https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics Discussions

Relating Climate Sensitivity Indices to projection uncertainty

Benjamin Sanderson[1,2]
[1]CERFACS, Toulouse, France
[2]NCAR, Boulder CO, USA
Correspondence: Benjamin Sanderson (sanderson@cerfacs.fr)

Abstract. Can we summarize uncertainties in global response to greenhouse gas forcing with a single number? Here we assess the degree to which traditional metrics are related to future warming indices using an ensemble of simple climate models together with results from CMIP5 and CMIP6. We consider Effective Climate Sensitivity (EffCS), Transient Climate Response at CO_2 quadrupling (T140) and a proposed simple metric of temperature change 140 years after a quadrupling of carbon dioxide (A140). In a perfectly equilibrated model, future temperatures under RCP(Representative Concentration Pathway)8.5 are almost perfectly described by T140, whereas in a mitigation scenario such as RCP2.6, both ECS and T140 are found to be poor predictors of 21st century warming, and future temperatures are better correlated with A140. However, we show that T140 and EffCS calculated in full CMIP simulations are subject to errors arising from control model drift and internal variability. Simulating these factors in the simple model leads to 30% relative error in the measured value of T140, but only a 10% error for EffCS. As such, if starting from a non-equilibrated state, measured values of Effective Climate Sensitivity can be better correlated with true TCR than measured values of TCR itself. We propose that this could be an explanatory factor in the previously noted surprising result that EffCS is a better predictor than TCR of future transient warming under RCP8.5.

Introduction

Summarizing the response of the Earth System to anthropogenic forcers with a metrics has long been practised as a way to illustrate uncertainty in Earth system response to greenhouse gases. For example, the concept of the Equilibrium Climate Sensitivity (ECS), the equilibrium global mean temperature increase which would be observed in response to a doubling of atmospheric carbon dioxide concentrations (Hansen et al., 1984) has existed for over 50 years (Charney et al., 1979) and significant amount of literature has been devoted to constraining its value (Knutti et al., 2017).

The Earth System responds to a step-change in forcing on a range of timescales ranging from days to millennia (Knutti and Rugenstein, 2015), so an 'Effective Climate Sensitivity' (EffCS hereon) is often used as a proxy for decadal to centennial feedbacks. EffCS is generally calculated in a coupled atmosphere-ocean model from the output of the 'abrupt4xCO2' simulation, a standard experiment in which CO2 concentrations are quadrupled instantaneously from pre-industrial levels and the model is allowed to evolve (Gregory et al., 2004).

EffCS is calculated by assuming that a model is associated with a single feedback parameter (i.e. a rate of change of top of atmosphere radiative flux per unit surface temperature increase), allowing the equilibrium temperature response to a step change forcing to be predicted by linear extrapolation (we refer to this approach henceforth as the Constant Feedback (CF)

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics Discussions

approximation, with EffCS referring to the estimate of ECS made using this approach). Another metric, the Transient Climate Response at CO_2 doubling (TCR) or quadrupling (T140) is calculated from an '1pctCO2' idealized experiment in which CO_2 concentrations are increased by 1 percent each year, starting from a pre-industrial state, resulting in linearly increasing forcing.

Although it was generally assumed that TCR would be a better predictor of transient warming under a high emissions scenario such as RCP8.5 (Riahi et al., 2011), a complication has arisen due to the fact that EffCS seems to be better correlated with 21st century warming from present day levels under a business-as-usual scenario than TCR in the CMIP5 ensemble (Grose et al., 2018). The reason for this is not yet well understood given the radiative pathway in RCP8.5 leading up to 2100 is relatively similar to that of the 1 percent annual increase experiment used to measure T140. Another pressing concern is that neither EffCS nor TCR is well correlated to end of century temperatures in a mitigation scenario (Grose et al., 2018) such as RCP2.6 (Van Vuuren et al., 2011).

Similarly, a number have studies have shown that the EffCS approximation does not well describe the true equilibrium behaviour of most models(Knutti et al., 2017). When GCM abrupt-4xCO2 simulations are continued for thousands of years, many are found to deviate significantly from the linear trend-line one would fit to a 150 year simulation (Andrews et al., 2015; Knutti et al., 2017; Senior and Mitchell, 2000; Rugenstein et al., 2016).

The conceptual models representing the evolving feedbacks as a function of timescales vary slightly between studies - either modulating the efficacy of deep ocean heat uptake (Geoffroy et al., 2013; Winton et al., 2010; Held et al., 2010) or by representing the climate system as sum of warming patterns which emerge on different adjustment timescales (Armour et al., 2013; Rugenstein et al., 2016), each associated with their feedback parameter. However, the analytical set of solutions for the temperature response to a step change in forcing is the same in either case - a superposition of decaying exponential modes with different timescales varying between a few years and a few centuries (Proistosescu and Huybers, 2017). It has been shown that the implications of these additional degrees of freedom, and ambiguity over contributions from different timescales of response might imply that equilibrium climate sensitivity may not be strongly constrained by temperature change over the last century (Proistosescu and Huybers, 2017; Andrews et al., 2018), and that the Long Term Equilibrium (LTE) sensitivity may be greater than that implied by estimates which use the CF framework (Otto et al., 2013; Lewis, 2013).

This state of understanding leads to a number of emerging critical questions which we discuss in this paper - can we explain the non-intuitive result that EffCS is a better predictor then T140 of end-of-century temperatures under RCP8.5, which summary metrics of global sensitivity to greenhouse gas forcing are most useful for effective policy decisions, and do the implicit structural assumptions underpinning the very existence or applicability of these metrics to the real world cause us to mis-categorize and potentially underestimate future warming risk?

1 A simple model example

We begin by considering an idealized ensemble of climate model simulations. We use a two timescale thermal response model, conceptually representing the deep ocean (with a response timescale of a century or more) and shallow ocean response timescales (with a response timescale of 10 to 50 years). Such a model, although simple, is capable of resolving evolving

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions
Open Access

feedback amplitudes and can emulate the climatological responses of complex Earth System Models to on a range of timescales (Proistosescu and Huybers, 2017; Geoffroy et al., 2013).

The physical parameters of this simple model are constrained by historical carbon dioxide concentrations together with observed global mean temperatures from 1870 to present day (together with aggregate forcing estimates representing other anthropogenic emissions (Meinshausen et al., 2011), which are not the focus of this study). A Markov-Chain Monte Carlo algorithm is used to produce a posterior parameter distribution for the model which can then be used to project the corresponding range of response in probabilistic projections of the future scenarios or in idealized experiments (see additional material)

The resulting ensemble produces model variants with Effective Climate Sensitivities (to a doubling of CO2) ranging from 2.4 to 4.6K (5th and 95th percentiles), and values of TCR from 1.6 to 2.2K (Figure 1(b,e)). This results in a range of 21st century warming under two scenarios considered, RCP2.6(RCP8.5) 2100 warming ranges from 1.4 and 2.4 K (3.8 to 5.1K) respectively (5th and 95th percentiles, see Figure 1(a)). We then consider in the context of this observationally constrained ensemble of simple models, what idealized metrics of system response are most informative for describing 21st century warming. We consider four metrics: the EffCS, TCR/T140 (transient warming under an annual compounded 1 percent increase in CO_2 concentrations at time of CO_2 doubling/quadrupling, corresponding to years 70 and 140 of the simulation). We also introduce A140 as a possible metric for consideration, defined as the global mean warming above pre-emission levels in the abrupt4xCO2 simulation calculated 140 years after time of CO_2 quadrupling (here and throughout estimated as the mean from years 131-150). Figure 2 illustrates how ensemble spread would be impacted for a set of different scenarios if each of these metrics were constrained to lie within a narrow range (nominally the 45-55th percentile range of values present in the entire observationally constrained ensemble).

In the high emissions, RCP8.5 scenario (Riahi et al., 2011), 2000-2100 warming is nearly perfectly described ($R^2 = 0.99$) by T140, the transient climate response after 140 years in a 1 percent CO2 simulation (Figure 1(c) and Figure 2(k)). The corresponding response after only 70 years, TCR, is a much poorer predictor at $R^2 = 0.31$).

These results are physically intuitive. The climate forcing and rate of change of forcing in RCP8.5 at the end of the 21st century are of similar magnitude to those in year 140 of the 1 percent CO2 simulation, and so it is unsurprising that T140 is an efficient predictor for RCP8.5. TCR is a poor predictor in the simple model ensemble largely because TCR itself is already highly constrained by historical warming (Figure 1(e)), and thus the ensemble is effectively conditioned on a value of TCR and it has little additional explanatory value in explaining the ensemble variance in the RCP projections (Figure 2(f,g)).

EffCS and A140 are also well correlated with the RCP8.5 warming ($R^2 = 0.77$ and 0.76 respectively), but less so than T140. For the mitigation scenario RCP2.6, the most effective predictor of 2000-2100 warming is A140 ($R^2 = 0.91$). Both EffCS and T140 are weakly correlated ($R^2 = 0.62$ and 0.65 respectively), and TCR shows no significant correlation.

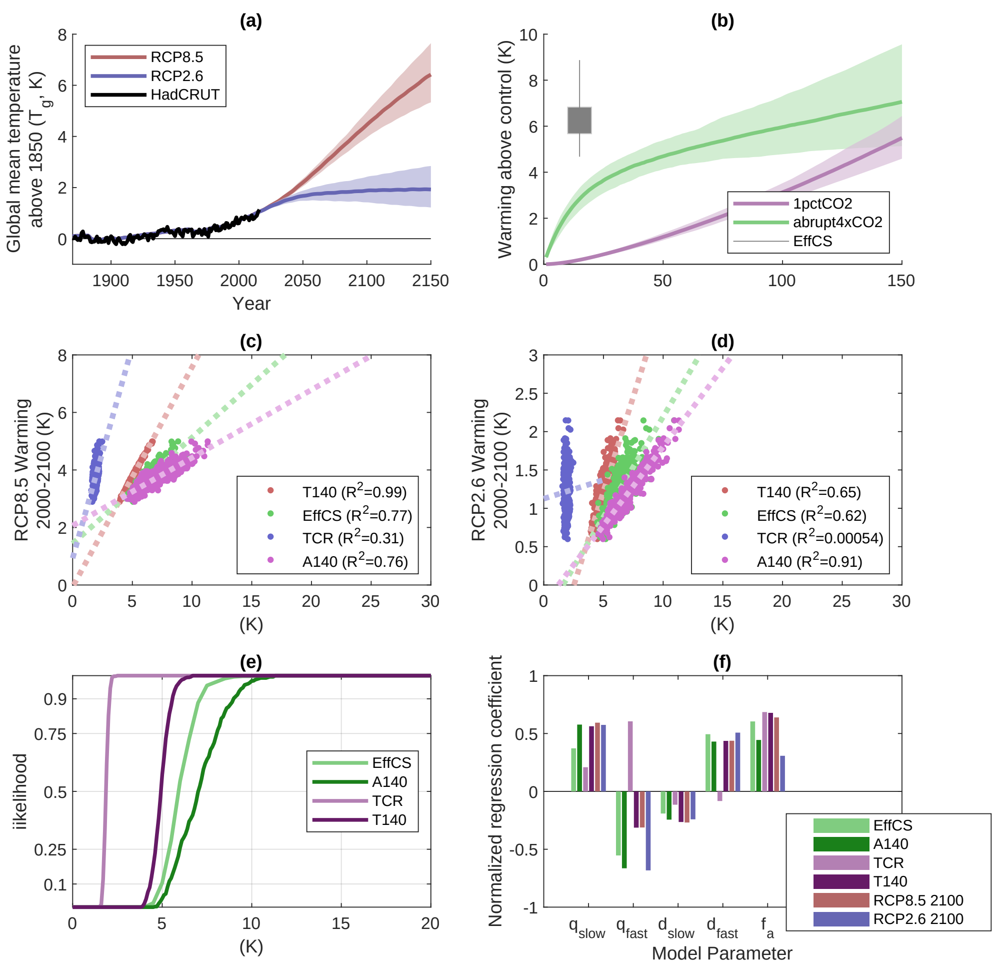
To help understand these relationships, we can perform a regression analysis of the metrics as a function of model ensemble parameters (Figure 1(f)) suggests that variance in both RCP8.5 warming and T140 are strongly controlled by the slow climate feedback parameter.

In a pulse response formulation, the response of the global temperature to forcing can be understood as a sum of a fast- and slow-equilibrating responses to the *change* in forcing in each timestep. Because the rate of change of forcing remains broadly

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

Figure 1. An observationally constrained ensemble of simple models. (a) shows the global mean temperature both historically and under the RCP2.6 and RCP8.5 scenarios. Black lines show the HadCRUT data used in calibration, whereas shaded regions show the 10-90% range of scenario projections in the posterior simple model ensemble distribution. (b) shows the corresponding time-series posterior distributions for the abrupt4xCO2 and 1pctCO2 simulated experiments, with grey errorbars showing range of EffCS for CO_2 quadrupling (boxes and whiskers show 25-75th and 1-99th percentiles respectively) . (c/d) show relationships between different sensitivity indicators and 2000-2100 temperature changes under RCP8.5/RCP2.6 respectively (e) shows the posterior cumulative probability density functions for the 4 sensitivity variables considered and (f) shows the parameter regression coefficients relating the 5 normalized model input parameters to the 4 normalized sensitivity metrics.



https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

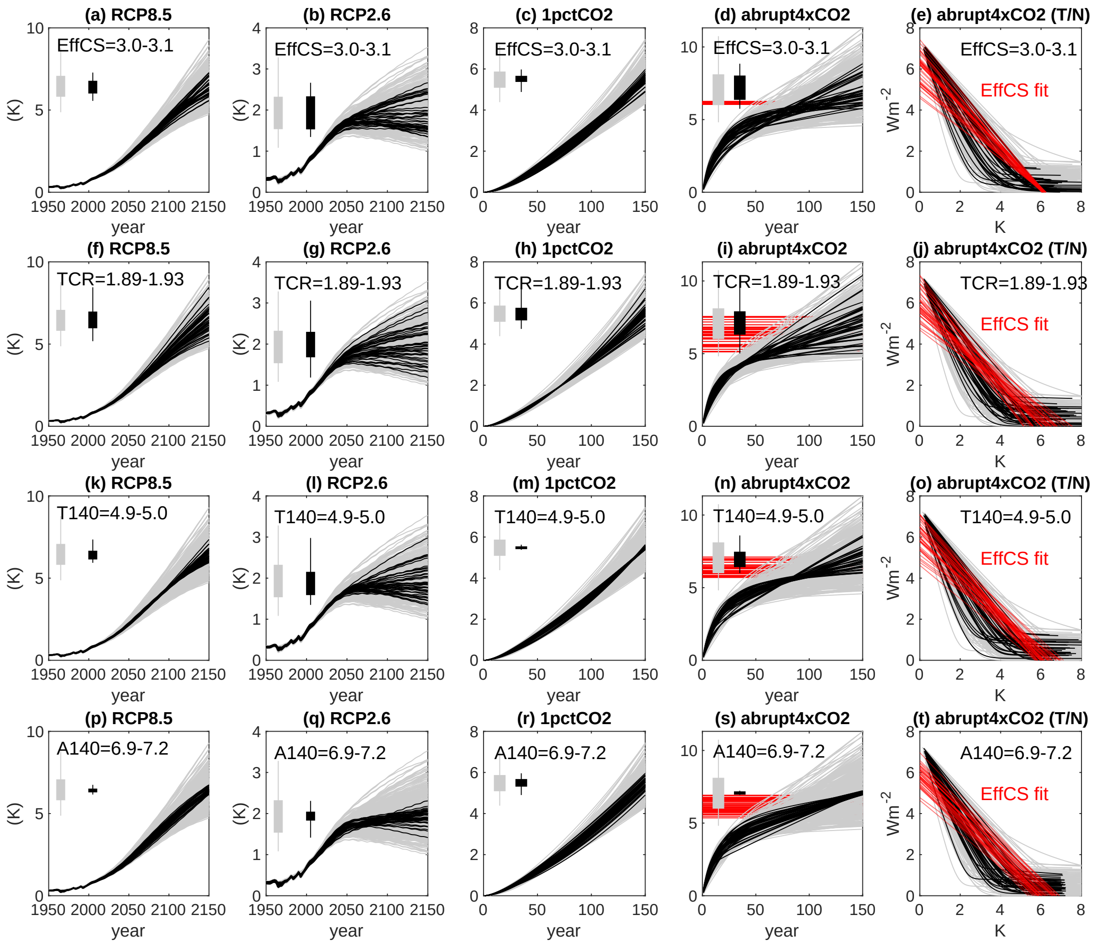


Figure 2. An illustration of how constraining different types of global sensitivity metric impact the idealized spread of global mean temperature evolution under different scenarios. Each row illustrates one constraint, Effective Climate Sensitivity to CO_2 doubling (EffCS), TCR (70 year, CO_2 doubling), T140 (140 year, CO_2 quadrupling) and A140. Lines in grey show the entire posterior distribution of models from Figure 1, while lines in black show the 45-55th percentiles of the distribution of the respective quantity. The first four columns show global mean temperature time-series of a scenario or idealized experiment - RCP8.5, RCP2.6, 1 percent ramping CO2, abrupt CO2 quadrupling (the 5th column shows energetic imbalance as a function of surface temperature in the abrupt4xCO2 experiment). Histograms show the resulting distribution of temperature in 2150 (RCP8.5/2.6) or year 140 (1pctCO2, abrupt4xCO2) for the complete distribution (grey) and 45-55th percentile range (black). Red lines show the distribution of values of effective climate sensitivity (4th column) and the trend lines used to compute it (5th column).

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

constant from 2000 until 2100, the fast feedback response associated with the shallow ocean is already saturated and the linear growth in temperature in the transient regime is governed mainly by the slow response (the rate of warming of the deep ocean).

RCP2.6 warming from 2000 to 2100, however is broadly defined by the *difference* between the slow and fast components of sensitivity. We can understand this in the context of the way the model is constrained by historical temperatures. There is a trade-off between fast and slow components of climate sensitivity in the posterior parameter distribution of the ensemble (see additional figure A3), which broadly determines the fraction of equilibrium warming associated with current forcing levels that has already been experienced.

If a greater fraction of today's observed warming is explained with the faster component of model response, there is less unrealized warming in a mitigation scenario later in the century. A140 shows similar parameter correlations and thus is well correlated to RCP2.6 end of century temperatures. Although the Effective Climate Sensitivity is a moderately good predictor of warming in both RCP2.6 and RCP8.5 in the simple model, A140 is more effective for predicting RCP2.6 temperatures due to its greater sensitivity to the slow feedback component (Figure 1(d,f))

2 Considering the multi-model ensemble

But how do the findings in the simple model framework reconcile with findings in the CMIP5 and CMIP6 multi-model ensembles? Firstly, it is plausible that there is some commonality in the lack of skill of TCR (the transient response after 70 years) in our simple model ensemble and in the CMIP ensembles. In our simple model case, the ensemble members were explicitly calibrated to reproduce the 20th and early 21st century warming - which is a very strong constraint on the value of TCR in this idealized setup.

Earth System Model calibration is conducted in a much larger parameter space by groups with a wide range of objectives which complicate interpretation (Mauritsen et al., 2012; Sanderson and Knutti, 2012), but simulations are generally only published using models which are able to adequately describe the 20th century and thus might be subject to a similar effective constraint on TCR which renders the metric ineffective for describing variance in the future evolution of the model. But there remains a direct contradiction for T140, where the simple model suggests T140 should be a better predictor than EffCS for non-mitigation warming in the 21st century whereas the opposite was found in the CMIP correlations (see additional material, Figure A2 and (Grose et al., 2018)).

To understand this, we need to consider how the properties of the simple model ensemble differ from the CMIP archive. Although the thermal response of the simple model is broadly able to represent the climatological response of CMIP models to step forcing and transient forcing in CO_2 over a century timescale ((Geoffroy et al., 2013; Proistosescu and Huybers, 2017)), it contains no internal climate variability and all experiments in section 1 are conducted from an idealized, perfectly spun up state.

Both of these assumptions are not true of CMIP5 or CMIP6. Measurement of EffCS and TCR are complicated by internal variability (Knutti and Rugenstein, 2015), and many models still exhibit some temperature drift in the control simulation from which the '1pctCO2' simulations and 'abrupt4xCO2' simulations are branched (Figure 3). This creates uncertainty from

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions
Open Access

two sources - firstly, it is not always apparent at what point during the control simulations the 1pctCO2 simulation has been branched, thus there is uncertainty in how the anomaly should be measured. Secondly, there is the potential for an unknown contribution of control drift to be erroneously included in the temperature evolution of the 1pctCO2 and abrupt4xCO2 simulations.

To assess the contribution of these two factors in metrics of climate sensitivity, we implement idealized representations of these sources of measurement error into our simple model from Section 1. We then create an idealized distribution of drift similar to that seen in the CMIP ensembles in the simple model ensemble by initializing the model 500 years before the experiment begins, defining an effective 'baseline' period from which anomalies are measured to be the average temperature between years 400 and 500. Climate internal variability is represented by a 2nd order autoregressive model, which is fitted to each CMIP model in turn. The ensemble-mean autoregressive parameters are used to create artificial 'noisy' simulations by linearly adding noise generated from the autoregressive model to the output of the simple model (see methods).

The results are illustrated in Figure 4(a), where the simple model ensemble is initialized in a non-equilibrium state with additive Gaussian noise. With these additional sources of error, both EffCS and A140 are not strongly impacted when measured in the noisy/unequilibrated model variants (Figure 4(b,c)), but the T140 measurement is strongly degraded (Figure 4(d)). Indeed, in this ensemble the biased measurements of EffCS or A140 are slightly better correlated with true T140 than the biased measurement of T140 itself. This provides a possible explanation for why T140 may be a poor predictor of RCP8.5 warming in CMIP.

In our simple framework, the reasons for the more accurate measurement of EffCS are primarily associated with the lack of equilibration. Simply adding noise from the autoregressive model has little effect on the accuracy of EffCS, T140 or A140 (where the both T140 and A140 are estimated using the average of years 131 to 150 in the simulation, see Table 1). However, both A140 and EffCS are less sensitive to non-equilibrated initial states than T140. The former experiences the same variance due to the uncertain climate drift, but the absolute value of A140 tends to be larger than T140, thus there is less relative error in its estimation. The effect on the drift on EffCS is muted because the near-linear climate drift primarily biases the estimation of slow rather than fast feedbacks (see Supplemental Figure A1). Because EffCS is primarily a measure of fast-mode feedback strength (see Figure 1(f)), its value is less impacted if experiments are started from a non-equilibrium state.

3 Conclusions

The question of which metric of climate sensitivity is most useful for summarizing uncertainty in future projections is conditional on a number of factors. Clearly, any single metric of sensitivity, even if known perfectly, will not constrain Earth System response on all timescales and scenarios. We have shown here that one can produce a number of model variants which can exhibit the same value of EffCS or TCR, but with a range of responses in a mitigation scenario such as RCP2.6.

In an idealized environment where models can be brought to a complete equilibrium control state, and ensemble sizes for '1pctCO2' simulations are large enough to avoid the effects of internal variability, the T140 metric would be the best idealized warming measure for century-scale warming under a high emissions scenario. However, the presence of even moderate control

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

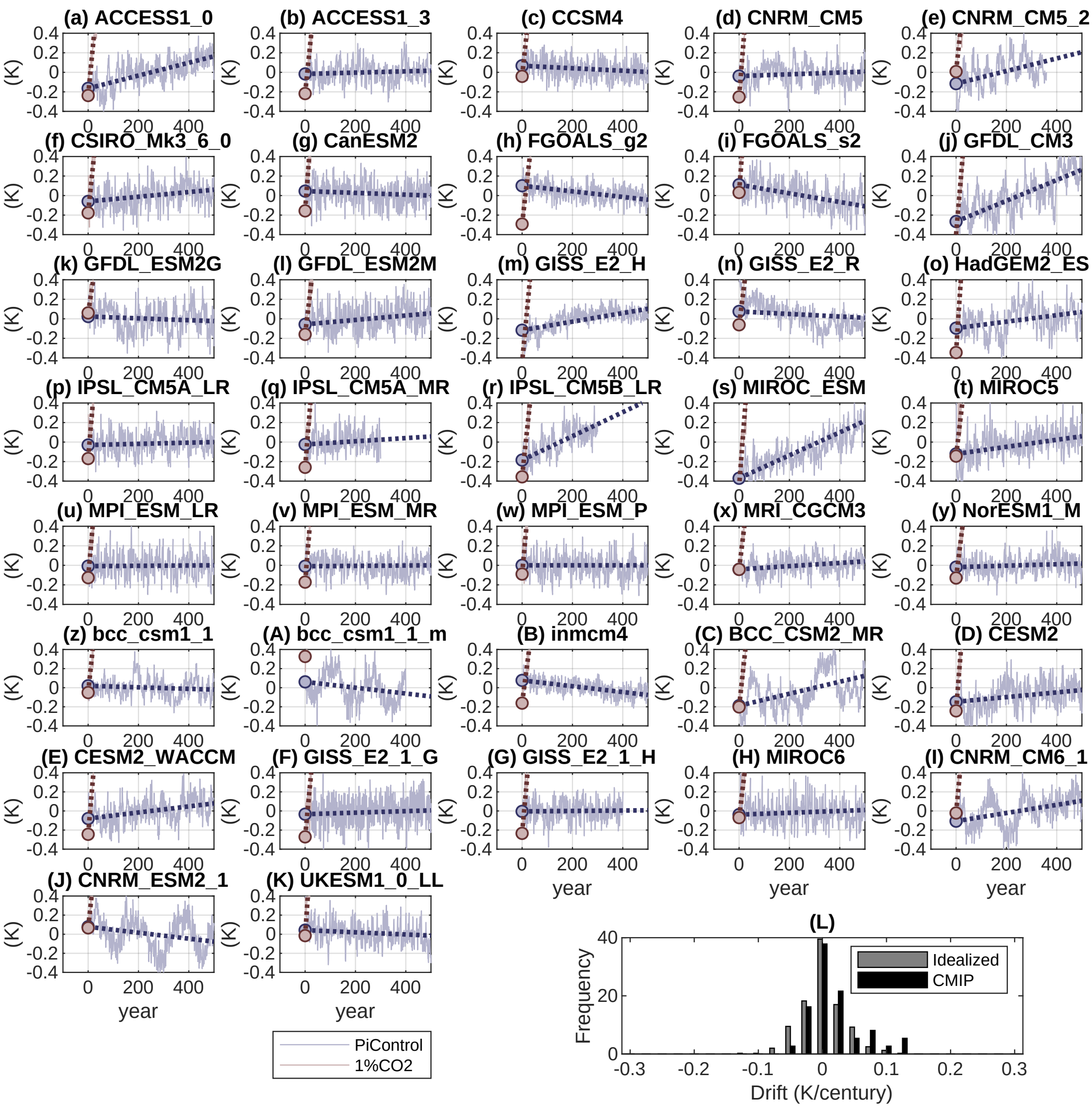


Figure 3. (a-K) Control simulation global mean temperatures from a selection of models in the CMIP5 and CMIP6 ensembles. Control simulations (blue) and initial years of 1pctCO2 simulations (pink) are plotted. Dotted lines show linear fit to the available timeseries. Blue and pink circles show the intersection of the linear temperature fit at the start of the simulation. (L) histogram showing the distribution of control model trend in CMIP (black) and in idealized ensemble of non-equilibrated simple models considered in Figure 4 (grey).

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

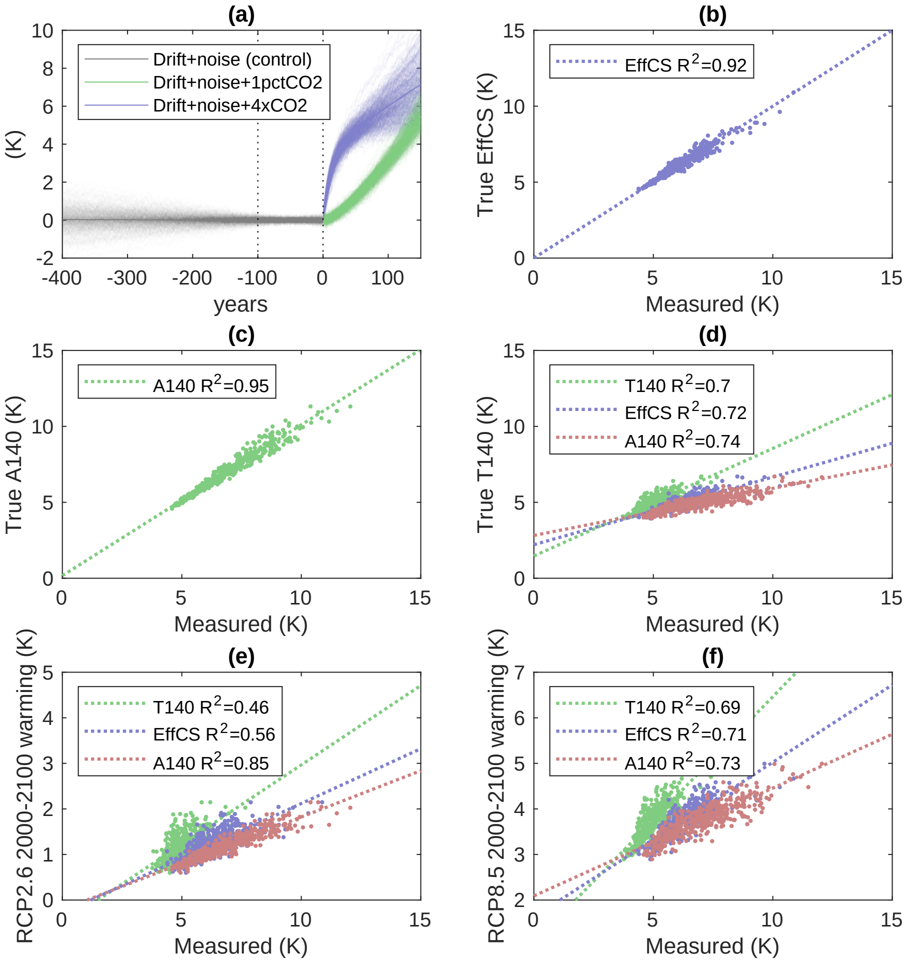


Figure 4. An idealized ensemble of simple models, where model parameters are identical to those considered in Figure 1(b), but models are initialized in a non-equilibrium state such that the baseline period is subject to some control drift, and model output is also subject to interannual variability of a similar magnitude to models in the CMIP archive. (a) shows global mean temperature evolution for the control period (gray), abrupt4xCO2 simulation (blue) and 1pctCO2 simulation (green). (b,c) show the true value of (EffCS,A140) as calculated in the noise-free, equilibrated simulations, plotted as a function of the measured value of (EffCS,A140) in a noisy, non-equilibrated simulations. (d,f,g) shows the true value of (T140,RCP2.6,RCP8.5 2000-2100 warming) plotted as a function of the measured values of T140, EffCS and A140 respectively.

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics Discussions

Predictor	T140	EffCS	A140	RCP8.5 2000-2100	RCP2.6 2000-2100
T140 (true)	1.00	0.78	0.77	0.99	0.65
EffCS (true)	0.78	1.00	0.70	0.77	0.62
A140 (true)	0.77	0.70	1.00	0.76	0.91
T140 (drift)	0.74	0.58	0.59	0.73	0.50
EffCS (drift)	0.73	0.94	0.67	0.73	0.59
A140 (drift)	0.74	0.67	0.95	0.73	0.86
T140 (noise)	0.99	0.77	0.76	0.98	0.65
EffCS (noise)	0.78	1.00	0.69	0.77	0.61
A140 (noise)	0.78	0.70	1.00	0.77	0.91
T140 (drift+noise)	0.70	0.55	0.55	0.69	0.47
EffCS (drift+noise)	0.72	0.93	0.65	0.71	0.58
A140 (drift+noise)	0.73	0.66	0.94	0.72	0.85

Table 1. A table showing R^2 regression statistics relating a set of predictors to a set of unbiased model properties. Predictors are Transient Climate Sensitivity at quadrupling of CO2 (T140), Effective Climate Sensitivity (EffCS) and warming 140 years after a quadrupling of CO2 (A140), additional rows show these values measured experiments conducted with unequilibrated base climates (drift), additive autoregressive noise (noise) and a combination of both factors (drift+noise). 'True' output model properties (T140, EffCS, A140, RCP8.5 and RCP2.6 warming from 2000 to 2100) are derived from the equilibrated model without noise.

drift can act as a significant source of error in the measurement of T140, and so here we find that EffCS is likely to be a better predictor of high emission warming in real-world applications.

EffCS itself has limitations, it is relatively insensitive to slow timescale feedbacks, which means that it poorly correlated with century-scale warming under RCP2.6 (where a large fraction of warming occurs due to slow feedback response to historical emissions), and for warming on multi-century timescales under a high emissions scenario. We find an simple, but useful alternative is to simply use the mean warming from the end of the abrupt-4xCO2 simulation - which is comparably skilled to EffCS in predicting RCP8.5 warming in 2100, but more sensitive to century timescale feedbacks than EffCS - so therefore it is better correlated with RCP2.6 end of century warming (though it is subject to greater fractional error due to control model drift than EffCS, but less so than T140).

Particularly concerning is that the two most common metrics of sensitivity, EffCS and TCR, provide very little guidance on peak warming expected under climate mitigation. The focus on these metrics has also given rise to the issue that slow feedbacks in Earth System Models are not well constrained by the set of experiments currently conducted by default in CMIP. The standard 150 year simulation used to calculate Effective Climate Sensitivity does not constrain true Equilibrium Climate Sensitivity (see Additional Material), and only a limited set of CMIP-class models have run models for long enough to be informative about equilibrium response (Rugenstein et al., 2019).

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Summary metrics may have value if the context of those metrics, and their range of applicability in relation to real-world futures is well understood, but their limitations should be kept in mind. Although it has been convincingly demonstrated that the diversity of simulated global mean dynamical response to greenhouse gas forcing over the coming centuries can be represented in simple models with a relatively small number of parameters (Smith et al., 2018; Meinshausen et al., 2011), this number is greater than one.

Data availability. CMIP5 and CMIP6 data are available through a distributed data archive developed and operated by the Earth System Grid Federation (ESGF).

Code and data availability. Code for this study is available on Github at https://github.com/benmsanderson/matlab_pulse

Appendix B: Methods

B1 Equilibrium sensitivity calculation

The two-timescale impulse response model follows the thermal feedback-timescale implementation from the FAIR simple climate model(Smith et al., 2018; Millar et al., 2017),resulting in a simple model for temperature and radiation response to a step change in forcing:

$$P(t) = F_{4xCO2} \sum_{n=1}^{2} q_n (1 - exp(-t/d_n)) \tag{B1}$$

$$R(t) = F_{4xCO2} \sum_{n=1}^{2} r_n (exp(-t/d_n)), \tag{B2}$$

where $P(t)$ is the annual global mean temperature and $R(t)$ is the net top-of atmosphere radiative imbalance, and F_{4xCO2} is the instantaneous global mean radiative forcing associated with a quadrupling of CO_2, taken here to be $3.7Wm^{-2}$(Myhre et al., 2013).

Constraining thermal parameters from historical temperatures and concentrations requires a consideration of other climate forcers. MCMC optimization of even a simple model of this form requires 10^7 or more calculations, so a very rapid model is required for computational tractability.

This study employs a fast pulse-response model to represent the response of surface global mean surface temperatures to forcing changes, where the model is implemented as a digital filter in MATLAB (see attached code) - allowing efficient computation and enabling Markov-Chain Monte Carlo parameter estimation for the physical parameters.

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

The thermal response is calculated by expressing the derivative of the forcing timeseries $F(t)$ as a series of step functions and using the CO2 quadrupling response T_p from equation B1 to calculate the integrated thermal response.

$$T(t) = \int_0^t \frac{\frac{dF}{dt}(t')}{F_{4xCO2}} T_p(t-t')dt'. \tag{B3}$$

Heat fluxes into the deep ($D(t)$) and shallow ($H(t)$) ocean components are estimated by the slow (n=1) and fast (n=2) components of $R(t)$.

B1.1 Model Optimization

The model input time-series for calibration are observed CO_2 concentrations, along with radiative estimates from (Meinshausen et al., 2011) of non-CO2 forcing agents. We optimize the thermal model parameters for 2 timescales $[\mathbf{q}, \mathbf{d}, \mathbf{r}]$ and the non-CO2 forcing factor f_r. Optimization is conducted with the (Goodman and Weare, 2010) MCMC implementation, using flat initial parameter distributions as shown in Table B1, 200 walkers and 50,000 iterations for each optimization. Cost functions are computed for global mean temperature and global CO_2 concentrations.

$$E_T = \sum_t \left(\frac{(T(t) - T_{GCM}(t))}{\sqrt{2}\sigma_T} \right)^2 \tag{B4}$$

$$E_H = \sum_t \left(\frac{(H(t) - H_{GCM}(t))}{\sqrt{2}\sigma_H} \right)^2, \tag{B5}$$

$$E_D = \sum_t \left(\frac{(D(t) - D_{GCM}(t))}{\sqrt{2}\sigma_D} \right)^2, \tag{B6}$$

where σ_T is defined as for the abrupt-CO2 case as the standard deviation of HadCRUT 1850-1950 values. Shallow and Deep Ocean heat fluxes are taken as the 0-300m and 300m+ heat content derivatives respectively in (Zanna et al., 2019), with σ_H and σ_D taken as 1850-1950 standard deviations from the same dataset.

Flat priors are used for all parameters, with an additional prior on true equilibrium climate sensitivity using the likely value and upper bound on Equilibrium Climate Sensitivity from (Goodman and Weare, 2010) fit the median and 90th percentile of a gamma distribution for equilibrium (i.e. warming as t$\rightarrow \infty$).

B1.2 Idealized Simulations

The posterior distribution of model configurations is then used to simulate a range of self-consistent values for various climate sensitivity metrics. Effective Climate Sensitivity is measured by implementing a step-change abrupt CO_2 quadrupling, and following (Gregory et al., 2004) to assess the linear extrapolation of warming at the point of net top of atmosphere energetic balance. A140 is calculated as the average of year 131-150 of the abrupt4xCO2 simulation. TCR and T140 are calculated as the average of years 61-80 and 131-150 respectively of the 1pctCO2 simulation, where CO_2 concentrations are increased

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Long name	Symbol	Min	Max
Thermal equilibration of deep ocean Sensitivity (KWm^{-2})	q_1	0	10*
Thermal adjustment of upper ocean Sensitivity (KWm^{-2})	q_2	0	10
Thermal equilibration of deep ocean timescale $(years)$	d_1	100	4000
Thermal adjustment of upper ocean timescale $(years)$	d_2	10	100
Fraction of forcing in deep ocean response	f_r	0.	1
Non-CO2 Forcing ratio	f_r	.7	1.3

Table B1. A table showing model parameter values and minimum and maximum values allowed in model optimization.

annually by 1pct resulting in a linear increase in climate forcing. RCP scenario temperature trajectories are calculated for each parameter set using concentration and forcing timeseries from (Meinshausen et al., 2011) from 1850 until 2300.

We consider the range of control drifts observed in the CMIP5 and CMIP6 ensembles (illustrated in Figure 3(L)) which range from -.3 to +.6K /century in the CMIP5 and CMIP6 models considered in this study. An idealized distribution of drift in the simple model ensemble is created by initializing the model 500 years before the abrupt4xCO2 or 1pctCO2 simulation with a non-zero, constant forcing drawn from a flat distribution ranging from -1 to $+1Wm^{-2}$, which results in a distribution of control drift of -.4K to +.4K per century (i.e. broadly comparable to the CMIP case). For each simulation we consider a baseline for temperature to be defined by the average global mean temperature in years 400-500.

To represent the first order effect of climate noise, we fit a 2nd order autoregressive model to the detrended global mean temperature timeseries in each available model in the CMIP5/6 ensemble. Taking CMIP mean parameters for the variance and autoregressive parameters, we generate noise for each realization of the simple model (though we note, in practise that the noise characteristics vary by model).

Author contributions. The author performed all analysis and writing for this project

Competing interests. The author declares no competing interests

Acknowledgements. This work is funded by the French National Research Agency, project number ANR-17-MPGA-0016. Benjamin Sanderson is an affiliate scientist with the National Center for Atmospheric Research, sponsored by the National Science Foundation.

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics Discussions

References

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, Journal of Climate, 28, 1630–1648, 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, 2018.

Armour, K. C., Bitz, C. M., and Roe, G. H.: Time-varying climate sensitivity from regional feedbacks, Journal of Climate, 26, 4518–4534, 2013.

Charney, J., Arakawa, A., Baker, D., Bolin, B., Dickinson, R., Goody, R., Leith, C., Stommel, H., and Wunsch, C.: Carbon Dioxide and Climate: A Scientific Assessment: Report of an Ad Hoc Study Group on Carbon Dioxide and Climate, Woods Hole, Massachusetts, July 23-27, 1979 to the Climate Research Board, Assembly of Mathematical and Physical Sciences, National Research Council, National Academies, 1979.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D., and Tytéca, S.: Transient climate response in a two-layer energy-balance model. Part II: Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs, Journal of Climate, 26, 1859–1876, 2013.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Communications in applied mathematics and computational science, 5, 65–80, 2010.

Gregory, J., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., Lowe, J., Johns, T., and Williams, K.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, 2004.

Grose, M. R., Gregory, J., Colman, R., and Andrews, T.: What Climate Sensitivity Index Is Most Useful for Projections?, Geophysical Research Letters, 45, 1559–1566, 2018.

Hansen, J., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., Ruedy, R., and Lerner, J.: Climate sensitivity: Analysis of feedback mechanisms, Climate processes and climate sensitivity, 29, 130–163, 1984.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K.: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing, Journal of Climate, 23, 2418–2427, 2010.

Knutti, R. and Rugenstein, M. A.: Feedbacks, climate sensitivity and the limits of linear models, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20150 146, 2015.

Knutti, R., Rugenstein, M. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, Nature Geoscience, 10, 727, 2017.

Lewis, N.: An objective Bayesian improved approach for applying optimal fingerprint techniques to estimate climate sensitivity, Journal of Climate, 26, 7414–7429, 2013.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., et al.: Tuning the climate of a global model, Journal of advances in modeling Earth systems, 4, 2012.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M., Lamarque, J.-F., Matsumoto, K., Montzka, S., Raper, S., Riahi, K., et al.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic change, 109, 213, 2011.

Millar, R. J., Nicholls, Z. R., Friedlingstein, P., and Allen, M. R.: A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions, Atmospheric Chemistry and Physics, 17, 7213–7228, 2017.

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Open Access
Discussions

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and natural radiative forcing, pp. 659–740, Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/CBO9781107415324.018, 2013.

Otto, A., Otto, F. E., Boucher, O., Church, J., Hegerl, G., Forster, P. M., Gillett, N. P., Gregory, J., Johnson, G. C., Knutti, R., et al.: Energy budget constraints on climate response, Nature Geoscience, 6, 415, 2013.

Proistosescu, C. and Huybers, P. J.: Slow climate mode reconciles historical and model-based estimates of climate sensitivity, Science Advances, 3, e1602 821, 2017.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP 8.5—A scenario of comparatively high greenhouse gas emissions, Climatic Change, 109, 33, 2011.

Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., Frölicher, T., Paynter, D., Danabasoglu, G., Yang, S., Dufresne, J.-L., Cao, L., SchmidtG.A., Abe-Ouchi, A., Geoffroy, O., and Knutti, R.: Equilibrium climate sensitivity estimated by equilibrating climate models, Geophysical Research Letters, in press, 2019.

Rugenstein, M. A., Caldeira, K., and Knutti, R.: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes, Geophysical Research Letters, 43, 9877–9885, 2016.

Sanderson, B. M. and Knutti, R.: On the interpretation of constrained climate model ensembles, Geophysical Research Letters, 39, 2012.

Senior, C. A. and Mitchell, J. F.: The time-dependence of climate sensitivity, Geophysical Research Letters, 27, 2685–2688, 2000.

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, Geoscientific Model Development, 11, 2273–2297, 2018.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M. M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.: Climate Change 2013 The Physical Science Basis, IPCC, 2013.

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

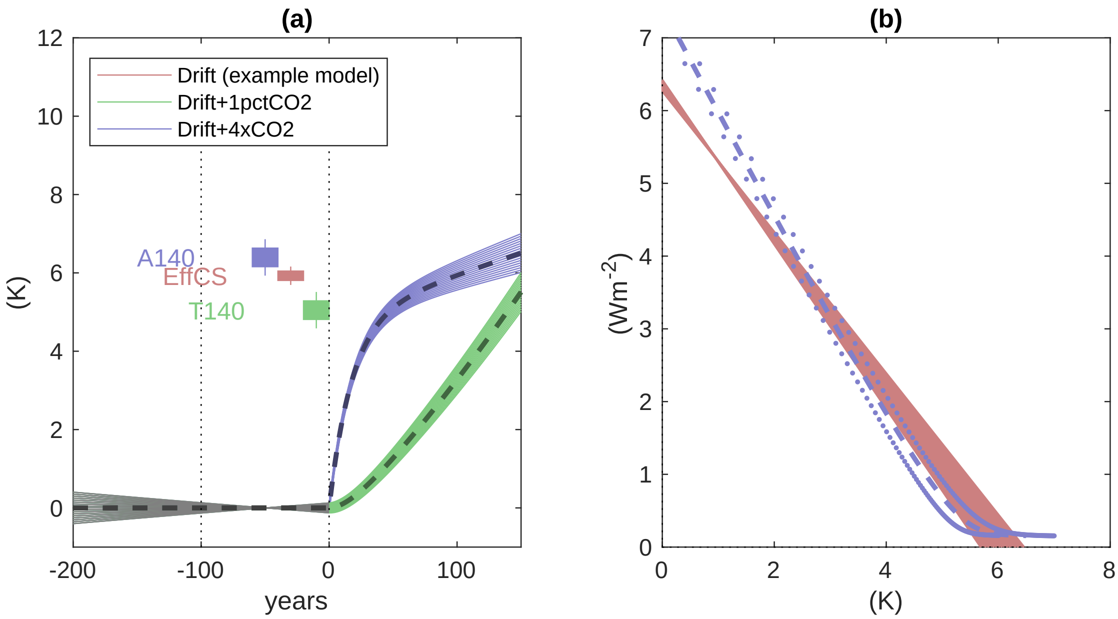
Winton, M., Takahashi, K., and Held, I. M.: Importance of ocean heat uptake efficacy to transient climate change, Journal of Climate, 23, 2333–2344, 2010.

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat storage and transport, Proceedings of the National Academy of Sciences, 116, 1126–1131, 2019.

https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

305 **Appendix : Supplementary Material**

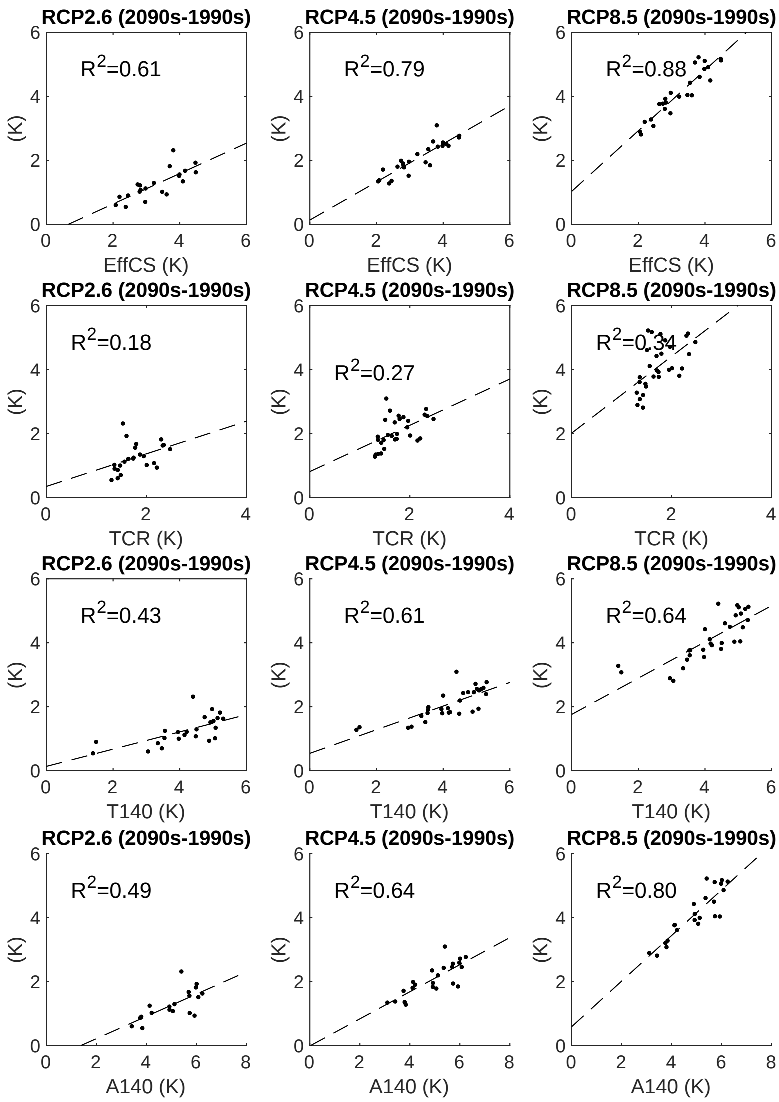
Figure A1. Plots illustrating how different types of sensitivity metric are influenced by climatological drift. Each line describes the evolution of the model (with default parameters), where the control simulation is initialized 500 years in advance of the senstivity experiment with a non-zero forcing ranging from -1 to 1 Wm^{-2}. (a) shows the global mean temperature time evolution of the abrupt4xCO2 simulations (blue) and the 1pctCO2 simulation (green), with box-whisker plots showing the range of biased values which are measured due to climate drift for A140, T140 and EffCS. (b) shows the trend lines used to compute the EffCS estimates from the simple model.



https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

Figure A2. Scatterplots of 21st century warming (difference between 20 year means in 2081-2100 and 1981-2000) and a range of sensitivity metrics for CMIP5. TCR, T140 and EffCS are reported values from (Stocker et al., 2013), A140 is calculated as the year 131-150 average global mean temperature above the control level (taken as the last 100 years of the relevant control simulation). Columns represent different RCPs, rows represent different sensitivity metrics considered in the text. Each point represents a single model from the archive. Only results from the 1st initial condition ensemble member are considered for each model.



https://doi.org/10.5194/esd-2019-77
Preprint. Discussion started: 10 December 2019

Earth System Dynamics
Discussions

Figure A3. A 'corner-plot' showing the posterior parameter distribution attained by MCMC calibration of the simple climate model. Diagonal plots show posterior histograms for parameter values optimized in the calibration, while the horizontal range indicates the bounding values of the initial flat prior distribution. Off-diagonal plots show pairwise distributions of parameters in the posterior distribution.