Supporting Information for "State-dependence of the Climate Sensitivity in Earth System Models of Intermediate Complexity"

Patrik L. Pfister[1,2] **and Thomas F. Stocker**[1,2]

[1]Climate and Environmental Physics, Physics Institute, University of Bern, 3012 Bern, Switzerland

[2]Oeschger Center for Climate Change Research, University of Bern, 3012 Bern, Switzerland

Contents of this file

1. Text Sections S1 to S3
2. Table S1
3. Figures S1 to S5

Corresponding author: Patrik Pfister, pfister@climate.unibe.ch

Introduction

This Supporting Information contains extended descriptions of our Methods (Sections S1, S2.1) and Results (Sections S2.2–S3, all Figures). Section S1 describes the restart offset correction we applied to calculate temperature anomalies in the EMIC ensemble. In Section S2, we describe the two different ECS extrapolation methods that were used (2.1). We further describe extrapolation problems arising for the $4\times CO_2$ (2.2) and for specific EMICs (2.3), and compare Gregory extrapolations from short and long time intervals (2.4). This is complemented by Table S1 that lists all ECS estimates, Figures S1–S3 that show the fits from both extrapolation methods, and Figure S4 that compares abrupt-forcing simulations to $1\%CO_2$ simulations in the Bern3D-LPX model. Section S3.1 describes additional Bern3D-LPX simulations with different initial conditions that are shown in Figure S5. Section S3.2 complements Section 3.4 by describing fast local feedback changes, which are visible in Figure 3 but not relevant for state-dependence.

S1 Calculation of EMIC temperature anomalies

The control simulation H_CTR available from EMICAR5 was run in parallel to the historic simulations that we do not discuss here [*Eby et al.*, 2013]. No published control simulations were run in parallel to the idealized experiments. Nevertheless, as the idealized experiments were supposed to be restarted from the end of H_CTR, subtracting the H_CTR mean from the idealized simulations is a reasonable approach to calculate anomalies (Michael Eby 2017, personal communication). For some of the EMICs however, the 1%CO_2 simulations start with a substantial temperature offset from both the mean and end of H_CTR, suggesting that these simulations were not restarted from H_CTR. To check this, we compared the first decadal average temperature of the 1%CO_2 simulation ($\bar{T}^{CO2}_{start}$) to the last decadal average temperature of H_CTR ($\bar{T}^{CTRL}_{end}$). If the difference $\bar{T}^{CO2}_{start} - \bar{T}^{CTRL}_{end}$ was in the interval $[0,\bar{R}^{CO2}_{start}/R(2\times CO_2)\cdot ECS(2\times CO_2)]$, we did not apply any correction to the ΔT anomaly; if the difference was larger (smaller), we corrected ΔT by the offset compared to the upper (lower) end of this interval. Note that this is a conservative correction: the lower limit corresponds to zero warming in presence of the CO_2 forcing, and the upper limit corresponds to a realized warming fraction of one, i.e., instant equilibration with the forcing. We choose this correction to allow some leeway for variability on multidecadal time scales, in addition to faster variability which is filtered out by the decadal averaging. An offset correction was applied to three EMICs: FAMOUS-XDBUA (+0.08 K), LOVECLIM1.2 (−0.22 K) and SPEEDO V2.0 (+0.28 K). The ECS and transient climate response (TCR) estimates for these models are therefore also corrected by these values. This slightly affects ECS state dependence of these models, as the temperature offset is halved for ECS(4×CO_2).

The anomalies of all other variables such as ocean heat uptake (N) were not corrected, based on the assumption that the simulations with temperature offsets were restarted from a different equilibrium simulation ($N \approx 0$). This is difficult to evaluate, due to the large N variability in those three models and the lack of straightforward definition for an upper limit of plausible $\bar{N}^{CO2}_{start}$.

S2 ECS estimation methods

S2.1 Extended description of extrapolation methods

In addition to the 1000-year temperatures, two complementary ECS estimates were obtained using different extrapolation methods. The second estimate is obtained using the Gregory method [*Gregory et al.*, 2004], by linearly fitting ΔT against N (equation 1). The ECS estimate ΔT_{lin} is the extrapolated temperature of the linear fit that corresponds to $N = 0$ (equilibrium). This method is only applicable for constant forcing, therefore we consistently fit years 150–1000 in $2\times CO_2$ and $4\times CO_2$. For comparison, we also perform a "Gregory fit" over years 20–150 for $2\times CO_2$, which is the time interval that was used for the latest CMIP5 ECS fits [*Andrews et al.*, 2015]. All fits are shown in Figure S1. No Gregory fits can be obtained for the UMD model, for which no N output was uploaded to EMICAR5.

The third estimate is an exponential fit on three different time scales τ_i (roughly decadal, centennial and millennial) to the ΔT time series:

$$\Delta T_{fit}(t) = \Delta T_{exp}(1 - a_1 e^{-t/\tau_1} - a_2 e^{-t/\tau_2} - a_3 e^{-t/\tau_3}), \tag{S1}$$

where ΔT_{exp} is the ECS estimate from this method. Together with the dampening amplitudes a_i and the timescales τ_i, we count seven free parameters that are optimized to best represent ΔT using least squares regression. Such a multi-exponential temperature response is only expected for constant forcing, therefore we fit years 150–1000 in $4\times CO_2$. We perform a consistent fit ($\Delta T_{exp}^{150-1000}$) in $2\times CO_2$ for comparison, but also obtain an additional ECS estimate from the whole time series (ΔT_{exp}^{1-1000}). The fits are shown in Figures S2 and S3. ECS extrapolation based on equation S1 has also been employed by *Steinacher and Joos* [2016]; *Voss and Mikolajewicz* [2001] used a similar extrapolation, based on *MacKay and Ko* [1997]. Equation S1 can also be simply used for timescale analysis in combination with an ECS estimate from a different method [*Caldeira and Myhrvold*, 2013; *Knutti and Rugenstein*, 2015].

We constrain the exponential dampening time scale to a maximum of 5000 years, long enough to represent all heat uptake processes relevant for the EMICs. As some EMICs include slow carbon reservoirs such as sediments, longer time scales may be required for emission-driven projections, but not for concentration-driven projections.

S2.2 Scenario dependence of ECS extrapolations

As motivated in Section 2.1 of the main text, our analysis focuses on $4\times CO_2$ simulations where the CO_2 concentration is not increased abruptly but by 1% per year. This Supporting Section summarizes ECS extrapolation difficulties arising from the analysis of this scenario. Section S2.3 describes examples from specific EMICs, $4\times CO_2$ in that section always refers to the 1%-scenario. We also obtain ECS estimates from the abrupt $4\times CO_2$ simulations for comparison (Table S1). These confirm the robustness of the non-exponential $4\times CO_2$ estimates (T^{1000} and $\Delta T_{lin}^{150-1000}$). This robustness is also illustrated by example of the Bern3Dnew model in Figure S4.

The causes of the problematic exponential fits from the $4\times CO_2$ simulation are twofold. Both arise from the use of a 1%/year stabilization scenario. Firstly, only simulation years 150–1000 can be used for the fit, because the warming evolution of earlier years is shaped by the forcing. The use of this time scale alone can falsify the extrapolation, as described in Section S2.3 for some $\Delta T_{exp}^{150-1000}$ estimates from the abrupt $2\times CO_2$ simulations. Secondly, although T^{1000} and $\Delta T_{lin}^{150-1000}$ are very similar between the 1%/year and abrupt scenarios (Table S1), $\Delta T_{exp}^{150-1000}$ estimates of the two scenarios differ notably. This is because the scenario influence on the early warming evolution, e.g., years 150–400 in Bern3Dnew (Fig. S4), impacts the fit parameters in Equation (S1) and thereby also the $\Delta T_{exp}^{150-1000}$ extrapolations. We suggest that the $\Delta T_{exp}^{150-1000}$ estimates from the abrupt scenario are generally more accurate: For most EMICs with notable $\Delta T_{exp}^{150-1000}$ differences between scenarios, the estimate from the abrupt scenario agrees better with the more scenario-independent Gregory estimate (B3, B3n, MI, LO, ML, C3). For Bern3Dnew, it is also closer to the true ECS ΔT_{5000}.

S2.3 Extrapolation difficulties for specific EMICs

Figure S1 shows the Gregory fits for both $2\times CO_2$ and $4\times CO_2$ simulations, Figures S2 and S3 show the exponential fits for the $2\times CO_2$ and $4\times CO_2$ simulation, respectively. In the following, we describe some problematic fits for single models that could not be discussed separately in the main text.

As described in Section 3.2, $\Delta T_{lin}^{150-1000}$ is larger than $\Delta T_{exp}^{150-1000}$ in most EMICs, mainly those with the strongest radiative imbalances ($N > 0$) after 1000 years (DC, ML, UV, C2, B3n; B3, FA, SP only for $4\times CO_2$, Figure S1). The opposite is true only for the

$4\times CO_2$ simulations of both Bern3D versions and MIROC-lite, where $\Delta T_{exp}^{150-1000}$ probably overestimates ECS (verified by ΔT^{5000} in case of the Bern3Dnew), and for both LOVECLIM simulations where $\Delta T_{lin}^{150-1000}$ probably underestimates ECS. These exceptions are explained below.

For some EMICs, the $\Delta T_{exp}^{150-1000}$ fit includes the maximum timescale of 5000 years (Section S2.1), indicating that the fit is limited to this maximum. This goes along with vastly overestimated ECS values due to the long warming extrapolations, therefore we discard these fits as unrealistic. One such example is the $2\times CO_2$ simulation of Bern3Dold, which yields a usable ECS fit for ΔT_{exp}^{1-1000}, but not for $\Delta T_{exp}^{150-1000}$ where some unforced warming near the end of the simulation has too much weight on the fit. This late warming is probably also the reason why $\Delta T_{exp}^{1-1000}(2\times CO_2)$ and $\Delta T_{exp}^{150-1000}(4\times CO_2)$ are larger than the corresponding Gregory estimates $\Delta T_{lin}^{150-1000}$. Similar short-lived warming accelerations are also found in MIROC-lite and UVic, leading to high $\Delta T_{exp}^{150-1000}$ from long fit timescales of 2000–5000 years, which probably overestimate ECS. The $2\times CO_2$ simulation of UVic is an exception, because its notable warming curvature between years 200 and 1000 is fitted by a maximum exponential time scale of about 350 years. Although no late warming acceleration is visible for the $4\times CO_2$ simulation of Bern3Dnew, $\Delta T_{exp}^{150-1000}$ also overestimates ECS by roughly 0.15°C in that simulation, again related to a long fitting timescale of $\sim$1900 years.

For LOVECLIM, both exponential fits yield a substantially larger ECS than the Gregory method, especially ΔT_{exp}^{1-1000}. LOVECLIM is far from equilibrium after 1000 years according to its slope in ΔT increase, although N is close to zero. We assume that the N anomaly we use is biased low and should be offset-corrected similar to ΔT, but the magnitude of this correction is unclear (Section S1). Therefore, the exponential estimates of ECS are probably more realistic than the Gregory estimates for LOVECLIM. While the ΔT_{exp}^{1-1000} fit better represents early temperature evolution, we cannot exclude the ECS estimate obtained from $\Delta T_{exp}^{150-1000}$.

SPEEDO's $2\times CO_2$ simulation exhibits a strong centennial oscillation on top of the forced warming, which is also seen in the control run (not shown), and makes an exponential fit impossible (last panel of Figure S2). The oscillation vanishes at higher forcing ($4\times CO_2$), enabling a reasonable fit there (last panel of Figure S3).

S2.4 Gregory estimates of ECS from 150-year simulations

How do the Gregory ECS estimates from years 20–150 compare to the ones from years 150–1000? For most EMICs, they are similar, because the ΔT to N relation is close to linear across the whole simulation duration, indicating that the global feedback strength λ is roughly constant. This does not necessarily indicate that the 20–150 ECS estimates obtained for CMIP5 [*Andrews et al.*, 2015] would be consistent with longer time scales, because the slope $-\lambda$ in most GCMs decreases with time during the simulated 150 years [*Winton et al.*, 2010; *Armour et al.*, 2013; *Geoffroy et al.*, 2013; *Andrews et al.*, 2015; *Knutti and Rugenstein*, 2015] and may change further thereafter.

The only EMICs that evidently show a decreasing $-\lambda$ in $2\times CO_2$ are LOVECLIM and Bern3Dnew, indicating that these models have time-dependent feedbacks and/or a high ocean heat uptake efficacy [*Winton et al.*, 2010; *Geoffroy et al.*, 2013]. This will be explored in a separate study. It explains the substantially smaller ΔT_{lin}^{20-150} compared to $\Delta T_{lin}^{150-1000}$ in these models. Other EMICs show smaller differences, some also of opposite sign (e.g., the DCESS model). SPEEDO's years 20–150 yield an obvious ECS misfit due to a slow initial change in N (Figure S1), related to the model's unforced centennial oscillations.

For models with strong variability (LOVECLIM, FAMOUS and SPEEDO), an ensemble of simulations would be required to ensure the robustness of differences between ΔT_{lin}^{20-150} and $\Delta T_{lin}^{150-1000}$ fits [*Knutti and Rugenstein*, 2015]. This is provided for LOVECLIM where the EMIC-AR5 output we use is an ensemble mean of 10 simulations [*Eby et al.*, 2013], but not for FAMOUS and SPEEDO where single simulations are used.

In Summary, Gregory fits from years 20–150 are in good agreement with the longer Gregory fits for most EMICs, but not for those with strong transient sensitivity changes. This indicates that a more sophisticated method is required in the presence of evolving sensitivities for the extrapolation of short integrations, for example using a 2-layer EBM [*Geoffroy et al.*, 2013].

Feedback changes on longer timescales (> 1000 years) can only be evaluated for Bern3Dnew, where a slight $-\lambda$ increase is simulated between year 1000 and 5000 due to the transiently weakening albedo feedback. $\Delta T_{lin}^{150-1000}$ does not account for this increase, leading to a slight ECS overestimation (3.1 K) compared to ΔT^{5000} (3.0 K).

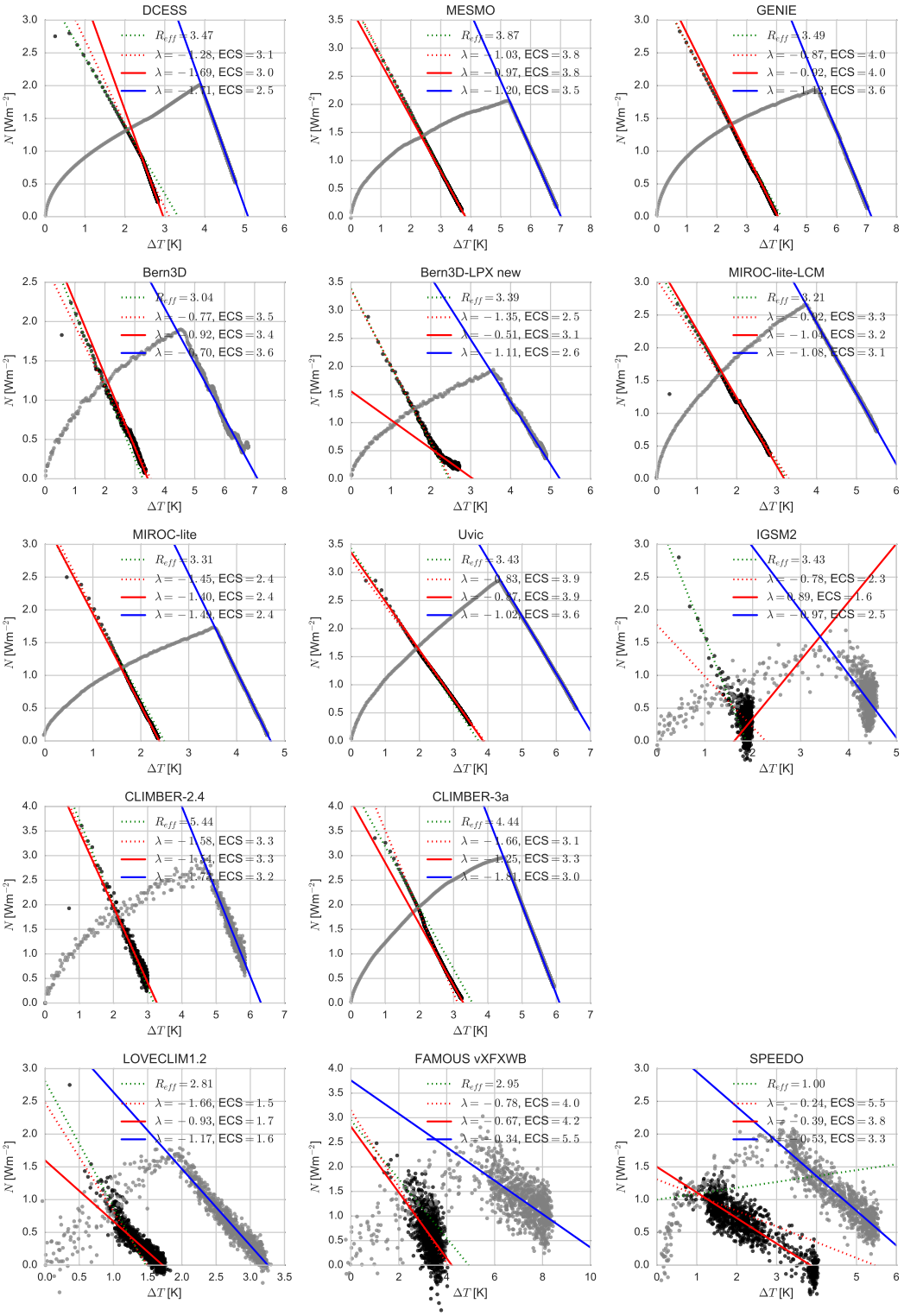

Figure S1. Estimates of ECS and effective radiative forcing R_{eff} of the EMIC ensemble using the Gregory method [*Gregory et al.*, 2004; *Andrews et al.*, 2015]. Raw model output is indicated by black ($2\times CO_2$) and grey ($4\times CO_2$) dots. All linear fits shown here are performed on this raw output (solid lines: years 150–1000, red dotted: years 20–150, green dotted: years 1–20). Therefore, they differ slightly from the ECS estimates given in Table 1 for years 150–1000, which were obtained from 50-year running means. The most notable difference is for the IGSM2 model, which is close to equilibrium already 200 years after the perturbation (Figure S2), thereafter the misfit shown here primarily acts on variability. R_{eff} is estimated from the first 20 years after N peaks. R_{eff} values in the legend are in Wm^{-2}, λ values are in Wm^{-2}K^{-1}, ECS values in K.

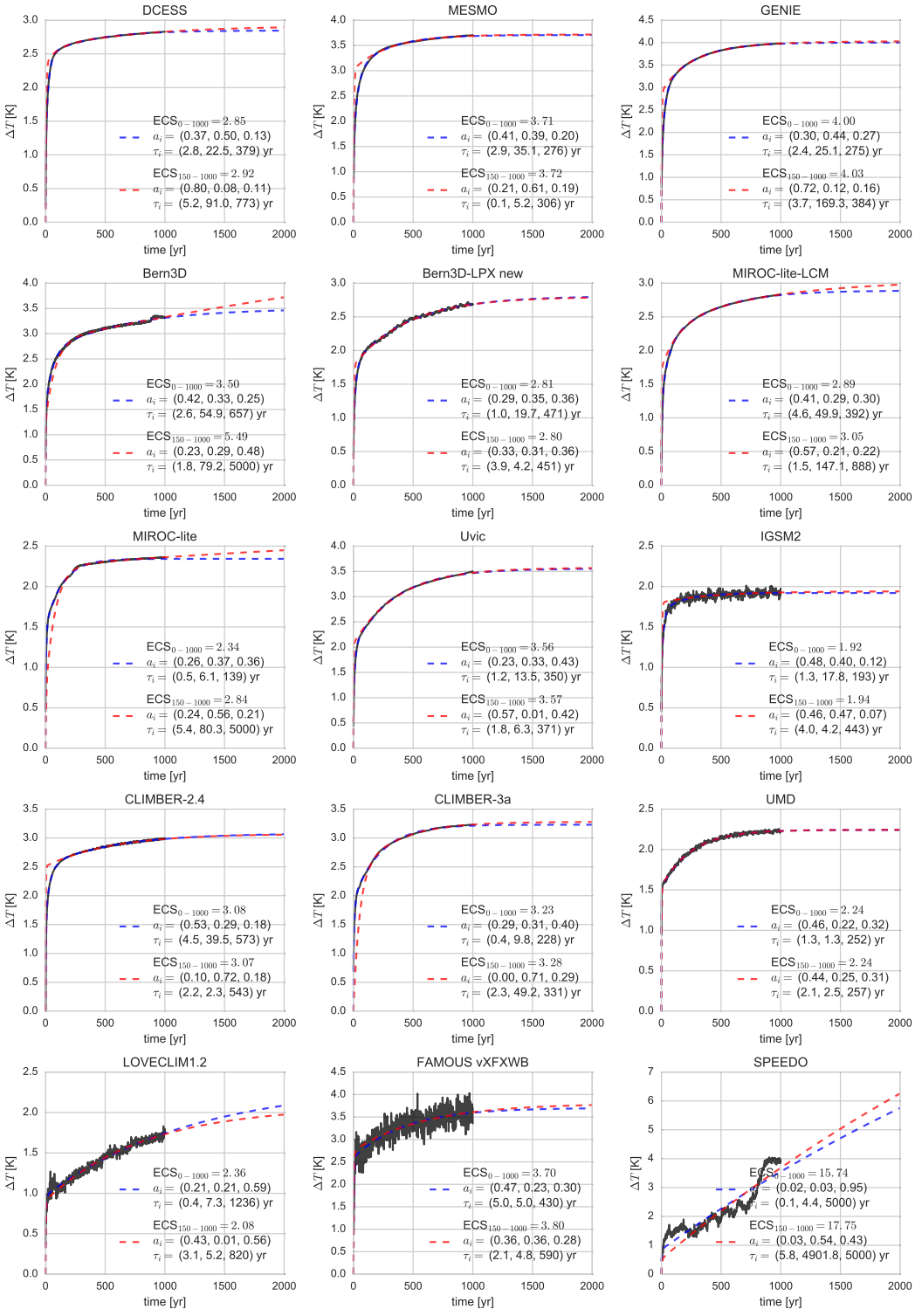

Figure S2. Estimates of ECS of the EMIC ensemble using an exponential fit (equation 3 in the main text), for the 2×CO_2 simulation. Blue fits are obtained from the whole temperature time series (ΔT_{exp}^{1-1000}), red fits only from years 150–1000 (ΔT_{exp}^{1-1000}), consistent with 4×CO_2 simulations. Fit amplitudes a_i and time scales τ_i (i=1,2,3), and fitted ECS estimates are listed in the legend. Fits that are constrained by the maximum time scale (τ_3 = 5000 years) are considered unrealistic (Section S2.3).

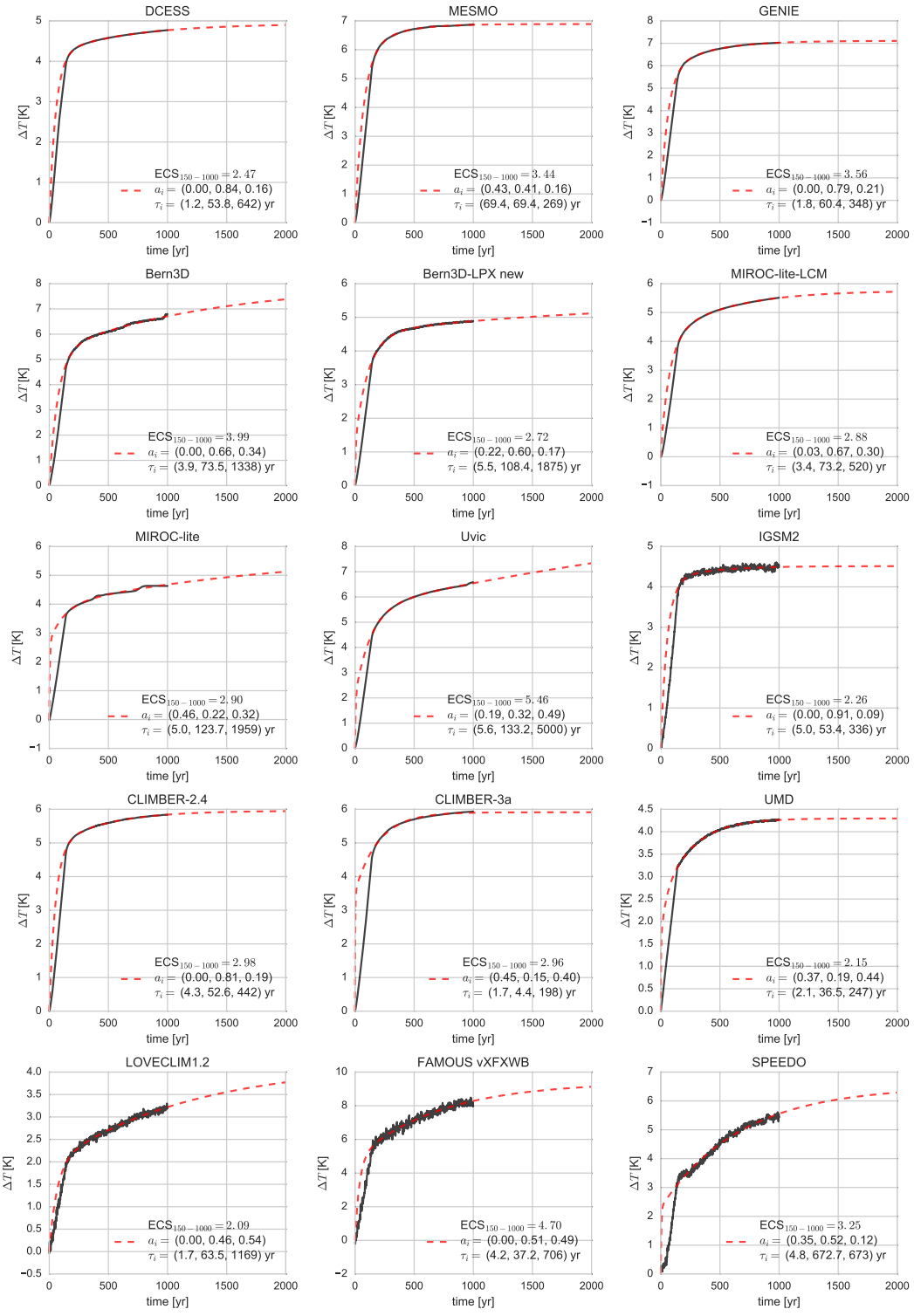

Figure S3. Like Figure S2, but for the $4\times CO_2$ simulations. No fit is obtained from the full time series duration, because the forcing is constant only from year 140 onwards.

(a) $2\times CO_2$

	DC	ME	GE	B3	B3n	ML	MI	UV	I2	C2	C3	UM	LO	FA	SP	mean
ΔT^{5000}					3.00											
ΔT^{1000}	2.83	3.70	3.98	3.32	2.68	2.83	2.36	3.50	1.92	2.99	3.23	2.22	1.74	3.60	3.89	2.95
ΔT_{lin}^{20-150}	3.11	3.80	4.05	3.53	2.50	3.30	2.38	3.90	2.27	3.27	3.12		1.49	4.04	5.46	3.09
$\Delta T_{lin}^{150-1000}$	2.96	3.80	4.04	3.43	3.10	3.20	2.40	3.86	2.11	3.27	3.32		1.73	4.44	3.82	3.20
$\Delta T_{exp}^{150-1000}$	2.92	3.72	4.03		2.80	3.05		3.57	1.94	3.07	3.28	2.24	2.08	3.80		3.07
ΔT_{exp}^{1-1000}	2.85	3.71	4.00	3.50	2.81	2.89	2.34	3.56	1.92	3.08	3.23	2.24	2.36	3.70		3.05

(b) $4\times CO_2$, 1%/year (divided by 2)

	DC	ME	GE	B3	B3n	ML	MI	UV	I2	C2	C3	UM	LO	FA	SP	mean
ΔT^{5000}					2.56											
ΔT^{1000}	2.38	3.44	3.51	3.39	2.44	2.76	2.32	3.29	2.25	2.92	2.97	2.13	1.61	4.04	2.72	2.83
$\Delta T_{lin}^{150-1000}$	2.54	3.50	3.56	3.56	2.63	3.10	2.35	3.60	2.42	3.15	3.06		1.63	5.62	3.26	3.12
$\Delta T_{exp}^{150-1000}$	2.47	3.44	3.56	3.99	2.72	2.88	2.90		2.26	2.98	2.96	2.15	2.09	4.70	3.25	3.01

(c) $4\times CO_2$, abrupt (divided by 2)

	DC	ME	GE	B3	B3n	ML	MI	UV	I2	C2	C3	UM	LO	FA	SP	mean
ΔT^{1000}	2.39	3.44	3.52	3.43	2.46	2.77	2.32	3.31	2.26	2.93	2.97	2.13	1.61	4.15	2.77	2.85
$\Delta T_{lin}^{150-1000}$	2.54	3.51	3.59	3.62	2.62	3.11	2.35	3.56	2.43	3.16	3.05		1.66	5.89	3.33	3.16
$\Delta T_{exp}^{150-1000}$	2.44	3.43	3.55	3.81	2.59	3.02	2.46		2.27	2.97	3.07	2.15	1.91	5.00	2.92	3.03

Table 1. ECS (in °C) of the EMICAR5 ensemble [*Eby et al.*, 2013], estimated from (a) $2\times CO_2$, (b) $4\times CO_2$ simulations with a 1%/year ramp-up (used in the main text), and (c) abrupt $4\times CO_2$ simulations for comparison. 5000-year B3new temperature in (c) is identical to (b) (not shown). Three ECS estimation methods are compared (Section S2.1): no extrapolation (ΔT), exponential extrapolation (ΔT_{exp}), and the Gregory method (ΔT_{lin}). Superscripts indicate the simulation years used for the estimations. Multi-model means only include the 10 models that have an ECS estimate for all methods. Model abbreviations roughly in order of ascending atmospheric complexity: DCESS v1 (DC), MESMO v1.0 (ME), GENIE 2-2-7 (GE), Bern3D-LPX (B3), MIROC-lite-LCM (ML), MIROC-lite (MI), UVic v2.9 (UV), IGSM v2.2 (I2), CLIMBER-2.4 (C2), CLIMBER-3α (C3), UMD v2.0 (UM), LOVECLIM v1.2 (LO), FAMOUS vXFXWB (FA), and SPEEDO (SP).

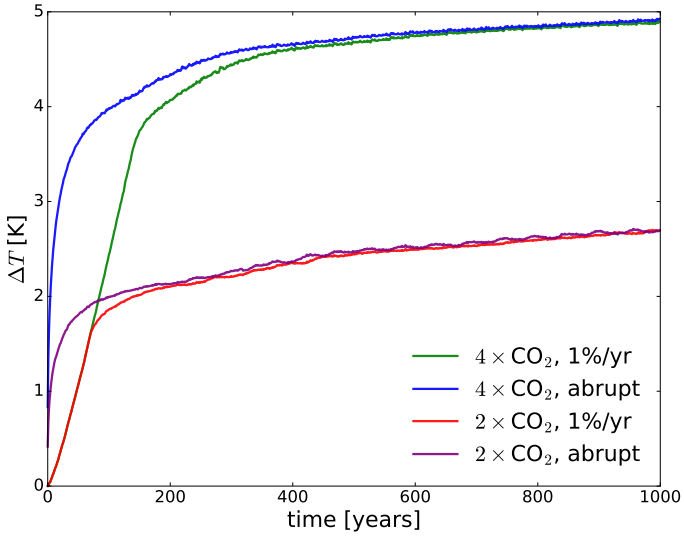

Figure S4. Comparison between temperature evolutions of 1%CO_2 and abrupt-forcing scenarios in the Bern3D-LPX model.

S3 Additional Bern3D-LPX results

S3.1 Simulations starting from different initial states

The term "state-dependence" is commonly used synonymously with "forcing-dependence". Climate model state-dependence assessments are therefore often carried out by imposing different levels of forcing on one given initial state [e.g., *Colman and McAvaney*, 2009; *Kutzbach et al.*, 2013, this study]. However, in more stringent terms "state-dependence" means "dependence on initial state", therefore some studies have instead doubled CO_2 starting from different initial states [e.g., *Weaver et al.*, 2007; *Jonko et al.*, 2013]. This Supporting Section assesses whether the two methods produce similar results in the Bern3D-LPX model, i.e., whether the equivalence of state-dependence and forcing-dependence of ECS holds this model.

We have performed additional abrupt-forcing simulations in the *standard* configuration, but starting from three different initial states with 1.1×, 1.5×, and 2× preindustial CO_2 concentration (Figure S5). Abrupt-forcing simulations were restarted from year 5000 of these initial simulations, to ensure that the initial states are in equilibrium. The abrupt-forcing simulations, with forcings amounting to 1.5×, 2× and 4× the initial CO_2 concentration, were also run for 5000 years, consistent with the simulations shown in Figure 2.

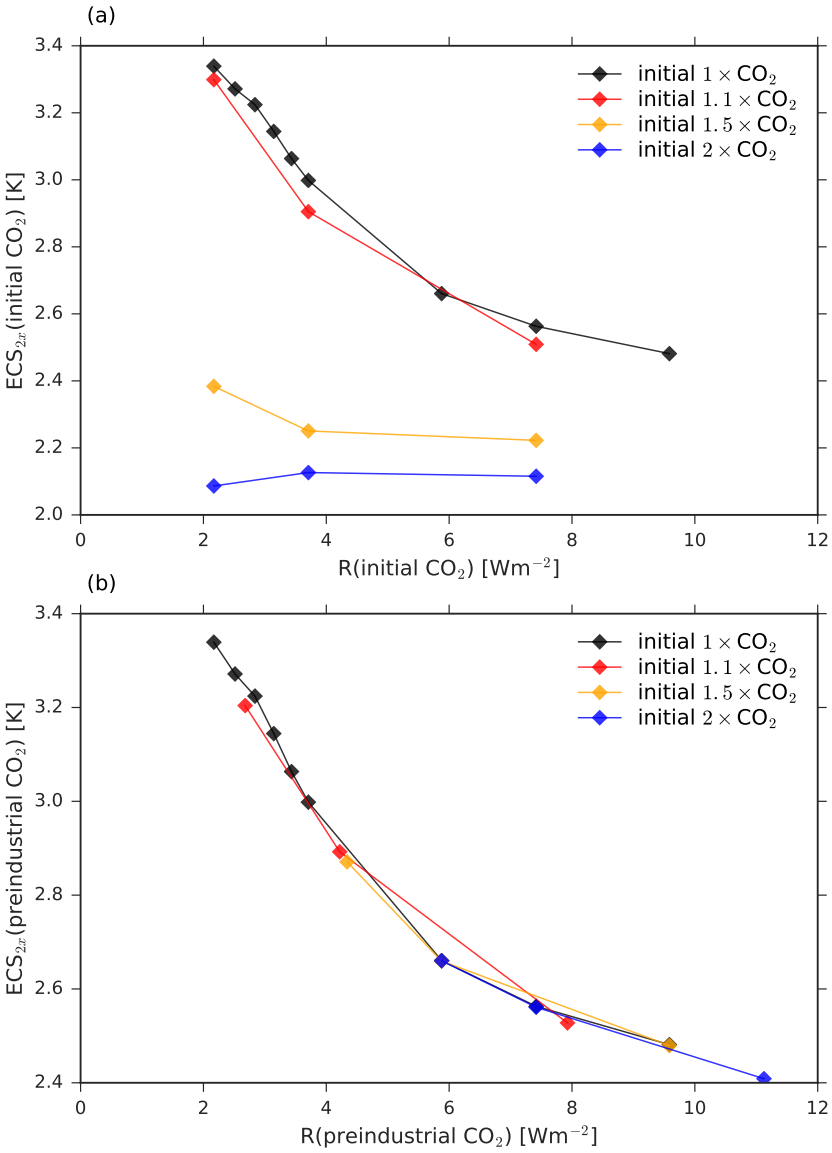

Figure S5. ECS state-dependence for experiments starting from different initial equilibrium states with elevated CO_2 concentrations (colors). The black dots represent the reference simulations starting from the preindustrial state, identical to the "standard" configuration in Figure 2. In a), ECS is calculated with reference to the concentration of the initial state; in b), with respect to the preindustrial concentration.

In Figure S5a, ECS is calculated with respect to the corresponding initial state. For example, the second dot of the blue line corresponds to a second doubling of the CO_2 concentration from an initial state with 2×preindustial CO_2 concentration. This second doubling causes an additional forcing of 3.7 Wm^{-2} and an additional warming of roughly 2.1 K, which is much less than the warming caused by the first doubling (black dot at the same forcing, 3.0 K), due to the reduced ice-albedo feedback (Sections 3.3, 3.4 in the main article). ECS for elevated initial states (1.5× and 2× preindustial CO_2) is much less state-dependent than ECS for the preindustrial initial condition, because sea-ice availability is drastically reduced in these initial states.

In Figure S5b, ECS is calculated with respect to preindustial, i.e., the forcing and warming of the abrupt-forcing simulation is added to the forcing and warming of the initial state compared to preindustrial. In this view, the second dot of the blue line corresponds to a 4×CO_2 simulation (with forcing raised in two steps), and it almost exactly matches the ECS of the abrupt 4×CO_2 simulation (underlying black dot). Similarly, all other simulations closely lie on the ECS state-dependence curve resulting from the preindustrial *standard* simulations (the black line is identical in Figures 2a, S5a and S5b).

These experiments confirm that the results shown in Figure 2 indeed correspond to state-dependence. In other words, Bern3D-LPX warming is independent of the concentration scenario path for the range of forcings and scenarios considered here.

S3.2 Initial increase in λ_{loc}(40–90°S)

Lastly, we describe the causes of the initial "ramp-up" of λ_{loc}(40–90°S) during the first 10-20 simulation years (Figure 3 c,d in the paper), which was not elaborated in the main text because it is not related to state-dependence. The ramp-up is due to a combination of delayed response in ΔT(40–90°S) compared to $(N - R)$(40–90°S), and a fast response of the atmospheric meridional heat transport F_{atm} across 40°S. The latter is an F_{atm} strengthening, because atmospheric warming over the Southern Ocean is slower than warming over the lower latitudes including land masses. Therefore, the atmospheric temperature gradient ∇T across 40°S) increases initially, and with it the sensible heat component of F_{atm}, which is parameterized as $K(\theta)\nabla T$ (where $K(\theta)$ is the atmospheric eddy-diffusivity, *Ritz et al.* [2011]). In terms of the local EBM, the change in meridional heat flux convergence ($-\Delta(\nabla \cdot \mathbf{F})$(40–90°S)) is initially negative in response to the abrupt forc-

ing. $-\Delta(\nabla \cdot \mathbf{F})(40–90°\mathrm{S})$ becomes positive after about 10 years and continues increasing rapidly up to about 20 years. This is due to the decreasing meridional heat transport, to which both the decreasing SOMOC and decreasing F_{atm} contribute, the latter due to decreasing ∇T by polar amplification.

References

Andrews, T., J. M. Gregory, and M. J. Webb (2015), The Dependence of Radiative Forcing and Feedback on Evolving Patterns of Surface Temperature Change in Climate Models, *J. Clim.*, *28*(4), 1630–1648, doi:10.1175/JCLI-D-14-00545.1.

Armour, K. C., C. M. Bitz, and G. H. Roe (2013), Time-Varying Climate Sensitivity from Regional Feedbacks, *J. Clim.*, *26*(13), 4518–4534, doi:10.1175/JCLI-D-12-00544.1.

Caldeira, K., and N. P. Myhrvold (2013), Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration, *Environ. Res. Lett.*, *8*(3), 034039, doi:10.1088/1748-9326/8/3/034039.

Colman, R., and B. McAvaney (2009), Climate feedbacks under a very broad range of forcing, *Geophys. Res. Lett.*, *36*(1), L01702, doi:10.1029/2008GL036268.

Eby, M., A. J. Weaver, K. Alexander, K. Zickfeld, A. Abe-Ouchi, A. A. Cimatoribus, E. Crespin, S. S. Drijfhout, N. R. Edwards, A. V. Eliseev, G. Feulner, T. Fichefet, C. E. Forest, H. Goosse, P. B. Holden, F. Joos, M. Kawamiya, D. Kicklighter, H. Kienert, K. Matsumoto, I. I. Mokhov, E. Monier, S. M. Olsen, J. O. P. Pedersen, M. Perrette, G. Philippon-Berthier, A. Ridgwell, A. Schlosser, T. Schneider von Deimling, G. Shaffer, R. S. Smith, R. Spahni, A. P. Sokolov, M. Steinacher, K. Tachiiri, K. Tokos, M. Yoshimori, N. Zeng, and F. Zhao (2013), Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity, *Clim. Past*, *9*(3), 1111–1140, doi:10.5194/cp-9-1111-2013.

Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. J. L. Olivie, and S. Tyteca (2013), 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, *J. Clim.*, *26*(6), 1859–1876, doi:10.1175/JCLI-D-12-00196.1.

Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative forcing and climate sensitivity, *Geophys. Res. Lett.*, *31*(3), L03205, doi:10.1029/2003GL018747.

Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu (2013), Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part II: Variation of Climate Feedbacks and Sensitivity with Forcing, *J. Clim.*, *26*(9), 2784–2795, doi:10.1175/JCLI-D-12-00479.1.

Knutti, R., and M. A. A. Rugenstein (2015), Feedbacks, climate sensitivity and the limits of linear models, *Phil. Trans. R. Soc. A*, *373*(2054), doi:10.1098/rsta.2015.0146.

Kutzbach, J. E., F. He, S. J. Vavrus, and W. F. Ruddiman (2013), The dependence of equilibrium climate sensitivity on climate state: Applications to studies of climates colder than present, *Geophys. Res. Lett.*, *40*(14), 3721–3726, doi:10.1002/grl.50724.

MacKay, R. M., and M. K. W. Ko (1997), Normal modes and the transient response of the climate system, *Geophys. Res. Lett.*, *24*(5), 559–562, doi:10.1029/97GL00286.

Ritz, S. P., T. F. Stocker, and F. Joos (2011), A Coupled Dynamical Ocean-Energy Balance Atmosphere Model for Paleoclimate Studies, *J. Clim.*, *24*(2), 349–375, doi:10.1175/2010JCLI3351.1.

Steinacher, M., and F. Joos (2016), Transient Earth system responses to cumulative carbon dioxide emissions: linearities, uncertainties, and probabilities in an observation-constrained model ensemble, *Biogeosciences*, *13*(4), 1071–1103, doi:10.5194/bg-13-1071-2016.

Voss, R., and U. Mikolajewicz (2001), Long-term climate changes due to increased CO2 concentration in the coupled atmosphere-ocean general circulation model ECHAM3/LSG, *Clim. Dyn.*, *17*(1), 45–60, doi:10.1007/PL00007925.

Weaver, A. J., M. Eby, M. Kienast, and O. A. Saenko (2007), Response of the Atlantic meridional overturning circulation to increasing atmospheric CO2: Sensitivity to mean climate state, *Geophys. Res. Lett.*, *34*(5), L05,708, doi:10.1029/2006GL028756.

Winton, M., K. Takahashi, and I. M. Held (2010), Importance of Ocean Heat Uptake Efficacy to Transient Climate Change, *J. Clim.*, *23*(9), 2333–2344, doi:10.1175/2009JCLI3139.1.