Atmospheric
Chemistry
and Physics
Discussions

1   **The influence of internal variability on Earth's energy balance framework and implications for**

2   **estimating climate sensitivity**

3   Andrew E. Dessler[1]*, Thorsten Mauritsen[2], Bjorn Stevens[2]

4   [1] Dept. of Atmospheric Sciences, Texas A&M University, College Station, TX 77843

5   [2] Max Planck Institute for Meteorology, Bundesstraße 53, 20146 Hamburg, Germany

6   *Correspondence to: adessler@tamu.edu, 979-862-1427

Atmospheric
Chemistry
and Physics
Discussions

Open Access

8   **Abstract:** Our climate is constrained by the balance between solar energy absorbed by the

9   Earth and terrestrial energy radiated to space. This energy balance has been widely used to

10   infer equilibrium climate sensitivity (ECS) from observations of 20th-century warming. Such

11   estimates yield lower values than other methods and these have been influential in pushing

12   down the consensus ECS range in recent assessments. Here we test the method using a 100-

13   member ensemble of the MPI-ESM1.1 climate model simulations of the period 1850-2005 with

14   known forcing. We calculate ECS in each ensemble member using energy balance, yielding

15   values ranging from 2.1 to 3.9 K. The spread in the ensemble is related to the central

16   hypothesis in the energy budget framework: that global average surface temperature

17   anomalies are indicative of anomalies in outgoing energy (either of terrestrial origin or reflected

18   solar energy). We find that assumption is not well supported over the historical temperature

19   record in the model ensemble or more recent satellite observations. We find that framing

20   energy balance in terms of 500-hPa tropical temperature better describes the planet's energy

21   balance.

22

Atmospheric
Chemistry
and Physics
Discussions

The problem

When an energy imbalance is imposed, such as by adding a greenhouse gas to the atmosphere, the climate will shift in such a way to eliminate the energy imbalance. This process is embodied in the traditional linearized energy balance equation:

$$R = F + \lambda\, T_s \tag{1}$$

where the forcing F is an imposed energy imbalance, T_S is the global average surface temperature, and λ relates changes in T_S to a change in net top-of-atmosphere (TOA) flux (Dessler and Zelinka, 2014). R is the resulting TOA flux imbalance from the combined forcing and response. All quantities are deviations from an equilibrium base state, usually the pre-industrial climate. Equilibrium climate sensitivity (hereafter ECS, the equilibrium warming in response to a doubling of CO_2) is equal to $-F_{2xCO2}/\lambda$, where F_{2xCO2} is the forcing from doubled CO_2.

Many investigators (e.g., Gregory et al., 2002; Annan and Hargreaves, 2006; Otto et al., 2013; Lewis and Curry, 2015; Aldrin et al., 2012; Skeie et al., 2014; Forster, 2016) have used Eq. 1 combined with estimates of R, F, and T_s to estimate λ:

$$\lambda = \Delta(R-F)/\Delta T_s \tag{2}$$

where Δ indicates the change between the start of the historical period (usually the mid to late nineteenth century) and a recent period. These calculations result in values of λ near -2 $W/m^2/K$ and appear to rule out ECS larger than ~4 K (Stevens et al., 2016). The substantial likelihood of an ECS below 2 K implied by these calculations led the IPCC Fifth Assessment Report to extend their lower bound on *likely* values of ECS to 1.5 K (Collins et al., 2013).

We test this energy balance methodology through a perfect model experiment consisting of an analysis of a 100-member ensemble of runs of the MPI Earth System Model, MPI-ESM1.1. This is the latest coupled climate model from the Max Planck Institute for Meteorology and consists of the ECHAM6.3 atmosphere and land model coupled to the MPI-OM ocean model. The atmospheric resolution is T63 spectral truncation, corresponding to about 200 km, with 47

Atmospheric
Chemistry
and Physics
Discussions
Open Access
EGU

49   vertical levels, whereas the ocean has a nominal resolution of about 1.5 degrees and 40 vertical

50   levels. MPI-ESM1.1 is a bug-fixed and improved version of the MPI-ESM used during CMIP5

51   (Giorgetta et al., 2013) and nearly identical to the MPI-ESM1.2 (Mauritsen et al., 2018) model

52   being used to provide output to CMIP6, except that the historical forcings are from the MPI-

53   ESM.

54   Each of the 100 members simulates the years 1850-2005 (Fig. 1) and use the same evolution of

55   historical natural and anthropogenic forcings. The members differ only in their initial

56   conditions —each starts from a different state sampled from a 2000-year control simulation.

57   We calculate effective radiative forcing F for the ensemble by subtracting top-of-atmosphere

58   flux R in a run with climatological sea surface temperatures (SSTs) and a constant pre-industrial

59   atmosphere from average R from an ensemble of three runs using the same SSTs but the time-

60   varying atmospheric composition used in the historical runs (Hansen et al., 2005; Forster et al.,

61   2016). The three-member ensemble begins with perturbed atmospheric states. We estimate

62   F_{2xCO2} using the same approach in a set of fixed SST runs in which CO_2 increases at 1% per year,

63   which yields a F_{2xCO2} value of 3.9 W/m^2.

64   We calculate λ using Eq. 2 for each ensemble member, producing values ranging from -1.88 to

65   -1.01 W/m^2/K, with an ensemble average of -1.43 W/m^2/K (Fig. 2a). In this calculation, $\Delta(R-F)$

66   and ΔT_S are the average difference between the first and last decade of each run. We also

67   calculate ECS = $-F_{2xCO2}/\lambda$ for each ensemble member, producing values ranging from 2.08 to

68   3.87 K (Fig. 2b), with an ensemble average of 2.76 K.

69   With respect to precision of the estimates, our analysis shows that λ and ECS estimated from

70   the historical record can vary widely simply due to internal variability. Given that we have only

71   a single realization of the 20th century, we should not consider estimates based on the historical

72   period to be precise — even with perfect observations. This supports previous work that also

73   emphasized the impact of internal variability on estimates of λ and ECS (Huber et al., 2014;

74   Andrews et al., 2015; Zhou et al., 2016; Gregory and Andrews, 2016).

Atmospheric
Chemistry
and Physics
Discussions

75   Previous researchers have questioned whether the historical record provides an accurate

76   measure of λ and ECS, and we can check this by comparing the ensemble average values to ECS

77   estimates from a 2xCO$_2$ run of the MPI-ESM1.2, which is physically very close to MPI-ESM1.1

78   and the changes between the MPI-ESM1.1 and MPI-ESM1.2 are not believed to be important

79   for its climate sensitivity. An abrupt 2xCO$_2$ run yields an ECS of 2.93 K in response to an abrupt

80   doubling of CO$_2$ (estimated by regressing years 100-1000 of a 1000-year run) — 6% larger than

81   the ensemble average. This is in line with the 10% difference in ECS estimated by Mauritsen and

82   Pincus (2017) to arise from the average CMIP5 model time-dependent feedback, but is smaller

83   than suggested in other recent studies of ECS in transient climate runs (e.g., Armour, 2017;

84   Proistosescu and Huybers, 2017).

85   Thus, there are a number of issues that need to be considered when interpreting estimates of λ

86   and ECS derived from the historical period. In addition to the precision and accuracy issues

87   discussed above, it also includes the large and evolving uncertainty in forcing over the 20th

88   century (Forster, 2016), different forcing efficacies of greenhouse gases and aerosols (Shindell,

89   2014; Kummer and Dessler, 2014), and geographically incomplete or inhomogeneous

90   observations (Richardson et al., 2016).

91   **Why are estimates using the traditional energy balance approach imprecise?**

92   In this section, we explain the physical process by which internal variability leads to the large

93   spread in λ and ECS estimated from the ensemble. We begin by observing that Eqs. 1 and 2

94   parameterize R-F in terms of T$_S$. In model runs with strong forcing driving large warming, such

95   as abrupt 4xCO$_2$ simulations, there is indeed a strong correlation between these variables (e.g.,

96   Gregory et al., 2004). However, because R-F in such runs is dominated by a monotonic trend,

97   correlations will exist with any geophysical field that also exhibits a monotonic trend, regardless

98   of whether there is a physical connection between the fields. Thus, one should not take the

99   correlation between R-F and T$_S$ in these runs as proving causality.

100   If T$_S$ is a good proxy for the response R-F, we would expect to also see a correlation in

101   measurements dominated by interannual variations. Observational data allow us to test this

102   hypothesis. We use observations of R from the Clouds and the Earth's Radiant Energy System

103   (CERES) Energy Balanced and Filled product (ed. 4) (Loeb et al., 2009), which cover the period

104   March 2000 to July. 2017. Our sign convention throughout the paper is that downward fluxes

105   are positive. Temperatures come from the European Centre for Medium Range Weather

106   Forecasts (ECMWF) Interim Re-Analysis (ERAi) (Dee et al., 2011). We assume forcing changes

107   linearly over this time period and account for it by detrending ΔR and ΔT anomaly time series

108   using a linear least-squares fit to remove the long-term trend.

109   These data show that ΔR is poorly correlated with ΔT_s in response to interannual variability (Fig.

110   3a), as has been noted many times in the literature; see, e.g., Sect. 5 of Forster (2016). In

111   particular, the low correlation coefficient tells us that ΔT_S explains little of the variance in ΔR.

112   Using explicit estimates of forcing or other temperature datasets (e.g., MERRA-2) yield the

113   same result.

114   GCMs that submitted output to the 5th phase of the Coupled Model Intercomparison Project

115   (CMIP5) (Taylor et al., 2012) also show this poor correlation. To demonstrate this, we have

116   calculated the correlation coefficient between ΔT_S and ΔR in CMIP5 pre-industrial control runs

117   (these are runs for which forcing F = 0). To facilitate comparison with the CERES data, as well as

118   avoid any issues with long-term drift in the control runs, we break each run into 16-year

119   segments and calculate the correlation coefficient of monthly anomalies of ΔR and ΔT_S for each

120   segment. Fig. 4 shows that the correlation between ΔR and ΔT_S in the models is similar to that

121   from the CERES analysis.

122   Recent work provides an explanation: the response of $\Delta(R-F)$ to a particular ΔT_S is determined

123   not only by the global average magnitude, but also by the pattern of warming (Armour et al.,

124   2013; Andrews et al., 2015; Gregory and Andrews, 2016; Zhou et al., 2016, 2017; Andrews and

125   Webb, 2018). During El Nino cycles that dominate the observations in Fig. 3, the spatial pattern

126   of warm and cool regions changes, leading to responses in $\Delta(R-F)$ that do not scale cleanly with

127   ΔT_S — something Stevens et al. (2016) refer to as "pattern effects"

Atmospheric
Chemistry
and Physics
Discussions

128   To demonstrate how this also generates the spread in λ in the model ensemble (Fig. 2a), we

129   calculate the local response λ_r in three equal-area regions (90°S-19.4°S, 19.4°S-19.4°N, 19.4°N-

130   90°N). We define λ_r as the regional analog to λ (Eq. 2):

131         $$\lambda_r = \Delta(R-F)_r / \Delta T_{S,r} \qquad\qquad\qquad (3)$$

132   where the "r" subscript indicates a regional average value.

133   We find that λ_r varies between the regions (Fig. 5). This means that different ensemble

134   members with similar global average ΔT_S but different patterns of surface warming produce

135   different values of global average $\Delta(R-F)$, thereby leading to spread in the estimated λ among

136   the ensemble members. We also see strong variability in λ_r within each region, suggesting that

137   how the warming is distributed within the region also drives some of the spread in estimated λ

138   in the ensemble.

139   This explanation is consistent with analyses showing that λ changes during transient runs as the

140   pattern of surface temperature evolves (Senior and Mitchell, 2000; Armour et al., 2013;

141   Andrews et al., 2015; Gregory and Andrews, 2016; Stevens et al., 2016). In our model

142   ensemble, however, the pattern changes are caused by internal variability rather than differing

143   regional heat capacities that cause some regions to warm more slowly than others during

144   forced warming.

145   **A better way to describe energy balance**

146   Our analysis demonstrates limitations of the conventional energy balance framework (Eq. 1). It

147   has been previously noted that ΔR correlates better with tropospheric temperatures than ΔT_S

148   (Murphy, 2010; Spencer and Braswell, 2010; Trenberth et al., 2015). Recent analyses have also

149   stressed the importance of atmospheric temperatures — through its influence on lapse rate —

150   as providing a fundamental control on the planet's energy budget (Zhou et al., 2016; Ceppi and

151   Gregory, 2017). Based on this, we test a new energy balance framework constructed using the

152   temperature of the tropical atmosphere:

Atmospheric
Chemistry
and Physics
Discussions

153        $R - F = \Theta\, T_A$ (4)

154   where T_A is the tropical average (30°N-30°S) 500-hPa temperature and Θ relates this quantity to

155   R-F. R and F are the same global average quantities they were in equation 1. ECS can be

156   expressed in terms of Θ:

157        $$ECS = -\frac{\Delta F_{2\times CO2}}{\Theta}\frac{\Delta T_S}{\Delta T_A}$$ (5)

158   where ΔT_S and ΔT_A are the equilibrium changes in these quantities in response to doubled CO_2;

159   the CMIP5 ensemble average ratio $\Delta T_S/\Delta T_A$ is 0.86±0.10 (±1σ), where Δ represents the average

160   difference between the first and last decades of the abrupt $4xCO_2$ runs.

161   Support for Eq. 4 can be found in the observations: ΔR shows a tighter correlation with ΔT_A than

162   with ΔT_S in observations (Figs. 3a vs. 3b). Given that the slope of these plots can be taken as

163   estimates of Θ and λ, the tighter correlation leads to more accurate estimates of Θ than λ,

164   both in absolute and relative terms.

165   Turning to the model ensemble, we next demonstrate that Θ is a more precise metric than λ.

166   We do this by calculating Θ [= $\Delta(R-F)/\Delta T_A$] in each ensemble member, yielding values ranging

167   from -1.18 to -0.89 W/m^2/K, with an ensemble average of -1.04 W/m^2/K (Fig. 2a). There is

168   clearly less variability in Θ among the ensemble members than for λ. This reflects less

169   variability in the regional response Θ_r (= $\Delta(R-F)_r/\Delta T_{A,r}$) than λ_r (Fig. 5), as well as less variability

170   within the regions. We therefore conclude that interannual variability has less of an impact on

171   Θ than λ. We show additional evidence for the superior precision of Θ in the Appendix.

172   As far as accuracy goes, we can compare Θ in the ensemble over the historical period to Θ in

173   response to much larger warming. The ensemble average Θ from the historic period, -1.04

174   W/m^2/K, is close to the value obtained from an analysis of the first 150 years of an abrupt

175    $4xCO_2$ run of the same model, Θ = -1.03 W/m^2/K, as well as Θ calculated from all 2600 years of

176   this run, Θ = -1.04 W/m^2/K. On the other hand, λ changes substantially in the $4xCO_2$ run as the

177   climate warms: λ = -1.36 W/m^2/K when calculated from the first 150 years, but λ = -0.95

178   W/m^2/K from all 2600 years of that run.

Atmospheric
Chemistry
and Physics
Discussions

179   We can verify this result in the CMIP5 abrupt $4xCO_2$ ensemble. It has been previously

180   demonstrated that plots of R-F vs. T_S do not trace straight lines as the climate warms (Andrews

181   et al., 2015; Rugenstein et al., 2016; Rose and Rayborn, 2016; Armour, 2017), so λ and ECS

182   calculated in a single model run may depend on the portion of the run selected. In the CMIP5

183   abrupt $4xCO_2$ ensemble, for example, average λ calculated by regressing years 10-30 (λ_{10-30}) is

184   more negative than λ calculated from years 30-150 (λ_{30-150}) by 0.50 $W/m^2/K$ (Fig. 6).

185   Several explanations for this have been advanced, most prominently that λ is function of the

186   pattern of surface warming (Senior and Mitchell, 2000; Armour et al., 2013; Andrews et al.,

187   2015; Gregory and Andrews, 2016; Zhou et al., 2016; Stevens et al., 2016). Using Θ largely

188   eliminates this pattern effect: Θ_{10-30} and Θ_{30-150} have an average difference of 0.16 $W/m^2/K$ for

189   the CMIP5 ensemble (Fig. 6). Thus, we find additional evidence that Θ tends to be similar for

190   different amounts and patterns of warming.

191   Finally, one of our ultimate goals for this revised framework is to help produce better estimates

192   of ECS. We are working on a detailed analysis of ECS based on this framework and will publish

193   that in a follow-on paper, but we briefly show here how the advantages of the revised energy

194   balance framework may be leveraged to do this. Fig. 7a shows Θ calculated from control runs

195   of 25 CMIP5 models. To calculate Θ in the control runs, we break each control run into 16-year

196   segments and calculate monthly anomalies of ΔR and ΔT_A during each segment. Then, we

197   calculate Θ for each segment as the slope of the regression of ΔR vs. ΔT_A for that segment.

198   Thus, for each control run, we generate a large number of estimates of Θ. The value in Fig. 7a is

199   the average of these individual values.

200   Fig. 7b shows the ECS of these models, calculated from the first 150 years of the abrupt $4xCO_2$

201   runs using the Gregory method (Gregory et al., 2004). If we assume that models with more

202   accurate simulation of short-term Θ produce more accurate estimates of ECS (Brown and

203   Caldeira, 2017; Wu and North, 2002), then we can use Figs. 7a and 7b to constrain ECS. We find

204   that the 15 models whose short-term Θ agrees with the CERES observations have ECS values

205   ranging from 2.0-3.9 K, with an average of 2.9 K. This excludes many of the highest ECS models.

206 It would not have been possible to draw this conclusion with the conventional energy balance

207 framework. Fig. 7c shows the comparison between λ from the control runs (calculated the

208 same way Θ was calculated) and CERES observations. Because of the much larger uncertainty

209 in the observational estimate of short-term λ, almost all models fall within the observational

210 range, thereby prohibiting any constraint on the ECS range.

211 **Conclusions**

212 We have estimated ECS in each of a 100-member climate model ensemble using the same

213 energy-balance constraint used by many investigators to estimate ECS from 20th-century

214 historical observations. We find that the method is imprecise — the estimates of ECS range

215 from 2.1 to 3.9 K (Fig. 2), with some ensemble members far from the model's true value of 2.9

216 K. Given that we only have a single ensemble of reality, this suggests that some skepticism is

217 appropriate when considering estimates of ECS derived from the historical record.

218 The source of the imprecision relates to the construction of the traditional energy balance

219 equation (Eq. 1). In it, the response of TOA net flux (R-F) is parameterized in terms of global

220 average surface temperature (T_S). Recent research has suggested that the response is not just

221 determined by the magnitude of T_S, but includes other factors, such as the pattern of T_S (e.g.,

222 Armour et al., 2013; Andrews et al., 2015; Gregory and Andrews, 2016; Zhou et al., 2017) or the

223 lapse rate (e.g., Zhou et al., 2017; Ceppi and Gregory, 2017). As a result, two ensemble

224 members with the same ΔT_S can have different climate responses, $\Delta(R-F)$, leading to spread in

225 the inferred λ.

226 The lack of a direct relationship between T_S and radiation balance suggests that it may be

227 profitable to investigate alternative formulations. We test parameterizing the response in terms

228 of 500-hPa tropical temperature (Eq. 4) and find that it is superior in many ways. Ultimately,

229 how investigators describe the energy balance of the planet will depend on the problem and

230 the available data. The surface temperature is indeed special, so the traditional framework

231 may be preferred for some problems. But investigators may find that the alternatives are

232 superior for certain problems, for instance constraining Earth's climate sensitivity.

Atmospheric
Chemistry
and Physics
Discussions

233 Appendix

234 It has been previously noted in analyses of the historical record that λ exhibits significant

235 interdecadal variability (Andrews et al., 2015; Gregory and Andrews, 2016; Zhou et al., 2016).

236 We can reproduce this in a 2000-year control run (a run with fixed pre-industrial boundary

237 conditions) of the MPI-ESM1.1 model. Fig. 8 shows λ calculated in a sliding 16-year window

238 and confirms significant temporal variability in λ. We can similarly calculate Θ and find that

239 temporal variability in Θ is substantially smaller (Fig. 8).

240 This result is reproduced in the CMIP5 control models. Fig. 9 plots the standard deviation of

241 each CMIP5 model's set of short-term λ divided by the standard deviation of that model's set of

242 short-term Θ (as described previously, we calculate time series of short-term λ and Θ values for

243 each model by regressing anomalies in a 16-year sliding window of the control runs). All of the

244 models fall above 1, demonstrating that there is less variability in the Θ time series than in the

245 λ time series in every climate model. This confirms that Θ is more robust with respect to

246 internal variability than λ. It also suggests that Θ estimated from the satellite data (Fig. 3)

247 should be considered a better estimate of the climate system's long-term value than λ

248 estimated from the same data set.

249

Atmospheric
Chemistry
and Physics
Discussions

Atmospheric
Chemistry
and Physics
Discussions

259   **References**

260   Aldrin, M., Holden, M., Guttorp, P., Skeie, R. B., Myhre, G., and Berntsen, T. K.: Bayesian
261   estimation of climate sensitivity based on a simple climate model fitted to observations of
262   hemispheric temperatures and global ocean heat content, Environmetrics, 23, 253-271,
263   10.1002/env.2140, 2012.
264   Andrews, T., Gregory, J. M., and Webb, M. J.: The dependence of radiative forcing and feedback
265   on evolving patterns of surface temperature change in climate models, J. Climate, 28, 1630-
266   1648, 10.1175/JCLI-D-14-00545.1, 2015.
267   Andrews, T., and Webb, M. J.: The Dependence of Global Cloud and Lapse Rate Feedbacks on
268   the Spatial Structure of Tropical Pacific Warming, J. Climate, 31, 641-654, 10.1175/jcli-d-17-
269   0087.1, 2018.
270   Annan, J. D., and Hargreaves, J. C.: Using multiple observationally-based constraints to estimate
271   climate sensitivity, Geophys. Res. Lett., 33, 10.1029/2005gl025259, 2006.
272   Armour, K. C., Bitz, C. M., and Roe, G. H.: Time-varying climate sensitivity from regional
273   feedbacks, J. Climate, 26, 4518-4534, 10.1175/jcli-d-12-00544.1, 2013.
274   Armour, K. C.: Energy budget constraints on climate sensitivity in light of inconstant climate
275   feedbacks, Nature Clim. Change, 7, 331-335, 10.1038/nclimate3278, 2017.
276   Brown, P. T., and Caldeira, K.: Greater future global warming inferred from Earth's recent
277   energy budget, Nature, 552, 10.1038/nature24672, 2017.
278   Ceppi, P., and Gregory, J. M.: Relationship of tropospheric stability to climate sensitivity and
279   Earth's observed radiation budget, Proc. Natl. Acad. Sci., 10.1073/pnas.1714308114, 2017.
280   Collins, M., et al.: Long-term climate change: Projections, commitments and irreversibility, in:
281   Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the
282   Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by:
283   Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y.,
284   Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and
285   New York, NY, USA., 2013.
286   Dee, D. P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data
287   assimilation system, Q. J. R. Meteor. Soc., 137, 553-597, 10.1002/qj.828, 2011.
288   Dessler, A. E., and Zelinka, M. D.: Climate feedbacks, in: Encyclopedia of Atmospheric Sciences,
289   edited by: North, G. R., Pyle, J., and Zhang, F., Elsevier, 18–25, 2014.
290   Forster, P. M.: Inference of climate sensitivity from analysis of Earth's energy budget, Annual
291   Review of Earth and Planetary Sciences, 44, 85-106, 10.1146/annurev-earth-060614-
292   105156, 2016.
293   Forster, P. M., et al.: Recommendations for diagnosing effective radiative forcing from climate
294   models for CMIP6, J. Geophys. Res., 121, 12460-12475, 10.1002/2016jd025320, 2016.
295   Giorgetta, M. A., et al.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM
296   simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances in
297   Modeling Earth Systems, 5, 572-597, 10.1002/jame.20038, 2013.
298   Gregory, J. M., Stouffer, R. J., Raper, S. C. B., Stott, P. A., and Rayner, N. A.: An observationally
299   based estimate of the climate sensitivity, J. Climate, 15, 3117-3121, 10.1175/1520-
300   0442(2002)015<3117:aobeot>2.0.co;2, 2002.

Atmospheric
Chemistry
and Physics
Discussions

301   Gregory, J. M., et al.: A new method for diagnosing radiative forcing and climate sensitivity,
302   Geophys. Res. Lett., 31, 10.1029/2003gl018747, 2004.
303   Gregory, J. M., and Andrews, T.: Variation in climate sensitivity and feedback parameters during
304   the historical period, Geophys. Res. Lett., 43, 3911-3920, 10.1002/2016GL068406, 2016.
305   Hansen, J., et al.: Efficacy of climate forcings, Journal of Geophysical Research: Atmospheres,
306   110, n/a-n/a, 10.1029/2005JD005776, 2005.
307   Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, Rev. Geophys.,
308   48, 10.1029/2010rg000345, 2010.
309   Huber, M., Beyerle, U., and Knutti, R.: Estimating climate sensitivity and future temperature in
310   the presence of natural climate variability, Geophys. Res. Lett., 41, 2086-2092,
311   10.1002/2013GL058532, 2014.
312   Kummer, J. R., and Dessler, A. E.: The impact of forcing efficacy on the equilibrium climate
313   sensitivity, Geophys. Res. Lett., 41, 3565-3568, 10.1002/2014gl060046, 2014.
314   Lewis, N., and Curry, J. A.: The implications for climate sensitivity of AR5 forcing and heat
315   uptake estimates, Climate Dynamics, 45, 1009-1023, 10.1007/s00382-014-2342-y, 2015.
316   Loeb, N. G., et al.: Toward optimal closure of the Earth's top-of-atmosphere radiation budget, J.
317   Climate, 22, 748-766, 10.1175/2008jcli2637.1, 2009.
318   Mauritsen, T., et al.: Developments in the MPI-M Earth System Model version 1.2 (MPI-
319   ESM1.2), in preparation, 2018.
320   Murphy, D. M.: Constraining climate sensitivity with linear fits to outgoing radiation, Geophys.
321   Res. Lett., 37, 10.1029/2010GL042911, 2010.
322   Otto, A., et al.: Energy budget constraints on climate response, Nature Geoscience, 6, 415-416,
323   10.1038/ngeo1836, 2013.
324   Richardson, M., Cowtan, K., Hawkins, E., and Stolpe, M. B.: Reconciled climate response
325   estimates from climate models and the energy budget of Earth, Nature Clim. Change, 6,
326   931-935, 10.1038/nclimate3066, 2016.
327   Rose, B. E. J., and Rayborn, L.: The effects of ocean heat uptake on transient climate sensitivity,
328   Current Climate Change Reports, 2, 190-201, 10.1007/s40641-016-0048-4, 2016.
329   Rugenstein, M. A. A., Caldeira, K., and Knutti, R.: Dependence of global radiative feedbacks on
330   evolving patterns of surface heat fluxes, Geophys. Res. Lett., 43, 9877-9885,
331   10.1002/2016GL070907, 2016.
332   Santer, B. D., et al.: Statistical significance of trends and trend differences in layer-average
333   atmospheric temperature time series, J. Geophys. Res., 105, 7337-7356,
334   10.1029/1999jd901105, 2000.
335   Senior, C. A., and Mitchell, J. F. B.: The time-dependence of climate sensitivity, Geophys. Res.
336   Lett., 27, 2685-2688, 10.1029/2000GL011373, 2000.
337   Shindell, D. T.: Inhomogeneous forcing and transient climate sensitivity, 4, 274,
338   10.1038/nclimate2136, 2014.
339   Skeie, R. B., Berntsen, T., Aldrin, M., Holden, M., and Myhre, G.: A lower and more constrained
340   estimate of climate sensitivity using updated observations and detailed radiative forcing
341   time series, Earth System Dynamics, 5, 139-175, 10.5194/esd-5-139-2014, 2014.
342   Spencer, R. W., and Braswell, W. D.: On the diagnosis of radiative feedback in the presence of
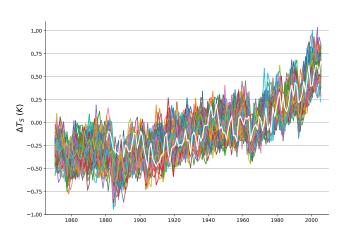343   unknown radiative forcing, J. Geophys. Res., 115, 10.1029/2009JD013371, 2010.

Atmospheric
Chemistry
and Physics
Discussions

Open Access

344 Stevens, B., Sherwood, S. C., Bony, S., and Webb, M. J.: Prospects for narrowing bounds on
345   Earth's equilibrium climate sensitivity, Earth's Future, 4, 512-522, 10.1002/2016EF000376,
346   2016.
347 Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design,
348   Bull. Am. Met. Soc., 93, 485-498, 10.1175/bams-d-11-00094.1, 2012.
349 Trenberth, K. E., Zhang, Y., Fasullo, J. T., and Taguchi, S.: Climate variability and relationships
350   between top-of-atmosphere radiation and temperatures on Earth, J. Geophys. Res.,
351   10.1002/2014JD022887, 2015.
352 Wu, Q., and North, G. R.: Climate sensitivity and thermal inertia, Geophys. Res. Lett., 29,
353   10.1029/2002GL014864, 2002.
354 Zhou, C., Zelinka, M. D., and Klein, S. A.: Impact of decadal cloud variations on the Earth/'s
355   energy budget, Nature Geosci, 9, 871-874, 10.1038/ngeo2828, 2016.
356 Zhou, C., Zelinka, M. D., and Klein, S. A.: Analyzing the dependence of global cloud feedback on
357   the spatial pattern of sea surface temperature change with a Green's function approach,
358   Journal of Advances in Modeling Earth Systems, 9, 2174-2189, 10.1002/2017MS001096,
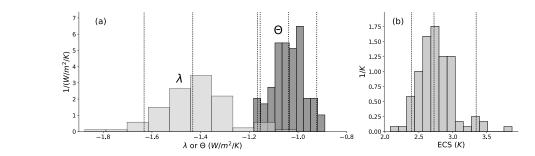359   2017.
360

361

362

363   Fig. 1. Plot of annual and global average surface temperature from the 100 members of the

364   MPI-ESM1.1 ensemble (colored lines), along with the GISTEMP measurements (Hansen et

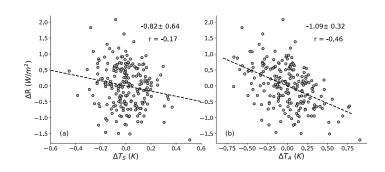365   al., 2010) (white line). Temperatures are referenced to the 1951-1980 average.

366

367

368   Figure 2. PDFs of (a) λ (lighter) and Θ (darker) and (b) ECS derived from the members of the

369   MPI-ESM1.1 historical ensemble. The vertical lines are the 5th, 50th, and 95th percentile of each

370   distribution.

Atmospheric
Chemistry
and Physics
Discussions

371

372   Figure 3. Scatter plot of detrended monthly anomalies of ΔR vs. (a) global average surface

373   temperature ΔT_S, (b) tropical average 500-hPa temperature ΔT_A. Observations cover the period

374   March 2000-Jan. 2017 and anomalies are deviations from the mean annual cycle. The dashed

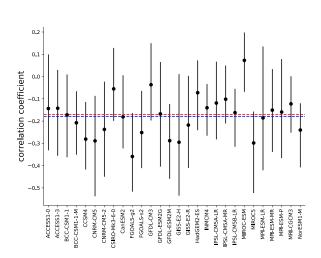375   lines are ordinary least-squares fits; the slope, 5-95% confidence interval, and correlation

376   coefficient are shown on each panel. Confidence intervals account for autocorrelation of the

377   time series (Santer et al., 2000).

378
379

Atmospheric
Chemistry
and Physics
Discussions

380

Fig. 4. Correlation coefficients between ΔR and ΔT_S in CMIP5 control runs. The dot is the

average of the correlation coefficients from the 16-year segments of the model run; the

bars indicate the maximum and minimum values from the control run. The blue dashed

line is the average of the CMIP5 models, while the red dashed line is the correlation

coefficient from the CERES regression in Fig. 2a.

386

Fig. 5. λ_r and Θ_r calculated as regional average $\Delta(R-F)$ divided by regional average temperature

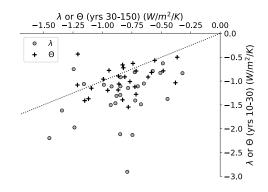(ΔT_S for λ and ΔT_A for Θ). The regions are 90°S-19.4°S (SH), 19.4°S-19.4°N (EQ), and 19.4°N-

90°N (NH). The values are calculated for each member of the 100-member ensemble; the solid

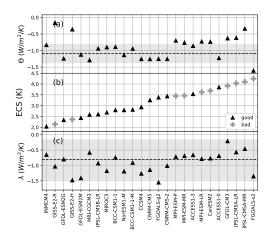symbols are the ensemble average while the bars show the 5-95% range.

391

Fig. 6. Scatterplot of λ_{10-30} vs. λ_{30-150} (gray circles) in CMIP5 abrupt4xCO$_2$ runs, as well as

Θ_{10-30} vs. Θ_{30-150} (black crosses) in the same models. Each point represents one model.

The dotted line is the 1:1 line. The subscripts (10-30, 30-150) indicate the years of the run
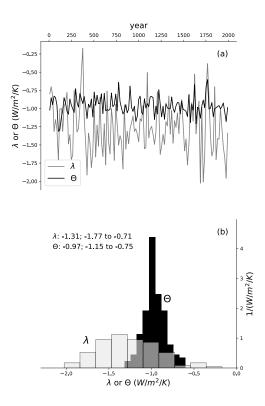
from which the values are calculated.

396

Figure 7. (a) Θ from individual CMIP5 control runs (calculation described in the Appendix).

The dotted line is the estimate from CERES observations; the gray region is the 5-95%

confidence band. (b) ECS from each CMIP5 model, estimated from the first 150 years of

abrupt 4xCO$_2$ runs using the Gregory method (Gregory et al., 2004). "Good" models are

those whose Θ agrees with observations, "bad" models are those that do not. (c) Same as
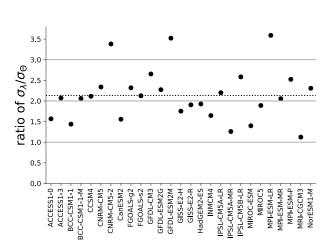
panel (a), but for λ.

403

404

Fig. 8. (a) Time series of λ (gray) and Θ (black) estimated in a 16-year sliding window of a 2000-year control run of the MPI-ESM1.1. (b) PDFs of the time series in panel a. Median and 5-95% confidence interval for each distribution is displayed on the plot.

Atmospheric
Chemistry
and Physics
Discussions

408

Fig. 9. The standard deviation of the λ time series divided by the standard deviation of the

Θ time series. Each time series is calculated from 16-year segments of CMIP5 control runs.

The dotted line is the ensemble average.

412