https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

Inferred effective climate sensitivity from a selection of CMIP6 models

Ragnhild Bieltvedt Skeie[1], Terje K. Berntsen[2], Ada Gjermundsen[2], Ragnar Bang Huseby[3], Gunnar Myhre[1], Dirk Olivié[4] and Trude Storelvmo[2]

[1]CICERO Center for International Climate Research, P.O. Box 1129 Blindern, 0318 Oslo, Norway

[2]Department of Geosciences, University of Oslo, P.O. Box 1047 Blindern, 0316 Oslo, Norway

[3]Norwegian Computing Center, P.O. Box 114 Blindern, 0314 Oslo, Norway

[4] Norwegian Meteorological Institute, Oslo, Norway

Correspondence to: Ragnhild Bieltvedt Skeie (r.b.skeie@cicero.oslo.no)

Abstract.
Climate sensitivity is a key metric in climate science, defined as the equilibrium global mean surface temperature response following a doubling of the atmospheric CO_2 concentration. Climate sensitivity, which is an idealized concept, can be quantified using different approaches, all with their own limitations. In this study, we employ a method previously used to infer climate sensitivity based on Earth system observations and radiative forcing estimates over the industrialized period, on historical climate model simulations. We estimate climate sensitivity using the transient response in historical simulations from four models contributing to the sixth phase of the Coupled Model Intercomparison Project (CMIP6), with a total of 40 ensemble members. The surface temperature and ocean heat content evolution from each ensemble member is known, effective radiative forcing time series for each model are diagnosed from targeted simulations within CMIP6 and the effective climate sensitivity (ECS_{eff}) for each model is calculated from the abrupt four times CO_2 experiment. The inferred effective climate sensitivity (ECS_{inf}) from the transient model response is generally lower than the models' ECS_{eff}, as expected due to time and state dependencies of climate sensitivity. For different ensemble members, the ECS_{inf} can be quite different. For CNRM-CM6-1, the posterior mean ECS_{inf} varies by 2.2 K among its 30 ensemble members, but the 90% confidence intervals all contain the model's ECS_{eff}. The posterior mean ECS_{inf} for HadGEM3 GC3.1-LL varies by as much as 4 K. For NorESM2-LM and NorESM2-MM where the ECS_{eff} is 2.5 K for both models, the posterior mean ECS_{inf} varies by 0.62 K and 1.2 K respectively, and only one of the in total six ensemble members has posterior mean ECS_{inf} larger than ECS_{eff}. Because the observed historical climate represents only one realization of the Earth's climate, these results highlight the importance of natural variability when estimating climate sensitivity using observations over the historical period and not only the uncertainty in the historical forcing due to aerosols.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

1 Introduction

A central quantity in climate science is the equilibrium climate sensitivity defined as the equilibrium temperature response following a doubling of the CO_2 concentration in the atmosphere. Climate sensitivity encapsulates the strength of climate feedbacks in the Earth system and plays a pivotal role for future climate change projections. Reducing the uncertainty in the climate sensitivity is crucial for constraining estimates of necessary global mitigation measures to meet the Paris agreement as climate sensitivity is a dominant source of uncertainty in long-term climate projections under given greenhouse gas emission scenarios (Hawkins and Sutton, 2009). Based on several lines of evidence - including paleoclimate reconstructions, instrumental records, process understanding and emergent constraints - the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) assessed equilibrium climate sensitivity to very likely be in the range of 2 K to 5 K (Forster et al., 2021).

One method to estimate climate sensitivity is to use historical observations of the Earth system over the industrialized period combined with estimates of changes in the Earth energy balance (Knutti et al., 2017; Sherwood et al., 2020). However, using historical observations to constrain climate sensitivity comes with limitations (Armour et al., 2024; Gregory et al., 2020; Marvel et al., 2018; Sherwood et al., 2020). Firstly, the historical forcing of the system is uncertain, mainly due to the uncertainty in the aerosol radiative forcing (Andreae et al., 2005; Bellouin et al., 2020; Forster et al., 2026). Secondly, climate feedbacks are expected to change over time (Armour, 2017; Bjordal et al., 2020; Rugenstein et al., 2020; Senior and Mitchell, 2000). Among the mechanisms contributing to this time dependence is the so-called pattern effect, where feedbacks depend on the spatial pattern of warming (Stevens et al., 2016), and that the temperature pattern evolves over time (Andrews et al., 2015). Lastly, the considerable internal variability of the climate system means that the observed climate change is only one single realization out of many possible historical climate trajectories. According to the climate models the observed sea surface temperature (SST) warming pattern results in more stabilizing feedbacks compared to the SST pattern that is expected for long-term warming due to increased CO_2 in the atmosphere (Andrews et al., 2022; Zhou et al., 2016). Hence, estimates of climate sensitivity based on changes over the historical period are expected to be lower than climate sensitivity estimated from the response to a quadrupling of CO_2 levels in the climate models (Andrews et al., 2018; Armour et al., 2024; Dong et al., 2021). However, it is currently unclear if the observed SST warming pattern is forced or a result of internal variability, as climate models struggle to reproduce it (Dhame et al., 2025; Wills et al., 2022). Therefore, observation-based estimates of climate sensitivities over the industrialized period will differ from the equilibrium climate sensitivity as assessed by AR6 (Forster et al., 2021) and we hereafter use the term inferred effective climate sensitivity (ECS_{inf}) for these estimates.

Climate models are useful tools for understanding the Earth System, the relationship between forcing and response being one example, but prior to the sixth phase of the Coupled Model Intercomparison Project (CMIP6), the historical radiative forcing in the models was not quantified. Within CMIP6 and the Radiative Forcing Model Intercomparison Project (RFMIP) there

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

were simulations targeted to quantify the effective radiative forcing (ERF) over the historical period (Pincus et al., 2016; Smith et al., 2020). For models that performed those simulations and quantified the ERF time evolution, output from the historical simulations can be used to investigate the usefulness of observation-based estimates of climate sensitivity. The first limitation to observation-based estimates of climate sensitivity mentioned above is thus eliminated as the ERF in the models will be better quantified than in the real world. Furthermore, by using results from different initial perturbation ensemble members, the role of internal variability in ECS_{inf} estimates can be investigated. Previously, using climate models, climate sensitivity estimates have been shown to vary widely due to internal natural variability (Dessler et al., 2018; Huber et al., 2014). Lastly, comparing ECS_{inf} to effective climate sensitivity (ECS_{eff}) calculated from coupled model experiments where atmospheric CO_2 concentrations are abruptly quadrupled from pre-industrial levels and held fixed (Zelinka et al., 2020), the influence of time and state dependencies on the ECS_{inf} can be assessed.

In this study we follow the same methodology as previously used to estimate ECS_{inf} from observations to infer the climate sensitivity from historical climate model simulations (1850-2014). We use the transient ocean heat content (OHC) and temperature response from the coupled historical simulations of four CMIP6 models and the corresponding ERF time series calculated from the individual models to estimate ECS_{inf} for each ensemble member. The ECS_{inf} is expected to be lower than the ECS_{eff} calculated from the abrupt increase in CO_2 concentrations (Forster et al., 2021), but how large is the difference between ECS_{eff} and ECS_{inf} for these models? As observed climate is only a single realization of the Earth's climate, how large is the spread in ECS_{inf} between the different ensemble members? Answering these questions can give new insights into the usefulness of observational based estimates of climate sensitivity.

2 Method

For each ensemble member of the historical simulations from four climate models, ECS_{inf} is estimated using the same approach as previously applied to historical observations over the industrialized period (Aldrin et al., 2012; Skeie et al., 2014, 2018, 2024). In the Bayesian framework the historical observations of surface temperature and OHC as well as estimates of the historical effective radiative forcing are replaced by the corresponding values from the historical climate model simulations.

2.1 Estimation method

The Bayesian estimation model consists of a dynamic process model, a data model that describes how various observations – in this case results from climate model simulations – are related to the process states, and a parameter model that expresses our prior knowledge of the parameters. The dynamic process model can be written as

$$\boldsymbol{g}_t = \boldsymbol{m}_t(\boldsymbol{x}_{1750:t}, \boldsymbol{\theta}) + \boldsymbol{n}_t, \tag{1}$$

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

where $\boldsymbol{m}_t(\boldsymbol{x}_{1750:t}, \boldsymbol{\theta})$ is a deterministic energy balance/upwelling-diffusion model (Sandstad et al., 2024) with idealized representation of the Earth's energy balance, $\boldsymbol{x}_{1750:t}$ are ERFs from start year until year t, $\boldsymbol{\theta}$ is a vector of seven parameters including the climate sensitivity parameter (λ) and parameters that represent mixing of heat into the ocean, and $\boldsymbol{n_t}$ is a stochastic process, consisting of three terms representing short-term internal variability where an ENSO index is used, long-term internal variability and model error. In this study $\boldsymbol{g}_t$ corresponds to climate model data from historical simulations (from 1850 to 2014) for hemispheric surface temperatures and OHC for both 0-700 meters and below 700 meters. The prior distribution of the model parameters $\boldsymbol{\theta}$ are equal for all the climate model simulations and the same as in Skeie et al. (2024), and we apply a Bayesian approach and use Markov Chain Monte Carlo techniques to sample from the posterior distribution as described in previous studies (Aldrin et al., 2012; Skeie et al., 2014, 2018, 2024).

2.2 CMIP6 data

Model output from coupled historical simulations within CMIP6 (Eyring et al., 2016) are used in this study. The models used are HadGEM3 GC3.1-LL and CNRM-CM6-1 that both share the same ocean model (Kuhlbrodt et al., 2023) and two versions of NorESM2 where the main difference is the atmospheric resolution. Model results from each ensemble member are used separately in the estimation, and the ensemble size ranges from 3 to 30 (Table 1).

Table 1: Overview of CMIP6 models used in this study, main model reference and the number of ensemble members used. The values for ECS_{eff} and ERF for doubling of CO_2 concentrations ($2xCO_2$) are both assessed in IPCC AR6 where ECS_{eff} values are the equilibrium climate sensitivity values from Table 7.SM.5 in Smith et al. (2021) and the ERF $2xCO_2$ from Table 7.2 in Forster et al. (2021).

Model	Ensemble members	ECS_{eff} [K]	ERF 2xCO2 [W m^{-2}]	Model reference
CNRM-CM6-1	30	4.83	4.01	Voldoire et al. (2019)
HadGEM3 GC3.1-LL	4	5.55	4.07	Williams et al. (2018)
NorESM2-LM	3	2.54	4.10	Seland et al. (2020)
NorESM2-MM	3	2.50	4.22	Seland et al. (2020)

For HadGEM3 GC3.1-LL and CNRM-CM6-1 the OHC for the individual ensemble members are taken from Kuhlbrodt et al. (2023) where OHC in the upper 700 meter, 700 meter to 2000 meter and below 2000 meter were calculated and detrended by subtracting the trend in the pre-industrial control simulation (piControl). In the estimation we combine the OHC for 700 meters to 2000 meters and the one below 2000 meters. For NorESM2-LM and NorESM2-MM the volume integrated OHC is calculated above and below 700 meters and the trend in the piControl for the corresponding segment to the historical simulation is subtracted. To be consistent with the method where historical observation-based data are used to estimate ECS_{inf} we use OHC data from 1950 and onwards in the estimation shown in Fig. S1.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

The surface temperature data used are from Nicholls et al. (2021) that have processed the gridded CMIP6 model output to global and regional averages. We have used the hemispheric temperature data over land (tas: near-surface air temperature) and ocean (ts: surface temperature, sea surface temperature over ocean) and combined these to produce blended mean surface temperature time series to be consistent with observed temperature time series. The hemispheric as well as the global temperature time series for the models and ensemble members are shown in Fig. S2. When historical observations have been used in the estimation, we add an observational error as there are uncertainties related to the method used to calculate global and hemispheric temperatures from observations with limited spatial coverage. Using climate model data, the temperature is known, but we add a small observational error of ±0.05 K. For OHC above and below 700 meter we add an observational error of ±0.25 10^{22} J.

To account for the temperature variability due to El Niño–Southern Oscillation (ENSO), the Southern Oscillation Index (SOI) is used. SOI is calculated as monthly mean sea level pressure difference between Tahiti and Darwin in the historical simulations. Values in the grid box nearest the two sites are used. The SOI-indexes for the four models are shown in Fig. S3.

The ERF time series for the individual models are diagnosed from CMIP6 RFMIP experiments (Smith et al., 2020), where effective radiative forcings are calculated from atmosphere-only simulations using sea-surface temperature and sea ice climatology from the models' own pre-industrial control simulation. For CNRM-CM6-1 and HadGEM3-GC3.1-LL ERF timeseries are available for all forces, greenhouse gases, aerosols and natural forcers as used in Kuhlbrodt et al. (2023). For NorESM2-LM ERF timeseries for all forcers, greenhouse gases, aerosols, natural forcers, land use and ozone are diagnosed. It is reasonable to assume that NorESM2-MM has similar forcing to NorESM2-LM as the two versions of the model differ mainly in atmospheric resolution and ERFs calculated for present day relative to pre-industrial as well as their 4xCO2 ERF are similar for both models (Seland et al., 2020; Smith et al., 2020). The sum of the individual ERF components do not add up to the ERF for all forcers as not all forcing components are individually calculated in RFMIP (Smith et al., 2020) in addition to possible non-linearities in the ERF when the individual ERFs are summed. Therefore, we introduce a residual component, which is the ERF for all forcers minus the sum of the individual forcing components included for the individual models. For aerosol ERF and land use, the global mean ERFs are split into hemispheric mean ERFs based on northern hemisphere and southern hemisphere contributions to the present day ERF in Smith et al. (2020). The ERF timeseries exhibit substantial variability, and as the estimation method is not sensitive to whether noise in the ERF time series is included or not (Skeie et al., 2024), the ERF time series (except natural forcing that includes volcanic eruptions) are smoothed using the locally weighted scatterplot smoothing function, with a smoothing fraction of 0.075. The ERF in the historical climate system is uncertain (Forster et al., 2024) while ERF diagnosed from model simulations should be more certain, within 0.1 W m^{-2} for a 30 year integration (Forster et al., 2016). For the individual forcing components, we therefore use a 90% uncertainty of ± 0.1 W m^{-2} at present day. The uncertainty scales linearly back in time to zero in 1849, the start year for the estimation where all the forcing

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

components are set to 0 W m^{-2}. Figure S4 shows the priors for the ERF time series from 1850 to 2014 used in the Bayesian estimation of the climate models' ECS_{inf}.

It is the climate sensitivity parameter (λ) that is estimated for each ensemble of the historical simulations. To present the results as ECS_{inf} the best estimate of the ERF for doubling of CO_2 concentrations (2xCO2) for the individual models (Table 1) is used. This value is compared to the models' ECS_{eff} (Table 1). The ECS_{eff} is calculated following the method by Gregory et al. (2004), using the first 150 years of a simulation with abrupt quadrupling of the CO_2 concentration. ECS_{eff} is derived from the linear regression of top-of-atmosphere net downward radiative flux anomalies against global mean surface air temperature anomalies and finally scaled by 0.5 to represent a doubling of the CO_2 concentration. The ECS_{inf} represent global mean surface temperature while ECS_{eff} represents global mean surface air temperature, and Skeie et al. (2024) discussed the implication of the different temperature definition that will slightly influence the climate sensitivity estimates.

3 Results

The estimated ECS_{inf} using results from coupled historical simulations combined with ERF time series from the four climate models are shown in Fig. 1.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

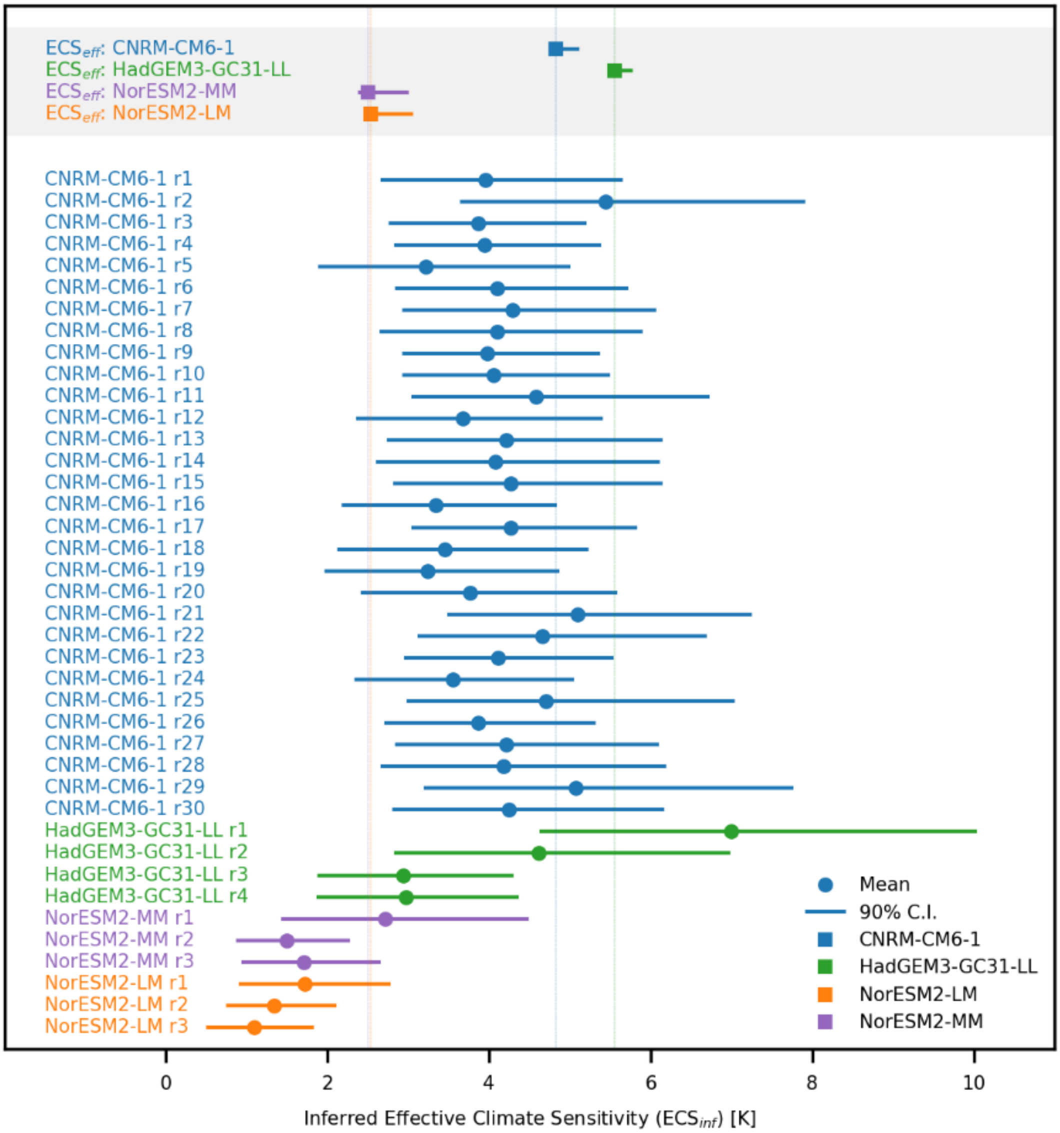


Figure 1: Posterior inferred effective climate sensitivity (ECS_{inf}) estimated using results from coupled historical simulations and effective radiative forcing time series over the historical period from four models. ECS_{inf} are estimated separately for each ensemble member. The posterior 5th to 95th percentile ranges are indicated by a solid line (90% C.I.) and the posterior mean values as a dot. The ECS_{eff} from IPCC AR6 (Table 1) for the individual models are shown as squares and a vertical line and the horizontal line indicate the 95% confidence interval using standard bootstrap approach (Zehrung et al., 2025).

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

For CNRM-CM6-1, the posterior mean ECS_{inf} ranges from 3.21 to 5.44 K. Out of the 30 ensemble members, only three of the members (r2, r21 and r29) have a posterior mean ECS_{inf} larger than the ECS_{eff} calculated from simulations with abrupt increase in CO_2 concentration of 4.83 K (Table 1). The ECS is within the estimated 90% C.I. posterior range of ECS_{inf} for all ensemble members, but for some of the members the ECS_{eff} is very close (less than 0.2 K difference) or at the upper 95th percentile of the posterior ECS_{inf} (r5, r16, r19). As expected, the model's ECS_{eff} are generally in the upper part of the posterior distribution of ECS_{inf} estimated from the historical model simulations. The mean of the differences between the posterior mean ECS_{inf} and ECS is -0.71 K (±0.53 K, one standard deviation) or -15 % lower than the ECS, but for 5 out of 30 ensemble members, the posterior mean ECS_{inf} are within ± 0.3 K (± 6%) of the ECS (r11, r21, r22, r25 and r29).

For HadGEM3 GC3.1-LL, the posterior distribution of ECS_{inf} varies widely between the four ensemble members. For two of the ensemble members (r3 and r4), the entire 90% C.I. is lower than the ECS_{eff} of 5.55 K (Table 1). For these two ensemble members the posterior mean ECS_{inf} is just below 3 K, around 2.5 K colder than the ECS_{eff}. For the two remaining ensemble members (r1 and r2), the posterior distributions are wider and contains the ECS_{eff} value. For r1 the posterior mean ECS_{inf} is 6.99 K, 1.44 K warmer than the ECS_{eff}, while for r2 the posterior mean ECS_{inf} is 0.94 K colder than the ECS_{eff}.

For the NorESM2 models, only the posterior mean ECS_{inf} for NorESM2-MM r1 is close to the ECS_{eff} calculated from the abrupt CO_2 simulation (Table 1). In two other ensemble members (r3 for NorESM2-MM and r1 for NorESM2-LM), the ECS_{eff} is within the ECS_{inf} 90% C.I., but for the remining three ensemble members the posterior distributions of ECS_{inf} are well below the models' ECS_{eff} values. As it takes a long time to equilibrate NorESM2, as greater storage of heat at depth in the Southern Ocean delays the surface warming and the cloud response (Gjermundsen et al., 2021), it is expected that ECS_{inf} is much lower than ECS_{eff} for NorESM2.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

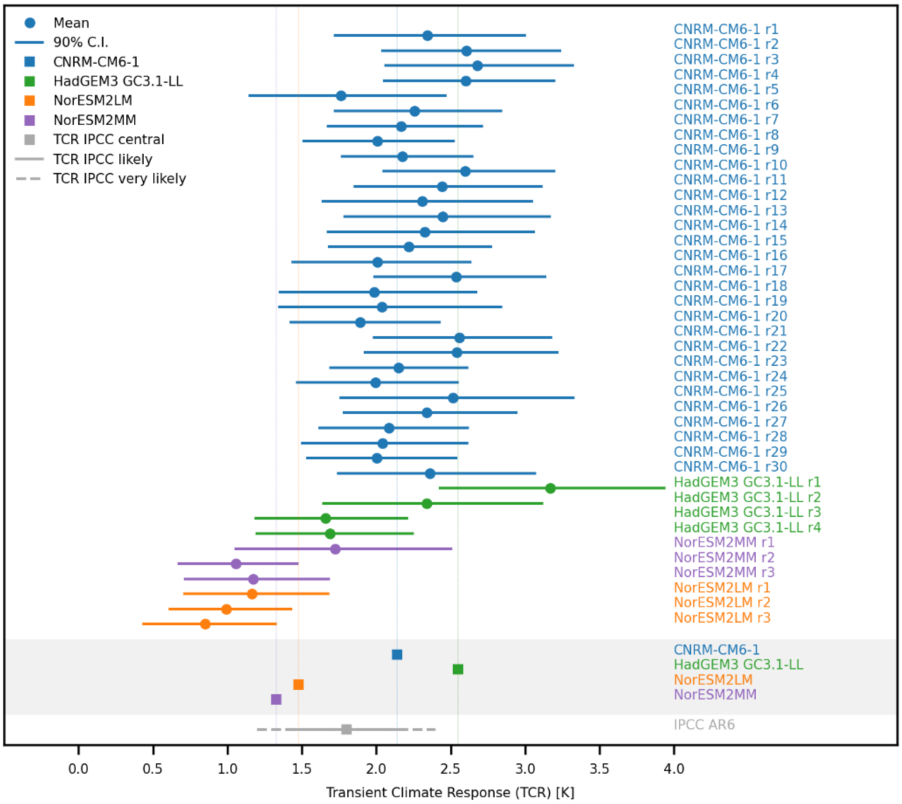


Figure 2: Posterior Transient Climate Response (TCR) calculated from the joint posterior distribution estimated using results from the coupled historical simulations and effective radiative forcing time series over the historical period from four models. TCR are estimated separately for each ensemble member. The 5th to 95th percentile ranges are indicated by a solid line (90% C.I.) and the mean values as a dot. The TCR from Table 7.SM.5 in IPCC AR6 (Smith et al., 2021) for the individual models are shown as squares and a vertical line. At the bottom the TCR central estimate, and likely (66% probability, solid line) and very likely (90% probability, dashed line) ranges from IPCC AR6 (Forster et al., 2021) are shown in grey colors.

The transient climate response (TCR), defined as the warming at the time of CO_2 doubling (after 70 years) following a 1% per year increase in CO_2 concentration, is expected to be less influenced by climate state and time than equilibrium climate sensitivity as it measures the transient response rather than an equilibrated state. From the joint posterior distribution of the model parameters and a time series of ERF for a 1% increase in CO_2 concentration, adjusted to the models own 2xCO2 ERF, the inferred transient climate response (TCR) can be calculated. For CNRM-CM6-1 the posterior mean inferred TCR estimate is more evenly distributed above and below the TCR calculated directly from climate model simulations (Fig. 2) than for ECS_{inf} and ECS_{eff} (Fig. 1). For 20 out of the 30 ensemble members from the CNRM model, the posterior mean TCRs were

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

larger than the models' TCR (Fig. 2) while for ECS_{inf} only three ensemble members had posterior means larger than the model's ECS_{eff}. For the other models, the general pattern of the difference between the posterior distributions of TCR and the TCR calculated from climate model simulations (Fig. 2) are like the difference between the posterior distribution of ECS_{inf} relative to the models' ECS_{eff} (Fig. 1).

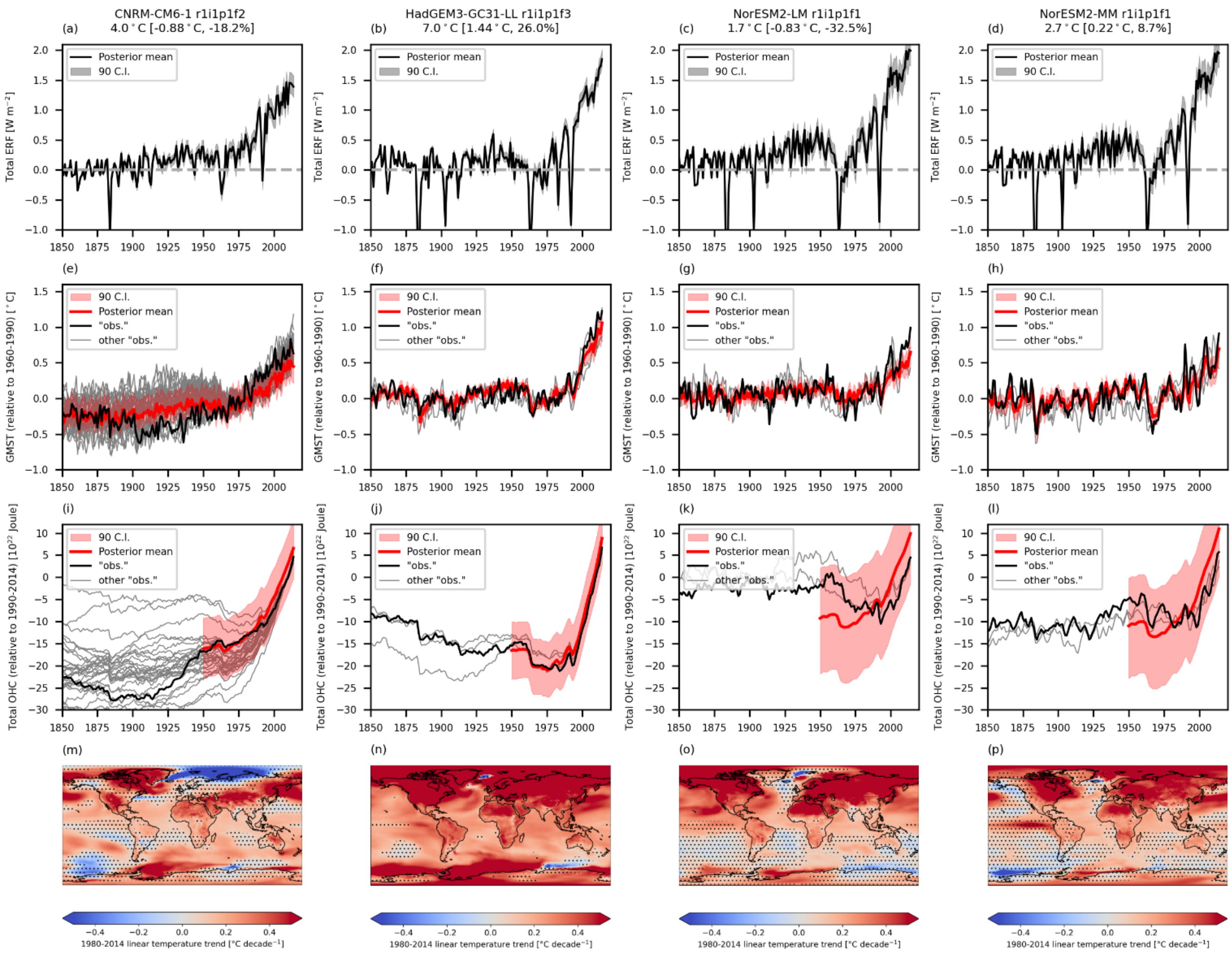


Figure 3: For the first ensemble member (r1) for the four models (each column) the posterior time series of total ERF (upper row), global mean surface temperature (second row), total OHC (third row) and the geographical distribution of the 1980 to 2014 linear temperature trend (bottom row). For the ERF (a to d), the posterior mean is shown as solid black line and the 90% C.I. as gray shading. For the global temperatures (e to h) the posterior mean is shown as red line and the 90% C.I. as red shading. The fitted values are the sum of the output from the deterministic energy balance model and the short term internal natural variability. The temperature is shown as anomalies relative to the 1960 to 1990 period. The black line ("obs.") is the corresponding temperature from the climate model simulation (r1), while the gray lines are the global mean temperature time series from the other ensemble members (other "obs."). The total global OHC is shown as anomalies relative to 1990 to 2014 (i to l). The total global OHC posterior mean is shown as red line and 90% C.I. as red shading, and the OHC from the corresponding ensemble member (r1) as black line. The other ensemble members are shown as gray lines. The 1980 to 2014 linear trend in near-surface air temperature (tas) (m to p)

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

are shown in the bottom row calculated using linear least-squares regression. The hatched areas mark where the linear trends are not significant at the 95 % level. The posterior mean ECS_{inf} for each ensemble member is given in the title for each column in addition to the absolute and relative deviation from the individual models ECS_{eff} (Table 1) in the parentheses. Similar figures for the rest of the ensemble members in Fig. S5 to S11.

For CNRM-CM6-1 with the largest number of ensemble members, there is a wide range of both global temperature evolution and OHC evolution among the individual ensemble members (black and gray lines in Fig. 3e,i). For temperature we use the entire time series from 1850 to 2014 in the estimation, while for OHC we use the period after 1950 to be consistent with the method using observations to estimate ECS_{inf}. For ensemble member r1, the temperature in the early 20th century is overestimated, while in the latter part of the century underestimated (Fig. 3e) and the posterior mean total OHC is slightly larger than the modelled total OHC (Fig. 3i). In this ensemble member the posterior mean ECS_{inf} is 0.88 K (18%) lower than the model's ECS_{eff}. For the ensembles with posterior mean ECS_{inf} like the ECS_{eff} (less than 0.3 K difference) the time evolution of the total OHC is well estimated (Fig. S5). For these ensemble members, the temperature evolution is quite well estimated except for r25 that has a quite different time evolution, without a temperature maximum in the first half of the 20th century as in the other six ensemble members (Fig. S5). For the other ensemble members where the posterior mean ECS_{inf} is lower than the ECS_{eff} (more than 0.3 K difference) how well the global mean temperature and OHC evolution is estimated varies (Fig. S6-S9).

For HadGEM3 GC3.1-LL, the rise in temperature after around year 2000 is underestimated but the temperature is well simulated prior to year 2000 for the first ensemble member (Fig. 3f). A similar pattern is seen for the other ensemble members, where posterior mean temperature increase since 1990s (r2 and r3) and since early 2000s (r4) are weaker than in the model (Fig. S10). The total OHC time evolution is well estimated in r1 (Fig. 3j), while for r3 and r4, the posterior mean OHC is overestimated in the 1980s and 1990s. These are the two ensemble members where the 90% C.I. for ECS_{inf} does not contain the model's ECS value (Fig. 1).

For NorESM2-LM, the total ERF increases following the volcanic eruption in 1963 (Fig. 3c), although for all the ensemble members, the total OHC continues to decline for decades, which is not captured in the estimated total OHC (Fig. S11m,n,o). The total OHC in one of the ensemble members (r2) has not returned to the 1950 level at the end of the simulation period (2014) (Fig. S11n). For NorESM2-MM, the decline in OHC is less pronounced, and the estimated total OHC is more similar to the total OHC from the model, but still, there are discrepancies (Fig. S11p,q,r). For the ensemble members with estimated ECS_{inf} similar to ECS (NorESM2-MM r1) the temperature evolution is well captured (Fig. 3h).

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

4 Discussion

One of the most central yet uncertain quantities in climate science is climate sensitivity, and one approach to estimate it is through the use of Earth system observations over the industrial era (Forster et al., 2021). In this study, we employ a method for observation-based estimates of climate sensitivity, but instead of Earth system observations we here use data from historical CMIP6 simulations. As expected, the posterior mean ECS_{inf} is generally lower than the models' own ECS_{eff} calculated from the abrupt quadrupling of atmospheric CO_2 following the method by Gregory et al. (2004), due to time and state dependence of the climate feedbacks (e.g. Armour, 2017; Senior and Mitchell, 2000). The results in this study also show that the spread in the ECS_{inf} between the different ensemble members is large, and the posterior mean differs by 0.62 K (NorESM2-LM), 1.2 K (NorESM2-MM), 2.2 K (CNRM-CM6-1) to as much as 4 K (HadGEM3-GC3.1-LL) for different ensemble members. For HadGEM3-GC3.1-LL the 90 % C.I. for different ensemble members did not even overlap. An important source of uncertainty in observation-based estimates of climate sensitivity arises from uncertainty in the radiative forcing, particularly associated with aerosols. Skeie et al. (2018) used observations up to 2014 and estimated a 90% C.I of 1.9K (97% of the best estimate of 1.95 K) which is smaller in absolute terms than the spread in CNRM-CM6-1 (2.2 K, 46% of ECS_{eff}) and HadGEM3-GC3.1-LL (4K, 73% of the ECS_{eff}) posterior means in this study which are solely driven by natural variability. Note that the 90% C. I. for ECS_{inf} might have been wider in absolute terms had the estimated central value been larger. As observed climate is only a single realization of the Earth's climate, the large spread in the posterior ECS_{inf} estimates for the different initial perturbation ensemble members highlight the importance of internal natural variability in addition to uncertainties in the historical forcing especially due to aerosols.

In the estimation, it is challenging to get a good fit for both the temperature evolution (1850 to 2014) and OHC (1950 to 2014) for all of the ensemble members (Fig. 3, Fig. S5-11) as all the ensembles, for a given climate model, are driven by the same prior forcing time series (Fig. S4). In addition, in the energy balance model used in the estimation (see Methods) neither the feedback nor the parameters controlling the storage of heat in the ocean change over time ($\boldsymbol{\theta}$ in equation 1 is not a function of time). In the following section we will discuss these three limitations: forcing, changes in the efficacy of ocean heat uptake and climate feedback over time.

The individual ERF time series are diagnosed from top-of-atmosphere radiation flux differences between two fixed SST and sea ice cover simulations, a control experiment with time evolving forcing components and a perturbation experiment where the emissions or the concentrations of a forcing agent are kept fixed at pre-industrial levels (Pincus et al., 2016). As not all possible forcing components have their own experiment, we have included a rest category, so the sum of the individual ERFs matches the total ERF (see Method). This category will also include possible non-linearities when ERFs are summed (Smith et al., 2020; Thornhill et al., 2021). Is the ERF, as implemented here for each model, representative of the forcing in all the ensemble members? The ERF may depend on the climate state (clouds etc.) and the different ensemble members do show

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

different temperature evolution (Fig. S2), but Forster et al. (2016) showed that the dependence on SST is small compared to the year-to-year random variation for the ERF. Due to natural variability, especially related to the cloud rapid adjustments there is noise in the ERF results (Forster et al., 2016). This is partially handled by smoothing the forcing time series, and the short term fluctuations in the ERF are shown to have minor impact on the posterior ECS_{inf} estimate using historical observations (Skeie et al., 2024). Calculations of ERFs in climate models for small perturbations are challenging (Forster et al., 2016), and the further back in time, the smaller the ERF signal. The uncertainty of 0.1 W m^{-2} added to year 2014 and scaled linearly back in time, may give a too certain ERF in the earlier period. Per definition the ERF is set to 0 W m^{-2} in 1849 and the anthropogenic ERF is small in the early period. Although we split the ERF into separate forcing components, the energy balance model in the estimation only sees the total ERF. For both the end year and 1950, the year where we start to include OHC in the estimation, the prior uncertainty in total ERF (90% C.I.) is greater than ±0.1 W m^{-2}. In 2014 the posterior estimated 90% C.I.s for the total ERF are ±0.20 W m^{-2}, ±0.20 W m^{-2} and ±0.23 W m^{-2} for CNRM-CM6-1, HadGEM3-GC3.1-LL and the NorESM2 models respectively, and for 1950 ±0.12 W m^{-2}, ±0.12 W m^{-2} and ±0.15 W m^{-2}. Despite these uncertainties, especially early in the record where the forcing signal is small, the total ERF applied in the estimation remains suitable.

In the estimation method, the efficiency of heat uptake does not change over time, as the parameter describing this is kept constant over the 1849 to 2014 period. In the estimation, we cannot reproduce the OHC time evolution in the NorESM2 models, with a positive model error for OHC above 700 meters prior to the 1980s and a negative model error after the 1980s (Fig. S12-13). There are limitations in the ability of the simplified model to reproduce the NorESM2 results, probably due to dynamical responses in the ocean circulation. The Atlantic meridional overturning circulation (AMOC) plays an important role in the heat transport in the Earth system and in NorESM2 AMOC slightly increases up to year 2000, and then rapidly decline (Seland et al., 2020). Also, in HadGEM3-GC3.1-LL, the AMOC varies systematically over the historical period where the AMOC declines after around year 1990 (Andrews et al., 2020). From the same time, the estimated warming in the northern hemisphere is underestimated, as less heat is transported down in the estimation compared to the full model, indicated by a positive model error term (Fig. S14). In the estimation we struggle to reproduce the systematic temporal distribution of OHC above and below 700 meters in all the ensembles for HadGEM3-GC3.1-LL, indicating structures that are not captured in our simple energy balance model. Also, for the total OHC, the posterior mean is larger than in the model from around 1960s (Fig. S10), especially for the two model ensembles with ECS_{inf} much lower than the models ECS_{eff} (r3, r4). For CNRM-CM6-1, there is little variability in AMOC and the ocean heat uptake variability is mainly linked to the Southern Ocean region (Kuhlbrodt et al., 2023). Over the entire period, the OHC shows large variability in this model (Fig. S1), and also for this model there are ensembles (r3, r4, r12) where the OHC decline after 1950s is not captured in the estimation (Fig. S7-S8). The long-term internal variability term also reflects large variability among the ensemble members (Fig. S15). The ocean-atmosphere system is coupled, where forcing, feedbacks and storage of heat in the system are interconnected. The change in AMOC may be linked to aerosol forcing in HadGEM3-GC3.1-LL (Andrews et al., 2020) and the climate feedbacks are dependent on the ocean heat uptake through changing temperature patterns and vice versa (Gjermundsen et al., 2021; Newsom

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

et al., 2020; Singh et al., 2022). Overall, these discrepancies highlight the limits of a simplified energy-balance model in capturing ocean-circulation-driven variations in heat uptake in the climate models.

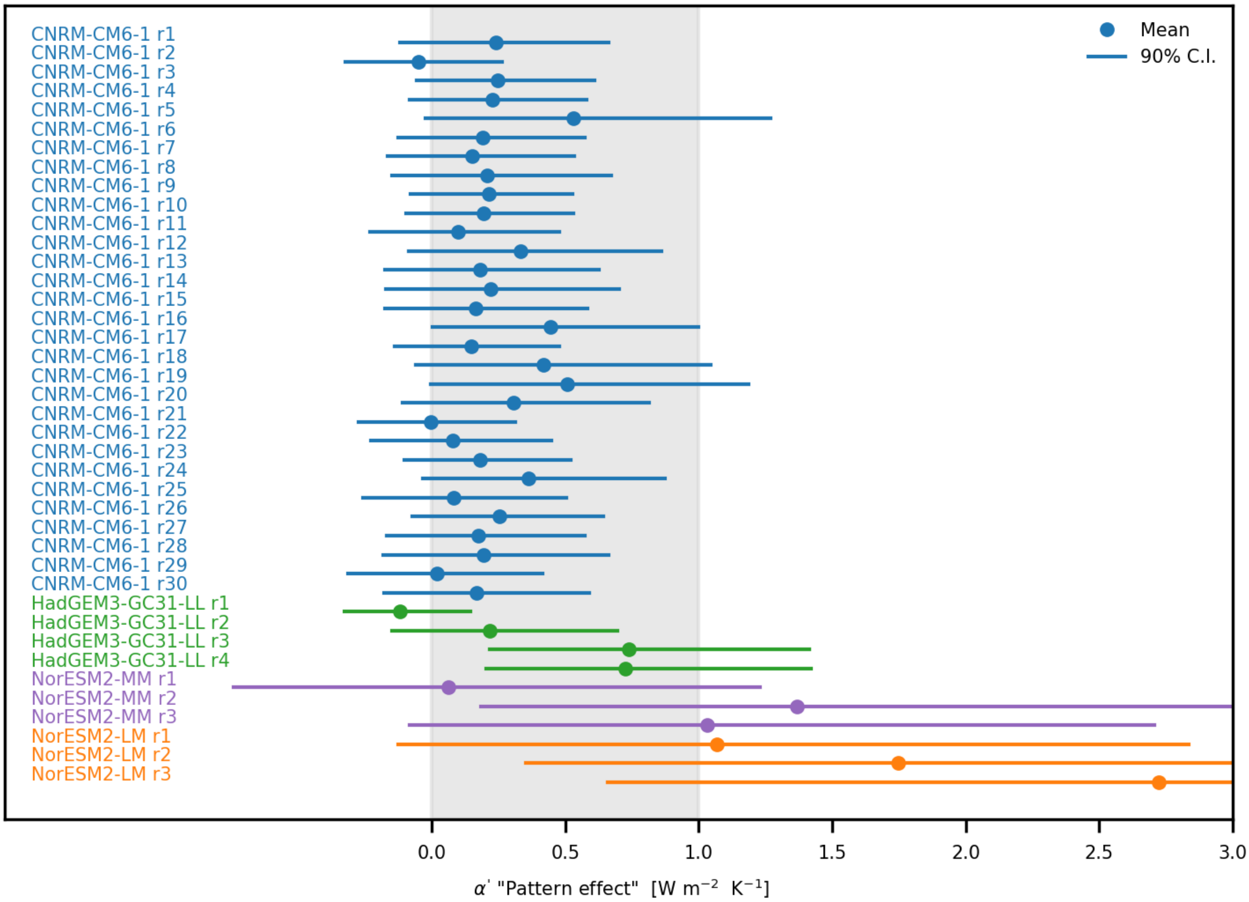
As the climate feedbacks may be state- and/or time dependent (Armour, 2017; Senior and Mitchell, 2000) and typically become more positive with time/warming, partly due to changes in temperature pattern (Andrews et al., 2015, 2022; Dong et al., 2021), it is expected that ECS_{inf} is lower than ECS_{eff} in the majority of the ensemble members included (Armour et al., 2024) as shown in Fig. 1. Over the last three to four decades, the observed temperature pattern differs from what the climate models simulate over the same period as well as what is expected for a CO_2 increase according to the climate models (Armour et al., 2024; Dong et al., 2021; Wills et al., 2022). In the climate models, the observed SST pattern is shown to dampen the feedbacks compared to the modeled temperature pattern (Dong et al., 2021), a reason for lower estimates of climate sensitivity using historical observations compared to other methods (Forster et al., 2021; Gregory and Andrews, 2016). We will therefore have a closer look at the geographical pattern of the temperature trend for the period 1980 to 2014 in the individual model ensembles (Fig. 3m-p, Fig. S5-11). For HadGEM3-GC3.1-LL, r1 and r2 have similar temperature patterns, a significant warming over most of the globe especially at high latitudes and warming in the eastern Pacific (Fig. S10). This is a pattern similar to what is expected for a long-term increase in CO_2 according to the climate models (Dong et al., 2021), and for these two ensembles, the ECS is within the posterior 90% C.I. of ECS_{inf} (Fig. 1). The two other ensembles (r3 and r4) where the estimated ECS_{inf} is much lower than ECS (Fig. 1), the temperature pattern shows an increase in the northern hemisphere particularly over land, while less warming in the Pacific and southern hemisphere (Fig. S10). This might indicate a larger role of the "pattern effect" in these two ensembles. For the NorESM2 models, NorESM2-MM r1 is the ensemble that has ECS_{inf} closest to the model's ECS_{eff} (Fig. 1). This is the ensemble with the largest temperature increase in central Pacific, although it is not significant, and a general warming trend at high latitudes, whereas the other two ensembles of NorESM2-MM exhibit cooling trends in parts of the high-latitude regions (Fig. S11). For CNRM-CM6-1 there are large variations between the ensemble members in the temperature pattern over the period 1980 to 2014, also within the ensembles where ECS_{inf} posterior mean is similar to ECS (Fig. S5) and where ECS_{inf} posterior mean is much lower than ECS (Fig. S7). The different temperature trend pattern among the different ensemble members (Fig. S5-11) and the resulting spread in the ECS_{inf} estimate (Fig. 1), highlight the importance of acknowledging that only one realization of the Earth's climate can be used in observation-based estimation method of climate sensitivity.

The "pattern effect" in each ensemble member can be quantified following the definition in IPCC AR6, where the "pattern effect" is presented as the change in the feedback parameter between the historical period and the time of equilibrium for a two times CO_2 forcing (α') (Forster et al., 2021). The equilibrium climate sensitivity can then be written as $ECS = -\Delta F_{2xco2}/(\alpha + \alpha')$, where α is the effective feedback parameter estimated over the historical period and ΔF_{2xco2} the 2xCO2 forcing. The change in the feedback parameter between the historical period and the time of equilibrium can be written as

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

$$\alpha' = \frac{1}{\lambda} - \frac{1}{\lambda_{model}}$$

where $\lambda_{model} = ECS_{eff}/\Delta F_{2xco2}$ can be calculated from Table 1 and λ is the estimated climate sensitivity parameter.



370 **Figure 4: The change in feedback parameter (α') between the estimated feedback parameter over the historical model in the individual ensemble members for the four climate models and the climate feedback calculated from ECS_{eff} and 2xCO2 forcing (using values in Table 1). The 5th to 95th percentile ranges are indicated by a solid line (90% C.I.) and the mean values as a dot. The IPCC AR6 assessed range for the pattern effect is indicated by gray shading.**

375 Figure 4 shows the posterior distribution of α' which is the change in the feedback parameter estimated over the historical period and at equilibrium. The IPCC AR6 assessed α' or the so called "pattern effect" to be in the range 0.0 to 1.0 W m^{-2} K^{-1} but with a low confidence in the upper end of this range (Forster et al., 2021). For CNRM-CM6-1, the posterior mean "pattern effect" is mostly within the lower part of the assessed range from IPCC AR6. For HadGEM3-GC3.1-LL, the posterior mean ranges from –0.1 to 0.7 Wm^{-2} K^{-1}, which is very similar to the range of –0.2 to +0.6 W m^{-2} K^{-1} previously found for the same

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

model (Andrews et al., 2019; Forster et al., 2021). This indicates that the estimation method captures features of the model response. For the NorESM2 models, the posterior distribution of the "pattern effect" span very large values, with posterior mean values ranging from close to zero to above 2.5 $Wm^{-2} K^{-1}$. For five out of six ensemble members, the posterior mean pattern effect is larger than 1 $Wm^{-2} K^{-1}$ at the upper end of the IPCC assessed range, which may be related to the very long time needed to equilibrate NorESM2. Gjermundsen et al. (2021) found that calculating ECS_{eff} by including 500 years rather than 150 years of the NorESM2-LM abrupt quadrupling of CO_2 simulation, increased the estimate by nearly 5 K.

Using the MPI-ESM1.1 model and 100 ensemble member simulations where only the initial condition differed, Dessler et al. (2018) found that inferred climate sensitivity ranged from 2.1 K to 3.9 K considering the time period 1850 to 2015 and the unforced pattern effect dominated over the forced pattern effect (Dessler, 2020). Internal variability plays an important role in model simulations over the historical period as also shown here in this study, as well as for uncertainties in climate change projections (Hawkins and Sutton, 2009; Jain et al., 2023) and evaluations of modelled hindcast trends (Thomas et al., 2025). As the historical evolution of the Earth's climate is only one out of many possible realizations of the Earth's transient climate, the internal variability has also played an important role over the industrialized period. This must be acknowledged when estimating climate sensitivity from historical observations of temperature and OHC.

5 Conclusions

In this study, results from historical free running simulations from four CMIP6 models are used to estimate the models' climate sensitivities from the transient response, to gain insights into estimates of climate sensitivity using historical Earth system observations. The estimated mean inferred climate sensitivity from different ensemble members is in general lower than the ECS_{eff} calculated from an abrupt quadrupling of CO_2 simulations, as expected due to time- and state dependencies of the climate feedbacks. Especially for the NorESM2 models, which are known to have a long equilibration time, the difference between the feedback estimated over the historical period and at equilibrium is larger than a factor of two for 5 out of 6 ensemble members. For CNRM-CM6-1, the posterior mean ECS_{inf} estimates differs by 2.2 K among the 30 ensemble members, but the 90% C.I.s all contain the model's ECS_{eff}. However, only 3 ensemble members have posterior mean ECS_{inf} greater than the ECS_{eff}. For HadGEM3-GC3.1-LL, the spread is even larger and the posterior mean ECS_{inf} differs by 4 K, and for two out of four ensemble members the model's ECS_{eff} is not included in the ECS_{inf} posterior 90% C.I. There are structural limitations with the simplified energy balance model used in the estimation with no time evolution of the climate feedback and ocean heat transport which limits the capability to reproduce the different evolution of temperature and OHC from all historical ensembles that are driven by the same forcing history. There is a tendency that if the estimated OHC and temperature response is more similar to the modeled response and the temperature pattern more in line with what is expected for a long-term CO_2 increase, the ECS_{inf} is in better agreement with ECS_{eff} from the model. The OHC evolution shows large spread in the possible outcome for the different ensemble members where only the initial conditions differ, and this highlights the importance of not only

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

relying on a single ensemble member in evaluation of historical and future ocean heat content (as in e.g. Cheng et al. (2022)) but rather use the full ensemble set. For CNRM-CM6-1, which includes 30 ensemble members and where OHC is relatively well estimated, the inclusion of the ECS_{eff} within the 90% C.I. of ECS_{inf} of all ensemble members, provides confidence in the method to constrain ECS_{inf} from historical observations. However, it is important to acknowledge the substantial spread in the mean ECS_{inf} of 2.2 K across the ensemble members.

This study used model data up to 2014 and a limited set of ensemble members for estimating the models' inferred climate sensitivity following the same method as previously used for observational constraints on climate sensitivity. The large spread in ECS_{inf} attributable to internal variability under known forcing conditions indicates that internal variability outweighs uncertainties in aerosol forcing in analyses of historical climate sensitivity, at least through 2014. Extending the time period of the analysis is expected to give more robust results for the estimated climate sensitivity (Urban et al., 2014). With time, the forced temperature response will emerge more clearly from internal variability and the signal-to-noise ratio for the ERF time series will increase in the future with a larger signal (Forster et al., 2016). The CO_2 ERF has already increased by 0.39 W m^{-2} since 2014 (Forster et al., 2026) and will continue to increase over the next decades. Although there are structural limitations with the simplified energy balance model and limitations within Earth System models, especially related to the limited capability of the models to capture the observed temperature trend pattern (Armour et al., 2024; Dhame et al., 2025; Wills et al., 2022), learning using Earth System Model experiments – including scenarios – may guide when observation-based estimates of ECS_{inf} from the single realization of the Earth's climate will be robust. Climate sensitivity is a key uncertainty in climate science and estimates from different angles, including Earth system observation, are important.

Code availability

The code to reproduce the figures in this manuscript can be found at https://github.com/ragnhibs/ecsinf_from_models/tree/main.

Data availability

The data needed to reproduce the figures in the manuscript can be found at https://github.com/ragnhibs/ecsinf_from_models/tree/main and will be made available in the NIRD Research Data Archive upon publication. The gridded CMIP6 temperature data used in Fig. 2 and surface pressure data for SOI calculations are publicly available from the Earth System Grid Federation (ESGF) at https://esgf-node.llnl.gov/search/cmip6/.

Author contribution

RBS wrote the paper and made the figures. RBS, TB, GM, RBH and TS discussed the design of the study. RBS prepared the data. RBH performed the Bayesian estimation. DO and AG contributed with NorESM2 input. All co-authors contributed to the writing.

Competing interests

No conflict of interests.

Acknowledgements

The work was funded through the Norwegian Research Council project (grant number 314997) and by the European Union's Horizon 2020 research and innovation program under grant agreement no. 821205 (FORCeS). Data will be uploaded and shared through the National Infrastructure for High-Performance Computing and Data Storage in Norway services (project no. NS9188K). We kindly acknowledge the climate modeling groups for contributing to these CMIP6 models and simulations and Till Kuhlbrodt for making the OHC and ERF data available for HadGEM3 GC3.1-LL and CNRM-CM6-1.

References

Aldrin, M., Holden, M., Guttorp, P., Skeie, R. B., Myhre, G., and Berntsen, T. K.: Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content, Environmetrics, 23, 253–271, https://doi.org/10.1002/env.2140, 2012.

Andreae, M. O., Jones, C. D., and Cox, P. M.: Strong present-day aerosol cooling implies a hot future, Nature, 435, 1187–1190, https://doi.org/10.1038/nature03671, 2005.

Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., Burke, E., Dittus, A. J., Florek, P., Gray, L. J., Haddad, S., Hardiman, S. C., Hermanson, L., Hodson, D., Hogan, E., Jones, G. S., Knight, J. R., Kuhlbrodt, T., Misios, S., Mizielinski, M. S., Ringer, M. A., Robson, J., and Sutton, R. T.: Historical Simulations With HadGEM3-GC3.1 for CMIP6, Journal of Advances in Modeling Earth Systems, 12, e2019MS001995, https://doi.org/10.1029/2019MS001995, 2020.

Andrews, T., Gregory, J. M., and Webb, M. J.: The Dependence of Radiative Forcing and Feedback on Evolving Patterns of Surface Temperature Change in Climate Models, https://doi.org/10.1175/JCLI-D-14-00545.1, 2015.

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, Geophysical Research Letters, 45, 8490–8499, https://doi.org/10.1029/2018GL078887, 2018.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

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.

Andrews, T., Bodas-Salcedo, A., Gregory, J. M., Dong, Y., Armour, K. C., Paynter, D., Lin, P., Modak, A., Mauritsen, T., Cole, J. N. S., Medeiros, B., Benedict, J. J., Douville, H., Roehrig, R., Koshiro, T., Kawai, H., Ogura, T., Dufresne, J.-L., Allan, R. P., and Liu, C.: On the Effect of Historical SST Patterns on Radiative Feedback, Journal of Geophysical Research: Atmospheres, 127, e2022JD036675, https://doi.org/10.1029/2022JD036675, 2022.

Armour, K. C.: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, Nature Clim Change, 7, 331–335, https://doi.org/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, https://doi.org/10.1073/pnas.2312093121, 2024.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. -L., Dufresne, J. -L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, e2019RG000660, https://doi.org/10.1029/2019RG000660, 2020.

Bjordal, J., Storelvmo, T., Alterskjær, K., and Carlsen, T.: Equilibrium climate sensitivity above 5 °C plausible due to state-dependent cloud feedback, Nat. Geosci., 13, 718–721, https://doi.org/10.1038/s41561-020-00649-1, 2020.

Cheng, L., von Schuckmann, K., Abraham, J. P., Trenberth, K. E., Mann, M. E., Zanna, L., England, M. H., Zika, J. D., Fasullo, J. T., Yu, Y., Pan, Y., Zhu, J., Newsom, E. R., Bronselaer, B., and Lin, X.: Past and future ocean warming, Nat Rev Earth Environ, 3, 776–794, https://doi.org/10.1038/s43017-022-00345-1, 2022.

Dessler, A. E.: Potential Problems Measuring Climate Sensitivity from the Historical Record, Journal of Climate, 33, 2237–2248, https://doi.org/10.1175/JCLI-D-19-0476.1, 2020.

Dessler, A. E., Mauritsen, T., and Stevens, B.: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity, Atmospheric Chemistry and Physics, 18, 5147–5155, https://doi.org/10.5194/acp-18-5147-2018, 2018.

Dhame, S., Olonscheck, D., and Rugenstein, M.: Higher-Resolution Climate Models Do Not Consistently Reproduce the Observed Tropical Pacific Warming Pattern, https://doi.org/10.1175/JCLI-D-24-0248.1, 2025.

Dong, Y., Armour, K. C., Proistosescu, C., Andrews, T., Battisti, D. S., Forster, P. M., Paynter, D., Smith, C. J., and Shiogama, H.: Biased Estimates of Equilibrium Climate Sensitivity and Transient Climate Response Derived From Historical CMIP6 Simulations, Geophysical Research Letters, 48, e2021GL095778, https://doi.org/10.1029/2021GL095778, 2021.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

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, edited by:

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

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., Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 923–1054, https://doi.org/10.1017/9781009157896.009, 2021.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., Andrews, T., Pincus, R., and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, Journal of Geophysical Research: Atmospheres, 121, 12,460-12,475, https://doi.org/10.1002/2016JD025320, 2016.

Forster, P. M., Smith, C., Walsh, T., Lamb, W. F., Lamboll, R., Hall, B., Hauser, M., Ribes, A., Rosen, D., Gillett, N. P., Palmer, M. D., Rogelj, J., von Schuckmann, K., Trewin, B., Allen, M., Andrew, R., Betts, R. A., Borger, A., Boyer, T., Broersma, J. A., Buontempo, C., Burgess, S., Cagnazzo, C., Cheng, L., Friedlingstein, P., Gettelman, A., Gütschow, J., Ishii, M., Jenkins, S., Lan, X., Morice, C., Mühle, J., Kadow, C., Kennedy, J., Killick, R. E., Krummel, P. B., Minx, J. C., Myhre, G., Naik, V., Peters, G. P., Pirani, A., Pongratz, J., Schleussner, C.-F., Seneviratne, S. I., Szopa, S., Thorne, P., Kovilakam, M. V. M., Majamäki, E., Jalkanen, J.-P., van Marle, M., Hoesly, R. M., Rohde, R., Schumacher, D., van der Werf, G., Vose, R., Zickfeld, K., Zhang, X., Masson-Delmotte, V., and Zhai, P.: Indicators of Global Climate Change 2023: annual update of key indicators of the state of the climate system and human influence, Earth System Science Data, 16, 2625–2658, https://doi.org/10.5194/essd-16-2625-2024, 2024.

Forster, P. M., Walsh, T., Smith, C., Lamb, W. F., Lamboll, R., Cassou, C., Hauser, M., Hausfather, Z., Lee, J.-Y., Palmer, M. D., Von Schuckmann, K., Slangen, A. B. A., Szopa, S., Trewin, B., Yun, J., Gillett, N. P., Jenkins, S., Matthews, H. D., Raghavan, K., Ribes, A., Rogelj, J., Rosen, D., Zhang, X., Allen, M., Andrew, R. M., Atkinson, C., Betts, R. A., Bombelli, A., Burgess, S. N., Cheng, L., Claxton, H. E., Friedlingstein, P., Frölicher, T. L., Domingues, C. M., Gasser, T., Gregory, C. H., Hoesly, R. M., Huppmann, D., Ishii, M., Kadow, C., Karwat, A., Kennedy, J., Killick, R. E., Kovilakam, M. V. M., Krummel, P. B., Lan, X., Lamarque, J.-F., Liné, A., Martín-Míguez, B., Monselesan, D. P., Morice, C., Mühle, J., Mussak, P., Peters, G. P., Pirani, A., Pongratz, J., Rigby, M., Rohde, R., Savita, A., Seneviratne, S. I., Smith, S. J., Taha, G., Tassone, C., Thorne, P., Wells, C., Western, L. M., Van Der Werf, G. R., Wijffels, S. E., Zecchetto, M., Zhong, J., Zhang, X.-Y., Masson-Delmotte, V., and Zhai, P.: Indicators of Global Climate Change 2025: annual update of key indicators of the state of the climate system and human influence, Earth Syst. Sci. Data, 18, 3889–3933, https://doi.org/10.5194/essd-18-3889-2026, 2026.

Gjermundsen, A., Nummelin, A., Olivié, D., Bentsen, M., Seland, Ø., and Schulz, M.: Shutdown of Southern Ocean convection controls long-term greenhouse gas-induced warming, Nat. Geosci., 14, 724–731, https://doi.org/10.1038/s41561-021-00825-x, 2021.

Gregory, J. M. and Andrews, T.: Variation in climate sensitivity and feedback parameters during the historical period, Geophysical Research Letters, 43, 3911–3920, https://doi.org/10.1002/2016GL068406, 2016.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, https://doi.org/10.1029/2003GL018747, 2004.

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?, Clim Dyn, 54, 129–157, https://doi.org/10.1007/s00382-019-04991-y, 2020.

Hawkins, E. and Sutton, R.: The Potential to Narrow Uncertainty in Regional Climate Predictions, https://doi.org/10.1175/2009BAMS2607.1, 2009.

Huber, M., Beyerle, U., and Knutti, R.: Estimating climate sensitivity and future temperature in the presence of natural climate variability, Geophysical Research Letters, 41, 2086–2092, https://doi.org/10.1002/2013GL058532, 2014.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

Jain, S., Scaife, A. A., Shepherd, T. G., Deser, C., Dunstone, N., Schmidt, G. A., Trenberth, K. E., and Turkington, T.: Importance of internal variability for climate model assessment, npj Clim Atmos Sci, 6, 68, https://doi.org/10.1038/s41612-023-00389-0, 2023.

Knutti, R., Rugenstein, M. A. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, Nature Geosci, 10, 727–736, https://doi.org/10.1038/ngeo3017, 2017.

Kuhlbrodt, T., Voldoire, A., Palmer, M. D., Geoffroy, O., and Killick, R. E.: Historical Ocean Heat Uptake in Two Pairs of CMIP6 Models: Global and Regional Perspectives, Journal of Climate, 36, 2183–2203, https://doi.org/10.1175/JCLI-D-22-0468.1, 2023.

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, 1595–1601, https://doi.org/10.1002/2017GL076468, 2018.

Newsom, E., Zanna, L., Khatiwala, S., and Gregory, J. M.: The Influence of Warming Patterns on Passive Ocean Heat Uptake, Geophysical Research Letters, 47, e2020GL088429, https://doi.org/10.1029/2020GL088429, 2020.

Nicholls, Z., Lewis, J., Makin, M., Nattala, U., Zhang, G. Z., Mutch, S. J., Tescari, E., and Meinshausen, M.: Regionally aggregated, stitched and de-drifted CMIP-climate data, processed with netCDF-SCM v2.0.0, Geoscience Data Journal, 8, 154–198, https://doi.org/10.1002/gdj3.113, 2021.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

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, e2019GL083898, https://doi.org/10.1029/2019GL083898, 2020.

Sandstad, M., Aamaas, B., Johansen, A. N., Lund, M. T., Peters, G. P., Samset, B. H., Sanderson, B. M., and Skeie, R. B.: CICERO Simple Climate Model (CICERO-SCM v1.1.1) – an improved simple climate model with a parameter calibration tool, Geoscientific Model Development, 17, 6589–6625, https://doi.org/10.5194/gmd-17-6589-2024, 2024.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Senior, C. A. and Mitchell, J. F. B.: The time-dependence of climate sensitivity, Geophysical Research Letters, 27, 2685–2688, https://doi.org/10.1029/2000GL011373, 2000.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of Geophysics, 58, e2019RG000678, https://doi.org/10.1029/2019RG000678, 2020.

Singh, H., Feldl, N., Kay, J. E., and Morrison, A. L.: Climate Sensitivity is Sensitive to Changes in Ocean Heat Transport, https://doi.org/10.1175/JCLI-D-21-0674.1, 2022.

Skeie, R. B., Berntsen, T., Aldrin, M., Holden, M., and Myhre, G.: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series, Earth System Dynamics, 5, 139–175, https://doi.org/10.5194/esd-5-139-2014, 2014.

Skeie, R. B., Berntsen, T., Aldrin, M., Holden, M., and Myhre, G.: Climate sensitivity estimates – sensitivity to radiative forcing time series and observational data, Earth System Dynamics, 9, 879–894, https://doi.org/10.5194/esd-9-879-2018, 2018.

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 System Dynamics, 15, 1435–1458, https://doi.org/10.5194/esd-15-1435-2024, 2024.

Smith, C., Nicolls. Z.R.J. Nicholls, K. Armour, W. Collins, P. Forster, M. Meinshausen, M.D. Palmer, and M., and Panel on Climate Change: 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, Cambridge University Press, 923–1054, https://doi.org/10.1017/9781009157896.009, 2021.

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J.-L., Nabat, P., Michou, 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., Olivié, D., Fiedler, S., Lewinschal, A., Mackallah, C., Dix, M., Pincus, R., and Forster, P. M.: Effective radiative forcing and adjustments in CMIP6 models, Atmospheric Chemistry and Physics, 20, 9591–9618, https://doi.org/10.5194/acp-20-9591-2020, 2020.

Stevens, B., Sherwood, S. C., Bony, S., and Webb, M. J.: Prospects for narrowing bounds on Earth's equilibrium climate sensitivity, Earth's Future, 4, 512–522, https://doi.org/10.1002/2016EF000376, 2016.

Thomas, R., Woollings, T., and Dunstone, N.: The Role of Internal Variability in Seasonal Hindcast Trend Errors, Journal of Climate, 38, 5541–5553, https://doi.org/10.1175/JCLI-D-24-0367.1, 2025.

Thornhill, G. D., Collins, W. J., Kramer, R. J., Olivié, D., Skeie, R. B., O'Connor, F. M., Abraham, N. L., Checa-Garcia, R., Bauer, S. E., Deushi, M., Emmons, L. K., Forster, P. M., Horowitz, L. W., Johnson, B., Keeble, J., Lamarque, J.-F., Michou, M., Mills, M. J., Mulcahy, J. P., Myhre, G., Nabat, P., Naik, V., Oshima, N., Schulz, M., Smith, C. J., Takemura, T., Tilmes, S., Wu, T., Zeng, G., and Zhang, J.: Effective radiative forcing from emissions of reactive gases and aerosols – a multi-model comparison, Atmospheric Chemistry and Physics, 21, 853–874, https://doi.org/10.5194/acp-21-853-2021, 2021.

Urban, N. M., Holden, P. B., Edwards, N. R., Sriver, R. L., and Keller, K.: Historical and future learning about climate sensitivity, Geophysical Research Letters, 41, 2543–2552, https://doi.org/10.1002/2014GL059484, 2014.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E., Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1, Journal of Advances in Modeling Earth Systems, 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W., Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M.

https://doi.org/10.5194/egusphere-2026-4634
Preprint. Discussion started: 14 August 2026

J., Scaife, A. A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I. G., Walters, D. N., West, A., Wood, R. A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations, Journal of Advances in Modeling Earth Systems, 10, 357–380, https://doi.org/10.1002/2017MS001115, 2018.

Wills, R. C. J., Dong, Y., Proistosecu, C., Armour, K. C., and Battisti, D. S.: Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change, Geophysical Research Letters, 49, e2022GL100011, https://doi.org/10.1029/2022GL100011, 2022.

Zehrung, A., King, A. D., Nicholls, Z., Zelinka, M. D., and Meinshausen, M.: Standardising the “Gregory method” for calculating equilibrium climate sensitivity, Geoscientific Model Development, 18, 9433–9450, https://doi.org/10.5194/gmd-18-9433-2025, 2025.

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, e2019GL085782, https://doi.org/10.1029/2019GL085782, 2020.

Zhou, C., Zelinka, M. D., and Klein, S. A.: Impact of decadal cloud variations on the Earth’s energy budget, Nature Geosci, 9, 871–874, https://doi.org/10.1038/ngeo2828, 2016.