ATMOSPHERIC SCIENCE

Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models

Gerald A. Meehl[1]*, Catherine A. Senior[2], Veronika Eyring[3,4], Gregory Flato[5], Jean-Francois Lamarque[1], Ronald J. Stouffer[6], Karl E. Taylor[7], Manuel Schlund[3]

For the current generation of earth system models participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6), the range of equilibrium climate sensitivity (ECS, a hypothetical value of global warming at equilibrium for a doubling of CO_2) is 1.8°C to 5.6°C, the largest of any generation of models dating to the 1990s. Meanwhile, the range of transient climate response (TCR, the surface temperature warming around the time of CO_2 doubling in a 1% per year CO_2 increase simulation) for the CMIP6 models of 1.7°C (1.3°C to 3.0°C) is only slightly larger than for the CMIP3 and CMIP5 models. Here we review and synthesize the latest developments in ECS and TCR values in CMIP, compile possible reasons for the current values as supplied by the modeling groups, and highlight future directions. Cloud feedbacks and cloud-aerosol interactions are the most likely contributors to the high values and increased range of ECS in CMIP6.

INTRODUCTION

One of the earliest concepts of the climate system response to increasing carbon dioxide (CO_2) concentration comes from a simple model of the relationship between forcing and response (*1*, *2*)

$$N = F + \lambda \Delta T \qquad (1)$$

where for the net top of atmosphere energy balance, N, and a given radiative forcing, F, there is a global surface temperature response, ΔT, multiplied by a feedback factor, λ. For a given forcing associated with a doubling of atmospheric CO_2 concentration (with a radiative forcing of about 3.7 W m^{-2}), at equilibrium $N = 0$, we can then solve for ΔT, a quantity known as the "equilibrium climate sensitivity" (ECS). This simple concept was first applied to gauge climate sensitivity in the earliest climate model experiments that were performed in the 1970s with models that included highly simplified oceans. In those experiments, CO_2 was instantaneously doubled, and the model was run to a statistical equilibrium state. The warming that occurred was defined as the ECS or just simply the climate sensitivity of the system (*1*). This simplistic view of the response of the Earth system to a change in external forcing from increasing CO_2 was viewed at that time as informing the magnitude of the relative climate change Earth could experience in the future and that range for the doubling of CO_2 was first assessed to be 1.5° to 4.5°C, based on physical understanding and results from two early idealized models (*1*).

Since then, ECS has been estimated from each generation of Earth system models as a standard metric of their response to increased CO_2. The transient climate response (TCR), defined as the global temperature change at the time of CO_2 doubling in a 1% per year compounded CO_2 increase experiment, has also become a standard

metric of model sensitivity as the response to increasing CO_2. Because modeling groups routinely calculate ECS and TCR for each new model version, the simulations required to calculate these metrics are now included in the standard Diagnostic, Evaluation and Characterization of Klima (DECK) experiments, which are requirements for participation in the Coupled Model Intercomparison Project Phase 6 (CMIP6) (*3*). The relationship between ECS and TCR, and how these metrics have varied over the generations of Earth system models, has been a subject of intense interest and is addressed here. The present generation of CMIP6 models has a greater range of ECS, with higher values at the upper end of the range than previous generations of models. This has elicited scrutiny, because there are implications for the magnitude of warming in future climate projections and associated policy-relevant mitigation strategies (*4*). There are also questions as to how the TCR range relates to corresponding values of ECS.

Here, we review the historical context for these metrics from previous generations of CMIP models in relation to the current generation in CMIP6 (*3*), discuss the relationship between TCR and ECS with regard to time scales of response, address factors at work that could be producing higher values of ECS in some of the CMIP6 models, as identified by the modeling groups, and point to unresolved questions surrounding ECS and TCR.

HISTORICAL CONTEXT OF ECS AND TCR

In the era of the Intergovernmental Panel on Climate Change (IPCC) assessments, starting in the 1990s, each assessment largely maintained the same assessed range for ECS (1.5° to 4.5°C), and the climate models used to estimate ECS were generally close to that range (Fig. 1). Early on, global atmospheric models coupled to simple nondynamic "slab" ocean (or mixed layer ocean) models fit into this paradigm and were assessed in the first IPCC report (*5*). CO_2 was instantaneously doubled in such a model configuration and then run to equilibrium, which was usually attained in about 20 years or so (*6*, *7*). Sea ice in these models was usually a simple thermodynamic formulation.
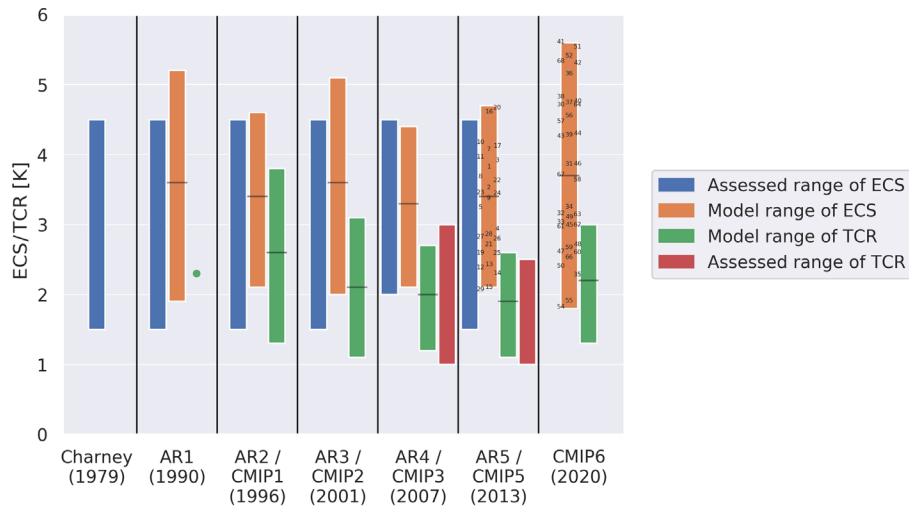
The estimates of the ECS using these models were likely comparable to later models using a full ocean model component if the latter

Fig. 1. Historical values of ECS and TCR. Assessed values of ECS (blue bars) and TCR (red bars), ranges from models of ECS (orange bars), and TCR (green bars; single value from the AR1 is green dot); numbers are individual model values of ECS from CMIP5 and CMIP6 (available on the ESGF as of March 2020). The numbers denoting individual models for CMIP5 are listed in Table 1 and those for CMIP6 in Table 2. Sources for values: AR1: table 3.2a of [IPCC First Assessment Report Ch. 3 (*5*)]; (ECS, 19 models with variable clouds; TCR, 1 model). AR2/CMIP1: figure 6.4 and table 6.3 of [IPCC Second Assessment Report Ch. 6 (*18*)] (ECS, 9 models; TCR, 13 models). AR3/CMIP2: table 9.1 of [IPCC Third Assessment Report, Ch. 9 (*20*)] (ECS, 14 models; TCR, 19 models). AR4/CMIP3: figure 10.25 of [IPCC Fourth Assessment Report Ch. 10 (*21*)] (ECS and TCR, 19 models). AR5/CMIP5: figure 9.42 and table 9.5 of [IPCC Fifth Assessment Report Ch. 9 (*25*)] (ECS, 23 models; TCR, 30 models; this differs somewhat from currently available CMIP5 models in the ESGF in Table 1). CMIP6: ECS (37 models) and TCR (37 models), with data available from a total of 39 models on the ESGF in March 2020 (Table 2).

were run to equilibrium (*8*, *9*). However, the transient response affects the estimate of ECS because the time response of the coupled model can change with warming, along with time-varying feedbacks and patterns of surface warming (*10–12*). This is because the response of the dynamic ocean can critically affect the transient response to CO_2 increases on time scales longer than a few years.

In the late 1980s, as soon as computer power allowed, global atmospheric models that previously had idealized continental outlines started to be run with realistic distributions of land and ocean and were synchronously coupled to coarse-grid (about 5° latitude-longitude) dynamical ocean models, still with simple sea-ice formulations (*13*, *14*). With this class of model [here referred to as atmosphere-ocean general circulation model (AOGCM)], it was possible to do a time-dependent or "transient" experiment where CO_2 could be gradually increased. This more realistically represented what could plausibly be expected to happen in the real climate system on the century time scale. Therefore, a common transient experiment was devised whereby CO_2 was increased 1% per year compounded, and the surface temperature increase at the time of CO_2 doubling (about year 70) was computed as the "climate change" due to a doubling of CO_2 (*14*, *15*). The transient warming in such an experiment was less than at equilibrium due to the large oceanic heat capacity. An early version of this type of model appeared in the first IPCC assessment (*16*), with a warming at the time of CO_2 doubling of 2.3°C (a single value given for IPCC 1990 in Fig. 1). Results from three more groups' transient experiments were included in the 1992 IPCC update (*17*), although these models used somewhat different experimental designs for their simulations and were only qualitatively comparable.

In the early 1990s, a growing number of international modeling groups were running such 1% CO_2 increase experiments to quantify the nonequilibrium response to a gradual increase of CO_2. This generation of models was assessed in the IPCC AR2 (*18*). Recognizing that there should be agreement on the experimental design used by all the modeling groups to facilitate intercomparison, an early phase of CMIP in the 1990s [CMIP2 (*19*)] specified that a 1% CO_2 increase experiment should be run to provide a standard measure of the transient response of the climate system. It was hoped that this intercomparison would lead to a greater understanding of the reasons for differences in the models' time-dependent responses. With the emergence of a range of plausible future emission scenarios furnished by the integrated assessment modeling community, climate models were additionally run in time-evolving climate change projections that were assessed in the IPCC AR3 with the Special Report on Emission Scenarios (SRES) scenarios (*20*) and in subsequent IPCC assessments (*21*). About this time, modeling groups started to run several standard experiments during the course of model development that included an instantaneous CO_2 doubling experiment with a mixed layer formulation to get an estimate of the ECS, and a 1% transient CO_2 increase experiment with the fully coupled model to obtain what was termed the TCR in the IPCC AR3 (*20*). These experiments provided baseline metrics to compare responses of coupled AOGCMs from one generation to the next.

In the IPCC AR3, it was argued that TCR, rather than ECS, was a more relevant metric of model response to increasing CO_2. It was straightforward to calculate in a modeling sense because the TCR is defined using the AOGCMs' temperature response itself, and the overall magnitude of TCR was thought at that time to be more comparable to the time scale and magnitude of the response in the real world over the 21st century. In addition, there were factors that complicated the calculation and interpretation of ECS that were emerging by the late 1990s. Modeling groups were struggling to

Table 1. ECS and TCR values (°C) calculated from CMIP5 model data available on the ESGF in March 2020. Model numbers denote individual models (in second column) in Figs. 1, 2, and 4. Model acronyms are defined at https://pcmdi.llnl.gov/mips/cmip5/availability.html.

CMIP5 model number	Model	ECS	TCR
1	ACCESS1-0	3.8	1.9
2	ACCESS1-3	3.5	1.6
3	BNU-ESM	3.9	2.5
4	CCSM4	2.9	1.7
5	CNRM-CM5	3.3	2.0
6	CNRM-CM5-2		1.8
7	CSIRO-Mk3-6-0	4.1	1.7
8	CanESM2	3.7	2.3
9	FGOALS-g2	3.4	1.4
10	FGOALS-s2	4.2	2.4
11	GFDL-CM3	4.0	1.9
12	GFDL-ESM2G	2.4	1.1
13	GFDL-ESM2M	2.4	1.4
14	GISS-E2-H	2.3	1.7
15	GISS-E2-R	2.1	1.5
16	HadGEM2-ES	4.6	2.5
17	IPSL-CM5A-LR	4.1	2.0
18	IPSL-CM5A-MR		2.0
19	IPSL-CM5B-LR	2.6	1.5
20	MIROC-ESM	4.7	2.2
21	MIROC5	2.7	1.4
22	MPI-ESM-LR	3.6	2.0
23	MPI-ESM-MR	3.5	2.0
24	MPI-ESM-P	3.5	2.0
25	MRI-CGCM3	2.6	1.6
26	NorESM1-M	2.8	1.4
27	bcc-csm1-1	2.8	1.7
28	bcc-csm1-1-m	2.9	2.1
29	inmcm4	2.1	1.3
	Multimodel mean	3.2	1.8
	SD	0.7	0.4

maintain nondynamic slab (or mixed layer ocean) model formulations coupled to the atmospheric model that were comparable to their fully coupled versions with a dynamical ocean coupled to the atmosphere (i.e., their AOGCMs). Sea-ice formulations in the AOGCMs were becoming more complex (including, for example, sea-ice dynamics). Sea surface temperature errors in the mixed layer models (due, in part, to their lack of ocean dynamics) were being corrected by different groups in different ways using a technique generally called "Q-flux" (22). A further complication related to the relevance of ECS was that paleoclimate studies were finding that other factors could influence the equilibrium response of the real world, such as vegetation, biogeochemistry, and dust, and most of these were not accounted for in the traditional definition of ECS (23).

A shortcut was proposed to estimate ECS (24). This resulted in a metric sometimes termed as the "effective climate sensitivity," but for our purposes here, we will refer to it as the "ECS calculated by the Gregory method." This will distinguish it from the ECS values obtained with earlier slab oceans coupled to atmospheric models and run to equilibrium. The ECS calculated by the Gregory method is derived from a fully coupled Earth system model and does not require equilibrium to actually be achieved. In the Gregory method, CO_2 is instantaneously quadrupled in a fully coupled Earth system model and run for 150 years. As the surface temperature asymptotes toward equilibrium, the slope of the time-evolving curve of the net top-of-atmosphere radiance against the surface temperature is calculated to extrapolate the eventual temperature increase at equilibrium some time far in the future for a doubling of CO_2, assuming that there is a roughly linear response that is half of the warming from a quadrupling of CO_2. In contrast to the ECS values in previous IPCC assessments using atmosphere models coupled to nondynamic slab oceans, the Gregory method was applied to CMIP5 coupled models assessed in the IPCC AR5 (25). The Gregory method is still the most frequently used approach to calculating ECS from AOGCM simulations, although complications arising from this method have led to various other alternatives to be proposed.

CURRENT ECS AND TCR RESULTS PUT INTO HISTORICAL CONTEXT

A compilation of ranges of TCR and ECS for six generations of climate models dating back to 1990 and through the current CMIP6 generation is shown in Fig. 1, along with the expert judgment–assessed ranges of ECS and TCR from the various IPCC assessments. The TCR and ECS values shown in Tables 1 and 2 have been consistently calculated with the Earth System Model Evaluation Tool (ESMValTool) version 2.0 (26, 27) for the individual CMIP5 and CMIP6 models that were available from the ESGF at the time of publication. All figures of the paper have been produced with ESMValTool v2.0. ECS uses the Gregory method from a 150-year run of an instantaneously quadrupled CO_2 simulation. TCR is calculated as the change in the 20-year average global mean surface temperature, centered around the time of CO_2 doubling (years 60 to 79) relative to the 140-year period in the pre-industrial (PI) control that includes the time period of the 1% per year experiment, with the global temperature in the PI control smoothed by applying a linear 140-year fit to account for residual drift. ECS and TCR values calculated by the modeling groups and reported in their papers may differ slightly from those in Tables 1 and 2 that are calculated by the ESMValTool. For example, the Community Earth System Model Version 2 (CESM2) has a value for ECS in Table 2 of 5.2°C, while there are published values of 5.3°C, with TCR values of 2.1° and 2.0°C, respectively (9).

As noted above, the original assessed range of ECS was 1.5° to 4.5°C and was based on a few early simplified models (1). The range of ECS diagnosed from model simulations has remained near the assessed range until the present CMIP6 generation where the range has expanded to 1.8° to 5.6°C from 37 models, the largest range of ECS in any generation of model to date (Fig. 1; although not all CMIP6 models are yet included). Table 2 shows ECS and TCR, along with the multimodel means and SDs for 39 CMIP6 models compiled for this study (inadequate data on the ESGF preclude calculating ECS and TCR from four of the models, so there are 37 models for ECS

Table 2. ECS and TCR values (°C) calculated from CMIP6 model data available on the ESGF in March 2020. Model numbers denote individual models (in second column) in Figs. 1, 2, and 4. Model acronyms are defined at https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html, and modeling groups at https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html.

CMIP6 model number	Model	ECS	TCR
30	ACCESS-CM2	4.7	2.1
31	ACCESS-ESM1-5	3.9	2.0
32	AWI-CM-1-1-MR	3.2	2.0
33	BCC-CSM2-MR	3.0	1.7
34	BCC-ESM1	3.3	1.8
35	CAMS-CSM1-0	2.3	1.7
36	CESM2	5.2	2.0
37	CESM2-WACCM	4.8	2.0
38	CNRM-CM6-1	4.8	2.1
39	CNRM-CM6-1-HR	4.3	2.5
40	CNRM-ESM2-1	4.8	1.9
41	CanESM5	5.6	2.7
42	E3SM-1-0	5.3	3.0
43	EC-Earth3	4.3	
44	EC-Earth3-Veg	4.3	2.6
45	FGOALS-f3-L	3.0	2.1
46	GFDL-CM4	3.9	2.1
47	GFDL-ESM4	2.6	1.6
48	GISS-E2-1-G	2.7	1.8
49	GISS-E2-1-H	3.1	1.9
50	GISS-E2-2-G	2.4	1.7
51	HadGEM3-GC31-LL	5.6	2.6
52	HadGEM3-GC31-MM	5.4	2.6
53	IITM-ESM		1.7
54	INM-CM4-8	1.8	1.3
55	INM-CM5-0	1.9	
56	IPSL-CM6A-LR	4.6	2.3
57	KACE-1-0-G	4.5	1.4
58	MCM-UA-1-0	3.7	1.9
59	MIROC-ES2L	2.7	1.6
60	MIROC6	2.6	1.6
61	MPI-ESM1-2-HR	3.0	1.7
62	MPI-ESM1-2-LR	3.0	1.8
63	MRI-ESM2-0	3.2	1.6
64	NESM3	4.7	2.7
65	NorCPM1		1.6
66	NorESM2-LM	2.5	1.5
67	SAM0-UNICON	3.7	2.3
68	UKESM1-0-LL	5.3	2.8
	Multimodel mean	3.7	2.0
	SD	1.1	0.4

and 37 models for TCR in Table 2). The SD of ECS in CMIP6 is 1.1°C, which can be compared to the lower value of 0.7°C in CMIP5 (Table 1), reflecting the previous smaller range. It could be possible that the SD is influenced by a single outlier, but the distribution of individual models in Fig. 1 and Table 2 shows two CMIP6 models with ECS values less than 2°C (1.8°and 1.9°C), while there are six models with ECS values above 5°C.

By contrast, the range in TCR of 1.3° to 3.8°C in CMIP1 (a range of 2.5°C) has shrunk to 1.3° to 3.0°C in CMIP6 (a range of 1.7°C), with an SD of 0.4°C (Fig. 1 and Table 2). This can be compared to the CMIP5 range of 1.5°C reported in the AR5 (in models available now that are compiled in Table 1, there is a range of 1.4°C, with a comparable SD value of 0.4°C). The CMIP6 range of 1.7°C indicates a modest upward shift from the range of 1.5°C in both CMIP3 and CMIP5. The multimodel mean values of ECS and TCR in CMIP5 are 3.2° and 1.8°C, respectively (Table 1), while comparable values of those in CMIP6 are 3.7° and 2.0°C (Table 2).

When comparing the ranges of ECS and TCR in Fig. 1, the changes in the range are larger than changes in mean values of ECS, with the latter being more consistent from generation to generation of model. For example, the highest ECS range is 3.8°C in CMIP6 compared to the lowest of 2.3°C in the CMIP3 models in the AR4. However, average values of ECS have remained near 3.5° ± 0.2°C for all generations of models. For TCR, starting with the CMIP2 models in the AR3, both the range and average values have remained relatively stable, with the average varying around 2.0° ±0.2°C and the range staying near 1.7° ± 0.2°C. It is also interesting to note that the high end of the assessed range of ECS, 4.5°C, has been exceeded by the high end of the multimodel range in every IPCC assessment except for the AR4. Similarly, the low end of the multimodel range has been higher than the low end of the assessed range of ECS in every IPCC assessment, which is similar to the two times that the TCR range has been assessed in the AR4 and AR5. Thus, over the generations of models, the multimodel ranges of ECS and TCR have been higher than the assessed ranges in the IPCC assessments, with the exception of the CMIP3 models in the AR4. It should be kept in mind that a range of model results cannot be interpreted as an estimate of uncertainty; the range might be too wide if unrealistic models are included, or it could be too small if many models are missing the same process or feedback.

RELATIONSHIP BETWEEN TCR AND ECS
There have been efforts to directly relate TCR to ECS. For example, in the IPCC AR4 from the CMIP3 multimodel ensemble, TCR was plotted as a function of ECS [figure 10.25 of (21)]. That TCR should increase with ECS is not unexpected because on all time scales, a more sensitive model can be expected to warm more. More refined analysis of simplified climate models suggests that the increase of TCR with ECS should be less than linear. A number of studies explained that there is a nonlinear relationship governed by a ratio involving two parameters, ECS and heat uptake efficiency. This implies that if all models have similar efficiency in sequestering heat, then the more sensitive models will, at any point in time, realize a smaller fraction of their eventual warming (28–32). The IPCC AR4 used the CMIP3 multimodel ensemble to show a nonlinear behavior consistent with this prediction [figure 10.25 of (21)]. For more realistic transient climate change experiments, a similar result holds (33). Although in the IPCC AR5 [figure 9.42b of (25)] a linear fit to
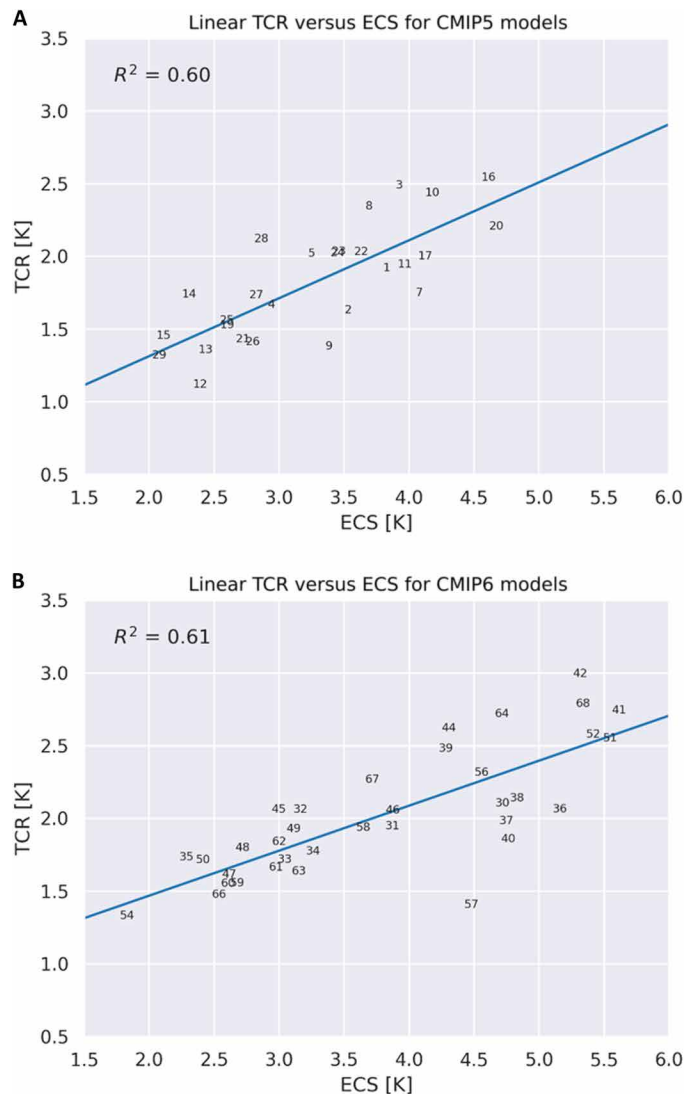
Fig. 2. ECS as a function of TCR. (**A**) From the CMIP5 models in the IPCC AR5 (black line is linear fit); (**B**) same as (A) except for CMIP6 models (black line is a linear fit). Note that 27 models are plotted for CMIP5 (Table 1) compared to a total of 23 and 30 models that supplied ECS and TCR values, respectively, to the IPCC AR5 used for the ranges in Fig. 1. The greater number of models plotted here denotes those with sufficient available data on the ESGF to perform corresponding ECS and TCR calculations, as defined in the ESMValTool discussed in the text. The R^2 values are given in the upper left parts of each panel. The numbers denoting individual models for CMIP5 in (A) are listed in Table 1 and those for CMIP6 in (B) in Table 2.

the TCR-ECS relationship intersects the ordinate well above 0, this violates basic physical considerations, which imply that the TCR should be 0 if the ECS is 0. If this constraint is applied, then the AR5 collection of model results are again consistent with a simple nonlinear relationship between TCR and ECS. A nonlinear relationship between TCR and ECS might also be explained by other factors, of course, and numerous studies have shown that warmer base states and higher CO_2 concentrations can produce proportionately larger feedbacks and thus contribute to this nonlinear relationship (*32, 34–36*).

Past studies also have indicated that global climate feedbacks likely vary as climate changes (*24, 37, 38*). Subsequently, this has been shown to be true in many models (*11, 39–41*) and is likely

related to variations in the patterns of surface temperature and radiative feedbacks (*12*). Thus, the increase in feedback (*36, 39*), due to either a transient pattern effect (*38*) or some kind of nonlinearity in the equilibrium response, such as feedback temperature dependence (*33*), can be demonstrated through a method that attempts to account for both types of phenomena in a ramp experiment (*42*).

Consequently, the ECS obtained by running fully cupled Earth system models to equilibrium over multimillennia is also likely larger than estimates of ECS computed using the Gregory method (*43*). This result has also been documented in studies that use paleoclimate data (*35*).

In addition, there are different time scales of response in the system, with some feedbacks and responses acting on the subcentury time scale and others operating on the multicentury time scale, with the nonlinear relationship between ECS and TCR indicative of those different time scales of response (*41, 44, 45*). The linear fit of ECS versus TCR in the CMIP6 models has a similar R^2 value of 0.61 compared to 0.60 in CMIP5 (Fig. 2; for sufficient available data on ESGF, see note in figure caption).

This time scale dependence of feedbacks and response can be demonstrated in Fig. 3 where for the 37 currently available CMIP6 models with sufficient data on the ESGF, the multimodel mean ECS from the Gregory method, calculated over the full 150-year period of the $4xCO_2$ experiment, is 3.7°C. For the first 20 years of the simulations, however, the sensitivity is 3.3°C, and for the last 130 years, the value is 4.0°C. This indicates a time-varying feedback strength, leading to different time scales of response in the climate system.

FACTORS CONTRIBUTING TO THE LARGER RANGE AND HIGHER VALUES OF ECS IN CMIP6

The question remains as to why the range of ECS has increased in CMIP6, with the upper end of the range extending well beyond the canonical 4.5°C value and higher than previous generations of models (Fig. 1). Various researchers, as well as the modeling groups, have been attempting to answer that question. One possibility is that the newer prognostic aerosol schemes that include aerosol-cloud interactions could have produced overly large negative radiative forcing, which then implied a need for a stronger model response to CO_2 if the model was to reproduce the historical temperature response. Such relationships between amplitude of aerosol forcing and ECS have been noted in previous generations of models (*46, 47*). However, because the aerosol forcing varies with time, it is difficult to tune the ECS based on responses to aerosols over different periods of the 20th century (*48*). On the basis of the information in Table 3 provided by a subset of CMIP6 modeling groups, Fig. 4 shows that there is a weak relationship between amplitude of aerosol forcing and ECS, with larger negative present-day aerosol forcing associated with larger ECS (linear fit with an R^2 of only 0.36 in Fig. 4). A simple explanation that these results might result from a relationship between forcing and aerosol-dependent cloud-radiative feedback responses that depend on whether aerosols are predicted or prescribed does not seem to hold. Some models have prognostic aerosol schemes with large negative forcing but low ECS (e.g., MRI-ESM2), while others have low values of negative aerosol forcing but high ECS (e.g., GFDL-CM4). For the latter, the composition (emissions) feedbacks present in GFDL-ESM4, but not in GFDL-CM4, can operate largely independently of the direct and indirect forcing;
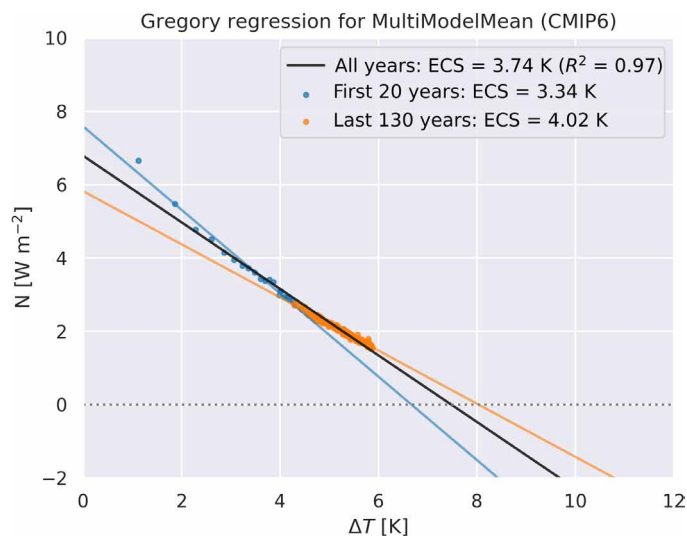
Fig. 3. ECS calculated for the CMIP6 models in Table 2 using the Gregory method over different time scales. Using the entire 150-year 4xCO2 experiment (black line), there is an ECS value of 3.7°C; using only the first 20 years (blue dots and blue line), there is an ECS of 3.3°C; and using the last 130 years, there is an ECS of 4.0°C (orange dots and orange line).

instead, they are feedbacks. In models with advanced treatments of biogeochemistry and vegetation, warming can induce changes in emissions, which alter the aerosol fields. In the GFDL-ESM4's case, these provide negative feedbacks and reduce ECS. One group with high ECS and large negative aerosol radiative forcing (CESM2) ascribes the high ECS value to cloud feedbacks and aerosol-cloud interactions related to the details of stratiform cloud microphysics and associated ice nucleation, turbulence, rain formation and evaporation processes, and SO_2 lifetime (49). Specifically, changes made to increase high-latitude supercooled liquid water, and to adjust warm rain susceptibility to aerosols in shallow clouds, have increased cloud feedbacks in that model. This would apply to anthropogenic and natural aerosols and contribute to higher values of both TCR and ECS. Another group in that category (HadGEM3-GC3.1-LL) has diagnosed the differences from an earlier model version with a lower ECS as arising from changes in the shortwave cloud-radiative feedback in the midlatitudes (mainly over the Southern Ocean), with the introduction of a new aerosol scheme and the development of a new mixed-phase cloud scheme (50). Meanwhile, two models with prognostic direct aerosol effect only but with no aerosol-cloud feedbacks (the two INM models) and two others with prescribed aerosols (the MPI and IPSL models), have similar relatively low values of aerosol forcing (around -0.6 W m^{-2}) but with ECS values that range from 1.8° to 4.6°C (Fig. 4).

One common theme that emerges in Table 3 from five of the six models with ECS values greater than 5°C [E3SM, CESM2, UK Earth System Model (UKESM1), HadGEM3, and Canadian Earth System Model version 5 (CanESM5)], all with prognostic aerosol direct and indirect effects, is the notable role played by cloud feedbacks and/or cloud-aerosol interactions in high ECS. Four of the five have relatively large negative values of present-day aerosol effective radiative forcing (ERF), ranging from -1.1 to -1.7 W m^{-2}. Two other models with prognostic aerosol direct and indirect effects and relatively large negative aerosol ERF values of -1.2 W m^{-2} (NorESM2-LM and MRI-

ESM2) have lower values of ECS (2.5° and 3.1°C, respectively), but neither cite cloud feedbacks as being a notable part of their lower ECS values (Table 3). However, a model from one group (GFDL-ESM4) has ECS values that decreased from CMIP5 to CMIP6, and of the six factors cited as playing a role in this decrease, three relate to negative cloud-aerosol feedbacks (Table 3).

Additional evidence for the role of cloud feedbacks in models with high ECS has been ascribed to decreasing extratropical low cloud coverage and albedo with increasing temperature. This produces stronger positive shortwave cloud-radiative feedbacks (i.e., warmer temperatures produce less low clouds, allowing more incoming solar and more warming, and so on) that are directly related to how clouds are represented in those models (51). There are indications from some paleoclimate and observational studies that the higher estimates of ECS are less credible than the lower estimates (36, 52–54), while other paleoclimate studies suggest a possibility of higher ECS values (35, 55). By applying observations in an emergent constraints context, there is evidence that the ECS is likely at the lower end of the IPCC range (53), with unlikely higher estimates of ECS (55), although there is other evidence supporting higher values of ECS. Thus, although there appears to be no single property in the current generation of CMIP6 models to which the increased range and higher values on the upper end of ECS can be attributed, cloud feedbacks and cloud-aerosol interactions in models with prognostic aerosol schemes seem to be playing an important role.

UNRESOLVED QUESTIONS

This brings us to the question of what is the most appropriate policy-relevant metric to assess the Earth system model response to an increase of CO_2. This inevitably depends on the time scale of interest, with the century time scale response represented by TCR being roughly consistent over multiple generations of models. Meanwhile, ECS, representing the century and longer time scales, is constrained less, because the moderating influence of ocean heat uptake does not ultimately affect it. Thus, we return to the point made in the IPCC AR3 (20). While ECS and TCR represent different aspects of the response of the Earth system, on time scales of the next several decades, TCR could be more relevant (56), although evidence has also been presented that for model simulations for the next 50 to 100 years, ECS is actually a better predictor (57). Because the climate system will never be in equilibrium, ECS is an abstract quantity that can never be observed. As noted above, ECS has proved to have numerous computational difficulties related to time scales of response, along with model-related complications, regarding how different feedbacks are modeled in different components of the climate system. It is still an open question as to how relevant ECS is to understanding historical climate change or the transient climate system response to increasing CO_2. The only partially understood relationship between ECS and TCR raises unresolved questions and research challenges regarding time scales of feedbacks in the climate system.

In addition, the decreasing range of TCR over generations of models, contrasted with the recent increase in range of ECS in the CMIP6 models, points to an interesting research question regarding why this has occurred. It likely involves processes connected to ocean heat uptake, a better quantification of which would require improved temperature observations through the full depth of the global ocean (58), as well as increased understanding of various feedbacks in the climate system. Nonlinearities in the feedbacks and

Table 3. Subsample of CMIP6 models shown in Fig. 1, with information supplied by the modeling groups regarding details of aerosol forcing and formulation and possible reasons for ECS values. For the GFDL models, the higher sensitivities in parentheses denoted by asterisks result from longer runs and attempts to filter out unforced variability. Model acronyms are defined at https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html, and modeling groups at https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html.

Model	ECS (K)	TCR (K)	Aerosol ERF (W m^2)	Aerosol scheme	Reasons for ΔECS from CMIP5?	Paper for more information
E3SM_1	5.3	3.0	−1.65	Prognostic—direct and indirect	No CMIP5 equivalent but unusually large positive SW cloud feedback	Golaz et al. (61)
CESM2 CESM2-WACCM	5.3 4.8	2.0 1.9	−1.67	Prognostic—direct and indirect	Increase by >1k from CESM1 related to cloud feedbacks and aerosol-cloud interactions	Gettelman et al. (49)
GFDL-CM4 GFDL-ESM4	3.9 (5.0*) 2.7 (3.4*)	2.1 1.6	−0.7 −0.7	Prognostic—direct and indirect	Preliminary investigation into the causes for this lower climate sensitivity in ESM4.1 compared to CM4.0 have indicated at least six drivers (3 −ve aerosol-climate feedbacks, −ve stratospheric ozone feedback, changes in ocean heat uptake, explicit representation of CO_2)	Held et al. (62) Winton et al. (63)
HadGEM3-GC3.1-LL UKESM1	5.5 5.4	2.6 2.8	−1.1 −1.17	Prognostic—direct and indirect	Cloud-aerosol interactions and cloud microphysics	Bodas-Salcedo et al. (50) Andrews et al. (64) Sellar et al. (65)
MIROC6	2.6	1.6	−0.76	Prognostic—direct and indirect(?) (SPRINTARS)	Very little change from CMIP5 to CMIP6	Tatebe et al. (66)
MRI-ESM2	3.1	1.6	−1.22	Prognostic—direct and indirect effects	Small increase in sensitivity (2.6–3.1) and many changes having a small impact, with largest impact possibly coming from changes to entrainment-detrainment rates (but not yet fully tested)	Yukimoto et al. (67)
MPI-ESM1.2	3.0		−0.6	Specified, direct only (MACv2-SP)	Tuned with cloud parameters to be the same as CMIP5. Pretuned version had ECS = 7 caused by a +ve low-cloud feedback in the tropics	Mauritsen et al. (68)
EC-Earth3 EC-Earth3-veg	4.2 4.3	2.6	Not yet known	Specified, direct only (MACv2Sp)	Early indications of the role of cloud-aerosol interactions	Wyser et al. (69)
INM-CM5 INM-CM4.8	1.9 1.8	1.3	−0.5	Prognostic—direct effect only	No change in ECS from CMIP5 although a lot of changes in parametrization of cloud and condensate	
ACCESS-CM2 ACCESS-ESM1.5	4.7 3.9		Not yet known	Prognostic—direct and indirect	Using HadGEM3-GC3.1 atmospheric component, so high ECS aligned to this Using ACCESS1.3 CMIP5 model physics—little change	
AWI-ESM	3.2	2.2	Not known	Specified—direct (MACv2Sp)	No CMIP5 model, but interesting from the "parent" model, MPI-ESM	
CanESM5	5.62	2.7		Prognostic—direct and indirect effect	Large increase since CMIP5 model (3.7–5.6)—at least half seems to be related to cloud feedback increase	Swart et al. (70) Paper on cloud feedbacks and ECS paper planned for 2020
NorESM2-LM	2.5		−1.2	Prognostic—direct and indirect	Small decrease since CMIP5 model (2.9–2.5), which is not yet understood	Paper hoped for in early 2020
IPSL-CM6A-LR	4.6	2.3	−0.6	Specified—direct and indirect		Lurton et al. paper planned for 2020 Servonnat et al. paper planned for 2020
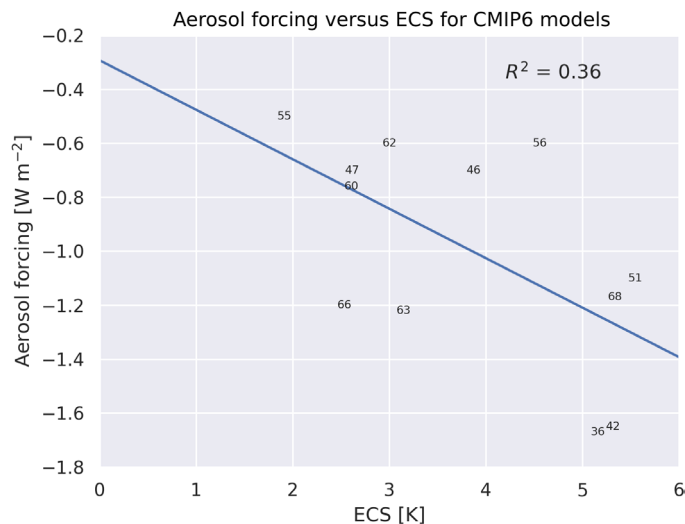
Fig. 4. Effective radiative forcing from aerosols versus ECS. Values supplied by the modeling groups (Table 3); black line is linear fit with R^2 of 0.36. The numbers denoting individual models are listed in Table 2.

pattern effects (e.g., Fig. 3) are far less severe in the transient case than close to equilibrium. Another factor that could play a role is that some models (e.g., HadGEM3-GC3.1 and ACCESS-CM2) interactively redistribute their climatological ozone concentrations during the abrupt 4xCO2 and scenario runs to retain the bulk of the ozone in the stratosphere as the tropopause height rises. This acts to reduce the ECS that would otherwise occur by approximately 10 to 15% (59).

Because the higher ECS values in some models are related to cloud feedbacks and cloud-aerosol interactions, a major research question that needs to be pursued is what is the actual nature and magnitude of cloud feedbacks in general and cloud-aerosol interactions in particular. Making progress will require enhanced observations that would provide new insights into the processes involved with cloud microphysics. Additional knowledge from new observations and improved modeling would be desired to quantify the details of how clouds interact with and are affected by aerosols, both natural and anthropogenic. This is because in calculations of ECS and TCR, there are no changes in anthropogenic aerosols, but natural aerosols in a given model's prognostic aerosol scheme can respond to the climate changes brought about by increases of CO_2. The World Climate Research Programme (WCRP) Grand Challenge on Clouds, Circulation, and Climate Sensitivity (https://wcrp-climate.org/component/content/article/61-gc-clouds-circulation?catid=30&Itemid=267) is attempting to do just that. The assessment of climate sensitivity undertaken by the WCRP (60) delivers a new range of ECS based on multiple lines of evidence and provides a new framework for model assessment. It will be critically important that all models are tested against this rigorous framework using new methods, observations, and process evaluation. If the high-ECS models fall outside of this new range, then the framework will enable process-based understanding of the reasons. Note also, however, that some models may fall within the ECS-assessed range but could do so through a cancelation of biases. It will be equally important to identify these and work to improve them if the modeling community is to be confident that models, whether inside or outside the newly assessed range of ECS, are really delivering reliable projections of future climate. In any case, ECS methods should be standardized through a more comprehensive comparison of ECS through both multimillennial climate perturbation simulations such as those conducted in LongRunMIP (42). In addition, slab ocean model comparisons would be useful to better understand the causes of these differences and to derive a more robust estimate of climate sensitivity from the current generation of models.

REFERENCES AND NOTES

1. J. G. Charney, B. Stevens, I. H. Held, J. F. Mitchell, J.-L. Dufresne, K. A. Emanuel, P. Friedlingstein, S. Griffies, C. Senior, *Carbon Dioxide and Climate: A Scientific Assessment* (U.S. National Academy of Sciences, Washington, DC, 1979).
2. R. D. Cess, G. L. Potter, J. P. Blanchet, G. J. Boer, S. J. Ghan, J. T. Kiehl, H. L. E. Treut, Z. X. Li, X. Z. Liang, J. F. Mitchell, J. J. Morcrette, D. A. Randall, M. R. Riches, E. Roeckner, U. Schlese, A. Slingo, K. E. Taylor, W. M. Washington, R. T. Wetherald, I. Yagai, Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science* **245**, 513–516 (1989).
3. V. Eyring, S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, K. E. Taylor, Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
4. P. M. Forster, A. C. Maycock, C. M. McKenna, C. J. Smith, Latest climate models confirm need for urgent mitigation. *Nat. Clim. Change* **10**, 7–10 (2020).
5. J. F. B. Mitchell, S. Manabe, V. Meleshko, T. Tokioka, Equilibrium climate change and its implications for the future, in *Climate Change: The IPCC Scientific Assessment*, J. T. Houghton, G. J. Jenkins, J. J. Ephraums, Eds. (Cambridge Univ. Press, Cambridge, UK, 1990).
6. S. Manabe, R. J. Stouffer, Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere. *J. Geophys. Res.* **85**, 5529–5554 (1980).
7. W. M. Washington, G. A. Meehl, Seasonal cycle experiment on the climate sensitivity due to a doubling of CO_2 with an atmospheric general circulation model coupled to a simple mixed-layer ocean model. *J. Geophys. Res.* **89**, 9475–9503 (1984).
8. R. J. Stouffer, S. Manabe, Response of a coupled ocean-atmosphere model to increasing atmospheric carbon dioxide: Sensitivity to the rate of increase. *J. Climate* **12**, 2224–2237 (1999).
9. G. Danabasoglu, P. R. Gent, Equilibrium climate sensitivity: Is it accurate to use a slab ocean model? *J. Climate* **22**, 2494–2499 (2009).
10. R. Knutti, M. A. A. Rugenstein, Feedbacks, climate sensitivity and the limits of linear models. *Phil.Trans.R.Soc.* **A373**, 20150146 (2015).
11. M. Rugenstein, J. Bloch-Johnson, J. Gregory, T. Andrews, T. Mauritsen, C. Li, T. L. Frölicher, D. Paynter, G. Danabasoglu, S. Yang, J.-L. Dufresne, L. Cao, G. A. Schmidt, A. Abe-Ouchi, O. Geoffroy, R. Knutti, Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophys. Res. Lett.* **47**, e2019GL8398 (2020).
12. J. M. Gregory, T. Andrews, Variation in climate sensitivity and feedback parameters during the historical period. *Geophys.Res. Lett.* **43**, 3911–3920 (2016).
13. R. J. Stouffer, S. Manabe, K. Bryan, Interhemispheric asymmetry in climate response to a gradual increase of atmospheric CO_2. *Nature* **342**, 660–662 (1989).
14. W. M. Washington, G. A. Meehl, Climate sensitivity due to increased CO_2: Experiments with a coupled atmosphere and ocean general circulation model. *Clim. Dynam.* **4**, 1–38 (1989).
15. S. Manabe, R. J. Stouffer, M. J. Spelman, K. Bryan, Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part I. Annual mean response. *J. Climate* **4**, 785–818 (1991).
16. F. P. Bretherton, K. Bryan, J. D. Woods, Time-dependent greenhouse-gas-induced climate change, in *Climate Change: The IPCC Scientific Assessment*, J. T. Houghton, G. J. Jenkins, J. J. Ephraums, Eds. (Cambridge Univ. Press, Cambridge, UK, 1990).
17. W. L. Gates, et al., Climate modelling, climate prediction and model validation, in *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment*, J. T. Houghton, B. A. Callander, S. K. Varney, Eds. (Cambridge Univ. Press, Cambridge, UK, 1992).
18. A. Kattenberg, F. Giorgi, H. Grassl, G. A. Meehl, J. F. B. Mitchell, R. J. Stoeffer, T. Tokioka, A. J. Weaver, T. M. L. Wigley, Climate models–projections of future climate, in *Climate Change 1995, The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, J. T. Houghton, BL. G. Meira Filbo, B. A. Callander, N. Harris, A. Kattenberg, K. Maskell, Eds. (Cambridge Univ. Press, 1996), pp. 285–357.
19. G. A. Meehl, G. J. Boer, C. Covey, M. Latif, R. J. Stouffer, Intercomparison makes for a better climate model. *EOS* **78**, 445–446 (1997).
20. U. Cubasch, G. A. Meehl, G. J. Boer, R. J. Stouffer, M. Dix, C. A. Senior, S. Raper, K. S. Yap, Projections of future climate change, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, C. I. Johnson, Eds. (Cambridge Univ. Press, 2001), pp. 525–582.

21. G. A. Meehl, T. F. Stocker, W. D. Collins, P. Friedlingstein, T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. G. Watterson, A. J. Weaver, Z. C. Zhao, Global Climate Projections, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, H. L. Miller, Eds. (Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA, 2007), pp. 747–845.

22. G. L. Russell, J. R. Miller, L.-C. Tsang, Seasonal oceanic heat transports computed from an atmospheric model. *Dyn. Atmos. Oceans* **9**, 253–271 (1985).

23. PALEOSENS Project Members, Making sense of palaeoclimate sensitivity. *Nature* **491**, 683–691 (2012).

24. J. M. Gregory, W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, K. D. Williams, A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.* **31**, L03205 (2004).

25. G. Flato, J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, M. Rummukainen, Evaluation of climate models, in *Climate Change 2013: The Physical Scientific Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, Cambridge, UK, 2013).

26. M. Righi, B. Andela, V. Eyring, A. Lauer, V. Predoi, M. Schlund, J. Vegas-Regidor, L. Bock, B. Brötz, L. de Mora, F. Diblen, L. Dreyer, N. Drost, P. Earnshaw, B. Hassler, N. Koldunov, B. Little, T.S. Loosveldt, K. Zimmermann, Earth System Model Evaluation Tool (ESMValTool) v2.0 – technical overview. *Geosci. Model Dev.* **13**, 1179–1199 (2020).

27. V. Eyring, L. Bock, A. Lauer, M. Righi, M. Schlund, B. Andela, E. Arnone, O. Bellprat, B. Brötz, L.-P. Caron, N. Carvalhais, I. Cionni, N. Cortesi, B. Crezee, E. Davin, P. Davini, K. Debeire, L. de Mora, C. Deser, D. Docquier, P. Earnshaw, C. Ehbrecht, B. K. Gier, N. Gonzalez-Reviriego, P. Goodman, S. Hagemann, S. Hardiman, B. Hassler, A. Hunter, C. Kadow, S. Kindermann, S. Koirala, N. V. Koldunov, Q. Lejeune, V. Lembo, T. Lovato, V. Lucarini, F. Massonnet, B. Müller, A. Pandde, N. Pérez-Zanón, A. Phillips, V. Predoi, J. Russell, A. Sellar, F. Serva, T. Stacke, R. Swaminathan, V. Torralba, J. Vegas-Regidor, J. von Hardenberg, K. Weigel, K. Zimmermann, ESMValTool v2.0 – Extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geosci. Model Dev.* 10.5194/gmd-2019-291 (2020).

28. J. Hansen, A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, J. Lerner, Climate Sensitivity: Analysis of feedback mechanisms, in *Climate Processes and Climate Sensitivity* (American Geophysical Union, 1984), Geophysical Monograph Series, vol. 29, pp. 130–163.

29. J. Hansen, G. Russell, A. Lacis, I. Fung, D. Rind, P. Stone, Climate response times: Dependence on climate sensitivity and ocean mixing. *Science* **229**, 857–859 (1985).

30. U. Siegenthaler, H. Oeschger, Transient temperature changes due to increasing CO_2 using simple models. *Ann. Glaciol.* **5**, 153–159 (1984).

31. T. M. L. Wigley, M. E. Schlesinger, Analytical solution for the effect of increasing CO_2 on global mean temperature. *Nature* **315**, 649–652 (1985).

32. S. C. B. Raper, J. M. Gregory, R. J. Stouffer, The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *J. Climate* **15**, 124–130 (2002).

33. R. Knutti, F. Joos, S. A. Müller, G.-K. Plattner, T. F. Stocker, Probabilistic climate change projections for CO_2 stabilization profiles. *Geophys. Res. Lett.* **32**, L20707 (2005).

34. J. Bloch-Johnson, R. T. Pierrehumbert, D. S. Abbot, Feedback temperature dependence determines the risk of high warming. *Geophys. Res. Lett.* **42**, 4973–4980 (2015).

35. T. Friedrich, A. Timmermann, M. Tigchelaar, O. Elison Timm, A. Ganopolski, Nonlinear climate sensitivity and its implications for future greenhouse warming. *Sci. Adv.* **2**, e1501923 (2016).

36. K. C. Armour, Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nat. Clim. Chang.* **7**, 331–335 (2017).

37. C. A. Senior, J. F. B. Mitchell, The time-dependence of climate sensitivity. *Geophys. Res. Lett.* **27**, 2685–2688 (2000).

38. M. Winton, K. Takahashi, I. M. Held, Importance of ocean heat uptake efficacy to transient climate change. *J. Climate* **23**, 2333–2344 (2010).

39. K. C. Armour, C. M. Bitz, G. H. Roe, Time-varying climate sensitivity from regional feedbacks. *J. Climate* **26**, 4518–4534 (2013).

40. T. Andrews, J. M. Gregory, M. J. Webb, The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate* **28**, 1630–1648 (2015).

41. T. Andrews, M. J. Webb, The dependence of global cloud and lapse rate feedbacks on the spatial structure of tropical pacific warming. *J. Climate* **31**, 641–654 (2018).

42. O. Geoffroy, D. Saint-Martin, Equilibrium- and transient-state dependencies of climate sensitivity: Are they important for climate projections? *J. Climate* **33**, 1863–1879 (2020).

43. D. Paynter, T. L. Frölicher, L. W. Horowitz, L. G. Silvers, Equilibrium climate sensitivity obtained from multimillennial runs of two GFDL climate models. *J. Geophys. Res.* **123**, 1921–1941 (2018).

44. B. Sanderson, Constraints on long term warming in a climate mitigation scenario. *Earth Syst. Dynam. Discuss.* 10.5194/esd-2019-82 (2020).

45. B. Sanderson, Relating climate sensitivity indices to projection uncertainty. *Earth Syst. Dynam. Discuss.* 10.5194/esd-2019-77 (2019).

46. J. T. Kiehl, Twentieth century climate model response and climate sensitivity. *Geophys. Res. Lett.* **34**, L22710 (2007).

47. P. M. Forster, T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, M. Zelinka, Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res. Atmos.* **118**, 1139–1150 (2013).

48. A. J. Dittus, E. Hawkins, L. J. Wilcox, R. Sutton, C. J. Smith, M. B. Andrews, P. M. Forster, Sensitivity of historical climate simulations to uncertain aerosol forcing. *Geophys. Res. Lett.* 10.1029/2019GL085806 (2020).

49. A. Gettelman, C. Hannay, J. T. Bacmeister, R. B. Neale, A. G. Pendergrass, G. Danabasoglu, J.-F. Lamarque, J. T. Fasullo, D. A. Bailey, D. M. Lawrence, M. J. Mills, High climate sensitivity in the Community Earth System Model Version 2 (CESM2). *Geophys. Res. Lett.* **46**, 8329–8337 (2019).

50. A. Bodas-Salcedo, J. P. Mulcahy, T. Andrews, K. D. Williams, M. A. Ringer, P. R. Field, G. S. Elsaesser, Strong dependence of atmospheric feedbacks on mixed-phase microphysics and aerosol-cloud interactions in HadGEM3. *J. Adv. Model. Earth Syst.* **11**, 1735–1758 (2019).

51. M. D. Zelinka, T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, K. E. Taylor, Causes of higher climate sensitivity in CMIP6 models. *Geophys. Res. Lett.* **47**, e2019GL085782 (2020).

52. A. Schmittner, N. M. Urban, J. D. Shakun, N. M. Mahowald, P. U. Clark, P. J. Bartlein, A. C. Mix, A. Rosell-Melé, Climate sensitivity estimated from temperature reconstructions of the Last Glacial Maximum. *Science* **334**, 1385–1388 (2011).

53. P. M. Forster, Inference of climate sensitivity from analysis of Earth's energy budget. *Annu. Rev. Earth Planet. Sci.* **44**, 85–106 (2016).

54. D. Jiménez-de-la-Cuesta, T. Mauritsen, Emergent constraints on Earth's transient and equilibrium response to doubled CO_2 from post-1970s global warming. *Nat. Geosci.* **12**, 902–905 (2019).

55. T. Schneider von Deimling, A. Ganopolski, H. Held, S. Rahmstorf, How cold was the Last Glacial Maximum? *Geophys. Res. Lett.* **33**, L14709 (2006).

56. R. Knutti, M. A. A. Rugenstein, G. C. Hegerl, Beyond equilibrium climate sensitivity. *Nat. Geosci.* **10**, 727–736 (2017).

57. M. R. Grose, J. Gregory, R. Colman, T. Andrews, What climate sensitivity index is most useful for projections? *Geophys. Res. Lett.* **45**, 1559–1566 (2018).

58. G. A. Meehl, A. Hu, J. M. Arblaster, J. Fasullo, K. E. Trenberth, Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *J. Climate* **26**, 7298–7310 (2013).

59. S. C. Hardiman, M. B. Andrews, T. Andrews, A. C. Bushell, N. J. Dunstone, H. Dyson, G. S. Jones, J. R. Knight, E. Neininger, F. M. O'Connor, J. K. Ridley, M. A. Ringer, A. A. Scaife, C. A. Senior, R. A. Wood, The impact of prescribed ozone in climate projections run with HadGEM3-GC3.1. *J. Adv. Model. Earth Syst.* **11**, 3443–3453 (2019).

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

61. J.-C. Golaz, P. M. Caldwell, L. P. Van Roekel, M. R. Petersen, Q. Tang, J. D. Wolfe, G. Abeshu, V. Anantharaj, X. S. Asay-Davis, D. C. Bader, S. A. Baldwin, G. Bisht, P. A. Bogenschutz, M. Branstetter, M. A. Brunke, S. R. Brus, S. M. Burrows, P. J. Cameron-Smith, A. S. Donahue, M. Deakin, R. C. Easter, K. J. Evans, Y. Feng, M. Flanner, J. G. Foucar, J. G. Fyke, B. M. Griffin, C. Hannay, B. E. Harrop, M. J. Hoffman, E. C. Hunke, R. L. Jacob, D. W. Jacobsen, N. Jeffery, P. W. Jones, N. D. Keen, S. A. Klein, V. E. Larson, L. R. Leung, H.-Y. Li, W. Lin, W. H. Lipscomb, P.-L. Ma, S. Mahajan, M. E. Maltrud, A. Mametjanov, J. L. McClean, R. B. McCoy, R. B. Neale, S. F. Price, Y. Qian, P. J. Rasch, J. E. J. Reeves Eyre, W. J. Riley, T. D. Ringler, A. F. Roberts, E. L. Roesler, A. G. Salinger, Z. Shaheen, X. Shi, B. Singh, J. Tang, M. A. Taylor, P. E. Thornton, A. K. Turner, M. Veneziani, H. Wan, H. Wang, S. Wang, D. N. Williams, P. J. Wolfram, P. H. Worley, S. Xie, Y. Yang, J.-H. Yoon, M. D. Zelinka, C. S. Zender, X. Zeng, C. Zhang, K. Zhang, Y. Zhang, X. Zheng, T. Zhou, Q. Zhu, The DOE E3SM Coupled Model Version 1: Overview and evaluation at standard resolution. *J. Adv. Model. Earth Syst.* **11**, 2089–2129 (2019).

62. I. M. Held, H. Guo, A. Adcroft, J. P. Dunne, L. W. Horowitz, J. Krasting, E. Shevliakova, M. Winton, M. Zhao, M. Bushuk, A. T. Wittenberg, B. Wyman, B. Xiang, R. Zhang, W. Anderson, V. Balaji, L. Donner, K. Dunne, J. Durachta, P. P. G. Gauthier, P. Ginoux, J.-C. Golaz, S. M. Griffies, R. Hallberg, L. Harris, M. Harrison, W. Hurlin, J. John, P. Lin, S.-J. Lin, S. Malyshev, R. Menzel, P. C. D. Milly, Y. Ming, V. Naik, D. Paynter, F. Paulot, V. Ramaswamy, B. Reichl, T. Robinson, A. Rosati, C. Seman, L. G. Silvers, S. Underwood,

N. Zadeh, Structure and performance of GFDL's CM4.0 climate model. *J. Adv. Model. Earth Syst.* **11**, 3691–3727 (2019).

63. M. Winton, A. Adcroft, J. P. Dunne, I. M. Held, E. Shevliakova, M. Zhao, H. Guo, W. Hurlin, J. Krasting, T. Knutson, D. Paynter, L. G. Silvers, R. Zhang, Climate sensitivity of GFDL's CM4.0. *J. Adv. Model. Earth Syst.* **12**, e2019MS001838 (2020).

64. T. Andrews, M. B. Andrews, A. Bodas-Salcedo, G. S. Jones, T. Kuhlbrodt, J. Manners, M. B. Menary, J. Ridley, M. A. Ringer, A. A. Sellar, C. A. Senior, Y. Tang, Forcings, feedbacks, and climate sensitivity in HadGEM3-GC3.1 and UKESM1. *J. Adv. Model. Earth Syst.* **11**, 4377–4394 (2019).

65. A. A. Sellar, C. G. Jones, J. P. Mulcahy, Y. Tang, A. Yool, A. Wiltshire, F. M. O'Connor, M. Stringer, R. Hill, J. Palmieri, S. Woodward, L. Mora, T. Kuhlbrodt, S. T. Rumbold, D. I. Kelley, R. Ellis, C. E. Johnson, J. Walton, N. L. Abraham, M. B. Andrews, T. Andrews, A. T. Archibald, S. Berthou, E. Burke, E. Blockley, K. Carslaw, M. Dalvi, J. Edwards, G. A. Folberth, N. Gedney, P. T. Griffiths, A. B. Harper, M. A. Hendry, A. J. Hewitt, B. Johnson, A. Jones, C. D. Jones, J. Keeble, S. Liddicoat, O. Morgenstern, R. J. Parker, V. Predoi, E. Robertson, A. Siahaan, R. S. Smith, R. Swaminathan, M. T. Woodhouse, G. Zeng, M. Zerroukat, UKESM1: Description and evaluation of the U.K. Earth System Model. *J. Adv. Model. Earth Syst.* **11**, 4513–4558 (2019).

66. H. Tatebe, T. Ogura, T. Nitta, Y. Komuro, K. Ogochi, T. Takemura, K. Sudo, M. Sekiguchi, M. Abe, F. Saito, M. Chikira, S. Watanabe, M. Mori, N. Hirota, Y. Kawatani, T. Mochizuki, K. Yoshimura, K. Takata, R. O'ishi, D. Yamazaki, T. Suzuki, M. Kurogi, T. Kataoka, M. Watanabe, M. Kimoto, Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geosci. Model Dev.* **12**, 2727–2765 (2019).

67. S. Yukimoto, H. Kawai, T. Koshiro, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T. Tanaka, M. Hosaka, S. Yabu, H. Yoshimura, E. Shindo, R. Mizuta, A. Obata, Y. Adachi, M. Ishii, The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *J. Meteorol. Soc. Jpn. Ser. II* **97**, 931–965 (2019).

68. T. Mauritsen, J. Bader, T. Becker, J. Behrens, M. Bittner, R. Brokopf, V. Brovkin, M. Claussen, T. Crueger, M. Esch, I. Fast, S. Fiedler, D. Fläschner, V. Gayler, M. Giorgetta, D. S. Goll, H. Haak, S. Hagemann, C. Hedemann, C. Hohenegger, T. Ilyina, T. Jahns, D. Jimenéz-de-la-Cuesta, J. Jungclaus, T. Kleinen, S. Kloster, D. Kracher, S. Kinne, D. Kleberg, G. Lasslop, L. Kornblueh, J. Marotzke, D. Matei, K. Meraner, U. Mikolajewicz, K. Modali, B. Möbis, W. A. Müller, J. E. M. S. Nabel, C. C. W. Nam, D. Notz, S.-S. Nyawira, H. Paulsen, K. Peters, R. Pincus, H. Pohlmann, J. Pongratz, M. Popp, T. J. Raddatz, S. Rast, R. Redler, C. H. Reick, T. Rohrschneider, V. Schemann, H. Schmidt, R. Schnur, U. Schulzweida, K. D. Six, L. Stein, I. Stemmler, B. Stevens, J.-S. Storch, F. Tian, A. Voigt, P. Vrese, K.-H. Wieners, S. Wilkenskjeld, A. Winkler, E. Roeckner, Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).

69. K. Wyser, T. van Noije, S. Yang, J. von Hardenberg, D. O'Donnell, R. Döscher, On the increased climate sensitivity in the EC-Earth model from CMIP5 to CMIP6. *Geophys. Model Dev.* 10.5194/gmd-2019-282 (2020).

70. N. C. Swart, J. N. S. Cole, V. V. Kharin, M. Lazare, J. F. Scinocca, N. P. Gillett, J. Anstey, V. Arora, J. R. Christian, S. Hanna, Y. Jiao, W. G. Lee, F. Majaess, O. A. Saenko, C. Seiler, C. Seinen, A. Shao, M. Sigmond, L. Solheim, K. von Salzen, D. Yang, B. Winter, The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.* **12**, 4823–4873 (2019).

Citation: G. A. Meehl, C. A. Senior, V. Eyring, G. Flato, J.-F. Lamarque, R. J. Stouffer, K. E. Taylor, M. Schlund, Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Sci. Adv.* **6**, eaba1981 (2020).