framework relative to a range of 1.19 to 5.20 °C for the Baseline Framework (Fig. 3e–f). The lower limit is found to be near identical between the two Frameworks; however, the upper limit of EffCS is substantially smaller within the AR6 Framework. This difference originates from the fact that the HadCRUT5 GMST record can be fit with combinations of high climate feedback and strong aerosol cooling only for the Baseline Framework, as described above. Fits with very strong aerosol cooling correspond to high λ_E and yield high values of EffCS (Eq. (6)) in the Baseline framework, resulting in a high upper bound for the range of EffCS.

—The median EffCS is also found to be near identical between the Baseline and AR6 Frameworks at 2.26 and 2.29 °C, respectively, with corresponding 5–95% ranges being [1.45 to 4.37 °C] and [1.54 to 3.11 °C] (Fig. 4b). The 95th percentile value of EffCS is considerably smaller within the AR6 Framework relative to the Baseline, due to the aforementioned phenomenon of not being able to obtain good fits to the GMST anomaly in the AR6 Framework for aerosol scenarios with ERF_{AER,2019} being more negative than about -1.2 W m^{-2} . Given that in the AR6 Framework the Gaussian used to weight the λ_E —ERF_{AER,f} grid is centered around the IPCC best estimate of -1.1 W m^{-2} , the few combinations of λ_E —ERF_{AER,2019} that yield good fits for which ERF_{AER,2019} is between -1.1 and -1.2 W m^{-2} are still assigned high weights. Within the Baseline Framework, good fits with strong aerosol cooling scenarios (as negative as -1.6 W m^{-2} in 2011) are far from the AR5 best estimate of -0.9 W m^{-2} cooling and are therefore assigned small weights (Fig. 3a). The high cooling scenarios correspond to the low probability tail of the PDF with EffCS greater than about 3.5 °C shown in Fig. 4b for the Baseline Framework. The

range of EffCS found for the AR6 Framework is considerably narrower than for the Baseline, which is well represented by the difference in the 95th percentiles for EffCS for the two frameworks of 3.11 °C and 4.37 °C, respectively (Fig. 4b).

— A recent study by Skeie *et al.* (2024b) found the best estimate and 90% uncertainty range for EffCS to be 2.2 [1.6 to 3.0 °C] using a Bayesian estimation model and ERF datasets from AR6. Our estimate of 2.29 [1.54 to 3.11 °C] within the AR6 Framework exhibits close agreement with this study. Furthermore, Skeie *et al.* (2024b) assume climate feedback to be constant over the historical period, similar to our work. However, climate feedbacks may not be constant over time, partly due to their dependence on the spatial pattern of warming, often termed “pattern effect” (Andrews *et al.*, 2015; Armour, 2017; Stevens *et al.*, 2016; Skeie *et al.*, 2024b). Since climate sensitivity is derived from the magnitude of climate feedback (i.e. (Forster *et al.*, 2021)), the magnitude of the pattern effect governs how much climate sensitivity inferred from historical records (EffCS) differs from ECS. The sensitivity of GMST projections found using our model to time-dependent feedback is quantified in Section 3.3.6 of McBride *et al.* (2021).

— Next, we quantify the effect of time-dependent feedback on ECS. Similar to Skeie *et al.* (2024b), we apply the AR6 formulation for the impact of the pattern effect on the ensemble output of the AR6 Framework. The impact of time-dependent feedback on ECS is quantified as in Eq. (7), where α' represents the difference between climate feedback inferred from historical warming and the climate feedback corresponding to that found for an abrupt doubling of the concentration of CO₂.

$$ECS = \frac{1}{\lambda_p - \lambda_{2\alpha} - \alpha'} \times \Delta ERF_{2 \times CO_2} \quad (7)$$

The formulation shown in Eq. (7) is equivalent with the methodology for the computation of ECS with the inclusion of α' described in Section 7.5.2 of AR6 (Forster *et al.*, 2021). The coefficient α' is assessed by Section 7.4.4.3 of AR6 to have a value of 0.5 ± 0.5 W m⁻² K⁻¹, at a low confidence level. Figure 5 shows ECS for values of α' within the AR6 assessed range of 0 to 1 W m⁻² K⁻¹, in increments of 0.1 W m⁻² K⁻¹ (black circles, median), where the error bars represent the 5th and 95th probabilistic range of ECS. The dashed horizontal line and shaded area correspond to the central value and very likely range of ECS from Table 7.13 of AR6 (Forster *et al.*, 2021).

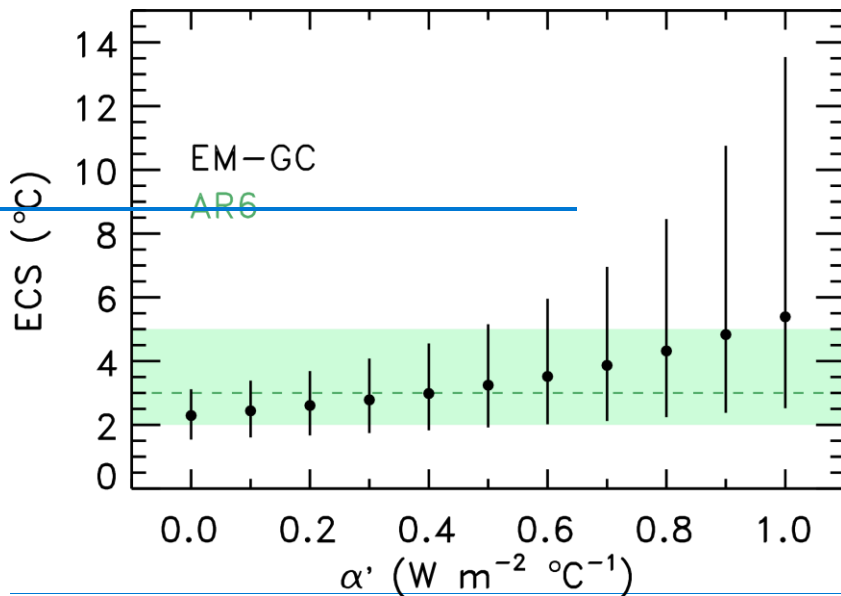


Figure 5: Equilibrium Climate Sensitivity (ECS) as the function of the pattern effect (α' , see text). Black vertical bars and circles correspond to the EM-GC 5–95% range, and 50% probability, respectively. The green shaded area and horizontal dashed line represent the AR6 very likely range, and central estimate of ECS respectively, from Table 7.13 of Forster et al. (2021). All results shown in this figure are based on inputs from the AR6-ERF Framework.

For $\alpha' = 0 \text{ W m}^{-2} \text{ K}^{-1}$ (corresponding to no pattern effect), EffCS is equal to ECS and we obtain a median and range for ECS of 2.29 [1.54 to 3.11 $^\circ\text{C}$]. Using the AR6 best estimate of $\alpha' = 0.5 \text{ W m}^{-2} \text{ K}^{-1}$, we find ECS to be 3.24 [1.92 to 5.15 $^\circ\text{C}$], while the AR6 upper limit of $\alpha' = 1.0 \text{ W m}^{-2} \text{ K}^{-1}$ yields a median and 5–95% range of 5.39 [2.52 to 13.54 $^\circ\text{C}$]. Our estimate of ECS agrees extremely well with the AR6 central estimate and likely range of ECS, for the AR6 best estimate of α' , which equals $0.5 \text{ W m}^{-2} \text{ K}^{-1}$. This conclusion is consistent with Skeie *et al.* (2024b), who also found their estimate of climate sensitivity to be “almost identical” to the AR6 central value and very likely range of 3.0 [2.0 $^\circ\text{C}$ to 5.0 $^\circ\text{C}$], upon using the AR6 best estimate of α' . Our range of ECS of [1.92 to 5.15 $^\circ\text{C}$] at $\alpha' = 0.5 \text{ W m}^{-2} \text{ K}^{-1}$ is also consistent with the recent assessment of Cooper et al. (2024), who found ECS to be between 1.4 $^\circ\text{C}$ to 5.0 $^\circ\text{C}$ based on analysis of data acquired during the Last Glacial Maximum (LGM). Our median estimate of 3.24 $^\circ\text{C}$ for $\alpha' = 0.5 \text{ W m}^{-2} \text{ K}^{-1}$ is, however, higher than the LGM-based best estimate of 2.4 $^\circ\text{C}$. However, when combining the LGM-based assessment with other lines of evidence, Cooper et al. (2024) finds ECS to be 2.9 $^\circ\text{C}$ [2.1 to 4.1 $^\circ\text{C}$], which agrees well with our estimate.

Additionally, Fig. 5 shows that the 5th percentile of ECS varies between about 1.5 to 2.5 $^\circ\text{C}$, depending on the magnitude of α' , consistent with the AR6 assessment of ECS being greater than 1.5 $^\circ\text{C}$ at a virtually certain level of confidence (Forster

et al., 2021). The 50th percentile estimates for ECS also all fall into the range of ECS assessed by AR6. Conversely, our 95th percentile estimates for ECS vary greatly with α' , and carry a substantially larger level of uncertainty than the 5th and 50th percentile estimates. These findings are consistent with Chapter 7 of AR6, which stated “warming over the instrumental record provides robust constraints on the lower end of the ECS range (high confidence), but owing to the possibility of future feedback changes it does not, on its own, constrain the upper end of the range, in contrast to what was reported in AR5.”

—We now discuss how the rate of warming in recent decades relates to EffCS/ECS and the spatial pattern of global warming. Armour *et al.* (2024) found the central estimate and 5–95% range for the rate of warming between 1981 and 2014 to be 0.18 [0.15 to 0.21 °C decade⁻¹], which corresponds to ECS of 2.7 [1.5 to 3.9 °C] and EffCS of 2.3 [1.9 to 2.7 °C]. The rate of warming obtained by Armour *et al.* (2024) is in close agreement with our AAWR estimate of 0.18 [0.13 to 0.21 °C decade⁻¹] between 1974 and 2014. Furthermore, our EffCS estimate of 2.29 [1.54 to 3.11 °C] closely matches that of Armour *et al.* (2024), albeit with a wider 5–95% range. Finally, our ECS estimate was found to be 3.24 [1.92 to 5.15 °C] for $\alpha' = 0.5 \text{ W m}^{-2} \text{ K}^{-1}$, which is broadly consistent with Armour *et al.* (2024) value. Samset *et al.* (2023) found that CMIP6 models with ECS greater than about 3.0 °C tend to provide an overly high estimate of the trend of observed warming in recent decades. Armour *et al.* (2024) suggest the overly high estimates of recent warming trends are related to CMIP5/6 historical simulations failing to reproduce observed SST patterns, particularly a cooling of the eastern tropical Pacific and a warming of the western Pacific. Weaver *et al.* (2024) reached the same conclusion based on analysis of top of the atmosphere albedo of clouds and aerosols from 340 nm radiances observed by NASA and NOAA satellite instruments. Armour *et al.* (2024) found that CMIP5/6 models with high ECS can nonetheless provide estimates of warming trends consistent with observations, after correcting for the biases in SST trend patterns. Weaver *et al.* (2024) showed that a significant number of CMIP6 models are unable to simulate the observed latitudinal pattern of cloud albedo trends over the Pacific Ocean, again tied to the observed cooling trend in the eastern tropical Pacific.

—Our multiple linear regression—energy balance RCM, despite not considering the spatial pattern of warming, is able to reproduce estimates of EffCS and ECS consistent with the Armour *et al.* (2024) study, who explicitly consider the spatial pattern of global warming. We find that the range of observed warming in recent decades is consistent with the AR6 estimate of ECS and that the CMIP5/6 models that estimate a more rapid rate of warming may indeed fall into the “hot model” category (Tokarska *et al.*, 2020b; McBride *et al.*, 2021; Hausfather *et al.*, 2022), potentially due to an overestimation of historical EffCS (Armour *et al.*, 2024).

3.2 Probabilistic forecast on future warming

Here we quantify the magnitude of future warming based on projections of ERF from four SSP scenarios, for both the Baseline and AR6 frameworks. We first briefly discuss the simulated GMST anomalies at the end of the 21st century, followed by the analysis of the projected temporal evolution of warming in this century. Finally, we quantify the likelihood of accomplishing the goals of the Paris Agreement under the four SSP scenarios, based on the simulations within both the Baseline and AR6 frameworks. Unless otherwise stated, all GMST anomalies are relative to an 1850–1900 pre-industrial baseline.

1075

1080

1085

1090

1095

Figures S4 and S5 show the GMST anomaly in year 2100 (ΔT_{2100}) as a function of climate feedback and ERF_{AER} , in a manner similar to Fig. 1, for the Baseline and AR6 frameworks, respectively. Probabilistic projections of ΔT_{2100} are obtained using the same weighting technique that was used for AAWR and EffCS, as described in Section 2.2.2. Table 1 provides median as well as 5th and 95th percentile values of ΔT_{2100} , for the four SSP scenarios considered throughout. Median projections of ΔT_{2100} within the AR6 framework are about 0.2 °C (SSP1–1.9, SSP1–2.6), 0.3 °C (SSP4–3.4), and 0.4 °C (SSP2–4.5) greater than found using the Baseline framework. This difference originates from the fact that projected ERF at the end of the century is higher in the AR6 framework than in Baseline, for all four SSPs, which is driven by higher end-of-century atmospheric concentrations of CO₂ and CH₄ in AR6 (Fig. S1i).

Table 1: Probabilistic projections of ΔT_{2100} for the four SSP scenarios studied.

	SSP Scenario	Baseline Framework (°C)	AR6 Framework (°C)
ΔT_{2100} Median [5–95% range]	SSP1–1.9	1.14 [0.81 to 1.87]	1.34 [0.93 to 1.72]
	SSP1–2.6	1.46 [1.05 to 2.28]	1.67 [1.18 to 2.13]
	SSP4–3.4	1.80 [1.31 to 2.77]	2.10 [1.54 to 2.62]
	SSP2–4.5	2.18 [1.62 to 3.14]	2.60 [1.92 to 3.20]

Next, we examine the temporal evolution of ΔT under the four SSP scenarios. Figure 3 shows time-dependent, probabilistic forecasts of ΔT for the AR6 framework. Time-dependent projections of ΔT for the Baseline framework are shown in Fig. S6. The colors in Fig. 3 correspond to the probability that ΔT would be equal to or greater than the given numerical value, as indicated by the color bar. Each panel of Fig. 3 also shows the multi-model mean (solid line), and minimum and maximum (dashed lines) for ΔT obtained from the CMIP6 ESM archive, normalized to zero over 1850 to 1900. These CMIP6 values of ΔT are based on our analysis of output from 10, 34, 6, and 32 models, for SSP1–1.9, SSP1–2.6, SSP4–3.4, and SSP2–4.5, as described in Section 3.3.1 of McBride et al. (2021). Figure 3 also shows the HadCRUT5 GMST observations in black, as well as a green trapezoid that represents the likely range of warming between 2016 and 2035 provided by Chapter 11 of the AR5 report (Kirtman et al., 2013). This trapezoid was placed on Fig. 11.25b of AR5 due to the recognition, by the chapter authors, that the CMIP5 models central to AR5 tended to overestimate the observed rate of global warming. Gold horizontal lines in Fig. 3 represent the 1.5 °C and 2.0 °C GMST anomalies relative to pre-industrial, while gold circles correspond to the years where the 1.5 °C and 2 °C thresholds are crossed with 5, 50 and 95% probability (hereafter termed crossover years). Table 2 provides the years at which the temperature anomaly thresholds are projected to be crossed for the median as well as the 5th and 95th percentiles.

Our probabilistic projections of ΔT are in excellent agreement with the assessed likely range of global warming provided by Chapter 11 of AR5 (Kirtman et al., 2013). Our projections of ΔT fall within the bottom half of those obtained from the CMIP6 ESMs. Numerous studies have similarly concluded that many of the ESMs central to CMIP6 tend to provide estimates

of the rate of global warming due to human activity (that is, AAWR) that exceeds empirically based estimates of AAWR (Tokarska et al., 2020b; Nijssen et al., 2020; McBride et al., 2021; Chylek et al., 2024), which Hausfather et al. (2022) have termed the “hot model problem”. Armour et al. (2024) suggested that the high estimates of AAWR exhibited by some of the CMIP6 historical simulations are due to the inability of these models to reproduce observed SST patterns, particularly an observed cooling of the eastern tropical Pacific and a warming of the western Pacific that affects the distribution of clouds in the tropics. Weaver et al. (2024) reached the same conclusion based on an analysis of top of the atmosphere albedo of clouds and aerosols, from radiances observed by NASA and NOAA satellite instruments.

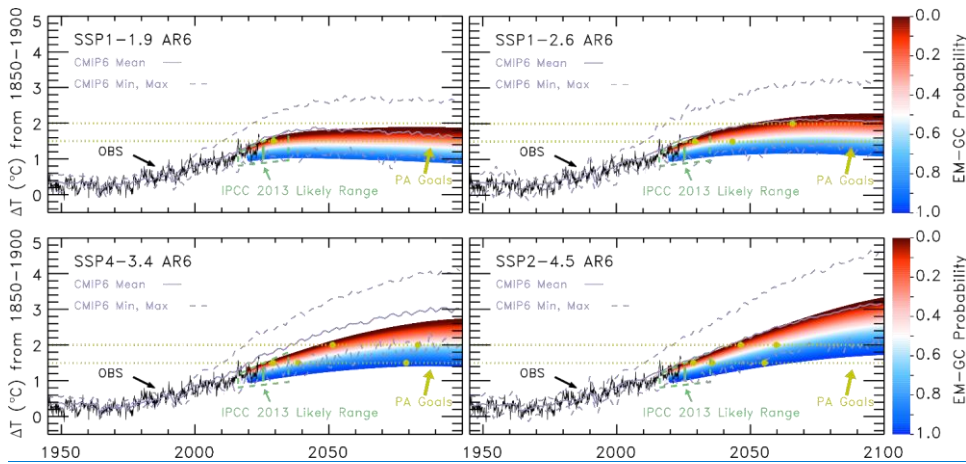


Figure 3: Time-dependent probabilistic forecasts of the GMST anomaly within the AR6 framework. The black line shows the HadCRUT5 GMST anomaly, which was used (along with OHC) to constrain the simulations. Colors represent the probability of reaching a certain temperature or higher at a given time, as indicated by the color bars on the right. The green trapezoid represents the likely range of warming as shown in Fig. 11.25b of the IPCC AR5 report (Kirtman et al., 2013). The target and upper limit of the Paris Agreement are shown by gold-colored horizontal dotted lines. Circle markers on these lines correspond to the projected GMST anomaly crossing these thresholds, with the probability indicated by the colors. The grey lines denote the multi-model mean (solid), as well as the minimum and maximum (dashed) projections of ΔT from CMIP6 ESMs as described in Sect. 3.3.1 of McBride et al. (2021). All values of ΔT shown in this figure are with respect to an 1850–1900 pre-industrial baseline. (a) GMST projections for SSP1–1.9, (b) GMST projections for SSP1–2.6, (c) GMST projections for SSP4–3.4, (d) GMST projections for SSP2–4.5. Results for the Baseline Framework are shown in the same fashion in Fig. S6.

For all four SSP scenarios, the 1.5 °C and 2.0 °C thresholds are projected to be crossed much earlier based on the simulations of the AR6 framework, relative to those of the Baseline. For example, for SSP2–4.5, the 2.0 °C threshold is projected to be crossed in the years 2059 and 2082 within the AR6 and Baseline frameworks, respectively. The AR6 updates to the ERF from GHGs and aerosols result in nearly a quarter of a century shift forward in crossover year. This finding is consistent with the increases in end-of-century warming (ΔT_{2100}) obtained from the simulations using the AR6 framework.

relative to those of Baseline (Table 1). Comparing to the literature, our crossover years found using the AR6 framework fall on the latter end of the projected crossover years for the 1.5 °C and 2.0 °C thresholds given in Table 4.5 of AR6 (Lee et al., 2021), and are much later than projected based on the analysis of CMIP5 and CMIP6 output shown in Table 1 of Tebaldi et al. (2021). The values in their table are based on an analysis of an unconstrained ESM ensemble, which projects higher levels of future warming than observationally constrained models, as evidenced by Table A6 of Tebaldi et al. (2021).

Table 2: Years of crossing the 1.5 °C and 2.0 °C GMST anomaly thresholds for the four SSP scenarios studied. For each entry, we present the 50% probability as our central estimate, as well as the 5–95% range. The label “n.c” is used in a manner similar to Table 4.5 of AR6 (Lee et al., 2021) and corresponds to a given threshold not being crossed in the 2020–2100 period.

	<u>1.5 °C Crossover</u>		<u>2.0 °C Crossover</u>	
	<u>Baseline</u>	<u>AR6</u>	<u>Baseline</u>	<u>AR6</u>
<u>SSP1–1.9</u>	n.c [2028 to n.c]	n.c [2029 to n.c]	n.c [n.c to n.c]	n.c [n.c to n.c]
<u>SSP1–2.6</u>	n.c [2029 to n.c]	2043 [2029 to n.c]	n.c [2059 to n.c]	n.c [2065 to n.c]
<u>SSP4–3.4</u>	2048 [2030 to n.c]	2038 [2029 to 2079]	n.c [2050 to n.c]	2083 [2051 to n.c]
<u>SSP2–4.5</u>	2042 [2030 to 2080]	2035 [2028 to 2055]	2082 [2047 to n.c]	2059 [2046 to n.c]

Figure 4 shows the PDF of ΔT_{2100} found with EM–GC for the four SSP scenarios, using the AR6 and Baseline frameworks. The height of the bars corresponds to the probability of ΔT_{2100} being in the range defined by the width of each column. Figure 4 also shows PDFs derived from a CMIP6 ESM ensemble, as detailed by McBride et al., (2021). As expected, based on the “hot model problem” described above, our projections of ΔT_{2100} within both the Baseline and AR6 frameworks fall on the lower end of the projections from the CMIP6 ensemble. Furthermore, the EM–GC based PDF for the AR6 framework tends to be shifted towards higher values of ΔT_{2100} than found for Baseline, with a smaller tail, behaviors that are consistent with higher end of century RF of the climate within the AR6 framework (Fig. S1i), as well as the ability to fit the climate record with higher values of climate feedback (model parameter λ_3) in the Baseline framework (Fig. 1).

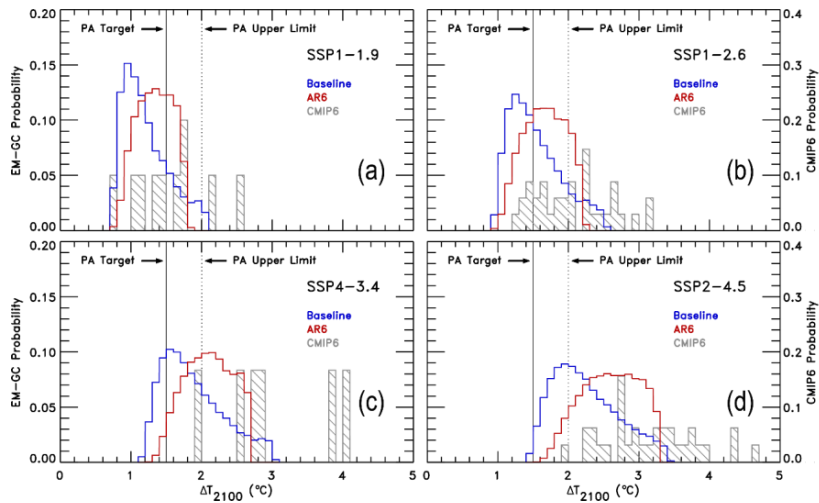


Figure 4: Probability Distribution Functions (PDFs) for ΔT_{2100} obtained from EM-GC simulations trained on the HadCRUT5 temperature dataset. Model runs for the Baseline and AR6 frameworks are shown in blue and red, respectively. Grey color represents the PDFs obtained from a CMIP6 multi-model ensemble as described in Sect. 3.3.1 of McBride et al. (2021), and are shown for comparison with EM-GC results. The left-hand y axis corresponds to the EM-GC probabilities, and the right-hand y axis is for the CMIP6 probabilities. The PA target and upper limit are shown as solid and dashed vertical lines, respectively. (a) PDFs for SSP1-1.9, (b) PDFs for SSP1-2.6, (c) PDFs for SSP4-3.4, (d) PDFs for SSP2-4.5.

We now further compare our projections of ΔT_{2100} from the AR6 framework with results based on CMIP6 model output. Tokarska et al. (2020b) reported that observationally constrained CMIP6 projections of end-of-century warming are 9% to 13% lower than unconstrained CMIP6 projections for SSP1-2.6 and SSP2-4.5, respectively. Tokarska et al. (2020b) found the median and 5–95% ranges of end-of-century warming relative to a 1995–2014 baseline to be 0.94 [0.41 to 1.46] °C and 1.84 [1.15 to 2.52] °C for SSP1-2.6 and SSP2-4.5, respectively, using observationally constrained CMIP6 models. Our values of ΔT_{2100} in Table 1, relative to 1995–2014, are 0.81 [0.32 to 1.27] °C and 1.74 [1.06 to 2.34] °C for SSP1-2.6 and SSP2-4.5. Consequently, our quantification of ΔT_{2100} is in very good agreement with the empirically constrained CMIP6 projections of Tokarska et al. (2020b). Chylek et al. (2024) recently found end-of-century warming to be 2.41 °C relative to pre-industrial conditions, for SSP2-4.5 using a set of CMIP6 models that accurately reproduce the 2014–2023 warming, which is about 0.5 °C smaller than the value of ΔT_{2100} obtained from their unconstrained CMIP6 ensemble. Consequently, their empirically constrained value of ΔT_{2100} is in good agreement with our median estimate of 2.6 °C for SSP2-4.5 (Table 1).

Table 3: EM–GC computed probabilities of achieving the Paris Agreement target (1.5 °C) and upper limit (2.0 °C). Columns with the “Baseline” header represent EM–GC simulations using the Baseline Framework, while “AR6” represents simulations of the AR6 Framework. The values presented in this table are derived from the PDFs shown in Fig. 4.

	Paris 1.5 °C		Paris 2.0 °C	
	Baseline	AR6	Baseline	AR6
SSP1–1.9	81%	70%	98%	100%
SSP1–2.6	54%	32%	87%	85%
SSP4–3.4	21%	3%	64%	40%
SSP2–4.5	1%	0%	35%	8%

We conclude by evaluating the probability of achieving both the target (1.5 °C) and upper limit (2.0 °C) of the Paris Agreement (PA). Table 3 provides the probability that end-of-century warming will be below either the target or the upper limit of the PA, relative to pre-industrial conditions. These estimates were obtained from our probabilistic forecasts of ΔT , for both the Baseline and AR6 frameworks (Fig. 4). As shown in Table 1, median projections of ΔT_{2100} are larger within the AR6 framework relative to the Baseline for all four SSPs, which leads to a decline in the probability of accomplishing the PA within the AR6 framework (Table 3). For the SSP1–1.9 and SSP1–2.6 scenarios, the probability of limiting global warming to 2.0 °C is high (at least 85%) for both model frameworks. For SSP4–3.4, the probability of limiting warming to 2.0 °C falls from 64% (Baseline) to 40% (AR6). Most notably, the 2.0 °C probability drops from 35% to 8% for the SSP2–4.5 scenario. The 1.5 °C warming probabilities for the AR6 framework are all uniformly lower than found for the Baseline framework, with the SSP1–2.6 scenario dropping from 54% (Baseline) to 32% (AR6). The takeaway message from Table 3 is that, for society to have high confidence in achieving at least the upper limit of the PA, the radiative forcing of climate due to GHGs must be placed close to the SSP1–2.6 pathway over the coming decades. More aggressive reductions in GHG radiative forcing are needed to achieve the target of the PA, such as those of the SSP1–1.9 scenario. This message stands in stark contrast to our prior statement in McBride et al. (2021), that SSP4–3.4 would provide about a two-thirds chance of limiting global warming to 2 °C by end-of-century, since our earlier work relied upon the Baseline framework.

Here we quantify the magnitude of future warming based on projections of ERF for four SSP scenarios. Here, we show projections of GMST for only constant feedback because the GMST anomaly can be fit quite well for a wide range of time invariant feedback, albeit spanning a large range of values, particularly for the Baseline Framework (Fig. 3c–d). We compare forecasts of GMST using the AR6 Framework with forecasts found using the Baseline Framework for the same SSPs, where “Baseline” refers to the state of knowledge prior to the release of AR6. Similarly to our analysis of AAWR and EffCS in Section 3.1, forecasts of the GMST anomaly for the λ_{Δ} –ERF_{AER,F} grid, weighted by an asymmetrical Gaussian function (Fig. 3a–b) that describes the AR5-based (Baseline) or AR6-based assessments of the likelihood of how much global warming by GHGs has been offset by tropospheric aerosols, is central to these probabilistic forecasts.

Formatted: Normal

Figure 6 shows computed values of the GMST anomaly in the year 2100 (ΔT_{2100}) for the AR6 Framework, for four SSPs. The colors correspond to values of ΔT_{2100} for model runs that provide a good fit to the observed GMST anomaly over the 1850–2019 training period, as defined by all three χ^2 metrics described in Section 2.7 yielding values less than or equal to 2.

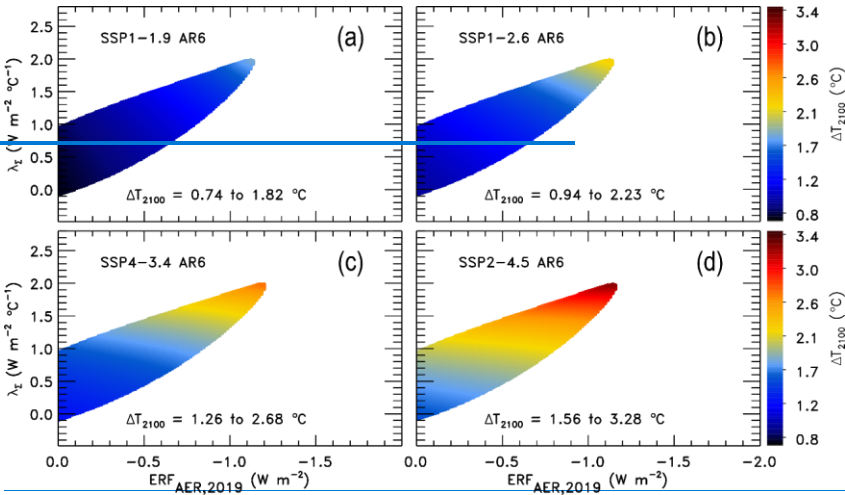


Figure 6: GMST anomaly at the end of the century (ΔT_{2100}) relative to a pre-industrial baseline for combinations of λ_E and $ERF_{AER,2019}$. Colors denote values of ΔT_{2100} as indicated by the color bars on the right. Plots generated from EM-GC simulations trained on the HadCRUT5 GMST record, and use the AR6 ERF Framework. Only those combinations are colored where a good fit on the HadCRUT5 GMST record was found. (a) ΔT_{2100} for SSP1=1.9, (b) ΔT_{2100} for SSP1=2.6, (c) ΔT_{2100} for SSP4=3.4, (d) ΔT_{2100} for SSP2=4.5. Values of ΔT_{2100} for the $\lambda_E = ERF_{AER}$ grid within the Baseline Framework are shown in Fig. S2.

Table 1 details numerical values for ΔT_{2100} projections for the four SSPs. As described in Sect. 3.1, within the Baseline Framework the historical GMST record can be fit with combinations of high λ_E and strong aerosol cooling. This difference results in a small shift in the upper boundary of ΔT_{2100} , with values associated with the Baseline Framework being larger than those for the AR6 Framework. Even though the climate record can be fit with much higher values of λ_E within the Baseline Framework, ERF projections within the AR6 Framework are much higher than in the Baseline Framework (Fig. 1i), which offsets some of the λ_E -driven difference. For SSP2=4.5, the upper estimate of ΔT_{2100} being 3.47 °C corresponds to a value of λ_E of about 2.5 $W m^{-2} °C^{-1}$ (Fig. S2) within the Baseline Framework. The upper boundary of $\Delta T_{2100} = 3.28$ °C is found at a smaller climate feedback of about $\lambda_E = 2.0 W m^{-2} °C^{-1}$ for the AR6 Framework.

Table 1: ΔT_{2100} for the four SSP scenarios studied. The first four rows correspond to data for the entire $\lambda_E = ERF_{AER}$ grid as shown in Figs. 6 and S2. The last four rows are associated with the EM-GC probabilistic assessment, with values derived from PDFs in Fig. 8 shown using the format of Median [5–95% range].

	SSP Scenario	Baseline Framework (°C)	AR6 Framework (°C)
--	--------------	-------------------------	--------------------

Full $\lambda_S = \text{ERF}_{\text{AER}, \text{grid}}$	SSP1-1.9	0.73 to 2.06	0.74 to 1.82
	SSP1-2.6	0.93 to 2.58	0.94 to 2.23
	SSP4-3.4	1.15 to 3.01	1.26 to 2.68
	SSP2-4.5	1.42 to 3.47	1.56 to 3.28
Median [5–95% range]	SSP1-1.9	1.14 [0.81 to 1.87]	1.34 [0.93 to 1.72]
	SSP1-2.6	1.46 [1.05 to 2.28]	1.67 [1.18 to 2.13]
	SSP4-3.4	1.80 [1.31 to 2.77]	2.10 [1.54 to 2.62]
	SSP2-4.5	2.18 [1.62 to 3.14]	2.60 [1.92 to 3.20]

— Figure 7 shows our probabilistic forecasts of GMST in a time dependent fashion for the AR6 Framework. Time-dependent projections of GMST for the Baseline Framework are in Fig. S3. Similar to AAWR and EffCS discussed above, the probabilistic forecasts of the GMST anomaly are provided by weighting model results that provide good fits to the climate record within the $\lambda_S = \text{ERF}_{\text{AER}, \text{grid}}$. The weighting procedure uses the Gaussian functions that describe the assessed likelihood of various values of RF of climate due to aerosols shown on the top panels of Fig. 3. The colors on Fig. 7 correspond to the probability of the ΔT anomaly being equal to or greater than a given value. The figure also shows the HadCRUT5 GMST observations in black, as well as the likely range of warming for 2016 to 2035 provided by the authors of Chapter 11 of the AR5 report (Kirtman et al., 2013) in recognition of the fact that the CMIP5 models tended to over-estimate observed warming. Gold horizontal lines represent the 1.5 °C and 2.0 °C GMST anomalies relative to pre-industrial, while gold circles correspond to the years where the 1.5 °C and 2 °C thresholds are crossed with 5, 50 and 95% probability (hereafter termed crossover years). The probabilistic GMST anomaly forecast using SSP1-1.9 in the AR6 Framework lies below the 2 °C threshold until the end of century throughout the ensemble (Fig. 7a), whereas the projection of GMST for SSP2-4.5 exceeds the 1.5 °C threshold in all cases (Fig. 7d).

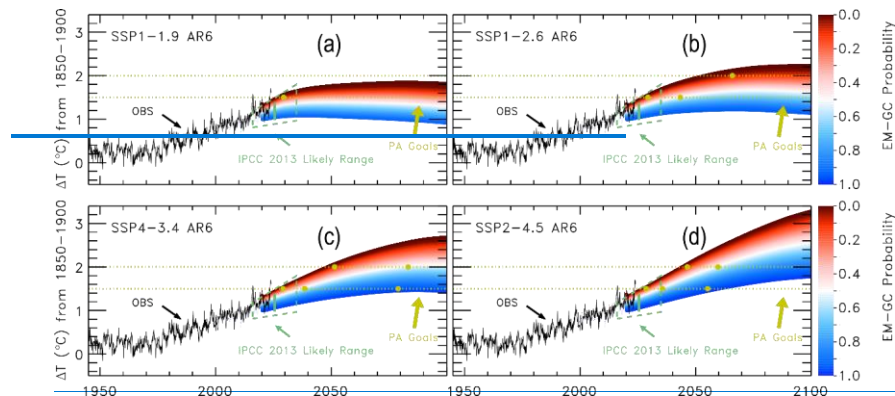


Figure 7: Time-dependent probabilistic forecasts of the GMST anomaly beyond 2019 within the AR6 Framework. The black line represents the HadCRUT5 GMST record on which EM-GC was trained for these simulations, relative to an 1850–1900 baseline. Colors represent the probability of reaching a certain temperature or higher at a given time, as indicated by the color bars on the right. The green trapezoid represents the likely range of warming as shown in Fig. 11.25b of the IPCC AR5 report (Kirtman et al., 2013). The target and upper limit of the Paris Agreement are shown with horizontal lines in gold. Circle markers on these lines correspond to the projected GMST anomaly crossing these thresholds with the probability indicated by the colors. (a) GMST projections for SSP1–1.9. (b) GMST projections for SSP1–2.6. (c) GMST projections for SSP4–3.4. (d) GMST projections for SSP2–4.5. Results for the Baseline Framework are shown in the same fashion in Fig. S3.

Table 2 provides the year that temperature anomaly thresholds are crossed for the median as well as the 5th and 95th percentiles of the probabilistic forecasts of ΔT . Figure 8 shows the probability distribution functions (PDFs) of ΔT_{2100} . Results for both the Baseline and AR6 Framework are shown in Table 2 and Fig. 8, although we generally discuss only the results found for the latter. For the 5th percentile of the AR6 Framework, the 1.5 °C threshold is crossed in either 2028 or 2029 for all four SSP scenarios. For the median member of our probabilistic forecasts, the 1.5 °C threshold is crossed in the years 2035, 2038, and 2043 for the SSP2–4.5, SSP4–3.4, and SSP1–2.6 scenarios, respectively, whereas the 1.5 °C threshold is not crossed for the median member of the SSP1–1.9 scenario. For the 95th percentile of the forecasts within the AR6 Framework, the 1.5 °C warming threshold is crossed in 2055 and 2079 for the SSP2–4.5 and SSP4–3.4 scenarios, respectively, and is not crossed for the other two scenarios. Here, the 5th percentile has more warming than the 95th percentile, due to the fact that Fig. 7 shows probabilistic forecasts that ΔT will exceed a given amount.

Median projections of ΔT_{2100} within the AR6 Framework are found to be between 0.2 °C (SSP1–1.9, SSP1–2.6) and 0.4 °C (SSP2–4.5) greater than in the case of the Baseline Framework, mainly due to the larger projected anthropogenic ERF for AR6 (Fig. 11). The PDF of ΔT_{2100} for the AR6 Framework is narrower and shifted a bit to the right (higher chance of more warming), compared to the Baseline Framework (Fig. 8), primarily due to two factors: the ability to obtain good fits to the GMST anomaly at the highest values of climate feedback only for the Baseline Framework, and the larger weight given to ensemble members with large, contemporary aerosol cooling in the AR6 Framework (Fig. 3).

1255

1260

1265

—Our median, 5th and 95th percentile estimates of ΔT_{2100} (Table 1) agree well with the long-term warming assessed by AR6, for years 2081–2100. Table 4.5 of AR6, which uses Global Surface Air Temperatures (GSAT) rather than GMST, projects warmings (median and 5–95% range) of 1.4 [1.0 to 1.8 °C], 1.8 [1.3 to 2.4 °C], and 2.7 [2.1 to 3.5 °C] for SSP1–1.9, SSP1–2.6 and SSP2–4.5, respectively. As described in Chapters 1, 2 and 4 of AR6 (Chen et al., 2021; Gulev et al., 2021; Lee et al., 2021), GMST and GSAT differ by up to 10%, but the sign of the difference is highly uncertain. Our projections of end-of-century warming and for crossing the 1.5 °C and 2.0 °C GMST thresholds are generally consistent with the GSAT-based assessments of AR6 (Lee et al., 2021).

—Finally, our study relies on SSP-based ERF data and does not factor in the significant decline of maritime sulfur emissions (and connected change in ERF_{AER}) caused by updated regulations on the sulfur content of ship fuels, introduced by the International Maritime Organization (IMO) in 2020. Varying estimates exist on the magnitude of ΔERF_{AER} from the new IMO regulations and their impact on GMST/GSAT (Yuan et al., 2024; Skeie et al., 2024a; Quaglia and Visioni, 2024; Yoshioka et al., 2024). The reduction in shipping emissions may have brought forward global warming by 2–3 years (Gettelman et al., 2024; Jordan and Henry, 2024). Such a change would also help explain the unexpectedly high GMST anomalies of the year 2023, which have not otherwise been captured by climate model projections (Schmidt, 2024). The impact of IMO regulations on GMST in our model framework will be quantified in a future study.

Table 2: Years of crossing the 1.5 °C and 2.0 °C GMST anomaly thresholds for the four SSP scenarios studied. For each entry, we present the 50% probability as our central estimate, as well as the 5–95% range. The label “n.e” is used consistent with Table 4.5 of AR6 (Lee et al., 2021) and corresponds to a given threshold not being crossed in the 2020–2100 period.

	1.5 °C Crossover		2.0 °C Crossover	
	Baseline	AR6	Baseline	AR6
SSP1–1.9	n.e [2028 to n.e]	n.e [2029 to n.e]	n.e [n.e to n.e]	n.e [n.e to n.e]
SSP1–2.6	n.e [2029 to n.e]	2043 [2029 to n.e]	n.e [2059 to n.e]	n.e [2065 to n.e]
SSP4–3.4	2048 [2030 to n.e]	2038 [2029 to 2079]	n.e [2050 to n.e]	2083 [2051 to n.e]
SSP2–4.5	2042 [2030 to 2080]	2035 [2028 to 2055]	2082 [2047 to n.e]	2059 [2046 to n.e]

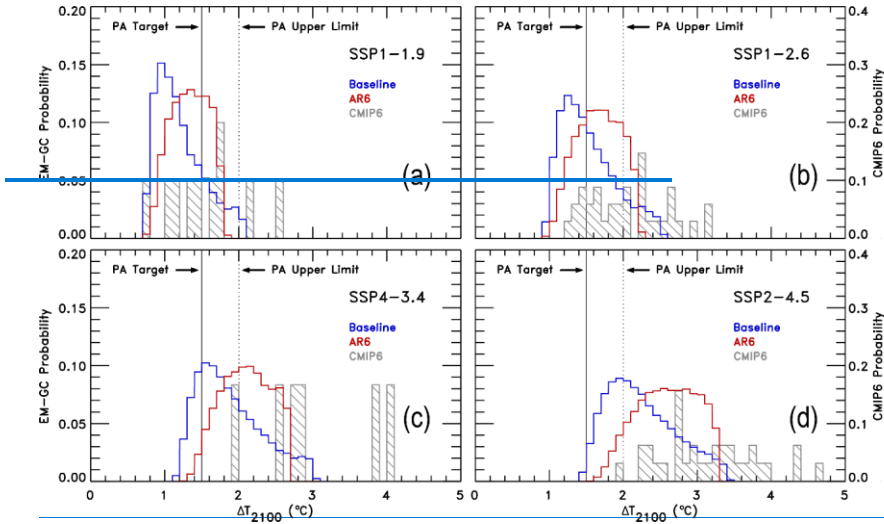


Figure 8: Probability Distribution Functions (PDFs) for ΔT_{2100} obtained from EM-GC simulations trained on the HadCRUT5 temperature dataset. Model runs for the Baseline and AR6 Frameworks are shown in blue and red, respectively. Grey color represents the PDFs obtained from a CMIP6 multi-model ensemble as described in Section 3.3.1 of McBride et al. (2021), and are shown for comparison with EM-GC results. The left-hand y axis corresponds to the EM-GC probabilities, and the right-hand y axis is for the CMIP6 probabilities. The PA target and upper limit are shown as solid and dashed vertical lines, respectively. (a) PDFs for SSP1-1.9, (b) PDFs for SSP1-2.6, (c) PDFs for SSP4-3.4, (d) PDFs for SSP2-4.5.

Figure 8 also shows PDFs of ΔT_{2100} from CMIP6 models. Consistent with many prior studies, including the Reduced Complexity Model Intercomparison Project (Nicholls et al., 2020; Nicholls et al., 2021), the free-running CMIP6 models exhibit larger end-of-century warming than projected by our model.

Tokarska et al. (2020b) reported that observationally-constrained CMIP6 projections of end-of-century warming are 9% to 13% lower than unconstrained CMIP6 projections for SSP1-2.6 and SSP2-4.5, respectively. Tokarska et al. (2020b) found the median and 5–95% ranges of end-of-century warming relative to a 1995–2014 baseline to be 0.94 [0.41 to 1.46] °C and 1.84 [1.15 to 2.52] °C for SSP1-2.6 and SSP2-4.5, respectively, using observationally-constrained CMIP6 models. Our values of ΔT_{2100} in Table 1, relative to 1995–2014, are 0.81 [0.32 to 1.27] °C and 1.74 [1.06 to 2.34] °C for SSP1-2.6 and SSP2-4.5. Consequently, our quantification of ΔT_{2100} is in good agreement with the findings of Tokarska et al. (2020b). As for the comparison with Fig. 4.2a of AR6 (Lee et al., 2021), it is important to highlight that our estimates use GMST, whereas Tokarska et al. (2020b) presents (adjusted) GSAT. Therefore, some of the differences between the forecasts can be explained by the difference in GSAT and GMST. Further, Chylek et al. (2024) recently found GSAT to be 2.41 °C in the year 2100 for SSP2-4.5 using a set of CMIP6 models that accurately reproduce the 2014–2023 warming, about 0.5 °C smaller than the average obtained from an unconstrained ensemble, in reasonably good agreement with our median estimate for ΔT_{2100} of 2.6 °C.

1295 —In summary, our projections of ΔT_{2100} using the EM–GC model, trained within the AR6 Framework, produces estimates consistent with the output of observationally constrained CMIP6 models presented in numerous other studies. These estimates were obtained assuming climate feedback is constant between 1850 and 2100 (that is, $\alpha' = 0 \text{ W m}^{-2} \text{ K}^{-1}$). The close consistency between our ΔT_{2100} projections and the prior literature provides confidence in the virtue of the assumption that climate feedback has been mainly constant over the past century and a half, even though the actual values of climate feedback is not well known. The interested reader is directed towards Sect. 3.3.6 of McBride *et al.* (2021), for projections of ΔT in our model framework that allow the value of climate feedback to vary over time.

1300 **Table 3: EM–GC computed probabilities of achieving the Paris Agreement target (1.5 °C) and upper limit (2.0 °C). Columns with the “Baseline” header represent EM–GC simulations using the Baseline Framework, while “AR6” represents simulations utilizing the AR6 Framework. The values presented in this table are derived from the PDFs shown in Fig. 8.**

	Paris 1.5 °C		Paris 2.0 °C	
	Baseline	AR6	Baseline	AR6
SSP1–1.9	81%	70%	98%	100%
SSP1–2.6	54%	32%	87%	85%
SSP4–3.4	21%	3%	65%	40%
SSP2–4.5	1%	0%	35%	8%

1305 —We conclude this section by evaluating the probability of achieving both the target (1.5 °C) and upper limit (2.0 °C) of the Paris Agreement (PA). Table 3 provides the probability that end-of-century warming will be below either the target or the upper limit, relative to pre-industrial conditions. These estimates were obtained from our probabilistic forecasts of ΔT , for both the Baseline and AR6 Frameworks (Fig. 8). As shown above, median projections of ΔT_{2100} are larger within the AR6 Framework relative to the Baseline for all four SSPs, which leads to a decline in the probability of accomplishing the PA within the AR6 Framework (Table 3). For the SSP1–1.9 and SSP1–2.6 scenarios, the probability of limiting global warming to 2.0 °C is high (at least 85%) for both model frameworks. For SSP4–3.4, the probability of limiting warming to 2.0 °C falls from 65% (Baseline) to 40% (AR6). Most notably, the 2.0 °C probability drops from 35% to 8%, for the SSP2–4.5 scenario. 1310 The 1.5 °C warming probabilities for the AR6 Framework are all uniformly lower than for the Baseline Framework, with the SSP1–2.6 scenario dropping from 54% (Baseline) to 32% (AR6). The takeaway message from Table 3 is that, for society to have high confidence in achieving at least the upper limit of the PA, the radiative forcing of climate due to GHGs must be placed close to the SSP1–2.6 pathway over the coming decades. More aggressive reductions in GHG radiative forcing are needed to achieve the target of the PA, such as those of the SSP1–1.9 scenario.

4. Conclusions

The extent of global warming is proportional to the ERF from greenhouse gases and tropospheric aerosols. In this work, we use a multiple linear regression energy balance model (EM–GC) to quantify how updates to the projections of effective radiative forcing (ERF) due to GHGs and tropospheric aerosols given by the IPCC AR6 report impact estimates of attributable anthropogenic warming, climate sensitivity, and projected future increases in GMST. We focus on four policy-relevant SSP scenarios: SSP1–1.9, SSP1–2.6, SSP4–3.4 and SSP2–4.5 (O'Neill et al., 2014; O'Neill et al., 2016).

We show that projected total anthropogenic ERF (ERF_{ANTH}) computed from AR6 forecasts of the atmospheric concentrations of GHGs, AR6 changes to the parameterizations of ERF due to GHGs, and AR6 updates to the radiative effects of tropospheric aerosols (termed the AR6 framework) results in considerably larger values of projected ERF_{ANTH} compared to values in our Baseline framework, which represents the state of knowledge prior to AR6. It is important to note that ERF_{ANTH} found in both the AR6 and Baseline frameworks relies on the same emission time series for GHGs. The higher values of ERF_{ANTH} in the AR6 framework relative to Baseline is driven by updates to future atmospheric concentrations of CO_2 and CH_4 due to the use by AR6 of a new version of the Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC) model (Meinshausen et al., 2011a; Meinshausen et al., 2011b; Meinshausen et al., 2020), updates to the mathematical formulations of ERF due to GHGs in AR6 that results in about a 6% increase in the radiative forcing of climate for “doubled CO_2 ” relative to the AR5 parameterization (Forster et al., 2021), as well as major changes in the assessed values of the magnitude and temporal evolution of ERF due to tropospheric aerosols by AR6. The combination of these factors results in about a 0.9 W m^{-2} increase in the end-of-century value of ERF_{ANTH} for the SSP2–4.5 scenario, for AR6 compared to Baseline. While some minor deviation of end-of-century ERF_{ANTH} from the nameplate RF value of SSP scenarios is expected (that is, 4.5 W m^{-2} radiative forcing for SSP2–4.5) (van Vuuren et al., 2014), the AR6 updates increase ERF_{ANTH} such that end-of-century values are now considerably larger than the nameplate RF, for each of the four SSP scenarios we have examined.

The rate of human-induced warming between 1974 and 2014 (AAWR) was found to be $0.18 [0.13 \text{ to } 0.21] \text{ } ^\circ\text{C decade}^{-1}$ within the AR6 framework, a slight increase relative to the central estimate and range of $0.16 [0.12 \text{ to } 0.20] \text{ } ^\circ\text{C decade}^{-1}$ for the Baseline framework (the range reflects the 5th and 95th percentiles). Our estimate of AAWR within the AR6 framework was shown to be consistent with other recent studies that adopt various means to separate human and natural influence on GMST (Gulev et al., 2021; Forster et al., 2023; Samset et al., 2023). Most importantly, our estimate of AAWR is lower than values found by many of the free-running (that is, unconstrained by observed GMST) CMIP6 Earth System Models (ESMs) (Tokarska et al., 2020b; McBride et al., 2021; Tebaldi et al., 2021; Hausfather et al., 2022; Chylek et al., 2024).

The magnitude of Effective Climate Sensitivity (EffCS) inferred from the historical GMST record was found to be $2.29 [1.54 \text{ to } 3.11] \text{ } ^\circ\text{C}$ within the AR6 framework and $2.26 [1.45 \text{ to } 4.37] \text{ } ^\circ\text{C}$ for the Baseline framework. Although the median value of EffCS is nearly identical between the two frameworks, there is a narrower range within the AR6 framework. The narrower range is driven by the ability to obtain good fits to the historical GMST record for a larger values of climate feedback within

the Baseline framework compared to AR6, which we have shown is driven by large differences in the assessed temporal evolution of cooling by tropospheric aerosols in AR6 compared to Baseline. Using the AR6 best estimate for the pattern effect (α' in Eq. (2)) of $0.5 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ (Forster et al., 2021), we find values for Equilibrium Climate Sensitivity (ECS) of 3.24 [1.92 to 5.15] $^{\circ}\text{C}$ for the AR6 framework. This estimate of ECS is quite similar to the AR6 assessment of 3.0 [2.0 to 5.0] $^{\circ}\text{C}$ given in Table 7.13 of Forster et al. (2021). Overall, our estimates of EffCS and ECS within the AR6 framework compare quite well with values reported by several other recent analyses of the climate system (Rugenstein et al., 2020; Skeie et al., 2024; Armour et al., 2024).

The computational efficiency of the EM-GC allows for probabilistic projections on future warming that rely on consideration of the uncertainty in the magnitude of ERF from tropospheric aerosols as well as climate feedback. Median projections on end-of-century warming (ΔT_{2100}) found using our model are higher, by 0.2 $^{\circ}\text{C}$ to 0.4 $^{\circ}\text{C}$, for the AR6 framework relative to projections found in the Baseline framework. Both frameworks rely upon the same time series for the future emissions of GHGs. The larger projected warming within the AR6 framework is driven by AR6 updates to projections of the atmospheric concentrations of GHGs, the ERF of GHGs, and the temporal evolution of aerosol cooling.

Finally, we evaluate the likelihood of achieving either the goal (1.5 $^{\circ}\text{C}$) or upper limit (2 $^{\circ}\text{C}$) warming thresholds of the Paris Agreement (PA). The AR6 updates to GHGs and aerosols result in a decline in the likelihood of limiting warming to either 1.5 or 2 $^{\circ}\text{C}$, compared to Baseline. Below, we give numerical results for only the AR6 framework. Model simulations using SSP2-4.5 (designed to reflect trends in the absence of further climate policies) and SSP1-2.6 (full implementation of currently proposed emission targets (Meinshausen et al., 2024)), are found to offer no chance and a 32% chance of limiting the rise in GMST to 1.5 $^{\circ}\text{C}$ by 2100. Model simulations conducted using SSP4-3.4 (intermediate scenario between SSP1-2.6 and SSP2-4.5) and SSP1-1.9 (aggressive future reductions in GHG emissions) show probabilities of 3% and 70% of limiting warming to 1.5 $^{\circ}\text{C}$ by end-of-century. The likelihood of limiting warming to 2.0 $^{\circ}\text{C}$ is found to be 8%, 40%, 85%, and 100% for the SSP2-4.5, SSP4-3.4, SSP1-2.6, and SSP1-1.9 scenarios, respectively.

In our earlier work using the EM-GC that relied upon the Baseline framework, we had concluded that the SSP4-3.4 scenario provided a 64% probability of limiting global warming to 2 $^{\circ}\text{C}$ by end-of-century (McBride et al., 2021). The AR6-based updates to ERF_{ANTH} considered here result in a lower, 40% probability of achieving this upper limit of the PA. Nonetheless, the EM-GC based estimates of limiting end-of-century warming to 2 $^{\circ}\text{C}$ for both the AR6 and Baseline frameworks are more optimistic than is provided by many free-running CMIP6 ESMs. Our results suggest that SSP1-2.6 is the “two degree pathway”, since this scenario provides an 85% probability of limiting global warming to the upper limit of the PA.

4 Conclusions

The extent of global warming is proportional to the ERF from greenhouse gases and tropospheric aerosols. In this work, we use a multiple linear regression energy balance model (EM–GC) to quantify how updates to the ERF formulations of GHGs and tropospheric aerosols adapted by the IPCC AR6 report (termed AR6 Framework in our study) impact estimates of climate sensitivity and projected future warming compared to forecasts made based on pre–AR6 datasets (Baseline Framework). Our study focuses on four policy-relevant SSP scenarios: SSP1–1.9, SSP1–2.6, SSP4–3.4 and SSP2–4.5 (O'Neill et al., 2014; O'Neill et al., 2016). Our model framework has two key elements: numerical representation of ocean heat export as well as the impact of uncertainty in radiative forcing due to tropospheric aerosols. The numerical values given below are results based on 160,000 member ensembles for each SSP.

—First, we compare ERF for GHGs and tropospheric aerosols between the two Frameworks. We find that the projected ERF due to GHGs—particularly CO_2 and CH_4 —is considerably higher in the AR6 Framework relative to the Baseline, for all four of the SSP scenarios. This increase in ERF_{GHG} originates from the updated assessments of future concentrations as well as updates to the formulations of ERF due to GHGs in AR6, in combination with updated adjustments to the RF of GHGs that now include responses throughout the troposphere, in addition to the stratosphere (Smith et al., 2018b; Hodnebrog et al., 2020; Meinshausen et al., 2020; Forster et al., 2021; Smith et al., 2021a; IPCC, 2021b). End-of-century ERF is found to be up to 0.7 W m^{-2} higher within the AR6 Framework relative to the Baseline for SSP2–4.5. While some minor deviation of end-of-century ERF from the nameplate RF value of SSP scenarios is expected (Van Vuuren et al., 2014), the AR6 Framework results in a level of increase in ERF such that end-of-century values are considerably larger than the nameplate RF for all four SSP scenarios.

—The magnitude of Effective Climate Sensitivity (EffCS) inferred from the historical GSMT record was found to be $2.29 [1.54 \text{ to } 3.11 \text{ }^\circ\text{C}, 5\text{--}95\% \text{ range}]$ within the AR6 Framework, and $2.26 [1.45 \text{ to } 4.37 \text{ }^\circ\text{C}]$ for the Baseline Framework. The median value of EffCS is nearly identical between the two Frameworks, and there is a much narrower range within the AR6 Framework. Equilibrium Climate Sensitivity (ECS) is estimated to be $3.24 [1.92 \text{ to } 5.15 \text{ }^\circ\text{C}]$ within the AR6 Framework, using the AR6 best estimate for the pattern effect ($\alpha' = 0.5 \text{ W m}^{-2} \text{ K}^{-1}$). This estimate of ECS is quite similar to the AR6 assessment of $3.0 [2.0 \text{ to } 5.0 \text{ }^\circ\text{C}]$ given in Table 7.13 of Forster *et al.* (2021). Overall, our estimates of EffCS and ECS within the AR6 Framework compare very well with values reported by several other recent studies (Armour et al., 2024; Cooper et al., 2024; Skeie et al., 2024b). The rate of human-induced warming between 1974 and 2014 (AAWR) was found to be $0.18 [0.13 \text{ to } 0.21 \text{ }^\circ\text{C decade}^{-1}]$ within the AR6 Framework, a slight increase relative to the central estimate and range of $0.16 [0.12 \text{ to } 0.20 \text{ }^\circ\text{C decade}^{-1}]$ for the Baseline Framework. Our estimate of AAWR within the AR6 Framework is also consistent with other recent studies (Gulev et al., 2021; Forster et al., 2023; Samset et al., 2023).

—Our projections of GSMT show good consistency with GSAT estimates from observationally constrained CMIP6 ensembles but fall below the estimates of global warming when considering unconstrained CMIP6 ensembles. Overall, our estimates of EffCS, ECS, AAWR and end-of-century warming using the AR6 Framework are all consistent with values found

by CMIP6 models that accurately reproduce observed warming trends over the past three or four decades. Our work highlights the importance of ensuring that CMIP6 models used for policy purposes succeed in reproducing observed trends in GMST.

—Our probabilistic projections on future warming consider the uncertainty in the magnitude of ERF from tropospheric aerosols, as assessed by AR6 and AR5, for the AR6 and Baseline Frameworks, respectively. Global warming projections that utilize the uncertainty in aerosol ERF is a common product of RCMs (Nicholls et al., 2020; Nicholls et al., 2021; Hope et al., 2017; McBride et al., 2021) and is not usually found using GCMs due to the computational demand of global models. Median projections on end-of-century warming using our RCM, are higher by 0.2 °C to 0.4 °C within the AR6 Framework relative to projections found in the Baseline Framework. This increase corresponds to a general decline in the likelihood of limiting warming to the Paris Agreement (PA) thresholds, for the AR6 Framework compared to the Baseline. Notably, model simulations using SSP2-4.5 and SSP1-2.6—which reflect trends in the absence of further climate policies and full implementation of currently proposed emission targets (Meinshausen et al., 2024)—are found to offer 32% and 0% chance of accomplishing the PA target of limiting the rise in GMST to 1.5 °C by 2100. Model simulations conducted using the SSP4-3.4 (intermediate scenario between SSP1-2.6 and SSP2-4.5) and SSP1-1.9 (aggressive future reductions in GHG emissions) show probabilities of 3% and 70% of limiting warming to 1.5 °C by end-of-century. The likelihood of limiting warming to the PA upper limit of 2.0 °C is found to be 40% and 8% within the AR6 Framework for SSP4-3.4 and SSP2-4.5, respectively. There is high confidence (85% and 100% probabilities) that global warming could be limited to 2.0 °C if the radiative forcing of climate due to GHGs could be placed on either the SSP1-2.6 or SSP1-1.9 pathways.

5. Data and Code availability

All data used as inputs of EM-GC are available from online resources. We have provided the links to these datasets below.

The compiled input files used by EM-GC are also provided on Zenodo.org at 10.5281/zenodo.14720490 (Farago et al., 2025).

The EM-GC output data is also provided in this Zenodo repository.

- SSP database (Baseline Framework): <https://tntcat.iiasa.ac.at/SspDb/>
- Tropospheric O₃ RF (Baseline Framework): <https://www.pik-potsdam.de/~mmalte/rcps/>
- AR6 Radiative Forcing (AR6 Framework): <https://doi.org/10.5281/zenodo.5705391>
- MEIv2 and MEI.ext: <https://psl.noaa.gov/enso/mei/> and <https://psl.noaa.gov/enso/mei.ext/>
- PDO: <http://research.jisao.washington.edu/pdo/PDO.latest.txt>
- COBE SST data used to construct the IOD time series is available at: <https://psl.noaa.gov/data/gridded/data.cobe.html>
- GloSSAC SAOD: <https://asdc.larc.nasa.gov/project/GloSSAC>
- TSI: <https://lasp.colorado.edu/sorce/data/tsi-data/>
- OHC Records:
 - Balmaseda: <https://www.cgd.ucar.edu/cas/catalog/ocean/oras4.html>
 - Carton: https://www2.atmos.umd.edu/~ocean/soda3_readme.htm

- Cheng: <http://www.ocean.iap.ac.cn/pages/dataService/dataService.html?navAnchor=dataService>
- Ishii: https://www.data.jma.go.jp/gmd/kaiyou/english/ohc/ohc_global_en.html
- Levitus: <https://www.ncei.noaa.gov/access/global-ocean-heat-content/>

6. Author contribution

EF updated the EM–GC model, performed the model simulations, conducted the data analysis and wrote the first draft of the manuscript. LM, AH and TC developed earlier versions of the EM–GC model and assisted in the review and editing of the manuscript. BB led the compilation of the IOD and SAOD datasets, and participated in the review and editing of the manuscript. RS supervised the project and participated in the review and editing of the manuscript.

7. Competing interests

The authors declare that they have no conflict of interest.

8. Acknowledgements

We thank the World Climate Research Programme for coordinating and promoting CMIP6 through its Working Group on Coupled Modelling. We also thank the climate modeling groups participating in CMIP6 for producing and making their model results available, the Earth System Grid Federation (ESGF) for archiving the data and providing access. We thank those who contributed to Chapter 7 and Annex III of AR6 for providing values of ERF and other quantities used in this manuscript. ~~Finally, w~~We appreciate the financial support of the NASA Climate Indicators and Data Products for Future National Climate Assessments program during the early phase of this research. Finally, we thank both reviewers for their very helpful constructive criticism of the original submitted version of this manuscript.

9. Financial support

This research was supported by the National Aeronautics and Space Administration (grant no. NNX16AG34G).

10. References

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour, K. C., Forster, P. M., and Titchner, H.: Accounting for changing temperature patterns increases historical estimates of climate sensitivity, *Geophys. Res. Lett.*, 45, 8490-8499, 10.1029/2018GL078887, 2018.

Formatted: Indent: Left: 0", Hanging: 0.2", Space After: 6 pt

- Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., Menary, M. B., Ridley, J., Ringer, M. A., Sellar, A. A., Senior, C. A., and Tang, Y.: Forcings, Feedbacks, and Climate Sensitivity in HadGEM3-GC3.1 and UKESM1, *Journal of Advances in Modeling Earth Systems*, 11, 4377-4394, <https://doi.org/10.1029/2019MS001866>, 2019.
- Arfeuille, F., Weisenstein, D., MacK, H., Rozanov, E., Peter, T., and Brönnimann, S.: Volcanic forcing for climate modeling: A new microphysics-based data set covering years 1600-present, *Climate of the Past*, 10, 10.5194/cp-10-359-2014, 2014.
- Armour, K. C.: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, *Nature Climate Change*, 7, 10.1038/nclimate3278, 2017.
- Armour, K. C., Proistosescu, C., Dong, Y., Hahn, L. C., Blanchard-Wrigglesworth, E., Pauling, A. G., Jnglin Wills, R. C., Andrews, T., Stuecker, M. F., Po-Chedley, S., Mitevski, I., Forster, P. M., and Gregory, J. M.: Sea-surface temperature pattern effects have slowed global warming and biased warming-based constraints on climate sensitivity, *Proceedings of the National Academy of Sciences*, 121, e2312093121, 10.1073/pnas.2312093121, 2024.
- Balmaseda, M. A., Trenberth, K. E., and Källén, E.: Distinctive climate signals in reanalysis of global ocean heat content, *Geophysical Research Letters*, 40, 10.1002/grl.50382, 2013.
- Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J. L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G., and Webb, M. J.: How well do we understand and evaluate climate change feedback processes?, 10.1175/JCLI3819.1, 2006.
- Calvin, K., Bond-Lamberty, B., Clarke, L., Edmonds, J., Eom, J., Hartin, C., Kim, S., Kyle, P., Link, R., Moss, R., McJeon, H., Patel, P., Smith, S., Waldhoff, S., and Wise, M.: The SSP4: A world of deepening inequality, *Global Environmental Change*, 42, 284-296, 2017.
- Canty, T., Mascioli, N. R., Smarte, M. D., and Salawitch, R. J.: An empirical model of global climate-Part 1: A critical evaluation of volcanic cooling, *Atmospheric Chemistry and Physics*, 13, 10.5194/acp-13-3997-2013, 2013.
- Carton, J. A., Chepurin, G. A., and Chen, L.: SODA3: A new ocean climate reanalysis, *Journal of Climate*, 31, 10.1175/jcli-d-18-0149.1, 2018.
- Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., and Zhu, J.: Improved estimates of ocean heat content from 1960 to 2015, 10.1126/sciadv.1601545, 2017.
- Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W. T., Stammer, D., and Unnikrishnan, A. S.: Sea Level Change. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)] doi:10.1017/CBO9781107415324.026, 2013.
- Chylek, P., Folland, C. K., Klett, J. D., Wang, M., Lesins, G., and Dubey, M. K.: Why Does the Ensemble Mean of CMIP6 Models Simulate Arctic Temperature More Accurately Than Global Temperature?, 10.3390/atmos15050567, 2024.
- Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., and Andrews, T.: Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models, *Journal of Climate*, 33, 7755-7775, <https://doi.org/10.1175/JCLI-D-19-1011.1>, 2020.
- Dudok de Wit, T., Kopp, G., Fröhlich, C., and Schödl, M.: Methodology to create a new total solar irradiance record: Making a composite out of multiple data records, *Geophysical Research Letters*, 44, 10.1002/2016GL071866, 2017.
- Edwards, D. P.: GENLN2: A general line-by-line atmospheric transmittance and radiance model, National Center for Atmospheric Research1992.
- Etminan, M., Myhre, G., Highwood, E. J., and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, *Geophysical Research Letters*, 43, 12,614-612,623, <https://doi.org/10.1002/2016GL071930>, 2016.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, 9, 10.5194/gmd-9-1937-2016, 2016.

Farago, E. Z., McBride, L. A., Hope, A. P., Canty, T. P., Bennett, B. F., and Salawitch, R. J.: EM-GC Input and Output files (1.0) [Data set], <https://doi.org/10.5281/zenodo.14720490>, 2025.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 923–1054, 10.1017/9781009157896.009, 2021.

Forster, P. M., Smith, C. J., Walsh, T., Lamb, W. F., Lamboll, R., Hauser, M., Ribes, A., Rosen, D., Gillett, N., Palmer, M. D., Rogelj, J., von Schuckmann, K., Seneviratne, S. I., Trewin, B., Zhang, X., Allen, M., Andrew, R., Birt, A., Borger, A., Boyer, T., Broersma, J. A., Cheng, L., Dentener, F., Friedlingstein, P., Gutiérrez, J. M., Gütschow, J., Hall, B., Ishii, M., Jenkins, S., Lan, X., Lee, J. Y., Morice, C., Kadow, C., Kennedy, J., Killick, R., Minx, J. C., Naik, V., Peters, G. P., Pirani, A., Pongratz, J., Schleussner, C. F., Szopa, S., Thorne, P., Rohde, R., Rojas Corradi, M., Schumacher, D., Vose, R., Zickfeld, K., Masson-Delmotte, V., and Zhai, P.: Indicators of Global Climate Change 2022: annual update of large-scale indicators of the state of the climate system and human influence, *Earth Syst. Sci. Data*, 15, 2295–2327, 10.5194/essd-15-2295-2023, 2023.

Foster, G. and Rahmstorf, S.: Global temperature evolution 1979–2010, *Environmental Research Letters*, 6, 10.1088/1748-9326/6/4/044022, 2011.

Fredriksen, H.-B., Smith, C. J., Modak, A., and Rugenstein, M.: 21st Century Scenario Forcing Increases More for CMIP6 Than CMIP5 Models, *Geophysical Research Letters*, 50, e2023GL102916, <https://doi.org/10.1029/2023GL102916>, 2023.

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., Rao, S., Schmid, E., Schoepp, W., and Riahi, K.: The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century, *Global Environmental Change*, 42, 10.1016/j.gloenvcha.2016.06.004, 2017.

Fu, B., Li, B., Gasser, T., Tao, S., Ciaia, P., Piao, S., Balkanski, Y., Li, W., Yin, T., Han, L., Han, Y., Peng, S., and Xu, J.: The contributions of individual countries and regions to the global radiative forcing, *Proceedings of the National Academy of Sciences*, 118, e2018211118, 10.1073/pnas.2018211118, 2021.

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

Gottelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., Lamarque, J. F., Fasullo, J. T., Bailey, D. A., Lawrence, D. M., and Mills, M. J.: High Climate Sensitivity in the Community Earth System Model Version 2 (CESM2), *Geophysical Research Letters*, 46, 10.1029/2019GL083978, 2019.

Gregory, J. M., Andrews, T., Ceppi, P., Mauritsen, T., and Webb, M. J.: How accurately can the climate sensitivity to CO₂ be estimated from historical climate change?, *Climate Dynamics*, 54, 129–157, 10.1007/s00382-019-04991-y, 2020.

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D., Kaufman, D. S., Nnamchi, H. C., Quaas, J., Rivera, J. A., Sathyendranath, S., Smith, S. L., Trewin, B., von Schuckmann, K., and Vose, R. S.: Changing State of the Climate System, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K.,

- 1555 Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 287–422, 10.1017/9781009157896.004, 2021.
- | Hansen, J., Sato, M., Kharecha, P., and Von Schuckmann, K.: Earth's energy imbalance and implications, *Atmospheric Chemistry and Physics*, 11, 10.5194/acp-11-13421-2011, 2011.
- | Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., and Zelinka, M.: Climate simulations: recognize the 'hot model' problem, *Nature*, 605, 26–29, 10.1038/d41586-022-01192-2, 2022.
- 1560 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-i., Li, M., Liu, L., Lu, Z., Moura, M. C. P., ampaposRourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geoscientific Model Development*, 11, 369–408, 2018.
- | Hope, A. P., Canty, T. P., Salawitch, R. J., Tribett, W. R., and Bennett, B. F.: Forecasting global warming, in: *Springer Climate*, 10.1007/978-3-319-46939-3_2, 2017.
- 1565 | IPCC: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.)], 2001.
- | IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)], 2007.
- 1570 | IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. doi:10.1017/CBO9781107415324, 2013.
- 1575 | IPCC: Annex VII: Glossary [Matthews, J.B.R., V. Möller, R. van Diemen, J.S. Fuglestedt, V. Masson-Delmotte, C. Méndez, S. Semenov, A. Reisinger (eds.)], in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2215–2256, 10.1017/9781009157896.022, 2021a.
- 1580 | IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 10.1017/9781009157896, 2021b.
- 1585 | IPCC: Annex III: Tables of historical and projected well-mixed greenhouse gas mixing ratios and effective radiative forcing of all climate forcers [Dentener F.J., B. Hall, C. Smith (eds.)], in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2139–2152, 10.1017/9781009157896.017, 2021c.
- 1590 | Ishii, M., Fukuda, Y., Hirahara, S., Yasui, S., Suzuki, T., and Sato, K.: Accuracy of Global Upper Ocean Heat Content Estimation Expected from Present Observational Data Sets, *SOLA*, 13, 10.2151/sola.2017-030, 2017.
- 1595 | Kennedy, J. J., Rayner, N. A., Atkinson, C. P., and Killick, R. E.: An Ensemble Data Set of Sea Surface Temperature Change From 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set, *Journal of Geophysical Research: Atmospheres*, 124, 10.1029/2018JD029867, 2019.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M.: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties, *Journal of Geophysical Research Atmospheres*, 116, 10.1029/2010JD015218, 2011.

1600 Kirtman, B., Power, S. B., Adedoyin, J. A., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F. J., Fiore, A. M., Kimoto, M., Meehl, G. A., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G. J., Vecchi, G., and Wang, H. J.: Near-term Climate Change: Projections and Predictability. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. doi:10.1017/CBO9781107415324.023, 2013.

1605 Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., Strefler, J., Baumstark, L., Bodirsky, B. L., Hilaire, J., Klein, D., Mouratiadou, I., Weindl, I., Bertram, C., Dietrich, J. P., Luderer, G., Pehl, M., Pietzcker, R., Piontek, F., Lotze-Campen, H., Biewald, A., Bonsch, M., Giannousakis, A., Kreidenweis, U., Müller, C., Rolinski, S., Schultes, A., Schwanitz, J., Stevanovic, M., Calvin, K., Emmerling, J., Fujimori, S., and Edenhofer, O.: Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century, *Global Environmental Change*, 42, 10.1016/j.gloenvcha.2016.05.015, 2017.

1610 Lean, J. L. and Rind, D. H.: How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006, *Geophysical Research Letters*, 35, 10.1029/2008GL034864, 2008.

Lean, J. L. and Rind, D. H.: How will Earth's surface temperature change in future decades?, *Geophysical Research Letters*, 36, 10.1029/2009GL038932, 2009.

1615 Lee, J. Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J. P., Engelbrecht, F., Fischer, E., Fyfe, J. C., Jones, C., Maycock, A., Mutemi, J., Ndiaye, O., Panickal, S., and Zhou, T.: Future Global Climate: Scenario-Based Projections and Near-Term Information, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 553–672, 10.1017/9781009157896.006, 2021.

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermocline sea level change (0-2000m), 1955-2010, *Geophysical Research Letters*, 39, 10.1029/2012GL051106, 2012.

1625 Marvel, K., Pincus, R., Schmidt, G. A., and Miller, R. L.: Internal Variability and Disequilibrium Confound Estimates of Climate Sensitivity From Observations, *Geophysical Research Letters*, 45, 10.1002/2017GL076468, 2018.

Mascioli, N. R., Canty, T., and Salawitch, R. J.: An empirical model of global climate – Part 2: Implications for future temperature, *Atmos. Chem. Phys. Discuss.*, 2012, 23913-23974, 10.5194/acpd-12-23913-2012, 2012.

1630 Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok De Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., Van De Kamp, M., Verronen, P. T., and Versick, S.: Solar forcing for CMIP6 (v3.2), *Geoscientific Model Development*, 10, 10.5194/gmd-10-2247-2017, 2017.

1635 McBride, L. A., Hope, A. P., Canty, T. P., Bennett, B. F., Tribett, W. R., and Salawitch, R. J.: Comparison of CMIP6 historical climate simulations and future projected warming to an empirical model of global climate, *Earth System Dynamics*, 12, 10.5194/esd-12-545-2021, 2021.

Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, *MAGICC6 - Part 1: Model description and calibration*, *Atmospheric Chemistry and Physics*, 11, 10.5194/acp-11-1417-2011, 2011a.

1640

- Meinshausen, M., Wigley, T. M. L., and Raper, S. C. B.: Emulating atmosphere-ocean and carbon cycle models with a simpler model, *MAGICC6 – Part 2: Applications, Atmos. Chem. Phys.*, 11, 1457-1471, 10.5194/acp-11-1457-2011, 2011b.
- Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., Van Den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, *Geoscientific Model Development*, 13, 10.5194/gmd-13-3571-2020, 2020.
- Meinshausen, M., Schleussner, C. F., Beyer, K., Bodeker, G., Boucher, O., Canadell, J. G., Daniel, J. S., Diongue-Niang, A., Driouech, F., Fischer, E., Forster, P., Grose, M., Hansen, G., Hausfather, Z., Ilyina, T., Kikstra, J. S., Kimutai, J., King, A. D., Lee, J. Y., Lennard, C., Lissner, T., Nauels, A., Peters, G. P., Pirani, A., Plattner, G. K., Pörtner, H., Rogelj, J., Rojas, M., Roy, J., Samset, B. H., Sanderson, B. M., Séférian, R., Seneviratne, S., Smith, C. J., Szopa, S., Thomas, A., Urge-Vorsatz, D., Velders, G. J. M., Yokohata, T., Ziehn, T., and Nicholls, Z.: A perspective on the next generation of Earth system model scenarios: towards representative emission pathways (REPs), *Geosci. Model Dev.*, 17, 4533-4559, 10.5194/gmd-17-4533-2024, 2024.
- 655 Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, *Journal of Geophysical Research: Atmospheres*, 126, 10.1029/2019JD032361, 2021.
- Myhre, G. and Stordal, F.: Role of spatial and temporal variations in the computation of radiative forcing and GWP, *Journal of Geophysical Research: Atmospheres*, 102, 11181-11200, <https://doi.org/10.1029/97JD00148>, 1997.
- 660 Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, *Geophysical Research Letters*, 25, 10.1029/98GL01908, 1998.
- Myhre, G., Shindell, D., Bréon, F. M., Collins, W. D., Fuglestad, J., Huang, J., Koch, D., Lamarque, J. F., Lee, D., Mendoza, B., Nakajima, T., Robock, a., Stephens, G., Takemura, T., and Zhan, H.: IPCC AR5 (2013) Chapter 8: Anthropogenic and Natural Radiative Forcing, in: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 2013.
- 665 Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., Gieseke, R., Hope, A. P., Leach, N. J., McBride, L. A., Quilcaille, Y., Rogelj, J., Salawitch, R. J., Samset, B. H., Sandstad, M., Shiklomanov, A., Skeie, R. B., Smith, C. J., Smith, S. J., Su, X., Tsutsui, J., Vega-Westhoff, B., and Woodard, D. L.: Reduced Complexity Model Intercomparison Project Phase 2: Synthesizing Earth System Knowledge for Probabilistic Climate Projections, *Earth's Future*, 9, 10.1029/2020EF001900, 2021.
- Nicholls, Z., Meinshausen, M., Lewis, J., Gieseke, R., Dommengot, D., Dorheim, K., Fan, C. S., Fuglestad, J. S., Gasser, T., Goluke, U., Goodwin, P., Hartin, C., P. Hope, A., Kriegler, E., J. Leach, N., Marchegiani, D., A. McBride, L., Quilcaille, Y., Rogelj, J., J. Salawitch, R., Samset, B. H., Sandstad, M., N. Shiklomanov, A., B. Skeie, R., J. Smith, C., Smith, S., Tanaka, K., Tsutsui, J., and Xie, Z.: Reduced Complexity Model Intercomparison Project Phase 1: Introduction and evaluation of global-mean temperature response, *Geoscientific Model Development*, 13, 10.5194/gmd-13-5175-2020, 2020.
- 675 Nijse, F. J. M. M., Cox, P. M., and Williamson, M. S.: Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models, *Earth System Dynamics*, 11, 10.5194/esd-11-737-2020, 2020.
- 680 O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and van Vuuren, D. P.: A new scenario framework for climate change research: The concept of shared socioeconomic pathways, *Climatic Change*, 122, 10.1007/s10584-013-0905-2, 2014.
- O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, 9, 3461-3482, 10.5194/gmd-9-3461-2016, 2016.
- 1685

- Proistosescu, C. and Huybers, P. J.: Slow climate mode reconciles historical and model-based estimates of climate sensitivity, 10.1126/sciadv.1602821, 2017.
- Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaserna, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Streffer, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, *Global Environmental Change*, 42, 10.1016/j.gloenvcha.2016.05.009, 2017.
- Ribes, A., Qasmi, S., and Gillett, N. P.: Making climate projections conditional on historical observations, *Science Advances*, 7, eabc0671, 10.1126/sciadv.abc0671, 2021.
- Rogelj, J., Popp, A., Calvin, K. V., Luderer, G., Emmerling, J., Gernaat, D., Fujimori, S., Streffer, J., Hasegawa, T., Marangoni, G., Krey, V., Kriegler, E., Riahi, K., Van Vuuren, D. P., Doelman, J., Drouet, L., Edmonds, J., Fricko, O., Harmsen, M., Havlik, P., Humenöder, F., Stehfest, E., and Tavoni, M.: Scenarios towards limiting global mean temperature increase below 1.5 °C, *Nature Climate Change*, 8, 10.1038/s41558-018-0091-3, 2018.
- Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., Frölicher, T. L., Paynter, D., Danabasoglu, G., Yang, S., Dufresne, J. L., Cao, L., Schmidt, G. A., Abe-Ouchi, A., Geoffroy, O., and Knutti, R.: Equilibrium Climate Sensitivity Estimated by Equilibrating Climate Models, *Geophysical Research Letters*, 47, 10.1029/2019GL083898, 2020.
- Rugenstein, M., Bloch-Johnson, J., Abe-Ouchi, A., Andrews, T., Beyerle, U., Cao, L., Chadha, T., Danabasoglu, G., Dufresne, J.-L., Duan, L., Foujols, M.-A., Frölicher, T., Geoffroy, O., Gregory, J., Knutti, R., Li, C., Marzocchi, A., Mauritsen, T., Menary, M., Moyer, E., Nazarenko, L., Paynter, D., Saint-Martin, D., Schmidt, G. A., Yamamoto, A., and Yang, S.: LongRunMIP: Motivation and Design for a Large Collection of Millennial-Length AOGCM Simulations, *Bulletin of the American Meteorological Society*, 100, 2551-2570, <https://doi.org/10.1175/BAMS-D-19-0068.1>, 2019.
- Salvi, P., Gregory, J. M., and Ceppi, P.: Time-Evolving Radiative Feedbacks in the Historical Period, *Journal of Geophysical Research: Atmospheres*, 128, e2023JD038984, <https://doi.org/10.1029/2023JD038984>, 2023.
- Samset, B. H., Zhou, C., Fuglestad, J. S., Lund, M. T., Marotzke, J., and Zelinka, M. D.: Steady global surface warming from 1973 to 2022 but increased warming rate after 1990, *Communications Earth & Environment*, 4, 400, 10.1038/s43247-023-01061-4, 2023.
- Skeie, R. B., Aldrin, M., Berntsen, T. K., Holden, M., Huseby, R. B., Myhre, G., and Storelvmo, T.: The aerosol pathway is crucial for observationally constraining climate sensitivity and anthropogenic forcing, *Earth Syst. Dynam.*, 15, 1435-1458, 10.5194/esd-15-1435-2024, 2024.
- Smith, C., Nicholls, Z. R. J., Armour, K., Collins, W., Forster, P., Meinshausen, M., Palmer, M. D., and Watanabe, M.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity Supplementary Material. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Available from <https://www.ipcc.ch/>, 2021a.
- Smith, C., Hall, B., Dentener, F., Ahn, J., Collins, W., Jones, C., Meinshausen, M., Dlugokencky, E., Keeling, R., Krummel, P., Mühle, J., Nicholls, Z., and Simpson, I.: IPCC Working Group I (WG1) Sixth Assessment Report (AR6) Annex III Extended Data (v1.0) [Data set]. Zenodo., <https://doi.org/10.5281/zenodo.5705391>, 2021b.
- Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, *Geosci. Model Dev.*, 11, 2273-2297, <https://doi.org/10.5194/gmd-11-2273-2018>, 2018a.

- Smith, C. J., Kramer, R. J., Myhre, G., Forster, P. M., Soden, B. J., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D.,
1730 Hodnebrog, Ø., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J. F., Mülmenstädt, J., Olivé, D., Richardson, T., Samset,
B. H., Shindell, D., Stier, P., Takemura, T., Voulgarakis, A., and Watson-Parris, D.: Understanding Rapid Adjustments to
Diverse Forcing Agents, *Geophysical Research Letters*, 45, 12,023-012,031, <https://doi.org/10.1029/2018GL079826>,
2018b.
- Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J. L., Nabat, P., Michou,
1735 M., Yukimoto, S., Cole, J., Paynter, D., Shiogama, H., O'Connor, F. M., Robertson, E., Wiltshire, A., Andrews, T., Hannay,
C., Miller, R., Nazarenko, L., Kirkevåg, A., Olivé, D., Fiedler, S., Lewinschal, A., Mackallah, C., Dix, M., Pincus, R., and
Forster, P. M.: Effective radiative forcing and adjustments in CMIP6 models, *Atmos. Chem. Phys.*, 20, 9591-9618,
10.5194/acp-20-9591-2020, 2020.
- Smith, S. J. and Bond, T. C.: Two hundred fifty years of aerosols and climate: The end of the age of aerosols, *Atmospheric*
1740 *Chemistry and Physics*, 14, 10.5194/acp-14-537-2014, 2014.
- Spencer, R. W. and Christy, J. R.: Effective climate sensitivity distributions from a 1D model of global ocean and land
temperature trends, 1970–2021, *Theoretical and Applied Climatology*, 10.1007/s00704-023-04634-7, 2023.
- Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B.,
1745 van Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z., Tokarska, K. B., Hurtt, G., Kriegler, E., Lamarque, J. F., Meehl,
G., Moss, R., Bauer, S. E., Boucher, O., Brovkin, V., Byun, Y. H., Dix, M., Gualdi, S., Guo, H., John, J. G., Kharin, S.,
Kim, Y., Koshiro, T., Ma, L., Olivé, D., Panickal, S., Qiao, F., Rong, X., Rosenbloom, N., Schupfner, M., Séférian, R.,
Sellar, A., Semmler, T., Shi, X., Song, Z., Steger, C., Stouffer, R., Swart, N., Tachiiri, K., Tang, Q., Tatebe, H., Voldoire,
A., Volodin, E., Wyser, K., Xin, X., Yang, S., Yu, Y., and Ziehn, T.: Climate model projections from the Scenario Model
Intercomparison Project (ScenarioMIP) of CMIP6, *Earth Syst. Dynam.*, 12, 253-293, 10.5194/esd-12-253-2021, 2021.
- Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J. P., Manney, G., Luo, B., Arfeuille, F., and Peter,
1750 T.: A global space-based stratospheric aerosol climatology: 1979-2016, *Earth System Science Data*, 10, 10.5194/essd-10-
469-2018, 2018.
- Tokarska, K. B., Hegerl, G. C., Schurer, A. P., Forster, P. M., and Marvel, K.: Observational constraints on the effective
climate sensitivity from the historical period, *Environmental Research Letters*, 15, 10.1088/1748-9326/ab738f, 2020a.
- 1755 Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains
future warming in CMIP6 models, *Science Advances*, 6, 10.1126/sciadv.aaz9549, 2020b.
- van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T.,
Mathur, R., and Winkler, H.: A new scenario framework for Climate Change Research: scenario matrix architecture,
Climatic Change, 122, 373-386, 10.1007/s10584-013-0906-1, 2014.
- 1760 van Vuuren, D. P., Stehfest, E., Gernaat, D. E. H. J., Doelman, J. C., van den Berg, M., Harmsen, M., de Boer, H. S., Bouwman,
L. F., Daioglou, V., Edelenbosch, O. Y., Girod, B., Kram, T., Lassaletta, L., Lucas, P. L., van Meijl, H., Müller, C., van
Ruijven, B. J., van der Sluis, S., and Tabeau, A.: Energy, land-use and greenhouse gas emissions trajectories under a green
growth paradigm, *Global Environmental Change*, 42, 10.1016/j.gloenvcha.2016.05.008, 2017.
- Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., Hu, Y., and Hostetler, C. A.: Fully
1765 automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products, *Laser*
Radar Techniques for Atmospheric Sensing, 10.1117/12.572024,
- Wang, C., Soden, B. J., Yang, W., and Vecchi, G. A.: Compensation Between Cloud Feedback and Aerosol-Cloud Interaction
in CMIP6 Models, *Geophysical Research Letters*, 48, e2020GL091024, <https://doi.org/10.1029/2020GL091024>, 2021.
- Weaver, C., Wu, D. L., Bhartia, P. K., Labow, G., Haffner, D. P., Borgia, L., McBride, L., and Salawitch, R.: Comparison of
1770 Proxy-Shortwave Cloud Albedo from SBUV Observations with CMIP6 Models, *Journal of Climate*, 37, 3093-3116,
<https://doi.org/10.1175/JCLI-D-23-0170.1>, 2024.

- Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., Guo, H., Hurlin, W., Krasting, J., Knutson, T., Paynter, D., Silvers, L. G., and Zhang, R.: Climate Sensitivity of GFDL's CM4.0, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001838, <https://doi.org/10.1029/2019MS001838>, 2020.
- 775 Wolter, K. and Timlin, M. S.: Monitoring ENSO in COADS with a seasonally adjusted principal component index, 1993.
- Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext), *International Journal of Climatology*, 31, 10.1002/joc.2336, 2011.
- Zelinka, M. D., Smith, C. J., Qin, Y., and Taylor, K. E.: Comparison of methods to estimate aerosol effective radiative forcings in climate models, *Atmos. Chem. Phys.*, 23, 8879-8898, 10.5194/acp-23-8879-2023, 2023.
- 780 Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, *Geophysical Research Letters*, 47, 10.1029/2019GL085782, 2020.
- Zhang, H.-M., Lawrimore, J., Huang, B., Menne, M., Yin, X., Sánchez-Lugo, A., Gleason, B., Vose, R., Arndt, D., Rennie, J., and Williams, C.: Updated Temperature Data Give a Sharper View of Climate Trends, *Eos*, 100, 10.1029/2019eo128229, 2019.
- 785 Zhou, J. and Tung, K. K.: Deducing multidecadal anthropogenic global warming trends using multiple regression analysis, *Journal of the Atmospheric Sciences*, 70, 10.1175/JAS-D-12-0208.1, 2013.