

1 **On the effect of historical SST patterns on radiative feedback**

2

3 **Timothy Andrews¹ and Alejandro Bodas-Salcedo**

4 Met Office Hadley Centre, Exeter, UK.

5 **Jonathan M. Gregory**

6 Met Office Hadley Centre, Exeter, UK and National Centre for Atmospheric Science, University of
7 Reading, Reading, UK.

8 **Yue Dong**

9 Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA.

10 Now at Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA.

11 **Kyle C. Armour**

12 Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA and School of
13 Oceanography, University of Washington, Seattle, WA, USA.

14 **David Paynter**

15 NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, NJ, USA.

16 **Pu Lin**

17 Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA.

18 **Angshuman Modak and Thorsten Mauritsen**

19 University of Stockholm, Department of Meteorology, Stockholm, Sweden.

20 **Jason N.S. Cole**

21 Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada,
22 Victoria, BC, Canada.

23 **Brian Medeiros**

24 National Center for Atmospheric Research, Climate and Global Dynamics Laboratory, Boulder, CO,
25 USA.

26 **James J. Benedict**

27 Rosenstiel School of Marine and Atmospheric Science, University of Miami, Coral Gables, FL, USA.

28 Now at Los Alamos National Laboratory, Los Alamos, NM, USA.

29 **Hervé Douville and Romain Roehrig**

30 CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France.

¹ Correspondence to Timothy Andrews (timothy.andrews@metoffice.gov.uk)

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Tsuyoshi Koshiro and Hideaki Kawai

Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan.

Tomoo Ogura

National Institute for Environmental Studies, Tsukuba, Japan.

Jean-Louis Dufresne

Laboratoire de Météorologie Dynamique/IPSL, CNRS, Sorbonne Université, École Normale Supérieure, PSL Research University, École Polytechnique, Paris, France.

Richard P. Allan

National Centre for Earth Observation, University of Reading, Reading, UK.

Department of Meteorology, University of Reading, Reading, UK.

Chunlei Liu

South China Sea Institute of Marine Meteorology, Guangdong Ocean University, Zhanjiang, China.

Submitted to *Journal of Geophysical Research*

21st February 2022; Revised June 2022 & August 2022.

48 **Abstract**

49 We investigate the dependence of radiative feedback on the pattern of sea-surface temperature
50 (SST) change in fourteen Atmospheric General Circulation Models (AGCMs) forced with observed
51 variations in SST and sea-ice over the historical record from 1871 to near-present. We find that over
52 1871-1980, the Earth warmed with feedbacks largely consistent and strongly correlated with long-
53 term climate sensitivity feedbacks (diagnosed from corresponding atmosphere-ocean GCM abrupt-
54 4xCO₂ simulations). Post 1980 however, the Earth warmed with unusual trends in tropical Pacific
55 SSTs (enhanced warming in the west, cooling in the east) and cooling in the Southern Ocean that
56 drove climate feedback to be uncorrelated with – and indicating much lower climate sensitivity than
57 – that expected for long-term CO₂ increase. We show that these conclusions are not strongly
58 dependent on the AMIP II SST dataset used to force the AGCMs, though the magnitude of feedback
59 post 1980 is generally smaller in nine AGCMs forced with alternative HadISST1 SST boundary
60 conditions. We quantify a ‘pattern effect’ (defined as the difference between historical and long-
61 term CO₂ feedback) equal to 0.48 ± 0.47 [5-95%] $W m^{-2} K^{-1}$ for the time-period 1871-2010 when the
62 AGCMs are forced with HadISST1 SSTs, or 0.70 ± 0.47 [5-95%] $W m^{-2} K^{-1}$ when forced with AMIP II
63 SSTs. Assessed changes in the Earth’s historical energy budget agree with the AGCM feedback
64 estimates. Furthermore satellite observations of changes in top-of-atmosphere radiative fluxes since
65 1985 suggest that the pattern effect was particularly strong over recent decades but may be waning
66 post 2014.

67 **1. Introduction**

68 *1.1. Background*

69 A common starting point for quantifying the sensitivity of the Earth’s climate to external
70 perturbations is consideration of the global-mean energy budget, $N = F + \lambda T$, where N is the net
71 downward radiative flux at the top-of-atmosphere (TOA) (units W m^{-2}), F the effective radiative
72 forcing (ERF, units W m^{-2}), λ the climate feedback parameter (units $\text{W m}^{-2} \text{K}^{-1}$, a negative number in
73 this paper, but the opposite convention is also used) and T the surface-air-temperature change
74 (units K) relative to an unperturbed steady state in which $N=F=0$. Applied to non-steady states, such
75 as the Earth’s historical record (since the 1800s), λ is determined via either (i) differences (denoted
76 by Δ) between two climate states (often present-day and pre-industrial) according to $\lambda = (\Delta N -$
77 $\Delta F)/\Delta T$ (e.g. Gregory et al., 2002; Otto et al., 2013; Sherwood et al., 2020), or (ii) regression in the
78 differential form $\lambda = d(N - F)/dT$ if the timeseries of N , F and T are known (Gregory et al. 2004;
79 Gregory et al. 2020).

80 Until recently it was often assumed that λ was - to a good approximation - a constant property of the
81 climate system, such that feedbacks that applied over the historical record also applied to the
82 Earth’s long-term response, as quantified by the canonical equilibrium climate sensitivity (ECS, units
83 K) to a forcing from a doubling of CO_2 (F_{2x}) over pre-industrial levels. Thus ECS was estimated directly
84 from historical changes in N , T and F , according to $\text{ECS} = -F_{2x} / \lambda = -F_{2x} \Delta T / (\Delta N - \Delta F)$ (e.g. Gregory et
85 al, 2002; Otto et al., 2013, amongst many others).

86 However, it is now recognised that λ varies in time since a forcing is applied and with the strength
87 and/or type of that forcing (e.g. Senior and Mitchell, 2000; Hansen et al., 2005; Andrews et al. 2012;
88 Armour et al., 2013; Geoffroy et al., 2013; Rose et al. 2014; Gregory et al. 2015; Andrews et al. 2015;
89 Marvel et al. 2016; Rugenstein et al. 2016; Richardson et al., 2019; Dong et al. 2020; Bloch-Johnson
90 et al., 2021; Rugenstein and Armour, 2021). Hence λ is an ‘effective feedback parameter’ that applies
91 only to the climate change over which it was calculated. More specifically, over the historical record
92 λ is thought to be more stabilizing (more negative, climate sensitivity smaller) than might operate in
93 the long-term future, and so λ estimated from historical climate change would understate ECS (e.g.
94 Gregory and Andrews, 2016; Zhou et al., 2016; Armour, 2017; Proistosescu & Huybers, 2017;
95 Andrews et al., 2018; Marvel et al., 2018; Silvers et al., 2018; Lewis and Curry, 2018; Gregory et al.
96 2020; Sherwood et al. 2020; Dong et al. 2021).

97 The reason for the underestimate of long-term ECS is that climate feedbacks setting λ , such as cloud
98 and lapse-rate changes, vary with the pattern of surface warming. Proxy reconstructions of past
99 equilibrium climates and atmosphere-ocean general circulation model (AOGCM) simulations of long-
100 term climate change show an ‘ENSO-like’ temperature pattern with strong temperature changes in
101 the eastern Pacific as well as the Southern Ocean, whereas observed historical warming shows more
102 pronounced warming in the western equatorial Pacific relative to the tropical mean and cooling in
103 the eastern Pacific and Southern Ocean over recent decades (e.g. Collins et al., 2013; Li et al., 2013;
104 Andrews et al., 2015; Gregory and Andrews, 2016; Zhou et al., 2016; Dong et al., 2019; Sherwood et
105 al., 2020; Rugenstein et al. 2020; Olonscheck et al., 2020; Fueglistaler and Silvers, 2021; Watanabe et
106 al. 2021; Power et al. 2021; Tierney et al. 2019; 2020).

107 Thus, more-stabilizing feedbacks have occurred over the historical record because enhanced
108 warming in the western Pacific warm pool – a region of deep ascent and convection – results in a
109 stronger negative lapse-rate feedback widely across the tropics due to efficient warming of the free
110 troposphere, which in turn causes increased cloudiness (a negative cloud feedback) in the eastern

111 tropical Pacific due to remotely controlled increased lower tropospheric stability. In contrast, less-
112 stabilizing feedbacks are expected in the future as enhanced warming in the eastern Pacific –
113 characterised by descending air and marine low cloud decks which are capped under a temperature
114 inversion and form over the relatively cool sea-surface-temperatures (SSTs) – results in a positive
115 cloud feedback, without an accompanying negative lapse-rate feedback since the warming is
116 ‘trapped’ in the boundary layer (e.g. Zhou et al., 2016, Andrews and Webb, 2018, Ceppi and Gregory,
117 2017; Dong et al. 2019).

118 The dependence of radiative feedback on the pattern of surface temperature change has been
119 termed a ‘pattern effect’ (Stevens et al., 2016), which distinguishes it from other feedback variations
120 that might occur for example as a function of the magnitude of ΔT (e.g. Block & Mauritsen, 2013;
121 Caballero and Huber, 2013; Bloch-Johnson et al., 2021). While the term ‘pattern effect’ could be
122 applied to any change in SST pattern and associated change in radiative feedback, here we will use it
123 to mean (unless explicitly stated) the pattern effect that arises due to the difference in warming
124 pattern between historical climate change and long-term ECS.

125 Armour (2017) and Andrews et al. (2018) proposed a method to account for the pattern effect in
126 estimates of ECS derived from historical climate changes via a modification of the energy budget
127 approach. Their method requires an estimate of the difference in feedback, $\Delta\lambda$, due to the pattern
128 effect that arises between historical climate change and long-term ECS, so that $ECS = -F_{2x} / (\lambda_{hist} + \Delta\lambda)$,
129 where λ_{hist} is the historical value. Since $\Delta\lambda$ is found to be positive, it increases the best estimate of
130 ECS and substantially lifts the upper uncertainty bound, but has only a small impact on the lower
131 bound (Armour, 2017; Andrews et al., 2018; Sherwood et al. 2020).

132 One way of defining the pattern effect, $\Delta\lambda$, is to contrast λ_{hist} in an Atmospheric GCM (AGCM)
133 simulation forced by observed historical SST and sea-ice variations (termed an *amip-piForcing*
134 simulation, see Section 2) with λ_{4xCO2} from 150 years of a coupled AOGCM *abrupt-4xCO2* simulation
135 with the same AGCM, so that $\Delta\lambda = \lambda_{4xCO2} - \lambda_{hist}$ (Andrews et al. 2018). Hence our quantification of $\Delta\lambda$
136 not only depends on λ_{hist} but also on the (somewhat arbitrary) time frame and method used to
137 calculate λ_{4xCO2} . Ideally we would use the feedback parameter directly associated with ECS rather
138 than λ_{4xCO2} , but this is difficult to calculate in AOGCMs due to the millennial timescales required to
139 equilibrate the deep ocean. Hence feedbacks calculated from 150 years of *abrupt-4xCO2* are often
140 used as a surrogate for long-term ECS feedbacks (Andrews et al., 2012). Technically this is still an
141 ‘effective feedback parameter’ and associated ‘effective climate sensitivity’ (EffCS), rather than
142 definitive ECS, but in practice it is found to provide a suitable analogue for long-term feedbacks in
143 climate projections (Grose et al., 2018) and ECS (Sherwood et al. 2020), hence the distinction
144 between EffCS and ECS is not considered further (see Rugenstein et al. (2020) and Rugenstein and
145 Armour (2021) for further discussion).

146 We assume other impacts on λ , such as the nature of the forcing agent – so called ‘efficacies’
147 (Hansen et al., 2005; Marvel et al. 2016; Richardson et al., 2019) – primarily occur due to forcing-
148 specific impacts on historical SST patterns that will be included in the historical record, rather than
149 any dependence on the actual forcing agent concentration in the atmosphere (which will be
150 excluded in our design, because forcing levels are fixed at pre-industrial levels in *amip-piForcing*)
151 (Haugstad et al., 2017). On the other hand, *abrupt-4xCO2* experiments contain larger warming than
152 the historical record, so any state dependence on T (e.g. Block & Mauritsen, 2013; Caballero and
153 Huber, 2013; Bloch-Johnson et al., 2021) might erroneously be diagnosed as a pattern effect using
154 our method. Bloch-Johnson et al. (2021) estimated that λ might vary with T by $\sim +0.029 \text{ W m}^{-2} \text{ K}^{-2}$
155 (multi-model-mean) in step CO_2 experiments relative to pre-industrial level temperature feedbacks,
156 but with substantial uncertainty in the both the magnitude and in some cases even the sign of this

157 state dependence (model range -0.14 to 0.109 to $\text{W m}^{-2} \text{K}^{-2}$). While this may play some role in our
158 diagnosed $\Delta\lambda$, we assume it to be small since both Gregory and Andrews (2016) and Andrews and
159 Webb (2018) showed that the pattern effect is large in experiments with identical T but contrasting
160 historical and *abrupt-4xCO2* SST patterns.

161 The principal advantage of using *amip-piForcing* simulations in the calculation of the pattern effect is
162 that λ_{hist} will be consistent with the SST patterns that occurred over the historical record. In contrast,
163 one could use AOGCM historical simulations for λ_{hist} , but when AOGCMs are free to simulate their
164 own historical SST patterns they struggle to reproduce the observed recent decadal trends in
165 tropical Pacific SST patterns (Gregory et al. 2020; Fueglistaler and Silvers, 2021; Watanabe et al.
166 2021; Dong et al., 2021) and the associated magnitude of λ_{hist} , thus underestimating the pattern
167 effect (Gregory et al., 2020; Dong et al. 2021). This AOGCM bias in the pattern effect has important
168 implications, which we return to in the Discussion, but our focus in this manuscript is on the
169 historical pattern effect as simulated by AGCMs *given* the observed SSTs, thus avoiding the issue of
170 AOGCM biases in historical SST patterns. Note that while our focus is on the atmospheric response
171 to a given SST pattern, causality can work in both directions. For example cloud feedback has been
172 shown to have an impact on the pattern of tropical Pacific SST changes in models (Chalmers et al.,
173 2022).

174 *amip-piForcing* simulations also show multi-decadal variations in λ_{hist} (Gregory and Andrews 2016;
175 Zhou et al., 2016; Andrews et al., 2018; Fueglistaler and Silvers, 2021; Dong et al. 2021). In particular
176 λ_{hist} is generally most negative (pattern effect largest) over the most recent decades. This is because
177 variations in atmospheric feedback are well explained by changes in SSTs in regions of tropical deep
178 convection relative to the tropical-mean (Fueglistaler and Silvers, 2021) or global-mean (Dong et al.
179 2019). Since the late 1970s, regions of deep convection have warmed by about +50% more than the
180 tropical-mean (Fueglistaler and Silvers, 2021), and the eastern Pacific has cooled despite
181 temperatures increasing globally (e.g. Hartmann et al. 2013; Power et al. 2021; and see our Figures 4
182 and 9). Hence under this configuration of tropical Pacific SST change, we would expect negative
183 feedback from the mechanisms described above (e.g. Zhou et al., 2016, Andrews and Webb, 2018,
184 Ceppi and Gregory, 2017; Dong et al. 2019).

185 A limitation of the *amip-piForcing* experiment for quantifying λ_{hist} is that it may include a structural
186 dependence on the underlying SST patterns and sea-ice in the Atmospheric Model Intercomparison
187 Project (AMIP) II boundary condition data set (Gates et al., 1999; Hurrell et al., 2008; Taylor et al.,
188 2000) used to force the *amip-piForcing* simulations (Andrews et al., 2018; Lewis and Mauritsen,
189 2021; Zhou et al., 2021; Fueglistaler and Silvers, 2021). Different SST reconstructions have slightly
190 different patterns of SST change over the historical period, and λ_{hist} may be affected. Indeed Lewis
191 and Mauritsen (2021) and Fueglistaler and Silvers (2021) showed that warming in the tropical
192 western Pacific relative to the tropical-mean is less pronounced in other SST datasets, and so we
193 might expect less negative feedbacks ($\Delta\lambda$ less positive) if the AGCMs were forced with non-AMIP II
194 datasets.

195 Consistent with this expectation, Andrews et al. (2018) noted that in one AGCM the magnitude of
196 λ_{hist} was reduced by $\sim 0.2 \text{ W m}^{-2} \text{K}^{-1}$ when the AMIP II SSTs were replaced by HadISST2.1 SSTs (sea-ice
197 remaining unchanged) in an *amip-piForcing* simulation. Partly because of this, Sherwood et al.
198 (2020) and Forster et al. (2021) assessed the historical pattern effect to be smaller and more
199 uncertain ($\Delta\lambda = 0.5 \pm 0.5 \text{ W m}^{-2}$) than simply taking the *amip-piForcing* based model distribution
200 reported by Andrews et al. (2018) ($\Delta\lambda = 0.64 \pm 0.40 \text{ W m}^{-2}$). Subsequently, Lewis and Mauritsen
201 (2021) and Zhou et al. (2021) also found λ_{hist} to be less negative ($\Delta\lambda$ smaller) when using other SST
202 datasets than AMIP II used in *amip-piForcing* simulations discussed here.

203 1.2. Aims and motivating questions

204 Andrews et al. (2018) provides much of the published quantitative analysis on λ_{hist} to observed SST
205 patterns and $\Delta\lambda$, but only six AGCMs from only four different modelling centres were considered.
206 Hence, a first motivation of this manuscript is to revisit their numbers with a broader set of models
207 by utilizing the new *amip-piForcing* simulations from the Cloud Feedback Model Intercomparison
208 Project phase 3 (CFMIP, Webb et al. 2017) contribution to the Coupled Model Intercomparison
209 Project phase 6 (CMIP6, Eyring et al., 2016). The larger ensemble totalling 14 models when
210 combined will provide a more robust quantification of the magnitude and spread of λ_{hist} and $\Delta\lambda$ to a
211 broader set of model physics and climate sensitivities (Zelinka et al. 2020; Meehl et al. 2020; Flynn
212 and Mauritsen, 2020).

213 Secondly, the limited set of models in Andrews et al. (2018) prevented them from robustly exploring
214 and quantifying the relationship between λ_{hist} and $\lambda_{4\times\text{CO}_2}$ across models. In other words, it is not
215 known whether feedbacks acting over the historical record in AGCMs are correlated to feedbacks
216 acting on long-term ECS. For example is there a relationship between the two that could form the
217 basis of an emergent constraint? Do different parts of the historical record relate better to
218 feedbacks acting on long-term ECS than other parts, and why? As we will show, feedbacks over
219 different parts of the historical record have different relationships to $\lambda_{4\times\text{CO}_2}$, and this is important for
220 understanding what can and cannot be directly constrained from the historical record.

221 Thirdly, λ_{hist} and $\Delta\lambda$ have been shown to vary substantially on decadal timescales with λ_{hist} being most
222 negative (pattern effect largest) over recent decades since ~ 1980 (Gregory and Andrews 2016; Zhou
223 et al., 2016; Andrews et al., 2018; Gregory et al. 2020; Dong et al. 2021). This is consistent with the
224 findings of Fueglistaler and Silvers (2021), who identified ~ 1980 as the point in which the Earth
225 begins to warm with a particular (even “*peculiar*”) configuration of tropical Pacific SSTs where
226 “*regions of deep convection warm about +50% more than the tropical average*” driving large
227 negative cloud feedbacks. Hence we are motivated to separate λ_{hist} and $\Delta\lambda$ into a ‘before’ and ‘after’
228 1980. This separation leads into our next motivating question.

229 Fourthly, are observations of recent decadal warming and TOA radiative fluxes since the 1980s in
230 agreement with the strongly negative λ values simulated by the AGCMs? If so, what would such a
231 strongly stabilizing feedback parameter (and large pattern effect) in the presence of a substantial
232 rate of observed global warming ($\sim 0.19 \text{ K dec}^{-1}$, Tokarska et al., 2020) imply for the efficiency of
233 ocean heat uptake and is there any relationship between them? Are any of these relationships
234 affected by the most recent data in which Loeb et al. (2020; 2021) identified a marked change in the
235 Earth’s radiation budget associated with the 2015/2016 El Niño event and a change in sign in the
236 Pacific Decadal Oscillation (PDO) index. Such a shift in tropical Pacific SST patterns (a shift to
237 warming in the eastern Pacific) should favour more positive feedbacks (Loeb et al., 2020).

238 Finally and fifthly, a limitation of the *amip-piForcing* approach, as discussed in Section 1.1, is that λ_{hist}
239 and $\Delta\lambda$ derived from these experiments includes a structural dependence on the SST patterns and
240 sea-ice in the AMIP II boundary condition data set used to force the AGCMs (Andrews et al., 2018;
241 Lewis and Mauritsen, 2021; Zhou et al., 2021; Fueglistaler and Silvers, 2021). To investigate this
242 further, we supplement the *amip-piForcing* simulations with sensitivity tests with nine AGCMs
243 forced with historical HadISST1 (Rayner et al., 2003) SSTs as per Lewis and Mauritsen (2021).

244 In summary, previous studies have shown that historical climate feedback (λ_{hist}) varies on decadal
245 timescales in *amip-piForcing* simulations and is larger in magnitude (climate sensitivity smaller) than
246 that seen in long-term *abrupt-4xCO2* simulations associated with ECS, giving rise to a ‘pattern

247 effect'. This is accentuated over recent decadal climate change. Here we make use of observations
248 of the Earth's energy budget from 1985 and a new suite of *amip-piForcing* simulations from
249 CFMIP3/CMIP6 (giving us a combined ensemble of 14 models), as well as targeted HadISST1 versus
250 AMIP II SST dataset sensitivity tests with nine AGCMs, to address the above questions.

251 The manuscript is organised as follows: Section 2 describes the model and observational data.
252 Section 3 presents the model results. Section 4 brings in the observational data. Section 5 presents a
253 summary, discussion and outlook.

254

255 **2. Methods and Data**

256 *2.1 amip-piForcing*

257 To provide estimates of λ_{hist} consistent with the observed variations in SST patterns we turn to
258 AGCMs forced with observed monthly variations in SSTs and sea-ice, while keeping all forcing agents
259 such as greenhouse gases and aerosols etc. constant at pre-industrial levels. Since the radiative
260 forcing is constant ($\Delta F=dF=0$) by construction, λ_{hist} can be diagnosed via $\lambda_{\text{hist}} = dN/dT$ (or $\Delta N/\Delta T$ if
261 using finite differences between climate states) (Andrews, 2014; Gregory and Andrews, 2016, Zhou
262 et al., 2016; Silvers et al., 2018; Andrews et al., 2018). Such an experimental design is now referred
263 to as *amip-piForcing* (Gregory and Andrews, 2016). The experimental protocol builds on the
264 Atmospheric Model Intercomparison Project (AMIP) design (Gates et al. 1999) that has long been
265 used in climate modelling, but extends back to 1870 (rather than 1979 in AMIP) and forcing agents
266 are kept at pre-industrial levels. As per AMIP, the underlying SST and sea-ice dataset used to force
267 the AGCMs is the AMIP II boundary condition data set (Gates et al., 1999; Hurrell et al., 2008; Taylor
268 et al., 2000). A description of the *amip-piForcing* protocol for CFMIP3/CMIP6 is given in Webb et al.
269 (2017). When forced with observed monthly SSTs and sea-ice, AGCMs generally reproduce the
270 observed relationships between surface temperature patterns, cloudiness and radiative fluxes well
271 (Allan et al., 2014; Loeb et al. 2020), lending some credibility to the radiative effects of their
272 simulated pattern effects to different SST patterns.

273 The *amip-piForcing* simulations used in this study are summarised in Table 1. They reflect a
274 combination of new CFMIP3/CMIP6 simulations with the latest generation of models archived in the
275 CMIP6 database and those used in Andrews et al. (2018) with some updates (see below). The
276 exception is MPI-ESM1-2-LR (Mauritsen et al., 2019); this is a CMIP6 generation model but its *amip-*
277 *piForcing* simulation is not currently included in the CMIP6 database. Note that this model contains
278 the ECHAM6.3 atmospheric model, so the results ought to be very similar to the older ECHAM6.3
279 simulations used in Andrews et al. (2018) and Lewis and Mauritsen (2021), though the models are
280 not identical owing to differences in atmospheric composition and land surface properties (see
281 Mauritsen et al., 2019, regarding the transition from MPI-ESM1.1 to MPI-ESM1.2). Furthermore, the
282 newer MPI-ESM1-2-LR simulations include a longer time-period than the ECHAM6.3 simulations
283 (Table 1).

284 The CFMIP3/CMIP6 *amip-piForcing* simulations begin in year 1870, but we discard the first year to
285 be consistent with the earlier Andrews et al. (2018) ensemble which started in January 1871. The
286 CFMIP3/CMIP6 simulations end in Dec 2014, whereas the simulations in the original Andrews et al.
287 (2018) ensemble (largely) ended in Dec 2010. In part to address this, some of the Andrews et al.
288 (2018) simulations have been rerun, including CAM4, GFDL-AM3 and GFDL-AM4 simulations, which
289 now end in Dec 2014 or later (see Table 1). Another difference to Andrews et al. (2018) is that we
290 now have an *abrupt-4xCO2* AOGCM simulation with GFDL-AM4 which they did not consider, to

291 permit a quantification of the pattern effect in that model. In contrast, we exclude the Andrews et
292 al. (2018) CAM5.3 simulation from our analysis since there is no *abrupt-4xCO2* AOGCM simulation to
293 compare against.

294 The models used, time-periods covered and number of ensembles are detailed in Table 1. Where
295 ensembles exist, an ensemble-mean dT and dN is created before analysis. Note that it makes little
296 difference to λ if, alternatively, individual members are first analysed and then the results ensemble-
297 meaned (Gregory and Andrews, 2016; Lewis and Mauritsen, 2021). All models share a common
298 1871-2010 time-period and so the principal analysis is restricted to this time-period, but we consider
299 the additional years to 2014 too. All data are global-annual-ensemble-means and expressed as
300 anomalies relative to an 1871-1900 baseline and the timeseries data has been made available (see
301 Data Availability Section).

302 Unless otherwise stated all uncertainties in multi model ensemble-mean results represent a 5-95%
303 confidence interval, calculated as 1.645σ across models assuming a gaussian distribution. We do not
304 attempt to adjust our uncertainty for the number of independent models, n , used in the ensemble
305 (i.e. dividing by square root of n). Our approach is similar to a "statistical indistinguishable ensemble"
306 approach (Annan and Hargraves, 2011; 2017) though likely overstates the uncertainty in the true
307 value if the ensemble shares characteristics of a "truth centred paradigm" (Sanderson and Knutti,
308 2012).

309 2.2 *HadSST-piForcing*

310 To test the sensitivity of the *amip-piForcing* results to the underlying SST dataset, we repeat the
311 *amip-piForcing* simulations with nine AGCMs (see Table 1) but replace the AMIP II boundary
312 condition SST dataset with HadISST1 (Rayner et al. 2003). All other aspects of the simulations,
313 including sea-ice, are identical to the *amip-piForcing* simulations. This is the same experimental
314 design as Lewis and Mauritsen (2021), and we include their ECHAM6.3 simulations here (which again
315 ought to be similar to the MPI-ESM1-2-LR simulations). The simulations cover a common time-period
316 across models of 1871-2010, like in *amip-piForcing*, but some models are also extended further (see
317 Table 1). We refer to these simulations as *hadSST-piForcing*, but note only the SSTs are from the
318 HadISST1 dataset (hence 'hadSST' rather than 'hadISST'), the sea-ice remains as per *amip-piForcing*.
319 Like *amip-piForcing*, all data are global-annual-ensemble-means and expressed as anomalies relative
320 to an 1871-1900 baseline, and the timeseries data has been made available (see Data Availability
321 Section).

322 Lewis and Mauritsen (2021) provide a summary of the source observational inputs used to construct
323 the AMIP II and HadISST1 SST datasets and how they differ. In addition, we note that AMIP II uses
324 HadISST1 SSTs (Rayner et al. 2003) prior to November 1981 and version 2 of the National Oceanic
325 and Atmospheric Administration (NOAA) weekly optimum interpolation (OI.v2) SST analysis
326 (Reynolds et al. 2002) thereafter. The merging procedure rebases the HadISST1 SSTs to avoid
327 discontinuities in the merged dataset (Hurrell et al. 2008). Hence AMIP II and HadISST1 might be
328 expected to be more similar before 1981, and diverge afterwards.

329 2.3 *abrupt-4xCO2*

330 A corresponding *abrupt-4xCO2* simulation using each AGCM's coupled AOGCM is used to determine
331 the model's long-term sensitivity metrics (F_{4x} , λ_{4xCO2} and $ECS = -0.5 * F_{4x} / \lambda_{4xCO2}$) from regression of
332 global-annual-mean dN against dT over 150 years of the simulations (see Andrews et al., 2012). We
333 also use λ_{4xCO2} diagnosed from years 1-20 and years 21-150 of the *abrupt-4xCO2* simulation following
334 Andrews et al. (2015), which approximately separates the two principal timescales of the climate

335 response: the mixed-layer and deep-ocean (see Geoffroy et al. 2013 and Andrews et al. 2015).
336 *abrupt-4xCO2 data* is available on the CMIP5 database (Taylor et al., 2012) for CCSM4, GFDL-CM3
337 and HadGEM2-ES. All other abrupt-4xCO2 data is available on the CMIP6 database (Eyring et al.,
338 2016), except for HadCM3 and MPI-ESM1.1. For ECHAM6.3/MPI-ESM1.1, *abrupt-4xCO2* global-
339 annual mean dN and dT timeseries data are provided by Andrews et al. (2018). HadAM3 data is
340 taken from Andrews et al. (2018) and Andrews et al. (2015); while a mean of seven realizations, this
341 simulation is only 100 years long so the calculations are over years 1-100 for λ_{4xCO2} and years 1-20 or
342 21-100 for the separation of timescales in this model.

343 Note when aligning each AGCM to its AOGCM, sometimes the AGCM and AOGCM model names
344 differ in the literature. We indicate where this is applicable in Table 1. This does not apply to the
345 newer CFMIP3/CMIP6 simulations which publish their AGCM and AOGCM simulations under
346 consistent names.

347 *2.4 Observations of recent decadal climate change*

348 To understand Earth's recent decadal climate change since ~1985 we turn to its observed global-
349 mean energy budget (i.e. dT , dN and dF). For dT we use the HadCRUT5 analysis dataset (Morice et al.
350 2021) (the current version is HadCRUT.5.0.1.0). This is an improvement on previous HadCRUT
351 products and extends coverage in data sparse regions (see Morice et al. 2021). For dF we use the
352 best estimate historical ERF timeseries produced by IPCC AR6 (Forster et al. 2021; Smith et al. 2021).
353 For dN we use various versions of the DEEP-C satellite based reconstruction of the Earth's radiation
354 balance from 1985 to near-present. These are described in detail in Allan et al. (2014) and Liu et al.
355 (2015; 2017; 2020), but as we will use various versions of this product we give a brief overview here.

356 The DEEP-C dataset is derived by merging satellite observations of top-of-atmosphere radiative flux
357 timeseries from ERBE WFOV (Earth Radiation Budget Experiment Satellite wide field of view) and
358 ECMWF reanalysis (ERA-Interim/ERA5) since 1985 with CERES (Clouds and the Earth's Radiant Energy
359 System) satellite observed fluxes since March 2000. Hence prior to March 2000 it is largely informed
360 by ERBE WFOV and ERA reanalysis, then aligns with CERES from March 2000. AMIP and high
361 resolution AGCM simulations and reanalyses are used in the merging process to bridge the gaps
362 between products and avoid discontinuities in the timeseries, including a gap in the satellite record
363 during 1993 and 1999 (Allan et al. 2014). It is important to note that substantial uncertainty in
364 decadal changes in dN associated with the merging process affects the record and this is
365 conservatively estimated to be as high as 0.5 Wm^{-2} for changes applying across the whole record (Liu
366 et al. 2020). However, uncertainty in the CERES period since March 2000 is much smaller based on
367 assessment of instrument drift (Loeb et al. 2021). Various versions of the DEEP-C dataset exist which
368 parallel updates to the underlying products and update the merging process. We use the latest
369 version (DEEP-C v5, Liu and Allan 2022) for our principal analysis, which is based on CERES EBAF v4.1
370 and ERBS WFOV v3, alongside ERA5 reanalysis and AMIP6 simulations (Liu and Allan, 2022). To
371 illustrate structural uncertainties in our analysis we also use previous versions (v2, v3 and v4) of the
372 DEEP-C datasets. The availability of datasets is provided in the Data Availability Section.

373

374 **3. Historical feedback and pattern effect in amip-piForcing and hadSST-piForcing simulations**

375 Figure 1a shows the multi-model ensemble mean dT timeseries in the *amip-piForcing* and *hadSST-*
376 *piForcing* simulations, alongside an observed estimate from HadCRUT5 analysis dataset. The AGCM
377 design reproduces the observed historical dT variability well (the correlation coefficient, r , between
378 observed and both simulated dT timeseries is 0.97). However the AGCMs do not reproduce the

379 observed trends precisely, notably omitting some observed warming particularly in the most recent
 380 decades (Figure 1a). This is because the AGCM design omits a small component of warming
 381 associated with land surface temperature change (which is not prescribed in the models) that arises
 382 as a direct consequence of increases in greenhouse gases and other forcing agents independent of
 383 SST change (this is often considered as part of the ERF rather than feedback) (see Andrews, 2014;
 384 Gregory and Andrews, 2016; Andrews et al., 2018). This will be included in the observed record but
 385 not in the simulated dT because greenhouse gases and other forcing agents are kept constant at pre-
 386 industrial levels in *amip-piForcing* and *hadSST-piForcing*.

387 As dT increases, dN reduces (Figure 1b), i.e. the climate loses more heat to space as a consequence
 388 of the climate response and feedbacks in the system. Figure 1c and 1d show the difference in the dT
 389 and dN timeseries between the *amip-piForcing* and *hadSST-piForcing* ensemble-mean response. For
 390 most of the time the differences vary approximately about zero. However, larger differences are
 391 evident from 1981 onwards, when the dN response in *amip-piForcing* is substantially larger than that
 392 in *hadSST-piForcing* (Figure 1b and 1d), up to $\sim 0.5 \text{ W m}^{-2}$ in some years (Figure 1d). This is consistent
 393 with 1981 being the year in which the AMIP II boundary condition source dataset switches from
 394 HadISST1 to OI.v2 SST (see Section 3.2). This motivates us to separate the historical record into two
 395 time-periods either side of 1980, i.e. 1871-1980 and 1981-2010 (Section 3.2).

396 However, we first consider feedback and the pattern effect that arises when calculated over the
 397 historical record as a whole, rather than any time-period within. This is useful for informing studies
 398 that use the entire observed historical record to estimate ECS via energy budget constraints (e.g.
 399 Andrews et al., 2018; Sherwood et al. 2020; Forster et al. 2021). It also allows a direct comparison of
 400 our results using a broad ensemble of models to the narrower range of model results reported by
 401 Andrews et al. (2018) and Lewis and Mauritsen (2021).

402 3.1 Considering the historical record as a whole

403 Figures 1e and 1f show the $\lambda_{\text{hist}} = dN/dT$ relationship in the ensemble-mean *amip-piForcing* and
 404 *hadSST-piForcing* simulation for 1871-2010. λ_{hist} is determined from ordinary least square linear
 405 regression on global-annual-mean dN and dT timeseries data. λ_{hist} values for individual models are
 406 given in Table 2 alongside their *abrupt-4xCO2* sensitivity metrics. Across the fourteen model
 407 ensemble of *amip-piForcing* simulations $\lambda_{\text{hist}} = -1.65 \pm 0.46 \text{ W m}^{-2} \text{ K}^{-1}$, slightly smaller in magnitude
 408 but with similar spread to the Andrews et al. (2018) ensemble (they reported $\lambda_{\text{hist}} = -1.74 \pm 0.48 \text{ W m}^{-2} \text{ K}^{-1}$).
 409 Like in Andrews et al. (2018), the spread in λ_{hist} is extremely similar to the spread in λ_{4xCO2} from
 410 the coupled AOGCM *abrupt-4xCO2* ensemble (Table 2) (this is also true for the individual feedback
 411 terms, see below). The pattern effect, $\Delta\lambda = \lambda_{4xCO2} - \lambda_{\text{hist}}$ between *amip-piForcing* and *abrupt-4xCO2*
 412 (with λ_{4xCO2} from years 1-150 of *abrupt-4xCO2*) is $\Delta\lambda = 0.70 \pm 0.47 \text{ W m}^{-2} \text{ K}^{-1}$ across the ensemble
 413 (Table 3), which is slightly larger in magnitude but with more spread than that reported by Andrews
 414 et al. (2018) ($0.64 \pm 0.40 \text{ W m}^{-2} \text{ K}^{-1}$).

415 Tables 2 and 3 also present the equivalent λ_{hist} and $\Delta\lambda$ values when the AGCMs are forced with
 416 HadISST1 SSTs instead (*hadSST-piForcing*) and Figure 2 shows the relationship to *amip-piForcing*. λ_{hist}
 417 $= -1.43 \pm 0.41 \text{ W m}^{-2} \text{ K}^{-1}$ in *hadSST-piForcing* (Table 2), which is smaller in magnitude but with similar
 418 spread to the *amip-piForcing* results above. Subsetting to the nine AGCMs with both simulations, λ_{hist}
 419 is $0.28 \pm 0.17 \text{ W m}^{-2} \text{ K}^{-1}$ smaller in magnitude in *hadSST-piForcing* but well correlated ($r=0.93$) with
 420 *amip-piForcing* values (Figure 2a, red points). The regression slopes of the red line in Figures 2a
 421 (slope = 0.84 ± 0.21) and 2b (slope = 0.84 ± 0.26) are statistically consistent with unity, implying
 422 there is little AGCM dependence in the difference between λ_{hist} from *amip-piForcing* and *hadSST-*
 423 *piForcing*. Hence, given the strong correlation and close approximation of being parallel to the one-

424 to-one line (Figure 2, red points), we suggest a simple offset given by the difference ($0.28 \pm 0.17 \text{ W}$
425 $\text{m}^{-2} \text{ K}^{-1}$, Table 3) well approximates the relationship between λ_{hist} over 1871-2010 in *amip-piForcing*
426 and *hadSST-piForcing*.

427 Despite λ_{hist} being smaller in magnitude in *hadSST-piForcing*, $\Delta\lambda = 0.48 \pm 0.36 \text{ W m}^{-2} \text{ K}^{-1}$ is still large
428 and positive across the *hadSST-piForcing* ensemble (Table 3). The smaller uncertainty than the *amip-*
429 *piForcing* pattern effect likely reflects the narrower diversity of model physics in the smaller *hadSST-*
430 *piForcing* ensemble, for example we do not have *hadSST-piForcing* experiments for the model
431 (MIROC6) with the smallest pattern effect in *amip-piForcing*. If we subset the *amip-piForcing*
432 ensemble to just those nine models with corresponding *hadSST-piForcing* experiments (Fig 2b, red
433 points), then the spread (as measured by 1.645σ) across models in $\Delta\lambda$ reduces from 0.47 to 0.38,
434 which is similar to the spread found in *hadSST-piForcing*.

435 That a large pattern effect is present in the *hadSST-piForcing* simulation over the historical record is
436 not in contradiction with the results of Lewis and Mauritsen 2021 (LM2021), who reported a
437 ‘negligible unforced historical pattern effect’ with ECHAM6.3 when forced with HadISST1 SSTs. This is
438 because LM2021 calculated their pattern effect by comparing λ from *hadSST-piForcing* to λ derived
439 from a coupled AOGCM historical simulation, or approximations of it from years 1-70 of 1%CO₂ or
440 years 1-50 of *abrupt-4xCO₂* simulations. This necessarily gives a smaller pattern effect because it
441 excludes many of the SST variations and patterns effects seen on longer timescales in CO₂ forced
442 simulations (Senior and Mitchell, 2000; Gregory et al. 2004; Andrews et al. 2012; Armour et al.,
443 2013; Geoffroy et al., 2013; Andrews et al. 2015; Rugenstein et al. 2016). While this might be useful
444 for trying to quantify different mechanisms of the pattern effect (e.g. forced or unforced, see
445 Dessler, 2020), it is a quantity we are less interested in, as we want to know the λ of relevance to
446 long-term ECS and projections of the late 21st century. Therefore contrasting to λ_{4xCO_2} from years 1-
447 150 is the most relevant metric (Sherwood et al., 2020), as we have done here.

448 Following Andrews et al. (2018) we decompose λ into its component longwave (LW) clear-sky,
449 shortwave (SW) clear-sky and cloud radiative effect (CRE, equal to all-sky minus clear-sky fluxes)
450 terms in Figure 3. Deviations away from the one-to-one line indicate a difference in *amip-piForcing*
451 and *abrupt-4xCO₂* λ (i.e. the pattern effect). Tables of the individual model results are given in the
452 Supplementary Tables 1 - 3. It confirms the basic premise that historical LW clear-sky and cloud
453 feedbacks are more stabilizing than under *abrupt-4xCO₂*, consistent with the mechanistic and
454 process understanding that the pattern effect arises predominantly from a lapse-rate (which affects
455 LW clear-sky fluxes) and cloud feedback dependence on SST patterns (e.g. Zhou et al., 2016,
456 Andrews and Webb, 2018, Ceppi and Gregory, 2017; Dong et al. 2019). Figure 3 and Supplementary
457 Tables 1 - 3 show that the inter-model spread in feedback in both *amip-piForcing* and *abrupt-4xCO₂*
458 is dominated by cloud rather than clear-sky feedbacks. Figure 3 also suggests there is a small
459 compensation to the total pattern effect from SW clear-sky feedbacks, likely from sea-ice. That is,
460 AGCMs forced with AMIP II boundary condition sea-ice changes have a slightly more positive
461 feedback than found in their coupled *abrupt-4xCO₂* simulations, though the difference is small
462 (Figure 3). Consequently, a simple attribution of the difference in total feedback between *amip-*
463 *piForcing* and *abrupt-4xCO₂* to an SST driven pattern effect (as we have done here) will slightly
464 understate the actual effect, though the term is small and we neglect it from now on. We discuss
465 sea-ice uncertainties further below.

466 MIROC6 is the only model in the *amip-piForcing* ensemble to have near zero pattern effect (Table 3
467 and note the single black dot on the one-to-one line in Figure 3). The reason for this different
468 behaviour remains unclear. One could speculate that there is a relationship between a model’s
469 climate sensitivity and its pattern effect, given that MIROC6 has the lowest ECS of all models

470 considered here (ECS=2.6K, Table 2). However, we note that there is little correlation between ECS
471 and $\Delta\lambda$ across models ($r=0.4$) and that several other models with low ECS have large $\Delta\lambda$.

472 Alternatively, it could be that MIROC6's atmospheric physics are largely insensitive to different SST
473 patterns and/or that its AOGCM *abrupt-4xCO2* warming pattern is more similar to the historical
474 record than other models. Both are potentially possible. For example, λ_{hist} for 1871-1980 and 1980-
475 2010 separately (next Section and Table 2) shows that MIROC6 does simulate a pattern effect, but
476 achieves a near zero pattern effect over the historical record as a whole by having a smaller (relative
477 to other models) pattern effect over recent decades, offset by a negative pattern effect over the
478 earlier period. In addition - and in contrast to other models - MIROC6 simulates a negative LW clear-
479 sky pattern effect (red dot below the one-to-one line, Figure 3) which offsets its positive cloud
480 feedback pattern effect.

481 The model with the largest pattern effect is CESM2 (Table 3). This occurs because of a particularly
482 large cloud feedback sensitivity to SST patterns (grey dot furthest from the one-to-one line, Figure
483 3). Zhu et al. (2022) argue that an issue in CESM2's cloud microphysics related to cloud ice number
484 leads to an unrealistically large cloud sensitivity to warming in this model. Whether this is
485 responsible for the model's large pattern effect is unclear. Mixed-phase clouds have not typically
486 been associated with the pattern effect, though might be of relevance to pattern effects over the
487 Southern Ocean (Dong et al. 2020; Bjordal et al. 2020). It would be interesting in future work to
488 identify the different cloud types associated with the pattern effect and conduct sensitivity
489 experiments with CESM2 to investigate which aspects of the cloud feedback change with different
490 cloud microphysics schemes.

491 Many of our *amip-piForcing* simulations (eleven models) continue to Dec 2014 (Table 1), and six
492 have corresponding *hadSST-piForcing* simulation, so we consider how this extended period affects
493 the overall assessment of the historical pattern effect. In the eleven *amip-piForcing* simulations, λ_{hist}
494 = $-1.65 \pm 0.48 \text{ W m}^{-2} \text{ K}^{-1}$ over 1871-2010, but this increases in magnitude so that $\lambda_{\text{hist}} = -1.71 \pm 0.51 \text{ W}$
495 $\text{m}^{-2} \text{ K}^{-1}$ if calculated over 1871-2014 (Supplementary Table 4). An increase occurs in every model and
496 the magnitude of change across the ensemble is $0.07 \pm 0.06 \text{ W m}^{-2} \text{ K}^{-1}$ (Supplementary Table 4). In
497 the six corresponding *hadSST-piForcing* simulations, $\lambda_{\text{hist}} = -1.48 \pm 0.41 \text{ W m}^{-2} \text{ K}^{-1}$ over 1871-2010, but
498 this increases in magnitude so that $\lambda_{\text{hist}} = -1.53 \pm 0.39 \text{ W m}^{-2} \text{ K}^{-1}$ if calculated over 1871-2014
499 (Supplementary Table 4). The magnitude of the increase ($0.05 \pm 0.05 \text{ W m}^{-2} \text{ K}^{-1}$) is thus slightly
500 smaller in this dataset (Supplementary Table 4).

501 While we have focused on the SST driven pattern effect, a remaining structural uncertainty in
502 assessing total feedback differences between $\lambda_{4\times\text{CO}_2}$ and λ_{hist} relates to the sea-ice dataset used to
503 force the AGCMs. Andrews et al. (2018) provided a sensitivity test (see their Supplementary
504 Material) by repeating the *amip-piForcing* simulation in two AGCMs but forced with HadISST2.1
505 (Titchner and Rayner, 2014) SSTs and sea-ice. They found that the historical feedback parameter
506 increased by $\sim 0.6 \text{ W m}^{-2} \text{ K}^{-1}$ when forced with HadISST2.1 compared to AMIP II, and attributed most
507 of this change to differences in the sea-ice datasets rather than SST. They noted that HadISST2.1 has
508 substantially more pre-industrial Antarctic sea-ice concentration (see Titchner and Rayner, 2014),
509 and so generated more sea-ice loss (more positive feedback) over the historical period (Andrews et
510 al. 2018), as well containing large discontinuities in the timeseries. The historical sea-ice trends and
511 associated feedbacks over the Southern Ocean in the HadISST2.1 dataset are difficult to reconcile
512 with those found in AOGCMs and our physical understanding of them (Schneider et al. 2018). We do
513 not pursue this further, but simply highlight that dataset assumptions made about pre-industrial sea-
514 ice concentrations in Antarctica can have substantial impacts on diagnosed feedbacks in AGCMs and
515 remains an outstanding uncertainty in assessing total feedback differences. Fortunately, in *amip-*

516 *piForcing* the difference in SW clear-sky feedback (which will be strongly impacted on by sea-ice
517 feedbacks) is similar to that seen in λ_{4xCO_2} (Figure 3) so this can be ignored if the focus is solely on SST
518 driven feedbacks in the atmosphere.

519 In summary, for warming since the 1800s (using either 1871-2010 or 1871-2014), both *amip-*
520 *piForcing* and *hadSST-piForcing* suggest a substantial pattern effect between radiative feedbacks
521 operating over historical climate change and long-term ECS.

522 3.2 Considering the historical record before and after 1980

523 We now return to the divergence in dN response between *amip-piForcing* and *hadSST-piForcing*
524 simulations around 1980 (Figure 1d). As well as the change in behaviour discussed above, 1980
525 provides a convenient separation of historical feedbacks and the pattern effect for two other
526 motivating reasons: (i) Fueglistaler and Silvers (2021) identify ~1980 as the point in which the Earth
527 begins to warm with a particular configuration of tropical Pacific SSTs where regions of deep
528 convection warm substantially more than the tropical mean, driving large negative cloud feedbacks
529 and consistent with a large pattern effect over this period (Gregory and Andrews 2016; Zhou et al.,
530 2016; Andrews et al., 2018; Gregory et al. 2020); and (ii) Fueglistaler and Silvers (2021) also identify
531 ~1980 as a useful approximation of when the satellite era was integrated into the global observing
532 system, and so developing an understanding of feedbacks and the pattern effect specifically from
533 1980 onwards will aid interpretation of our most comprehensive observations of climate change and
534 how they might relate to the future change (next Section).

535 Figure 4 compares the surface temperature pattern over the two time-periods 1871-1980 and 1981-
536 2010 in *amip-piForcing* and *hadSST-piForcing*. Differences between the two SST reconstructions are
537 extremely subtle. For the earlier 1871-1980 time period, warming is more uniform, in part because
538 of the longer time-period considered which will smooth out variability. Since 1981 there has been
539 western Pacific warming with cooling in the Southern Ocean and off equatorial eastern Pacific
540 (which are regions of marine low clouds), despite temperatures increasing in the global mean.
541 Hence, we might expect a small pattern effect prior to 1980 and a large pattern effect post 1980
542 (e.g. Gregory and Andrews, 2016; Zhou et al., 2016, Andrews and Webb, 2018, Ceppi and Gregory,
543 2017; Dong et al. 2019, Fueglistaler and Silvers 2021).

544 Figures 1g and 1h show the $\lambda_{\text{hist}} = dN/dT$ relationship in the ensemble-mean *amip-piForcing* and
545 *hadSST-piForcing* simulation for 1871-1980 (grey points) and 1981-2010 (blue points). Results for
546 individual models are given in Table 2. Figures 1g and 1h confirms the basic premise that λ_{hist}
547 strengthens in magnitude post 1980, consistent with the change in SST patterns (Figure 4).

548 For the earlier time-period, 1871-1980, $\lambda_{\text{hist}} = -1.14 \pm 0.33 \text{ W m}^{-2} \text{ K}^{-1}$ in *amip-piForcing* is similar to λ_{hist}
549 $= -1.21 \pm 0.38 \text{ W m}^{-2} \text{ K}^{-1}$ in *hadSST-piForcing* (Table 2) – suggesting little sensitivity of the results to
550 these two SST datasets over this time period. This is unsurprising given that the datasets are similar
551 (though not identical) prior to this period (Section 2.2 and Figure 4). For the nine AGCMs that
552 performed both simulations Figure 2a shows the relationship between λ_{hist} in *amip-piForcing* and
553 *hadSST-piForcing*. For all time-periods λ_{hist} in *amip-piForcing* and *hadSST-piForcing* is found to be well
554 correlated ($r \geq 0.87$, Figure 2a). For the earlier 1871-1980 results, the λ_{hist} values fall close to the one-
555 to-one line (blue dots, Figure 2) and within the range of λ_{4xCO_2} (grey shaded areas in Figure 2). This
556 suggests that for 1871-1980 λ_{hist} is broadly independent of the two SST datasets (consistent with
557 their common basis) and that the pattern effect is small for this time period. Indeed, the 1871-1980
558 pattern effect is small but positive ($\Delta\lambda = 0.19 \pm 0.35 \text{ W m}^{-2} \text{ K}^{-1}$ in *amip-piForcing* and $0.26 \pm 0.26 \text{ W m}^{-2}$
559 K^{-1} in *hadSST-piForcing*, Table 3 and Figure 2b).

560 In contrast, for 1981 onwards (i.e. 1981-2010), λ_{hist} is generally far from the $\lambda_{4\times\text{CO}_2}$ range (i.e. a large
561 pattern effect) and away from the one-to-one line (i.e. a dependence on the SST dataset) (Figure 2a;
562 grey points). Indeed, λ_{hist} over 1981-2010 is substantially stronger in magnitude than over 1871-1980
563 ($\lambda_{\text{hist}} = -2.33 \pm 0.72 \text{ W m}^{-2} \text{ K}^{-1}$ in *amip-piForcing* over 1981-2010, Table 2; Figure 2a) and the pattern
564 effect is large ($\Delta\lambda = 1.38 \pm 0.75 \text{ W m}^{-2} \text{ K}^{-1}$, Table 3; Figure 2b), although somewhat weaker in
565 magnitude in *hadSST-piForcing* ($\Delta\lambda = 1.24 \pm 0.88 \text{ W m}^{-2} \text{ K}^{-1}$, Table 3; Figure 2b). For 1981-2010, λ_{hist} is
566 generally weaker in *hadSST-piForcing* (Table 2; Figure 3a) by $0.24 \pm 0.46 \text{ W m}^{-2} \text{ K}^{-1}$ across the nine
567 AGCMs using both SST datasets.

568 These results are generally consistent with Fueglistaler and Silvers (2021) and Lewis and Mauritsen
569 (2021) who both point to the AMIP II SST dataset as having larger (relative) western tropical Pacific
570 warming than in other SST datasets, and hence from the process understanding we would expect a
571 more negative feedback (and larger pattern effect) in *amip-piForcing*, as found above. The one
572 exception is GFDL-AM4, which simulates a more negative λ_{hist} under HadISST1 SSTs than AMIP II
573 from 1981-2010, and so a larger pattern-effect over this period under HadISST1 SSTs (Tables 2 and 3
574 and the single grey dots in Figures 2a and 2b which sit on the other side of the one-to-one line from
575 the other models). The reasons for this remain unclear.

576 In summary we have shown that a division around 1980 usefully separates historical climate change
577 into two time-periods: (i) pre 1981 the Earth warmed over most of the historical record with an
578 averaged warming pattern that is relatively uniform, and feedbacks largely consistent with long-term
579 ECS feedbacks (i.e. a relatively small pattern effect), and (ii) post 1980 where the Earth warmed with
580 a particular configuration of strong SST gradients that drove feedbacks much more stabilizing than
581 those seen in long-term ECS feedbacks (i.e. large pattern effect), albeit with a sensitivity of the
582 magnitude of this result to the SST dataset considered.

583 3.3 Relationships between historical and ECS feedbacks

584 We now consider whether feedbacks over the historical period in *amip-piForcing* are related to
585 $\lambda_{4\times\text{CO}_2}$. This is in contrast to the previous sections which only quantified their difference (i.e. the
586 pattern effect).

587 Firstly, we note that the spread in feedback across models over the earlier (1871-1980) time-period
588 in *amip-piForcing* is well correlated with the spread in feedback across models in *abrupt-4xCO2*
589 ($r=0.69$, Figure 5a). In contrast, feedbacks over the most recent decades (1981-2010) are only weakly
590 correlated with $\lambda_{4\times\text{CO}_2}$ ($r=0.27$). Secondly, feedback over the full historical record (1871-2010) is only
591 weakly correlated with feedback from the 1871-1980 time-period ($r=0.45$, Figure 5b). In contrast,
592 1871-2010 feedback is strongly correlated with feedback over the most recent 1980-2010 decades
593 ($r=0.91$, Figure 5b). This strong correlation between 1981-2010 and the 1871-2010 feedback arises
594 because the spread for 1871-2010 is dominated by the spread for 1981-2010.

595 Given that the feedbacks applying in 1871-1980 and in 1981-2010 are different, we infer that the SST
596 patterns over these two periods are driven by different mechanisms. Because the feedbacks of
597 1871-1980 are correlated with *abrupt-4xCO2*, the difference between the two periods could be
598 explained by CO_2 being the dominant influence in 1871-1980 SST patterns, while something else (e.g.
599 perhaps variability, aerosol, volcanism) dominates during 1981-2010. This is only a hypothesis,
600 because these experiments do not provide a way to attribute the observed SST changes to causes.

601 The result is that the spread in feedbacks over the full historical record are only weakly correlated
602 with $\lambda_{4\times\text{CO}_2}$ ($r=0.51$, Figure 3), because of the strong pattern effect post 1980. Hence, we can say little
603 about future $\lambda_{4\times\text{CO}_2}$ directly from climate change post 1980 or even the full historical record without

604 adjusting for a pattern effect. In contrast, the feedbacks operating over the earlier 1871-1980 time-
605 period are correlated with $\lambda_{4\times\text{CO}_2}$ ($r=0.69$, Figure 5a).

606 That recent decadal feedbacks are the most unrepresentative of the long-term climate sensitivity is
607 unfortunate, not just because it coincides with the advent of the satellite record and so is extremely
608 well observed, but also because climate change since ~ 1980 ought to provide the best constraint on
609 ECS (e.g. Jiménez-de-la-Cuesta and Mauritsen, 2019). This is because it offers a strong global
610 warming signal, which AOGCMs attribute to greenhouse gas increases, while avoiding the large
611 uncertainty associated with global-mean aerosol radiative forcing in energy budget estimates of ECS.
612 However the role of aerosols should not be discounted entirely, since strong compensating regional
613 changes may have impacted on SST patterns (e.g. Smith et al. 2015; Takahashi & Watanabe, 2016;
614 Moseid et al., 2020). In contrast, although feedbacks operating over the earlier 1871-1980 part of
615 the historical record are correlated with long-term CO_2 induced feedbacks, a reliable observational
616 constraint is harder because the climate change signal is smaller and the observations poorer. We
617 discuss this further in the Discussion section.

618 Up to now we have only considered a comparison of *amip-piForcing* feedbacks to a single definition
619 of *abrupt-4xCO2* feedbacks (i.e. feedbacks diagnosed over years 1-150 in *abrupt-4xCO2*). Here we
620 briefly consider separating $\lambda_{4\times\text{CO}_2}$ into the two principal timescales of the *abrupt-4xCO2* response
621 following Andrews et al. (2015) by calculating $\lambda_{4\times\text{CO}_2}$ over years 1-20 (a fast timescale) and 21-150 (a
622 slow timescale) (Table 2). The rationale is that 20 years is approximately the timescale required for
623 the mixed-layer to equilibrate in response to step forcing, and any subsequent climate response
624 scaling with the slower deep-ocean timescale, as approximated by two-layer models (Held et al.,
625 2010; Geoffroy et al., 2013; Gregory et al., 2015).

626 Figure 5c shows λ_{hist} from 1871-1980 is largely scattered about the one-to-one line with $\lambda_{4\times\text{CO}_2}$ from
627 years 1-20, suggesting little to no pattern effect between these two. This is potentially consistent
628 with the historical record largely being the result of the faster timescale responses (Held et al. 2010;
629 Proistosescu & Huybers, 2017). In contrast, post-1980 λ_{hist} is far from the one-to-one line (i.e. large
630 pattern effect to years 1-20 of *abrupt-4xCO2*, Figure 5c) but is marginally correlated ($r=0.53$),
631 suggesting recent decades do contain some information relevant to the feedback seen in the fast
632 timescale response to CO_2 . However, the longer-term feedbacks associated with the slow timescale
633 response to CO_2 (years 21-150 of *abrupt-4xCO2*, Figure 5d) have no correlation with λ_{hist} post-1980
634 ($r=-0.06$, Figure 5d).

635 636 3.4 Decadal variability in feedbacks and the pattern effect 637

638 In this final section of GCM results we briefly comment how λ_{hist} and the pattern effect varies on
639 decadal timescales in the *amip-piForcing* and *hadSST-piForcing* simulations.

640 Following Gregory and Andrews (2016) we calculate $\lambda_{\text{hist}} = dN/dT$ over a moving 30 year window in
641 the *amip-piForcing* and *hadSST-piForcing* simulations (Figure 6a and b). For example λ_{hist} calculated
642 over the 30 year period 1925 to 1954 is presented at year 1939.5 in Figure 6. In Figures 6c-h the LW
643 and SW clear-sky and cloud radiative effect of the feedback are also shown. The correlation
644 coefficient between the *amip-piForcing* and *hadSST-piForcing* multi-model-mean λ_{hist} timeseries is
645 0.84, suggesting the broad features of the decadal λ_{hist} variations are robust to the SST datasets. In
646 particular λ_{hist} peaks (least negative, smallest pattern effect) around 1940 while generally being large
647 in magnitude (large pattern effect) over recent decades (see also Gregory and Andrews, 2016; Zhou
648 et al. 2016; Andrews et al. 2018; Gregory et al. 2020). The clear sky feedbacks (Figures 6c-f) are

649 largely stable, while the variation in λ_{hist} is almost entirely explained by variation in cloud feedback
650 (Figures 6g-h), consistent with previous findings (e.g. Zhou et al. 2016; Andrews et al. 2018).

651 In Section 5, we discuss further the reasons for the decadal variations in SST patterns and λ_{hist} , i.e.
652 whether they are the result of spatiotemporal changes in forcings such as aerosols or volcanic
653 forcing or due to unforced variability.

654

655 **4. Observed climate change**

656 We next consider whether the radiative feedback and pattern effects simulated by the GCMs are
657 consistent with observed variations in the Earth's energy budget. Gregory et al. (2020) asked a
658 similar question for the post 1980 period and suggested they are (see their Figure 5c), but here we
659 go a few steps further. Specifically, not only do we consider the post 1980 period, but also assess
660 changes in the Earth's energy budget back to the 1800s. Furthermore we investigate the implications
661 of a strongly negatively feedback parameter (large pattern effect) since 1985 on the observed rate of
662 global warming.

663 The observations also provide an opportunity to bring our λ_{hist} and pattern effect estimate up to date
664 with the most recently observed data (up to and including 2019), whereas our GCM analysis
665 generally finished in 2014. The observations post 2014 period are of particular interest given they
666 include the major El-Nino event of 2015/2016 that was associated with eastern-pacific warming and
667 marked changes in the observed radiation budget (Loeb et al. 2020; 2021). We expect these post
668 2014 years to have an impact λ_{hist} and the pattern effect, given the process understanding discussed
669 previously (e.g. Zhou et al., 2016, Andrews and Webb, 2018, Ceppi and Gregory, 2017; Dong et al.
670 2019).

671 *4.1 Comparison of AGCM results to observed estimates*

672 We first validate the AGCM λ_{hist} estimates over recent decades. To do this we use a merged satellite
673 dataset (ERBE WFOV + CERES) (Allan et al. 2014) that provides an observational estimate of dN
674 variations from 1985 to 2019. For dT we use the HadCRUT5 analysis dataset (Morice et al. 2021). For
675 dF we use the IPCC AR6 (Forster et al. 2021; Smith et al., 2021) best estimate historical ERF changes.
676 These datasets are described in further detail in Section 2.4. We first consider the 30-year period
677 1985 to 2014, consistent with many of the AGCMs.

678 Figure 7a and 7b show the dT , dN and dF timeseries over this period. The 1985-2014 'observed' $-\lambda_{\text{hist}}$
679 $= d(F - N)/dT \sim 2.0 \pm 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ relationship is shown in Figure 7d. Note the stated 5-95%
680 uncertainty is $\pm 1.645\sigma$ from the standard error of the linear fit, with no allowance for systematic
681 uncertainties. As discussed in Section 2.4, observed multi-decadal changes in dN are subject to a
682 substantial uncertainty (up to 0.5 W m^{-2}) primarily related to the breaks in the record prior to 2000,
683 though are considerably smaller afterwards (Liu et al. 2020). Note also that years 1991-2 are
684 excluded from the calculation as these years are identified as being strongly impacted by the
685 volcanic forcing from the Pinatubo eruption (Figure 7b). Whilst λ_{hist} is robust to this (we get just the
686 same $\lambda_{\text{hist}} \sim -2.0 \pm 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ if we include these years), including these years has an impact on the
687 ocean heat uptake efficiency estimate (see Section 4.3). The observed 1985-2014 λ_{hist} estimate is
688 shown on Figure 6a and 6b (red line) as an illustration in comparison to the AGCM decadal variations
689 in λ_{hist} . The observed λ_{hist} best estimate agrees exceptionally well with the AGCM multi-model mean,
690 and nearly all models are within the 5-95% uncertainty estimate as they approach the 1985-2014
691 value (Figure 6a and 6b).

692 A more rigorous comparison of individual AGCM results to the observed estimate is shown in Figure
693 8. Here the AGCM λ_{hist} estimates from *amip-piForcing* and *hadSST-piForcing* have been calculated in
694 the same way as the observations, i.e. over 1985-2014 excluding 1991-2. The overlap between the
695 model and observed estimates points to broad consistency between the models and observations in
696 the recent decadal value of λ_{hist} (Figure 8). The large uncertainties (which are likely underestimated
697 since we have not accounted for structural errors) inhibit a more precise validation of individual
698 models against the observed estimate.

699 For the full the historical record we estimate λ_{hist} from IPCC AR6 assessed changes in T , N and F .
700 Forster et al. (2021) give these as $\Delta T = 1.03 \pm 0.20$ K, $\Delta N = 0.59 \pm 0.35$ W m⁻² and $\Delta F = 2.20$ [1.53 to
701 2.91] W m⁻² for the time-period 1850-1900 to 2006-2019. For simplicity we assume $\Delta F = 2.20 \pm 0.7$ W
702 m⁻², where we have approximated the uncertainty in ΔF as a Gaussian. Randomly sampling (with
703 replacement) from the Gaussian distributions in ΔN , ΔF and ΔT gives $\lambda_{\text{hist}} = (\Delta N - \Delta F)/\Delta T = -1.6 \pm 0.8$
704 W m⁻² K⁻¹. This is again in agreement with the *amip-piForcing* ($\lambda_{\text{hist}} = -1.65 \pm 0.46$ W m⁻² K⁻¹, Table 2)
705 and *hadSST-piForcing* ($\lambda_{\text{hist}} = -1.43 \pm 0.41$ W m⁻² K⁻¹, Table 2) 1871-2010 ensembles, though an exact
706 match is not expected given the slightly different time-periods and methods (e.g. finite differences
707 versus regression) used. Still, the agreement provides further confidence in the GCM's simulated
708 radiative response to observed SST and sea-ice variations over the historical record, and strengthens
709 the conclusion that λ_{hist} has become more negative over recent decades compared to the longer
710 1871-2010 time-period.

711 Finally, IPCC AR6 assessed the long-term ECS relevant feedback parameter (analogous to our $\lambda_{4\times\text{CO}_2}$)
712 to be -1.16 ± 0.65 W m⁻² K⁻¹ (Forster et al., 2021) by combining lines of evidence from observations,
713 theory, process models and GCMs on individual climate feedback processes. Combining this with our
714 observed λ_{hist} estimates above gives an estimate of the pattern effect independently of our GCM
715 ensemble. This gives an estimated pattern effect of $\sim 0.8 \pm 1.0$ W m⁻² K⁻¹ for 1985-2015 and $\sim 0.4 \pm 1.1$
716 W m⁻² K⁻¹ for the full historical record (the 1850-1900 to 2006-2019 changes). While the uncertainties
717 are substantial, there is again agreement with our GCM results.

718 4.2 Recent observed trends and the efficiency of ocean heat uptake

719 We have seen that both models and observed variations in the Earth's energy budget agree on the
720 Earth having had strongly stabilizing feedbacks over recent decades relative to AOGCM feedbacks
721 under long-term CO₂ forced climate change. Quantifying this in a different way, a feedback
722 parameter of ~ -2.0 W m⁻² K⁻¹ suggests an EffCS = $-F_{2\times}/\lambda_{\text{hist}}$ as low as $\sim 4.0/2.0 \sim 2.0$ K operating over
723 1985-2014, assuming $F_{2\times} = 4.0$ W m⁻² (Sherwood et al. 2020). From this it seems possible that the
724 rate of global warming over this period (~ 0.19 K dec⁻¹, Tokarska et al., 2020) might have been larger
725 had the Earth warmed over this period with a pattern of SST associated with more positive
726 feedbacks, as found in earlier parts of the historical record (Section 3). However, we also investigate
727 the possibility that changes in ocean heat uptake efficiency may have compensated the changes in
728 feedbacks and low EffCS to maintain a higher warming rate over this period than would be expected
729 without this compensation.

730 To do this we turn to the 'climate resistance' (ρ , units W m⁻² K⁻¹) "zero-layer" model of Gregory and
731 Forster (2008) to analyse the ocean heat uptake efficiency (κ , units W m⁻² K⁻¹). This is expressed as
732 $dF = \rho dT$, where $\rho = \kappa - \lambda$, and κ is defined as $\kappa = dN/dT$ and is found to be strongly related to the
733 thermal coupling constant (γ , units W m⁻² K⁻¹) between the upper and lower ocean in the two-layer
734 model (Gregory et al. 2015; see their Figure 8). While initially proposed to describe scenarios with
735 steadily increasing forcing, it is also been applied to ~ 30 year timescales to usefully describe or
736 interpret the energy balance (Gregory and Forster, 2008; Watanabe et al., 2013). Despite being a

737 gross simplification of the climate system (we discuss potential limitations below), $dF = \rho dT$ is found
738 to be an excellent approximation ($r=0.86$) over 1985 – 2014 (excluding the 1991-2 Pinatubo years,
739 see below) in our data (Figure 7c). From this relationship we deduce $\rho = dF/dT \sim 2.4 \pm 0.5 \text{ W m}^{-2} \text{ K}^{-1}$
740 over 1985-2014 (Figure 7c) and similarly $\kappa = dN/dT \sim 0.4 \pm 0.8 \text{ W m}^{-2} \text{ K}^{-1}$. In contrast, AOGCM
741 simulations of steady increasing CO_2 generally have a larger ocean heat uptake efficiency ($\kappa = 0.73 \pm$
742 $0.18 \text{ W m}^{-2} \text{ K}^{-1}$ for years 61-80 of CMIP5 1% CO_2 AOGCM simulations, Gregory et al., 2015).

743 Another effect on surface temperature to consider is the possibility that the pattern of surface
744 warming and/or atmospheric circulation may change the efficiency of global heat uptake (and vice
745 versa), thus not only is λ inconstant, but κ may also vary. Using passive ocean
746 uptake experiments wherein ocean circulation cannot change, Newsom et al. (2020) find that ocean
747 heat uptake efficiency can be expected to be smaller when warming is enhanced in the tropics
748 (where deep ocean ventilation is small) and larger when warming is enhanced in the high latitudes
749 (where deep ocean ventilation is large). With relatively small warming in the southern high latitudes,
750 this suggests that the surface/ocean-mixed layer might have been less efficient at fluxing heat into
751 the deep ocean over the same period as the large pattern effect, potentially enhancing global
752 surface warming and muting some of the impact of feedback changes. However, stronger trade
753 winds, as have been observed over 1981-2010, can also be expected to accelerate subtropical cells,
754 enhancing ocean heat uptake efficiency and slowing global surface warming (England et al. 2014),
755 an effect not accounted for in the passive ocean heat uptake experiments of Newsom et al. (2020).
756 Thus, variations in both radiative feedbacks and ocean heat uptake appear to be physically
757 linked through SST patterns and may even to some extent co-vary (Newsom et al. 2020).

758 As our dN timeseries does not predate 1985 we cannot investigate whether κ has varied in a way
759 that would counter changes in λ_{hist} prior to 1985. Instead, we go forward in time exploiting the
760 datasets up to and including 2019. This includes the major El-Nino event of 2015/2016 and marked
761 changes in the observed radiation budget (Loeb et al. 2020; 2021). Figure 9 illustrates the impact of
762 this event on the pattern of decadal surface warming. Over 1985-2014 there is marked cooling over
763 the eastern Pacific (Figure 9a) which is much reduced when the pattern is calculated over 1987-2016
764 (Figure 9b) to include the peak 2015-16 El-Nino years. The difference (Figure 9c) shows the warming
765 event of the 2015-16 El-Nino on the eastern Pacific, while cooling in the western Pacific, as well as a
766 slight reduction in Southern Ocean cooling. This is precisely the pattern of SST change we'd expect to
767 have an impact on λ .

768 Table 4 shows the impact on 30-year derived ρ , λ and κ values moving forward in time from 2014, up
769 to and including 1990-2019. Figure 7 (red crosses) shows these additional 5 years in comparison to
770 the 1985-2014 ρ and λ relationships. Post 2014, λ reduces in magnitude (Table 4) and all the red
771 crosses fall below the 1985-2014 λ relationship in Figure 7d. λ is approximately 25% smaller in
772 magnitude over 1990-2019 compared to 1985-2014 (Table 4). This is consistent with process based
773 arguments that a shift to eastern Pacific warming post 2014 ought to drive more positive feedbacks
774 and consequently a reduction of the pattern effect over these years. It is also consistent with Loeb
775 et al. (2020) who performed a similar analysis but over 2001-2014 compared to 2001-2017. They
776 also showed that AGCMs were able to capture this change in radiative response. It would be useful
777 for future analysis if *amip-piForcing* type simulations were extended to at least 2019 to capture the
778 largest change in λ (Table 4), and ideally right up to the most recent SST and sea-ice data available.

779 In contrast to λ , ρ is relatively stable to these additional years (Table 4) and the 1985-2014 ρ
780 relationship is found to be an excellent predictor for 2015-2019 (red crosses fall on or close to the
781 line, Figure 7c). A consequence of ρ being well approximated as constant but λ not, is that κ (equal to
782 $\rho + \lambda$) must compensate for the change in λ . Thus beyond 2014, the pattern effect declines but its

783 impact on surface temperature is buffered by a change in ocean heat uptake efficiency. This is
784 consistent with the original hypothesis that variations in SST patterns affect both heat loss to space
785 (radiative feedbacks) and the efficiency of heat uptake into the deep-ocean in a way that might co-
786 vary (Newsom et al., 2020). However, the extent of any anti-correlation is unclear, it may simply
787 apply to short term variability. It clearly does not apply to longer-term forced changes, given that
788 Gregory et al. (2015) found substantial variations in ρ , which would not occur if κ and λ were
789 strongly anti-correlated.

790 While the zero-layer model appears to work well on this short timescale (Figure 7c) we caution
791 against assuming all changes in ocean heat content are driven by global T , as assumed by the $dN =$
792 κdT relationship. This is because, especially on short timescales, other influences that do not
793 correlate with global T , such as wind-driven ocean circulation changes perhaps, will also alter ocean
794 heat content (England et al., 2014). In such a situation, it would be reasonable to write $N = \kappa T + U$
795 where U is an additional term to the heat balance, not related to global T . This implies $\kappa = N/T - U/T$,
796 and including this term in the forced heat balance, $N = F + \lambda T + U$, gives $\lambda = (N-F)/T - U/T$. Thus U/T
797 would perturb the estimate of κ (a positive number) and λ (a negative number) in opposite
798 directions, as we see in our data. Hence our results are potentially evidence for variation in ocean
799 heat content not driven by global T , but we cannot say exactly what it is – other than it does not
800 scale with global T .

801 We caution that structural errors could impact on our diagnosis. Specifically, both κ and λ are related
802 to dN and so any bias or error in the observed dN trend would bias κ and λ in opposite directions.
803 Moreover $\rho = dF/dT$ would be unaffected by any bias or error in dN , and so the anti-correlation would
804 compensate to leave $\rho = \kappa - \lambda$ unaffected. We illustrate this in Table 4, which shows these quantities
805 calculated over 1985-2014 using 5 available different versions of the DEEP-C dN datasets (see
806 Section 2.4). Differences in the results emerge (λ reduces in magnitude from $\sim -2.2 \text{ W m}^{-2} \text{ K}^{-1}$ to ~ -2.0
807 $\text{W m}^{-2} \text{ K}^{-1}$, with a compensating increase in κ) as the DEEP-C datasets transition from v3 to v4 (i.e. v2
808 and v3 give the same results, as do v4 and v5), highlighting the impact of potential structural errors
809 in these results. We do not pursue the cause of the difference in the results, but it is likely due to
810 changes between v3 and v4 in how the DEEP-C method bridges the gap between satellite products in
811 the 1990s (a longer adjustment period and a different modelling ensemble is used) (Liu et al., 2020).
812 However it is also important to note that the observational record since 2000, applying the CERES
813 dataset, is subject to much smaller structural uncertainty than the earlier record implying a greater
814 confidence in our analysis of the anomalous N variations post 2014.

815 *4.3 Effect of the Pinatubo volcanic eruption*

816 Finally, we comment on the effect of the Pinatubo volcanic eruption on these results. There is a large
817 negative spike in dF and dN around 1991 and 1992 (Figure 7b). While we found no impact of these
818 years on our estimate of 1985 – 2014 λ_{hist} , they have a strong impact on ρ and κ . Including these
819 years in the regression analysis, we find $\rho = dF/dT \sim 2.9 \pm 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ and $\kappa = dN/dT \sim 0.8 \pm 0.9 \text{ W}$
820 $\text{m}^{-2} \text{ K}^{-1}$, much larger than when these years are excluded from the analysis as above. This is
821 consistent with Gregory et al. (2015) who found the ‘transient climate response parameter’ (equal to
822 $1/\rho$, units $\text{K W}^{-1} \text{ m}^2$) to explosive eruptions to be smaller (ρ larger) than that evaluated in AOGCMs
823 under steadily increasing CO_2 , principally because the surface/mixed-layer readily gives up heat (κ
824 larger) in response to a short-lived forcing like an explosive volcanic eruption. Hence if the time-
825 period under consideration contains large volcanic eruptions then the “zero-layer” model ($dF = \rho dT$)
826 is found to be a poor approximation (i.e. ρ not constant) over the entire time-period because it
827 neglects the importance of the upper-ocean heat capacity on short timescales (Gregory and Forster,

828 2008; Held et al. 2010; Gregory et al., 2016). This manifests itself as a sensitivity of ρ and κ to the
829 inclusion or exclusion of volcanic years, as we have found here.

830

831 **5. Summary, Discussion and Conclusions**

832 *5.1 Historical feedbacks and the pattern effect*

833 The dependence of radiative feedback on the pattern of SST change was investigated in fourteen
834 Atmospheric General Circulation Models (AGCMs) forced with observed variations in sea-surface-
835 temperature (SST) and sea-ice over the historical record from 1871 to near-present (*amip-piForcing*
836 experiment). We found that the pattern effect identified in a previous model intercomparison
837 (Andrews et al, 2018) is largely robust to a wider set of new generation AGCMs with a broader range
838 of atmospheric physics and climate sensitivities. Our qualitative conclusions were not strongly
839 dependent on the AMIP II SST dataset used to force the AGCMs; indeed, the feedbacks in nine
840 AGCMs using SSTs from HadISST1 (*hadSST-piForcing*) were found to be strongly correlated with
841 feedbacks in *amip-piForcing*, though the magnitude of the pattern effect post 1980 was found to be
842 smaller under HadISST1 SSTs (see also Andrews et al., 2018; Lewis and Mauritsen, 2021; Zhou et al.,
843 2021; Fueglistaler and Silvers, 2021).

844 Separating the historical record at 1980, we found that over 1871-1980 the Earth warmed with a
845 relatively uniform warming pattern and feedbacks largely consistent and strongly correlated with
846 long-term *abrupt-4xCO2* feedbacks (i.e. with relatively small pattern effect - Figures 2 and 5). In
847 contrast, post 1980 the Earth warmed with a strong tropical Pacific SST gradient (Figure 4) where
848 regions of deep convection warm substantially more than the tropical mean (Fueglistaler and Silvers,
849 2021). This drove large negative feedbacks and pattern effects in both our *amip-piForcing* and
850 *hadSST-piForcing* simulations, consistent with the physical understanding of how lapse-rate and
851 cloud feedbacks depend on tropical Pacific SST patterns (Zhou et al., 2016; Andrews and Webb,
852 2018; Ceppi and Gregory, 2017; Dong et al., 2019).

853 As well as a large pattern effect, feedbacks post 1980 were found to be uncorrelated with long term
854 CO₂ driven feedbacks (Figure 5). This is unfortunate, because the feedback inferred from this period
855 therefore does not constrain the CO₂ feedback or ECS. It is also surprising, because the period since
856 ~1980 contains a well observed large global temperature response, which AOGCMs attribute to
857 increasing greenhouse gases, and it avoids the aerosol forcing uncertainty issue which is small in
858 energy budget estimates of ECS over this period (at least in the global-mean; regional aerosol forcing
859 could still impact on SST patterns and feedbacks) (Jiménez-de-la-Cuesta and Mauritsen, 2019).
860 Despite this, it turns out to be the worst period for inferring the Earth's long-term CO₂ climate
861 sensitivity from the observed global energy balance. Conversely, feedbacks acting earlier in the
862 record (1871-1980) are representative of the long-term response (i.e. smaller pattern effect) and do
863 correlate with λ_{4xCO_2} across models, yet this period has a smaller climate change signal and is not as
864 well observed, containing much larger uncertainties relative to the climate change signal (e.g. Otto
865 et al., 2013), as well as a large forcing uncertainty. Hence the usefulness of this time-period is limited
866 for setting a constraint on λ_{hist} .

867 Considering the historical record as a whole is useful for informing studies that use the entire
868 observed record to estimate ECS via energy budget constraints (e.g. Sherwood et al. 2020). We
869 found that the pattern effect over 1871-2010 to be $\Delta\lambda = 0.70 \pm 0.47 \text{ W m}^{-2} \text{ K}^{-1}$ in our *amip-piForcing*
870 ensemble and $\Delta\lambda = 0.48 \pm 0.36 \text{ W m}^{-2} \text{ K}^{-1}$ in *hadSST-piForcing*, where the smaller uncertainty in
871 *hadSST-piForcing* likely reflects the narrower set of model physics in this smaller ensemble (for

872 example we do not have *hadSST-piForcing* experiments for the model (MIROC6) with the smallest
873 pattern effect in *amip-piForcing*). The question therefore arises as to which of these estimates ought
874 to be used for adjusting historical energy budget constraints on ECS for pattern effects.

875 Both Lewis and Mauritsen (2021) and Fueglistaler and Silvers (2021) showed that the AMIP II dataset
876 had the largest warm pool trends relative to the tropical-mean of all SST reconstructions they
877 considered. Hence one interpretation of our results is that the pattern effect in *amip-piForcing* might
878 usefully be regarded as an upper bound on the structural uncertainty of the experimental design to
879 observational uncertainty in SST reconstructions. A best estimate might place more weight on the
880 *hadSST-piForcing* pattern effects, which have warm pool trends (relative to the tropical-mean) closer
881 to the middle of the range of SST reconstructions (Fueglistaler and Silvers, 2021; Lewis and
882 Mauritsen, 2021). In that case, a best estimate of the historical pattern effect could be 0.48 ± 0.47
883 $\text{W m}^{-2} \text{K}^{-1}$ for the time-period 1871-2010, which represents the pattern effect from *hadSST-piForcing*
884 but retaining the larger uncertainty from the (larger ensemble) *amip-piForcing* results. If calculated
885 over 1871-2014 the pattern effect increases by $0.05 \pm 0.05 \text{ W m}^{-2} \text{K}^{-1}$ according to the *hadSST-*
886 *piForcing* ensemble. This best estimate of the historical pattern effect is close to that used in
887 Sherwood et al. (2020), who assumed a value of $0.5 \pm 0.5 \text{ W m}^{-2} \text{K}^{-1}$ (they were informed by Andrews
888 et al. (2018) who used *amip-piForcing* but allowed for a potentially smaller pattern effect than that
889 study based on expert judgement). On the other hand, just because the AMIP II SST trends are at
890 one end of the range of SST reconstructions does not necessarily mean they are more erroneous.
891 Indeed, Zhou et al. (2021) showed that TOA radiative fluxes simulated by CAM5.3 correlated better
892 with CERES observations when forced with AMIP II SSTs rather than HadISST SSTs, suggesting the
893 results from *amip-piForcing* may be more reliable. In this case, the 1871-2010 pattern effect is $0.70 \pm$
894 $0.47 \text{ W m}^{-2} \text{K}^{-1}$. In the future, a model intercomparison of the pattern effect to a broader range of
895 SST reconstructions would be useful to address any outstanding structural uncertainty to SST
896 reconstructions.

897 To provide independent evidence for the historical pattern effect, we used IPCC AR6 assessed
898 changes in T , N and F between 1850-1900 to 2006-2019 (Forster et al. 2021) to estimate a historical
899 feedback parameter of $\lambda_{\text{hist}} = (\Delta N - \Delta F)/\Delta T = -1.6 \pm 0.8 \text{ W m}^{-2} \text{K}^{-1}$. This was found to be in agreement
900 with the *amip-piForcing* and *hadSST-piForcing* ensembles. IPCC AR6 also assessed the long-term ECS
901 relevant feedback parameter ($-1.16 \pm 0.65 \text{ W m}^{-2} \text{K}^{-1}$, Forster et al., 2021) from combining lines of
902 evidence from observations, theory, process models and GCMs on individual climate feedback
903 processes. Contrasting this with the λ_{hist} estimate above gives an estimate of the pattern effect of 0.4
904 $\pm 1.1 \text{ W m}^{-2} \text{K}^{-1}$ for historical changes between 1850-1900 to 2006-2019. While the uncertainties are
905 substantial, this is in agreement with our GCM based estimate of the historical pattern effect.

906 5.2 Observed climate change since 1985 and ocean heat uptake efficiency

907 Satellite based reconstructions of the Earth's energy balance over 1985 to 2014 suggest a feedback
908 parameter of $\sim -2.0 \pm 0.7 \text{ W m}^{-2} \text{K}^{-1}$, in agreement with our *amip-piForcing* and *hadSST-piForcing*
909 ensembles. Evidence is also emerging from satellite records in support of the physical processes and
910 mechanisms of the pattern effect between surface temperature, atmospheric stability, cloudiness
911 and radiative fluxes over recent decades (e.g. Zhou et al., 2016; Ceppi and Gregory, 2017; Loeb et al.,
912 2020; Fueglistaler and Silvers, 2021; Ceppi and Fueglistaler, 2021).

913 Extending our analysis post 2014 included the major El-Nino event of 2015/2016 that was associated
914 with eastern-pacific warming and marked changes in the observed radiation budget (Loeb et al.
915 2020; 2021). Including these post 2014 years (up to and including 2019) reduced the magnitude of
916 the observed λ estimate by up to $\sim 25\%$, consistent with eastern Pacific warming driving more

917 positive feedbacks (as also suggested in Loeb et al., 2020). This suggests the pattern effect that has
918 existed over recent decades may be waning if a shift from western to eastern Pacific warming is
919 maintained in the longer term, as might be expected from a change in the PDO index identified by
920 Loeb et al. (2021).

921 Given the substantial rate of global warming since 1985, what does the presence of a large pattern
922 effect imply for ocean heat uptake efficiency (κ)? We estimated $\kappa = dN/dT \sim 0.4 \pm 0.8 \text{ W m}^{-2} \text{ K}^{-1}$ over
923 1985-2014, which is smaller (but not necessarily inconsistent) with AOGCM simulations of steady
924 increasing CO_2 ($\kappa = 0.73 \pm 0.18 \text{ W m}^{-2} \text{ K}^{-1}$ for years 61-80 of CMIP5 1% CO_2 AOGCM simulations,
925 Gregory et al. 2015). It raises the possibility that the pattern of surface warming and/or atmospheric
926 circulation may also change the efficiency of global heat uptake, thus both λ and κ might vary and to
927 some extent be related (Newsom et al., 2020). If an anti-correlation existed, it could buffer the
928 impact of a large pattern-effect on transient climate change.

929 We found that despite the change in radiative feedback post 2014 when the eastern Pacific warmed,
930 the climate resistance $\rho = dF/dT = \kappa - \lambda$ remained approximately constant, suggesting that κ and λ
931 co-varied. We showed that this result is potential evidence for a change in ocean heat content not
932 driven by global T . While this result is suggestive, the extent of this compensation and timescales it
933 applies to remains unclear. It may simply apply to short term variability and clearly does not apply to
934 longer-term forced changes (e.g. Gregory et al., 2015). Future research investigating how ocean
935 uptake and atmospheric radiative feedbacks are linked through patterns of SST change would be
936 useful.

937 *5.3 Outlook and Implications for AOGCMs*

938 Our results raise important questions for studies that have used emergent relationships from
939 AOGCMs to constrain ECS from recently observed decadal warming since ~ 1980 (e.g. Jiménez-de-la-
940 Cuesta and Mauritsen, 2019; Tokarska et al., 2020; Nijse et al., 2020).

941 Firstly, how is it possible that AOGCMs produce an emergent relationship between their recent
942 decadal warming trends and their ECS, while our results suggest that recent decadal feedbacks
943 ought to be unrelated to ECS? One solution to this conundrum is provided by Fueglistaler and Silvers
944 (2021), who showed that AOGCMs typically do not simulate the recent configuration of tropical
945 Pacific SST patterns that gave rise to the recent pattern effect (though some models do have broad
946 agreements, e.g. Olonscheck et al. 2021, Watanabe et al. 2021). Instead, the pattern of warming in
947 AOGCMs (and thus feedbacks) over recent decades is more similar to that seen in their *abrupt-*
948 *4xCO2* simulations (Gregory et al., 2020; Dong et al. 2021). Hence AOGCMs are generally biased in
949 their simulation of the recent decadal feedbacks and the pattern effect, compared to their
950 equivalent AGCMs forced with observed SST variations, as shown in Gregory et al. (2020) and Dong
951 et al. (2021).

952 If AOGCMs are biased in their simulation of recent decadal feedbacks and the pattern effect, it
953 suggests they may be biased toward simulating recent decadal temperature trends that are too high;
954 in turn, this would bias emergent constraints that use them toward values of ECS that are too low.
955 Alternatively, those models that do match the observed warming trend may do so via a
956 compensation of processes: too small a pattern effect balanced against too large a heat uptake into
957 the deep-ocean. Some evidence for the potential of this compensating behaviour is provided by
958 Hedemenn et al. (2017). Analysing the origins of decadal temperature variability in models, they
959 demonstrated an anti-correlation between the TOA radiative flux and deep-ocean (defined as below
960 100m) flux contributions to the model's surface layer and decadal temperature trends (see their

961 Figure 3). In other words, when the TOA radiative flux is in such a configuration to reduce its
962 contribution to the surface layer, then the surface/mixed-layer taps into the deep-ocean to
963 compensate for this loss, and vice versa. We speculate that such a configuration of TOA radiative flux
964 is potentially consistent with a large negative feedback, since in this configuration of atmospheric
965 feedbacks the surface efficiently radiates heat back to space. This again suggests a potential anti-
966 correlation between the ocean heat uptake efficiency and λ during unforced decadal variability
967 timescales as discussed previously.

968 Going forward, a critical question for future research is to understand what caused the particular
969 configuration of SST patterns over recent decades (e.g. strong warming in the western Pacific while
970 cooling in the eastern Pacific and Southern Ocean, despite temperature increasing in the global-
971 mean; Figure 4 and 9), and how might this pattern evolve in the future. For example, various
972 hypotheses have been put forward:

- 973 1. It could represent a mode of unforced coupled atmosphere-ocean variability (e.g.
974 Xie et al., 2016; Watanabe et al. 2021), albeit an unusual one is that is rarely
975 simulated by AOGCMs (Fueglistaler and Silvers, 2021). In this scenario, we might
976 expect the pattern effect to reduce in the near-future as the configuration of
977 tropical SST patterns shift to more warming in the east than the west. There is some
978 evidence (Loeb et al. 2020; 2021) this has already begun to happen in the most
979 recent years, as we have also shown. We might therefore expect an acceleration of
980 warming trends, unless the additional heat at the surface from the reduced pattern
981 effect is tempered by compensating heat exchanges with the deep-ocean
982 (Hedemann et al. 2017).
- 983 2. Spatiotemporal variations in anthropogenic forcings such as aerosols (e.g., Smith et
984 al., 2015; Takahashi & Watanabe, 2016; Moseid et al., 2020; Heede and Fedorov,
985 2021) or explosive volcanic eruptions (Smith et al. 2015; Gregory et al. 2020) have
986 been implicated in driving tropical Pacific SST patterns. In these scenarios, the
987 pattern effect may decline with the reduction in aerosol emissions in the future, or
988 continue to have decadal variations associated with future volcanism. Whether
989 changes in deep-ocean fluxes will be accompanied with such forced changes in the
990 pattern effect is unclear.
- 991 3. While not explaining the eastern Pacific cooling per se, a delayed warming in the
992 eastern Pacific relative to the west is an expected transient response to forcing due
993 to the upwelling of (as yet) unperturbed waters from below (Clement et al., 1993;
994 Held et al. 2010; Heede and Fedorov, 2021). The implication of this is that
995 eventually the eastern Pacific will warm, and hence we might expect the pattern
996 effect to reduce and the Earth to warm with stronger (positive) cloud feedbacks
997 (e.g. Dessler, 2020).
- 998 4. In contrast, AOGCMs may overstate the expected warming in the eastern Pacific
999 (e.g. Seager et al., 2020). Under this scenario, we might expect the pattern effect to
1000 reduce after the eastern Pacific stops cooling, but the full pattern effect according
1001 to AOGCMs may never materialise if they incorrectly simulate a strong 'ENSO-like'
1002 pattern in their long-term response to CO₂. However, a lack of eastern Pacific
1003 warming in the long-term seems unlikely according to paleoclimate records (Tierney
1004 et al. 2019; 2020).
- 1005 5. Teleconnections from either the Atlantic Ocean (McGregor et al. 2018) or Southern
1006 Ocean (Hwang et al. 2017) have potentially driven the tropical Pacific SST patterns.
1007 Under the scenario of an Atlantic influence, we might expect the pattern effect to

1008 reduce as Atlantic SST trends evolve over the next few decades. Under the scenario
1009 of a Southern Ocean influence, we might expect the pattern effect to reduce as the
1010 Southern Ocean surface warms; this could take years to decades if the Southern
1011 Ocean temperature trends have been largely mediated by internal variability (e.g.,
1012 Zhang et al. 2019) but could take centuries or longer if Southern Ocean cooling
1013 continues due, for instance, to freshwater input from ongoing Antarctic ice shelf
1014 melt (e.g., Sadai et al. 2020).

1015 These are merely some of the proposed hypotheses, and not meant to be an exhaustive list. But
1016 whatever the reason, the fact that AOGCMs rarely simulate this pattern (e.g. Watanbe et al., 2021;
1017 Fueglistaler and Silvers, 2021; Dong et al., 2021) is a concern, suggesting either that their unforced
1018 decadal variability is deficient, or that their forced response is biased, and in either case there is a
1019 serious systematic error which affects all AOGCMs. Moreover, each of the above interpretations
1020 imply different futures, and therefore untangling them is critical for informing both near-term and
1021 long-term climate projections. This is time critical because satellite evidence suggests the Pacific SST
1022 pattern that has dominated recent decades is currently shifting (Loeb et al., 2020) and indeed the
1023 Earth's energy balance is rapidly changing with it (Loeb et al. 2021; Raghuraman et al., 2021).
1024 Predicting the near future therefore depends on maintaining the continuity of the satellite record
1025 and untangling the above mechanisms.

1026 **Acknowledgements**

1027 TA and JMG thank Richard Wood and Mark Ringer for useful discussions. We thank Bosong Zhang,
1028 Zhihong Tan and Knut von Salzen for useful comments on an early draft version of the manuscript.
1029 We thank Maria Rugenstein and an anonymous reviewer for positive and constructive comments.
1030 TA, JMG and ABS were supported by the Met Office Hadley Centre Climate Programme funded by
1031 BEIS. TA, TM, AM and RR have received funding from the European Union’s Horizon 2020 research
1032 and innovation programme under grant agreement 820829. JMG’s work is also supported by the
1033 European Research Council (ERC, grant agreement No 786427, project “Couplet”). TM and AM
1034 received funding from the European Research Council grant 770765. The contribution of JB to this
1035 work was funded by the U. S. Department of Energy’s Regional and Global Modeling Analysis
1036 program area. TO was supported by the Integrated Research Program for Advancing Climate Models
1037 (TOUGOU) Grant Number JPMXD0717935457 from the Ministry of Education, Culture, Sports,
1038 Science and Technology (MEXT), Japan. CL was supported by the National Natural Science
1039 Foundation of China (42075036). KCA and YD acknowledge support from the National Science
1040 Foundation (Grant AGS-1752796) and from the National Oceanic and Atmospheric Administration
1041 MAPP Program (Award NA20OAR4310391). BM acknowledges support by the U.S. Department of
1042 Energy under Award Number DE-SC0022070 and National Science Foundation (NSF) IA 1947282; the
1043 National Center for Atmospheric Research, which is a major facility sponsored by the NSF under
1044 Cooperative Agreement No. 1852977; and the National Oceanic and Atmospheric Administration
1045 under award NA20OAR4310392. We acknowledge the World Climate Research Programme, which,
1046 through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the
1047 climate modeling groups for producing and making available their model output, the Earth System
1048 Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding
1049 agencies who support CMIP6 and ESGF.

1050

1051 **Data Availability**

1052 Global-annual-ensemble-mean dT and dN data from all *amip-piForcing*, *hadSST-piForcing* and
1053 *abrupt-4xCO2* simulations used in this study are provided at
1054 <https://doi.org/10.5281/zenodo.6799004> (Andrews et al. 2022). Raw data from CMIP6 *amip-*
1055 *piForcing* simulations (indicated in Table 1) are available at <https://pcmdi.llnl.gov/CMIP6/> (Eyring et
1056 al., 2016). *abrupt-4xCO2* raw data for most models is available at CMIP5 ([https://esgf-
node.llnl.gov/projects/cmip5/](https://esgf-
node.llnl.gov/projects/cmip5/)) (Taylor et al., 2012) or CMIP6 (<https://pcmdi.llnl.gov/CMIP6/>) (Eyring
1057 et al., 2016). The HadCRUT5 analysis dataset is available at
1058 <https://www.metoffice.gov.uk/hadobs/hadcrut5/> (Morice et al., 2021). IPCC AR6 ERF timeseries is
1059 available at <https://doi.org/10.5281/zenodo.5211358> (Smith et al., 2021). DEEP-C v5 dN radiative
1060 fluxes can be obtained from <https://doi.org/10.17864/1947.000347> (Lui and Allan, 2022) and
1061 previous versions described at <http://www.met.reading.ac.uk/~sgs02rpa/research/DEEP-C/GRL/>.
1062 The HadISST1 SSTs used to force the *hadSST-piForcing* simulations are available at
1063 <https://www.metoffice.gov.uk/hadobs/hadisst/> (Rayner et al. 2003).
1064

1065 **References**

- 1066 Allan, R. P., Liu, C., Loeb, N. G., Palmer, M. D., Roberts, M., Smith, D., and Vidale, P.-L. (2014),
1067 Changes in global net radiative imbalance 1985–2012, *Geophys. Res. Lett.*, 41, 5588– 5597,
1068 doi:[10.1002/2014GL060962](https://doi.org/10.1002/2014GL060962).
- 1069 Andrews, Timothy, Bodas-Salcedo, Alejandro, Gregory, Jonathan, Dong, Yue, Armour, Kyle, Paynter,
1070 David, Lin, Pu, Modak, Angshuman, Mauritsen, Thorsten, Cole, Jason, Medeiros, Brian, Benedict,
1071 James, Douville, Herve, Roehrig, Romain, Koshiro, Tsuyoshi, Kawai, Hideaki, Ogura, Tomoo,
1072 Dufresne, Jean-Loius, Allan, Richard, & Liu, Chunlei. (2022). amip-piForcing and HadSST-piForcing
1073 simulation data (v1.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.6799004>
- 1074 Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour,
1075 K. C., Forster, P. M., & Titchner, H. (2018). Accounting for changing temperature patterns increases
1076 historical estimates of climate sensitivity. *Geophysical Research Letters*, 45, 8490– 8499.
1077 <https://doi.org/10.1029/2018GL078887>
- 1078 Andrews, T., Gregory, J. M., & Webb, M. J. (2015). The dependence of radiative forcing and feedback
1079 on evolving patterns of surface temperature change in climate models. *Journal of Climate*, 28, 1630–
1080 1648.
- 1081 Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate
1082 sensitivity in CMIP5 coupled atmosphere–ocean climate models. *Geophysical Research Letters*, 39,
1083 L09712. <https://doi.org/10.1029/2012GL051607>
- 1084 Andrews, T., & Ringer, M. A. (2014). Cloud feedbacks, rapid adjustments, and the forcing–response
1085 relationship in a transient CO₂ reversibility scenario. *Journal of Climate*, 27(4), 1799– 1818.
1086 <https://doi.org/10.1175/JCLI-D-13-00421.1>
- 1087 Andrews, T., & Webb, M. J. (2018). The dependence of global cloud and lapse rate feedbacks on the
1088 spatial structure of tropical Pacific warming. *Journal of Climate*, 31. <https://doi.org/10.1175/JCLI-D-17-0087.1>
- 1090 Annan, J. D., & Hargreaves, J. C. (2017). On the meaning of independence in climate science. *Earth*
1091 *System Dynamics*, 8, 221– 224.
- 1092 Annan, J. D., & Hargreaves, J. C. (2011). Understanding the CMIP3 Multimodel Ensemble, *Journal of*
1093 *Climate*, 24(16), 4529–4538. <https://doi.org/10.1175/2011JCLI3873.1>.
- 1094 Armour, K. C. (2017). Energy budget constraints on climate sensitivity in light of inconstant climate
1095 feedbacks. *Natural Climate Change*, 7, 331– 335. <https://doi.org/10.1038/nclimate3278>
- 1096 Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-varying climate sensitivity from regional
1097 feedbacks. *Journal of Climate*, 26, 4518– 4534.
- 1098 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., & Bastrikov, V., et al. (2020).
1099 Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling*
1100 *Earth Systems*, 12, e2019MS002010. <https://doi.org/10.1029/2019MS002010>
- 1101 Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., & Gregory, J. M.
1102 (2021). Climate sensitivity increases under higher CO₂ levels due to feedback temperature
1103 dependence. *Geophysical Research Letters*, 48, e2020GL089074.
1104 <https://doi.org/10.1029/2020GL089074>

- 1105 Block, K., and Mauritsen, T. (2013), Forcing and feedback in the MPI-ESM-LR coupled model under
1106 abruptly quadrupled CO₂, *J. Adv. Model. Earth Syst.*, 5, 676– 691, doi:[10.1002/jame.20041](https://doi.org/10.1002/jame.20041).
- 1107 Bjordal, J., Storelvmo, T., Alterskjær, K. *et al.*, 2022: Equilibrium climate sensitivity above 5 °C
1108 plausible due to state-dependent cloud feedback. *Nat. Geosci.* **13**, 718–721 (2020).doi,
1109 [10.1038/s41561-020-00649-1](https://doi.org/10.1038/s41561-020-00649-1).
- 1110 Caballero, R., & Huber, M. (2013). State-dependent climate sensitivity in past warm climates and its
1111 implications for future climate projections. *Proceedings of the National Academy of Sciences of the*
1112 *United States of America*, 110, 14,162– 14,167. <https://doi.org/10.1073/pnas.1303365110>
- 1113 Ceppi, P., & Fueglistaler, S. (2021). The El Niño–Southern Oscillation pattern effect. *Geophysical*
1114 *Research Letters*, 48, e2021GL095261. <https://doi.org/10.1029/2021GL095261>.
- 1115 Ceppi, P., & Gregory, J. M. (2017). Relationship of tropospheric stability to climate sensitivity and
1116 Earth's observed radiation budget. *Proceedings of the National Academy of Sciences of the United*
1117 *States of America*, 114, 13,126– 13,131. <https://doi.org/10.1073/pnas.1714308114>
- 1118 Chalmers, J., Kay, J. E., Middlemas, E. A., Maroon, E. A., & DiNezio, P., 2022: Does disabling cloud
1119 radiative feedbacks change spatial patterns of surface greenhouse warming and cooling?, *Journal of*
1120 *Climate* (published online ahead of print 2022;
1121 <https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-21-0391.1>)
- 1122 Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and
1123 irreversibility. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge
1124 University Press, 1029–1136.
- 1125 Clement, A. C., Seager, R., Cane, M. A., & Zebiak, S. E. (1996). An ocean dynamical thermostat.
1126 *Journal of Climate*, 9, 2190– 2196. [https://doi.org/10.1175/1520-0442\(1996\)009<2,190:AODT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2,190:AODT>2.0.CO;2)
- 1128 Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al.
1129 (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling*
1130 *Earth Systems*, 12, e2019MS001916. <https://doi.org/10.1029/2019MS001916>
- 1131 Dessler, A. E. (2020). Potential Problems Measuring Climate Sensitivity from the Historical Record,
1132 *Journal of Climate*, 33(6), 2237–2248.
- 1133 Dong, Y., Armour, K. C., Proistosescu, C., Andrews, T., Battisti, D. S., Forster, P. M., et al. (2021).
1134 Biased estimates of equilibrium climate sensitivity and transient climate response derived from
1135 historical CMIP6 simulations. *Geophysical Research Letters*, 48, e2021GL095778.
1136 <https://doi.org/10.1029/2021GL095778>
- 1137 Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., & Andrews, T. (2020).
1138 Inter-model spread in the sea-surface temperature pattern effect and its contribution to climate
1139 sensitivity in CMIP5 and CMIP6 models. *Journal of Climate*, 33(18), 7755– 7775.
1140 <https://doi.org/10.1175/JCLI-D-19-1011.1>
- 1141 Dong, Y., Proistosescu, C., Armour, K. C., & Battisti, D. S. (2019). Attributing historical and future
1142 evolution of radiative feedbacks to regional warming patterns using a Green's Function approach:
1143 The preeminence of the western Pacific. *Journal of Climate*, 32(17), 5471– 5491.
1144 <https://doi.org/10.1175/JCLI-D-18-0843.1>

- 1145 Donner, L. et al.. (2011). The Dynamical Core, Physical Parameterizations, and Basic Simulation
 1146 Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3,
 1147 *Journal of Climate*, 24(13), 3484-3519.
- 1148 England, M. H., McGregor, S., Spence, P., Meehl, G. A., Timmermann, A., Cai, W., Gupta, A. S.,
 1149 McPhaden, M. J., Purich, A., & Santoso, A. (2014). Recent intensification of wind-driven circulation in
 1150 the Pacific and the ongoing warming hiatus. *Nature Climate Change*, 4(3), 222– 227.
 1151 <https://doi.org/10.1038/nclimate2106>
- 1152 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016).
 1153 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and
 1154 organizations. *Geoscientific Model Development*, 9, 1937– 1958. [https://doi.org/10.5194/gmd-9-](https://doi.org/10.5194/gmd-9-1937-2016)
 1155 [1937-2016](https://doi.org/10.5194/gmd-9-1937-2016)
- 1156 Fueglistaler, S., & Silvers, L.G. (2021). The peculiar trajectory of global warming. *Journal of*
 1157 *Geophysical Research: Atmospheres*, 126, e2020JD033629. <https://doi.org/10.1029/2020JD033629>
- 1158 Forster, P., T. Storelvmo, K. Armour, W. Collins, J. L. Dufresne, D. Frame, D. J. Lunt, T. Mauritsen, M.
 1159 D.Palmer, M. Watanabe, M. Wild, H. Zhang, 2021, The Earth's Energy Budget, Climate Feedbacks,
 1160 and Climate Sensitivity. In: Climate Change 2021: The Physical Science Basis. Contribution of Working
 1161 Group to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-
 1162 Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I.
 1163 Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R.
 1164 Yu and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY,
 1165 USA, pp. 923–1054, doi:10.1017/9781009157896.009.
- 1166 Flynn, C. M. and Mauritsen, T.: On the climate sensitivity and historical warming evolution in recent
 1167 coupled model ensembles, *Atmos. Chem. Phys.*, 20, 7829–7842, [https://doi.org/10.5194/acp-20-](https://doi.org/10.5194/acp-20-7829-2020)
 1168 [7829-2020](https://doi.org/10.5194/acp-20-7829-2020), 2020.
- 1169 Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., et al. (1999). An
 1170 overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the*
 1171 *American Meteorological Society*, 80(1), 29– 55. [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2)
 1172 [0477\(1999\)080<0029:AOOTRO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2)
- 1173 Geoffroy, O., et al. (2013). Transient climate response in a two-layer energy-balance model. Part II:
 1174 Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs.
 1175 *Journal of Climate*, 26, 1859– 1876.
- 1176 Gregory, J. M., & Andrews, T. (2016). Variation in climate sensitivity and feedback parameters during
 1177 the historical period. *Geophysical Research Letters*, 43, 3911– 3920.
 1178 <https://doi.org/10.1002/2016GL068406>
- 1179 Gregory, J. M., Andrews, T., & Good, P. (2015). The inconstancy of the transient climate response
 1180 parameter under increasing CO₂. *Philosophical Transactions of the Royal Society A*, 373, 140– 417.
 1181 <http://doi.org/10.1098/rsta.2014.0417>
- 1182 Gregory, J. M., Andrews, T., Ceppi, P., Mauritsen, T., & Webb, M. J. (2020). How accurately can the
 1183 climate sensitivity to CO₂ be estimated from historical climate change? *Climate Dynamics*, 54(1–2),
 1184 129– 157. <https://doi.org/10.1007/s00382-019-04991-y>

1185 Gregory, J. M., Stouffer, R. J., Raper, S. C. B., Stott, P. A., & Rayner, N. A. (2002). An observationally
1186 based estimate of the climate sensitivity. *Journal of Climate*, 15(22), 3117– 3121.
1187 [https://doi.org/10.1175/1520-0442\(2002\)015<3117:AOBEOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2)

1188 Gregory, J. M., et al. (2004). A new method for diagnosing radiative forcing and climate sensitivity.
1189 *Geophysical Research Letters*, 31, L03205. <https://doi.org/10.1029/2003GL018747>

1190 Hartmann, D. L., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013:*
1191 *The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 159–254.

1192 Haugstad, A. D., Armour, K. C., Battisti, D. S., and Rose, B. E. J. (2017), Relative roles of surface
1193 temperature and climate forcing patterns in the inconstancy of radiative feedbacks, *Geophys. Res.*
1194 *Letts.*, 44, 7455– 7463, doi:[10.1002/2017GL074372](https://doi.org/10.1002/2017GL074372).

1195 Hansen, J., Sato, M. K. I., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I.,
1196 Bauer, M., Bauer, S. & Bell, N. (2005). Efficacy of climate forcings, mathematical physical and
1197 engineering sciences 365, 1925–54. *Journal of Geophysical Research*, 110, D18104.
1198 <https://doi.org/10.1029/2005JD005776>

1199 Heede, U.K., Fedorov, A.V. Eastern equatorial Pacific warming delayed by aerosols and thermostat
1200 response to CO₂ increase. *Nat. Clim. Chang.* **11**, 696–703 (2021). <https://doi.org/10.1038/s41558-021-01101-x>

1202 Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., et al. (2019). Structure and
1203 performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, 11,
1204 3691– 3727. <https://doi.org/10.1029/2019MS001829>

1205 Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., & Vallis, G. K. (2010). Probing the fast
1206 and slow components of global warming by returning abruptly to preindustrial forcing. *Journal of*
1207 *Climate*, 23(9), 2418– 2427. <https://doi.org/10.1175/2009JCLI3466.1>

1208 Hedemann, C., Mauritsen, T., Jungclaus, J. *et al.* The subtle origins of surface-warming hiatuses.
1209 *Nature Clim Change* **7**, 336–339 (2017). <https://doi.org/10.1038/nclimate3274>

1210 Hurrell, J., Hack, J., Shea, D., Caron, J., & Rosinski, J. (2008). A new sea surface temperature and sea
1211 ice boundary dataset for the community atmosphere model. *Journal of Climate*, 21(19), 5145– 5153.
1212 <https://doi.org/10.1175/2008JCLI2292.1>

1213 Hwang, Y.-T., Xie, S.-P., Deser, C., and Kang, S. M. (2017), Connecting tropical climate change with
1214 Southern Ocean heat uptake, *Geophys. Res. Letts.*, 44, 9449– 9457, doi:[10.1002/2017GL074972](https://doi.org/10.1002/2017GL074972).

1215 Jiménez-de-la-Cuesta, D., & Mauritsen, T. (2019). Emergent constraints on Earth's transient and
1216 equilibrium response to doubled CO₂ from post–1970s global warming. *Nature Geoscience*, 12, 902–
1217 905. <https://doi.org/10.1038/s41561-019-0463-y>

1218 Kawai, H., S. Yukimoto, T. Koshiro, N. Oshima, T. Tanaka, H. Yoshimura, and R. Nagasawa, 2019:
1219 Significant Improvement of Cloud Representation in Global Climate Model MRI-ESM2. *Geosci. Model*
1220 *Dev.*, 12, 2875-2897.

1221 Lewis, N., & Mauritsen, T. (2021). Negligible Unforced Historical Pattern Effect on Climate Feedback
1222 Strength Found in HadISST-Based AMIP Simulations, *Journal of Climate*, 34(1), 39-55.
1223 <https://doi.org/10.1175/JCLI-D-19-0941.1>.

- 1224 Lewis, N., & Curry, J. A. (2018). The impact of recent forcing and ocean heat uptake data on
1225 estimates of climate sensitivity. *Journal of Climate*, 31, 6051– 6071.
- 1226 Li, C., von Storch, J.-S., & Marotzke, J. (2013). Deep-ocean heat uptake and equilibrium climate
1227 response. *Climate Dynamics*, 40, 1071– 1086.
- 1228 Liu, C., Allan, R. P., Berrisford, P., Mayer, M., Hyder, P., Loeb, N., Smith, D., Vidale, P.-L., and
1229 Edwards, J. M. (2015), Combining satellite observations and reanalysis energy transports to estimate
1230 global net surface energy fluxes 1985–2012, *J. Geophys. Res. Atmos.*, 120, 9374– 9389,
1231 doi:[10.1002/2015JD023264](https://doi.org/10.1002/2015JD023264).
- 1232 Liu, C., Allan, R. P., Mayer, M., Hyder, P., Loeb, N. G., Roberts, C. D., Valdivieso, M., Edwards, J. M.,
1233 and Vidale, P.-L. (2017), Evaluation of satellite and reanalysis-based global net surface energy flux
1234 and uncertainty estimates, *J. Geophys. Res. Atmos.*, 122, 6250– 6272, doi:[10.1002/2017JD026616](https://doi.org/10.1002/2017JD026616).
- 1235 Liu, C., Allan, R.P., Mayer, M. *et al.* Variability in the global energy budget and transports 1985–2017.
1236 *Clim Dyn* 55, 3381–3396 (2020). <https://doi.org/10.1007/s00382-020-05451-8>
- 1237 Liu, C. and R. Allan (2022): Reconstructions of the radiation fluxes at the top of atmosphere and net
1238 surface energy flux: DEEP-C Version 5.0 [Dataset]. University of Reading.
1239 <https://doi.org/10.17864/1947.000347>.
- 1240 Loeb, N. G., Johnson, G. C., Thorsen, T. J., Lyman, J. M., Rose, F. G., & Kato, S. (2021). Satellite and
1241 ocean data reveal marked increase in Earth’s heating rate. *Geophysical Research Letters*, 48,
1242 e2021GL093047. <https://doi.org/10.1029/2021GL093047>.
- 1243 Loeb, N. G., Wang, H., Allan, R., Andrews, T., Armour, K., Cole, J. N. S., et al. (2020). New generation
1244 of climate models track recent unprecedented changes in earth's radiation budget observed by
1245 CERES. *Geophysical Research Letters*, 47, e2019GL086705. <https://doi.org/10.1029/2019GL086705>.
- 1246 Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments
1247 in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂.
1248 *Journal of Advances in Modeling Earth Systems*, 11, 998– 1038.
1249 <https://doi.org/10.1029/2018MS001400>
- 1250 Marvel, K., Pincus, R., Schmidt, G. A., & Miller, R. L. (2018). Internal variability and disequilibrium
1251 confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, 45,
1252 1595– 1601. <https://doi.org/10.1002/2017GL076468>
- 1253 Marvel, K., Schmidt, G. A., Miller, R. L., & Nazarenko, L. (2016). Implications for climate sensitivity
1254 from the response to individual forcings. *Nature Climate Change*, 6, 386– 389).
1255 <https://doi.org/10.1038/nclimate2888>.
- 1256 Martin, G.M., et al., 2011: The HadGEM2 family of Met Office Unified Model climate configurations,
1257 *Geosci. Model Dev.*, 4, 723–757, <https://doi.org/10.5194/gmd-4-723-2011>, 2011.
- 1258 McGregor, S., Stuecker, M.F., Kajtar, J.B. *et al.* Model tropical Atlantic biases underpin diminished
1259 Pacific decadal variability. *Nature Clim Change* 8, 493–498 (2018). <https://doi.org/10.1038/s41558-018-0163-4>
- 1261 Meehl G A, Senior C A, Eyring V, Flato G, Lamarque J-F, Stouffer R J, Taylor K E and Schlund M,
1262 (2020), Context for interpreting equilibrium climate sensitivity and transient climate response from
1263 the CMIP6 Earth system models. *Sci. Adv.* 6, 26, <https://doi.org/10.1126/sciadv.aba1981>.

- 1264 Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., et al. (2021). An
 1265 updated assessment of near-surface temperature change from 1850: the HadCRUT5 data set.
 1266 *Journal of Geophysical Research: Atmospheres*, 126, e2019JD032361.
 1267 <https://doi.org/10.1029/2019JD032361>
- 1268 Moseid, K. O., Schulz, M., Storelvmo, T., Julsrud, I. R., Olivié, D., Nabat, P., Wild, M., Cole, J. N. S.,
 1269 Takemura, T., Oshima, N., Bauer, S. E., and Gastineau, G.: Bias in CMIP6 models as compared to
 1270 observed regional dimming and brightening, *Atmos. Chem. Phys.*, 20, 16023–16040,
 1271 <https://doi.org/10.5194/acp-20-16023-2020>, 2020.
- 1272 Neale, R. B., Richter, J., Park, S., Lauritzen, P. H., Vavrus, S. J., Rasch, P. J., & Zhang, M. (2013). The
 1273 Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled
 1274 Experiments, *Journal of Climate*, 26(14), 5150-5168. <https://doi.org/10.1175/JCLI-D-12-00236.1>.
- 1275 Newsom, E., Zanna, L., Khatiwala, S., & Gregory, J. M. (2020). The influence of warming patterns on
 1276 passive ocean heat uptake. *Geophysical Research Letters*, 47, e2020GL088429.
 1277 <https://doi.org/10.1029/2020GL088429>
- 1278 Nijse, F. J. M. M., Cox, P. M., & Williamson, M. S. (2020). An emergent constraint on Transient
 1279 Climate Response from simulated historical warming in CMIP6 models. *Earth System Dynamics*.
 1280 <https://doi.org/10.5194/esd-2019-86>.
- 1281 Olonscheck, D., Rugenstein, M., & Marotzke, J. (2020). Broad consistency between observed and
 1282 simulated trends in sea surface temperature patterns. *Geophysical Research Letters*, 47,
 1283 e2019GL086773. <https://doi.org/10.1029/2019GL086773>
- 1284 Otto, A., Otto, F. E. L., Boucher, O., Church, J., Hegerl, G., Forster, P. M., Gillett, N. P., Gregory, J.,
 1285 Johnson, G. C., Knutti, R., Lewis, N., Lohmann, U., Marotzke, J., Myhre, G., Shindell, D., Stevens, B., &
 1286 Allen, M. R. (2013). Energy budget constraints on climate response. *Nature Geoscience*, 6(6), 415–
 1287 416. <https://doi.org/10.1038/ngeo1836>
- 1288 Pope, D. V., M. Gallani, R. Rowntree, and A. Stratton (2000), The impact of new physical
 1289 parameterizations in the Hadley Centre climate model: HadAM3, *Clim. Dyn.*, 16(2–3), 123–146.
- 1290 Power, S., et al., (2021). Decadal climate variability in the tropical Pacific: Characteristics, causes,
 1291 predictability, and prospects. *Science*. 374. eaay9165. 10.1126/science.aay9165.
- 1292 Proistosescu, C., & Huybers, P. J. (2017). Slow climate mode reconciles historical and model-based
 1293 estimates of climate sensitivity. *Science Advances*, 3, 1– 7. <https://doi.org/10.1126/sciadv.1602821>
- 1294 Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and
 1295 Kaplan, A. (2003), Global analyses of sea surface temperature, sea ice, and night marine air
 1296 temperature since the late nineteenth century, *J. Geophys. Res.*, 108, 4407,
 1297 doi:[10.1029/2002JD002670](https://doi.org/10.1029/2002JD002670), D14.
- 1298 Raghuraman, S.P., Paynter, D. & Ramaswamy, V. Anthropogenic forcing and response yield observed
 1299 positive trend in Earth’s energy imbalance. *Nat Commun* 12, 4577 (2021).
 1300 <https://doi.org/10.1038/s41467-021-24544-4>
- 1301 Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An Improved In Situ
 1302 and Satellite SST Analysis for Climate, *Journal of Climate*, 15(13), 1609-1625.

1303 Richardson, T. B., Forster, P. M., Smith, C. J., Maycock, A. C., Wood, T., Andrews, T., et al. (2019).
1304 Efficacy of climate forcings in PDRMIP models. *Journal of Geophysical Research: Atmospheres*, 124,
1305 12824– 12844. <https://doi.org/10.1029/2019JD030581>

1306 Rose, B. E. J., Armour, K. C., Battisti, D. S., Feldl, N., & Koll, D. D. (2014). The dependence of transient
1307 climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophysical*
1308 *Research Letters*, 41, 1– 8.

1309 Rugenstein, M. A. A., Caldiera, K., & Knutti, R. (2016). Dependence of global radiative feedbacks on
1310 evolving patterns of surface heat fluxes. *Geophysical Research Letters*, 43, 9877– 9885.
1311 <https://doi.org/10.1002/2016GL070907>

1312 Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., et al. (2020).
1313 Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophysical Research*
1314 *Letters*, 47, e2019GL083898. <https://doi.org/10.1029/2019GL083898>

1315 Rugenstein, M. A. A., & Armour, K. C. (2021). Three flavors of radiative feedbacks and their
1316 implications for estimating equilibrium climate sensitivity. *Geophysical Research Letters*, 48,
1317 e2021GL092983. <https://doi.org/10.1029/2021GL092983>

1318 Sadai, S., Condrón, A., DeConto, R., & Pollard, D. (2020). Future climate response to Antarctic Ice
1319 Sheet melt caused by anthropogenic warming. *Science advances*, 6(39), eaaz1169.

1320 Sanderson, B. M., and Knutti, R. (2012), On the interpretation of constrained climate model
1321 ensembles, *Geophys. Res. Lett.*, 39, L16708, doi:[10.1029/2012GL052665](https://doi.org/10.1029/2012GL052665).

1322 Schneider, A., Flanner, M. & Perket, J. Multidecadal variability in surface albedo feedback across
1323 CMIP5 models. *Geophys. Res. Lett.* **45**, 1972–1980 (2018).

1324 Seager, R., Cane, M., Henderson, N., Lee, D.-E., Abernathey, R., & Zhang, H. (2019). Strengthening
1325 tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases.
1326 *Nature Climate Change*, 9, 517– 522.

1327 Senior, C. A., & Mitchell, J. F. B. (2000). The time dependence of climate sensitivity. *Geophysical*
1328 *Research Letters*, 27, 2685– 2688. <https://doi.org/10.1029/2000GL011373>

1329 Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al.
1330 (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of*
1331 *Geophysics*, 58, e2019RG000678. <https://doi.org/10.1029/2019RG000678>.

1332 Silvers, L. G., Paynter, D., & Zhao, M. (2018). The diversity of cloud responses to twentieth century
1333 sea surface temperatures. *Geophysical Research Letters*, 45, 391– 400.
1334 <https://doi.org/10.1002/2017GL075583>

1335 Smith, C., Piers Forster, Matt Palmer, Bill Collins, Nick Leach, Masa Watanabe, Sophie Berger, & Brad
1336 Hall, (2021). IPCC-WG1/Chapter-7: IPCC WGI AR6 Chapter 7 (v.1.0) [dataset]. Zenodo.
1337 <https://doi.org/10.5281/zenodo.5211358>.

1338 Smith, D. M., and Coauthors, 2015: Earth's energy imbalance since 1960 in observations and CMIP5
1339 models. *Geophys. Res. Lett.*, **42**, 1205–1213, <https://doi.org/10.1002/2014GL062669>.

1340 Stevens, B., Sherwood, S. C., Bony, S., & Webb, M. J. (2016). Prospects for narrowing bounds on
1341 Earth's equilibrium climate sensitivity. *Earth's Future*, 4, 512– 522.
1342 <https://doi.org/10.1002/2016EF000376>.

1343 Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V.,
1344 Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao,
1345 A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System
1346 Model version 5 (CanESM5.0.3), *Geosci. Model Dev.*, 12, 4823–4873, [https://doi.org/10.5194/gmd-](https://doi.org/10.5194/gmd-12-4823-2019)
1347 12-4823-2019, 2019.

1348 Takahashi, C., and M. Watanabe, 2016: Pacific trade winds accelerated by aerosol forcing over the
1349 past two decades. *Nature Climate Change*, 6, 768-772, doi: 10.1038/nclimate2996.

1350 Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M.,
1351 Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K.,
1352 Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.:
1353 Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity
1354 in MIROC6, *Geosci. Model Dev.*, 12, 2727–2765, <https://doi.org/10.5194/gmd-12-2727-2019>, 2019.

1355 Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design.
1356 *Bulletin of the American Meteorological Society*, 93, 485– 498.

1357 Taylor, K. E., Williamson, D., & Zwiers, F. (2000). The sea surface temperature and sea-ice
1358 concentration boundary conditions for AMIP II simulations, PCMDI Report No. 60, Program for
1359 Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory.

1360 Tierney, J. E., Haywood, A. M., Feng, R., Bhattacharya, T., & Otto-Bliesner, B. L. (2019). Pliocene
1361 warmth consistent with greenhouse gas forcing. *Geophysical Research Letters*, 46, 9136– 9144.
1362 <https://doi.org/10.1029/2019GL083802>.

1363 Tierney, J. E., Zhu, J., King, J., Malevich, S.B., Hakim, G.J., & Poulsen, C.J. (2020). Global cooling and
1364 climate sensitivity revisited. <https://doi.org/10.31223/osf.io/me5uj>.

1365 Titchner, H. A., & Rayner, N. A. (2014). The Met Office Hadley Centre sea ice and sea surface
1366 temperature data set, version 2: 1. Sea ice concentrations. *Journal of Geophysical Research:*
1367 *Atmospheres*, 119, 2864– 2889. <https://doi.org/10.1002/2013JD020316>

1368 Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., & Knutti, R. (2020).
1369 Past warming trend constrains future warming in CMIP6 models. *Science Advance*, 6, eaaz9549.
1370 <https://doi.org/10.1126/sciadv.aaz9549>.

1371 Voltaire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al. (2019).
1372 Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *Journal of Advances in Modeling Earth*
1373 *Systems*, 11, 2177– 2213. <https://doi.org/10.1029/2019MS001683>

1374 Watanabe, M., Dufresne, J.L., Kosaka, Y. *et al.*, (2021): Enhanced warming constrained by past trends
1375 in equatorial Pacific sea surface temperature gradient. *Nat. Clim. Chang.* **11**, 33–37 (2021).
1376 <https://doi.org/10.1038/s41558-020-00933-3>

1377 Watanabe, M., Kamae, Y., Yoshimori, M., Oka, A., Sato, M., Ishii, M., Mochizuki, T., and Kimoto, M.
1378 (2013), Strengthening of ocean heat uptake efficiency associated with the recent climate hiatus,
1379 *Geophys. Res. Lett.*, 40, 3175– 3179, doi:[10.1002/grl.50541](https://doi.org/10.1002/grl.50541).

1380 Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H.,
1381 Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Siebesma, A. P., Skinner, C.
1382 B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.: The Cloud Feedback Model

1383 Intercomparison Project (CFMIP) contribution to CMIP6, *Geosci. Model Dev.*, 10, 359–384,
1384 <https://doi.org/10.5194/gmd-10-359-2017>, 2017.

1385 Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., ... Xavier, P. K.
1386 (2017). The Met Office Global Coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal*
1387 *of Advances in Modeling Earth Systems*, 10, 357– 380. <https://doi.org/10.1002/2017MS001115>

1388 Yukimoto, S., H. Kawai, T. Koshiro, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T.
1389 Tanaka, M. Hosaka, S. Yabu, H. Yoshimura, E. Shindo, R. Mizuta, A. Obata, Y. Adachi, and M. Ishii,
1390 2019: The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0:
1391 Description and basic evaluation of the physical component. *J. Meteor. Soc. Japan*, 97, 931-965.

1392 Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., &
1393 Taylor, K. E. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research*
1394 *Letters*, 47, e2019GL085782. <https://doi.org/10.1029/2019GL085782>

1395 Zhang, L., Delworth, T.L., Cooke, W. *et al.* Natural variability of Southern Ocean convection as a
1396 driver of observed climate trends. *Nature Clim Change* 9, 59–65 (2019).
1397 <https://doi.org/10.1038/s41558-018-0350-3>

1398 Zhou, C., Zelinka, M. D., & Klein, S. A. (2016). Impact of decadal cloud variations on the Earth's
1399 energy budget. *Nature Geoscience*, 9, 871– 875.

1400 Zhou, C., Zelinka, M.D., Dessler, A.E. *et al.* Greater committed warming after accounting for the
1401 pattern effect. *Nat. Clim. Chang.* 11, 132–136 (2021). <https://doi.org/10.1038/s41558-020-00955-x>

1402 Zhu, J., B.L. Otto-Bliesner, E.C. Brady, C. Poulson, J.K. Shaw, J.E. Kay (2022), LGM paleoclimate
1403 constraints inform cloud parameterizations and equilibrium climate sensitivity in CESM2. *Earth and*
1404 *Space Science Open Archive*, <https://doi.org/10.1002/essoar.10507790.1>.

1405

1406 **Table1: Summary of the Atmospheric General Circulation Model (AGM) simulations used in this study.** *amip-piForcing* refers to an AGCM simulation
 1407 forced with time-varying observed monthly SSTs and sea-ice using the AMIP II boundary condition SST and sea-ice dataset, forcing agents such greenhouse
 1408 gases, aerosol emission etc. are kept at pre-industrial levels. *hadSST-piForcing* is identical in all aspects except SSTs are taken from the HadISST1 database
 1409 (sea-ice remains the same as *amip-piForcing*). The ensemble size and time-periods covered for each experiment and AGCM is indicated. *amip-piForcing*
 1410 simulations included in the CFMIP3 (Webb et al. 2017) contribution to CMIP6 are indicated by a y/n. The corresponding name of each AGCMs parent
 1411 AOGCM is indicated. Global-annual-ensemble-mean dT and dN timeseries data are available for all *amip-piForcing* and *hadSST-piForcing* AGCM simulations
 1412 (see Data Availability Statement).

AGCM	Corresponding AOGCM name	Model description	<i>amip-piForcing</i>			<i>hadSST-piForcing</i>	
			CMIP6? (y/n)	Ensemble size	Time-period covered	Ensemble size	Time-period covered
CAM4	CCSM4	Neale et al. (2013)	n	3	1870 – 2014	3	1870 – 2014
CESM2	unchanged	Danabasoglu et al. (2020)	y	1	1870 – 2014	1	1870 - 2015
CNRM-CM6-1	unchanged	Voltaire et al. (2019)	y	1	1870 – 2014	-	-
CanESM5	unchanged	Swart et al. (2019)	y	3	1870 – 2014	-	-
ECHAM6.3	MPI-ESM1.1	Mauritsen et al. (2019)	n	5	1871 – 2010	5	1871 – 2015
GFDL-AM3	GFDL-CM3	Donner et al. (2011)	n	1	1870 – 2014	1	1870 – 2014
GFDL-AM4	GFDL-CM4	Held et al. (2019)	n	1	1870 – 2016	1	1870 – 2016
HadAM3	HadCM3	Pope et al. (2000)	n	4	1871 – 2012	4	1871 – 2012
HadGEM2	HadGEM2-ES	Martin et al. (2011)	n	4	1871 – 2012	1	1871 – 2012
HadGEM3-GC31-LL	unchanged	Williams et al. (2017)	y	1	1870 – 2014	1	1871 – 2016
IPSL-CM6A-LR	unchanged	Boucher et al. (2020)	y	1	1870 – 2014	-	-
MIROC6	unchanged	Tatebe et al. (2019)	y	1	1870 – 2014	-	-
MRI-ESM2-0	unchanged	Yukimoto et al. (2019), Kawai et al. (2019)	y	1	1870 – 2014	-	-
MPI-ESM1-2-LR	unchanged	Mauritsen et al. (2019)	n	3	1871 – 2017	3	1871 – 2017

1413

1414

1415 **Table 2: Feedback parameter in *amip-piForcing* and *hadsst-piForcing* simulations over various historical time-periods, as well as *abrupt-4xCO2***
1416 **sensitivity parameters.** λ values from *amip-piForcing* and *hadsst-piForcing* are calculated from OLS regression ($\lambda = dN/dT$) over the relevant time-periods
1417 using global-annual-mean timeseries data. $F_{2\times CO_2}$ is calculated as $F_{4\times CO_2}/2$ and $ECS = -F_{2\times CO_2}/\lambda_{4\times CO_2}$ from 150 years of *abrupt-4xCO2* experiments ($\lambda_{4\times CO_2}$ calculated
1418 over years 1-20 and 21-150 is also shown) (see Andrews et al., 2012; 2015).

	abrupt-4xCO2					$\lambda_{1871-2010}$ (W m ⁻² K ⁻¹)		$\lambda_{1871-1980}$ (W m ⁻² K ⁻¹)		$\lambda_{1981-2010}$ (W m ⁻² K ⁻¹)	
	ECS (K)	$F_{2\times}$ (W m ⁻²)	$\lambda_{4\times CO_2}$ (W m ⁻² K ⁻¹)	$\lambda_{4\times CO_2, 1-20}$ (W m ⁻² K ⁻¹)	$\lambda_{4\times CO_2, 21-150}$ (W m ⁻² K ⁻¹)	AMIP	HadISST1	AMIP	HadISST1	AMIP	HadISST1
CAM4	2.95	3.64	-1.23	-1.52	-0.94	-2.14	-1.77	-1.22	-1.45	-2.84	-2.70
CESM2	5.16	3.39	-0.66	-1.17	-0.49	-1.93	-1.49	-0.87	-0.95	-3.08	-2.92
CNRM-CM6-1	4.88	3.66	-0.75	-0.93	-0.87	-1.23	-	-1.10	-	-1.64	-
CanESM5	5.61	3.64	-0.65	-0.70	-0.59	-1.44	-	-0.93	-	-1.83	-
ECHAM6_3	3.01	4.10	-1.36	-1.47	-1.08	-1.92	-1.57	-1.43	-1.38	-2.69	-2.42
GFDL-AM3	3.99	2.97	-0.74	-1.13	-0.61	-1.44	-1.35	-0.72	-0.99	-1.90	-1.41
GFDL-AM4	3.84	3.32	-0.86	-1.54	-0.60	-1.84	-1.66	-1.33	-1.40	-2.57	-2.93
HadAM3	3.37	3.52	-1.04	-1.25	-0.75	-1.65	-1.44	-1.35	-1.40	-2.19	-1.86
HadGEM2	4.62	2.90	-0.63	-0.81	-0.33	-1.39	-1.04	-1.12	-1.08	-2.26	-1.54
HadGEM3-GC31-LL	5.54	3.49	-0.63	-0.81	-0.60	-1.28	-1.01	-0.95	-0.84	-1.87	-1.55
IPSL-CM6A-LR	4.56	3.41	-0.75	-0.98	-0.61	-1.59	-	-1.17	-	-2.50	-
MIROC6	2.58	3.72	-1.44	-1.61	-1.60	-1.42	-	-1.21	-	-1.87	-
MRI-ESM2-0	3.13	3.44	-1.10	-1.68	-0.78	-1.93	-	-1.23	-	-2.79	-
MPI-ESM1-2-LR	3.02	4.21	-1.39	-1.61	-1.34	-1.88	-1.58	-1.30	-1.45	-2.55	-2.42
MEAN	4.02	3.53	-0.95	-1.23	-0.80	-1.65	-1.43	-1.14	-1.21	-2.33	-2.19
1.645 σ	1.64	0.57	0.49	0.54	0.55	0.46	0.41	0.33	0.38	0.72	0.95

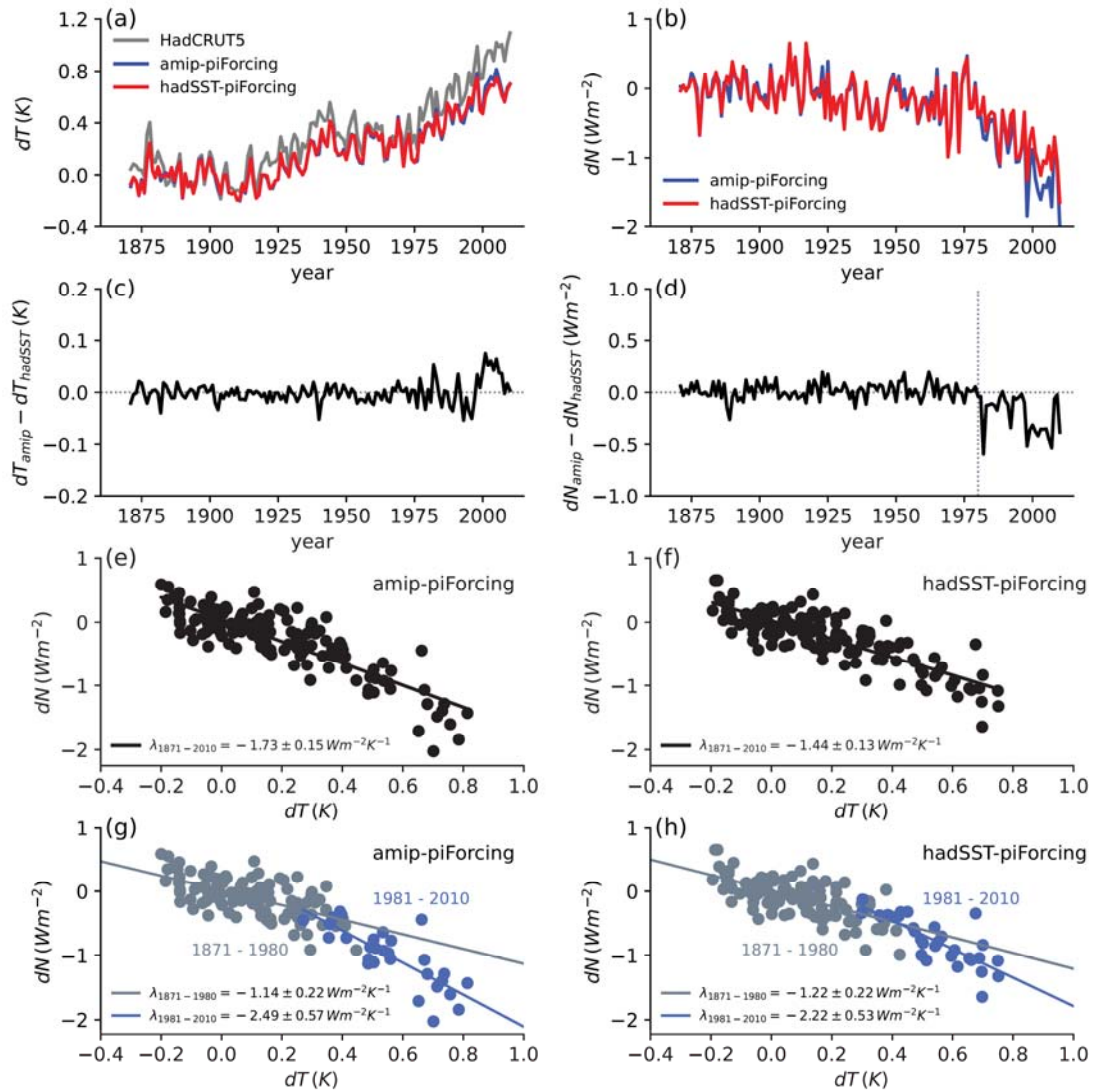
1419 **Table 3: The pattern effect ($\Delta\lambda = \lambda_{4xCO_2} - \lambda_{hist}$ with λ_{4xCO_2} from years 1-150 of *abrupt-4xCO2*)**
1420 **between *abrupt-4xCO2* radiative feedback and radiative feedback calculated over different**
1421 **historical periods (i.e. λ_{hist} from 1871-2010, and its separation into 1871-1980 and 1981-2010) in**
1422 ***amip-piForcing* and *hadSST-piForcing*, as well as their difference.**

	1871 – 2010 (W m ⁻² K ⁻¹)			1871 – 1980 (W m ⁻² K ⁻¹)			1981 – 2010 (W m ⁻² K ⁻¹)		
	AMIP	HadSST	Diff	AMIP	HadSST	Diff	AMIP	HadSST	Diff
CAM4	0.90	0.53	0.37	-0.01	0.22	-0.23	1.60	1.47	0.13
CESM2	1.27	0.84	0.43	0.21	0.29	-0.08	2.43	2.26	0.17
CNRM-CM6-1	0.48			0.35			0.89		
CanESM5	0.80			0.28			1.19		
ECHAM6_3	0.56	0.21	0.35	0.07	0.02	0.05	1.32	1.06	0.26
GFDL-AM3	0.69	0.61	0.08	-0.03	0.24	-0.27	1.15	0.67	0.48
GFDL-AM4	0.97	0.80	0.17	0.47	0.53	-0.06	1.70	2.07	-0.37
HadAM3	0.61	0.40	0.21	0.31	0.35	-0.04	1.15	0.82	0.33
HadGEM2	0.76	0.41	0.35	0.49	0.45	0.04	1.63	0.91	0.72
HadGEM3-GC31-LL	0.65	0.38	0.27	0.32	0.21	0.11	1.24	0.92	0.32
IPSL-CM6A-LR	0.84			0.43			1.76		
MIROC6	-0.02			-0.23			0.42		
MRI-ESM2-0	0.83			0.14			1.69		
MPI-ESM1-2-LR	0.49	0.19	0.30	-0.09	0.06	-0.15	1.16	1.03	0.13
MEAN	0.70	0.48	0.28	0.19	0.26	-0.07	1.38	1.24	0.24
1.645σ	0.47	0.36	0.17	0.35	0.26	0.20	0.75	0.88	0.46

1423

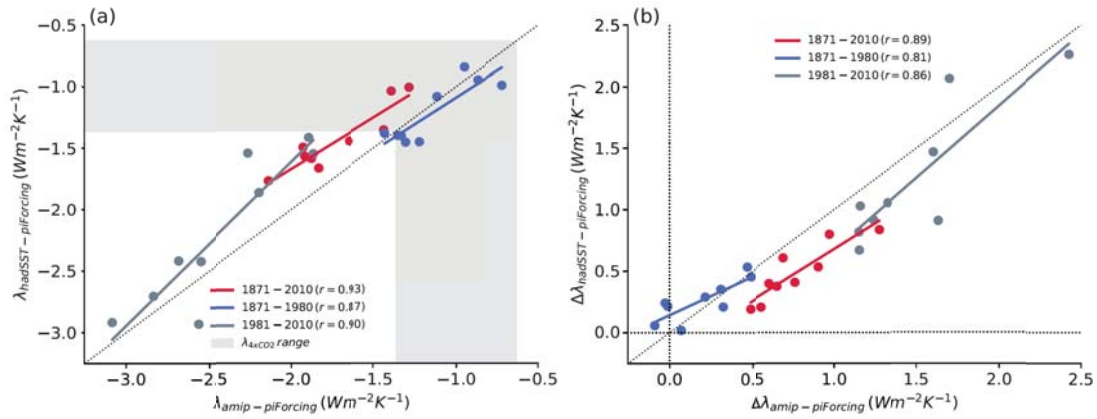
1424 Table 4: Comparison of the 1985-2014 climate resistance ($\rho = dF/dT$), feedback parameter ($-\lambda = -$
1425 $d(N - F)/dT$ and ocean heat uptake efficiency ($\kappa = dN/dT$) using different versions of the DEEP-C
1426 (Allan et al., 2014) satellite based reconstruction of dN (see Section 2.4). The lower half of the
1427 table shows how ρ , λ and κ estimates change as the 30 year moving window advances to 1990-
1428 2019. In all calculations HadCRUT5 analysis dT (Morice et al. 2021) and IPCC AR6 dF (Forster et al.,
1429 2021; Smith et al., 2021) are used. Years 1991-2 are excluded from the calculation as these years
1430 are identified as being strongly impacted by the volcanic forcing from the Pinatubo eruption
1431 (Section 4).

dN dataset version	Start year	End year	ρ ($W m^{-2} K^{-1}$)	$-\lambda$ ($W m^{-2} K^{-1}$)	κ ($W m^{-2} K^{-1}$)
DEEP-C v2G			2.38	2.24	0.14
DEEP-C v3			2.38	2.24	0.14
DEEP-C v3G	1985	2014	2.38	2.24	0.14
DEEP-C v4			2.38	1.98	0.41
DEEP-C v5			2.38	1.98	0.41
DEEP-C v5	1986	2015	2.38	1.75	0.63
DEEP-C v5	1987	2016	2.25	1.55	0.70
DEEP-C v5	1988	2017	2.21	1.62	0.59
DEEP-C v5	1989	2018	2.23	1.66	0.57
DEEP-C v5	1990	2019	2.30	1.44	0.86

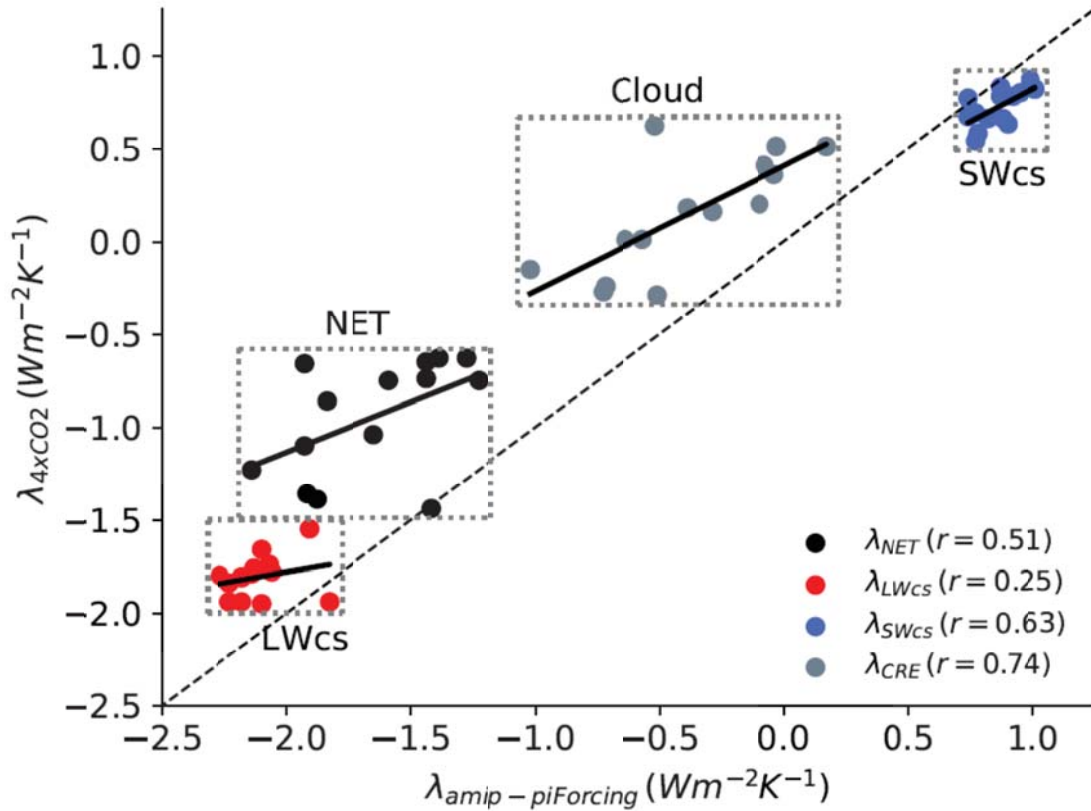


1432

1433 **Figure 1: Comparison of multi-model ensemble-annual-mean (a) dT and (b) dN in the $amip-piForcing$ and $hadSST-piForcing$ simulations. (c) and (d) shows the difference in dT and dN**
 1434 **respectively, highlighting 1980 as a key year where the dN response diverges according to the SST**
 1435 **dataset. In (a) the HadCRUT5 observed dT evolution is shown for comparison. (e) and (f) show the**
 1436 **relationship between global-annual-mean dT and dN in $amip-piForcing$ and $hadSST-piForcing$**
 1437 **respectively, where $\lambda=dN/dT$ is calculated from OLS regression on the global-annual-mean data**
 1438 **points. The stated 5-95% uncertainty is $\pm 1.645\sigma$ from the standard error of the linear fit. (g) and**
 1439 **(h) show the dT and dN relationship separated into two time-periods: years 1871-1980 (grey) and**
 1440 **years 1981-2010 (blue). The multi-model ensemble-means are restricted to the nine AGCMs that**
 1441 **performed both simulations (see Table 1).**
 1442

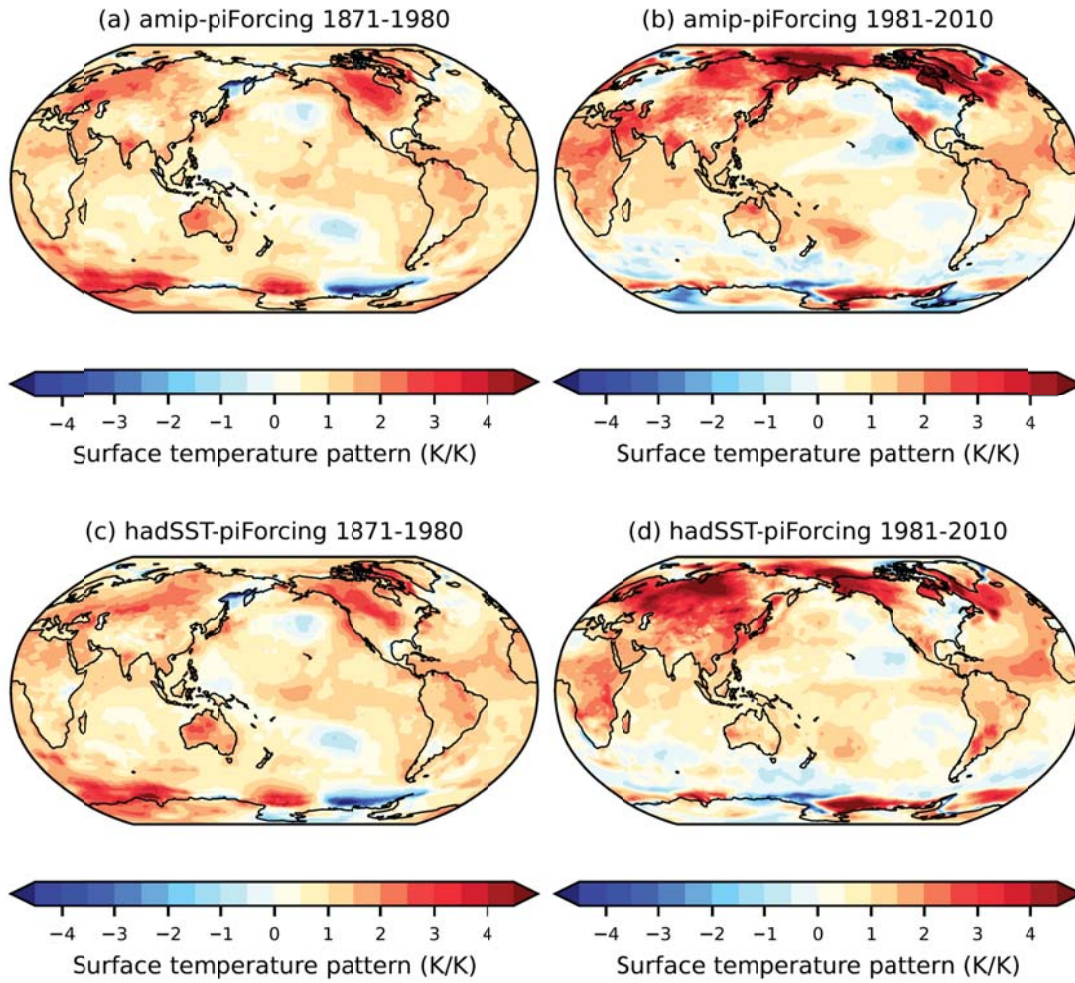


1440
 1441 Figure 2: (a) Relationship between the feedback parameter, λ , in the *amip-piForcing* and *hadSST-*
 1442 *piForcing* simulations over various historical time-periods. Each point is a single AGCM. The
 1443 shaded grey region shows the range of λ_{4xCO2} from the AGCMs corresponding parent AOGCM
 1444 *abrupt-4xCO2* simulation. The one-to-one line (dotted) is shown. (b) Relationship between the
 1445 pattern effect, $\Delta\lambda = \lambda_{4xCO2} - \lambda_{hist}$, diagnosed from the *amip-piForcing* and *hadSST-piForcing*
 1446 simulations over various historical time-periods.



1450

1455 **Figure 3: Relationship across models (dots) between the feedback parameter in *amip-piForcing***
 1455 **(calculated over years 1871-2010) and *abrupt-4xCO2* simulation (calculated over years 1-150). The**
 1457 **net feedback parameter is decomposed into its longwave clear-sky, SW clear-sky and cloud**
 1458 **radiative effect components.**



1455

1452

1453

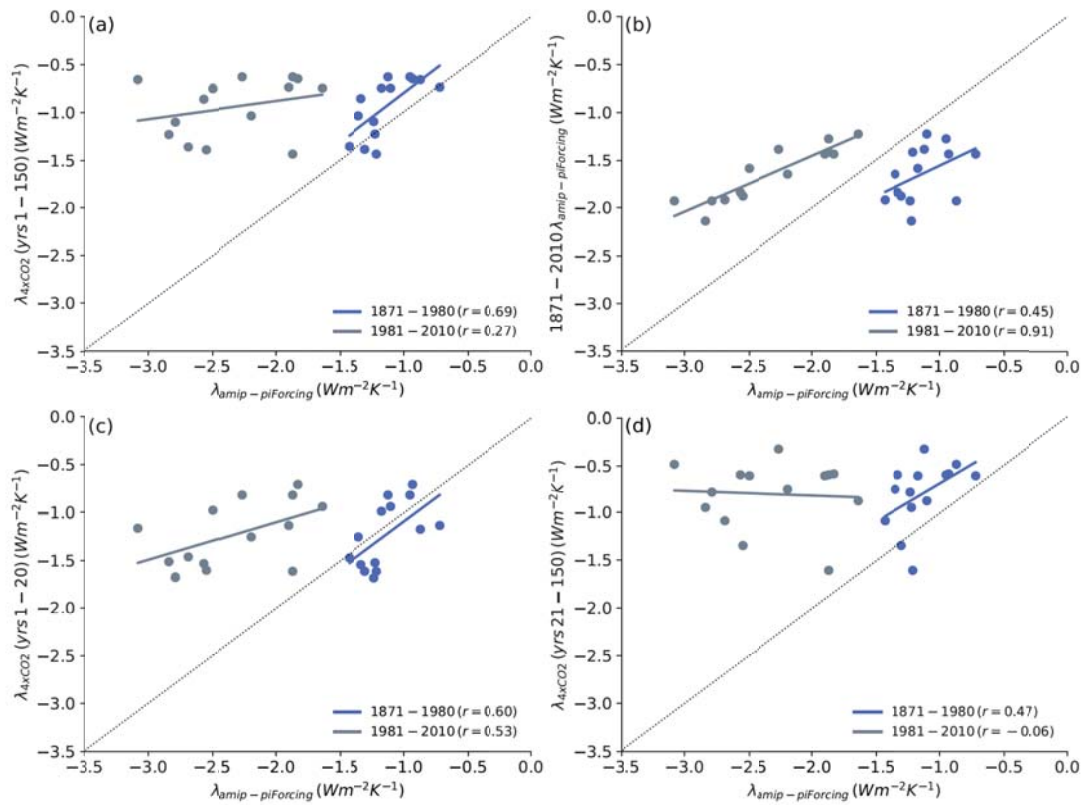
1454

1455

1466

1467

Figure 4: Pattern of near-surface temperature change (local dT per global-mean dT) for the time-periods 1870-1980 and 1981-2010 in (a) and (b) *amip-piForcing* and (c) and (d) *hadSST-piForcing*. Patterns are calculated from the slope of the linear regression of local temperature change against global-mean temperature change using annual-mean data points. Note that by definition the global-means are unity. Data from HadGEM3-GC31-LL simulations have been used for this illustration.



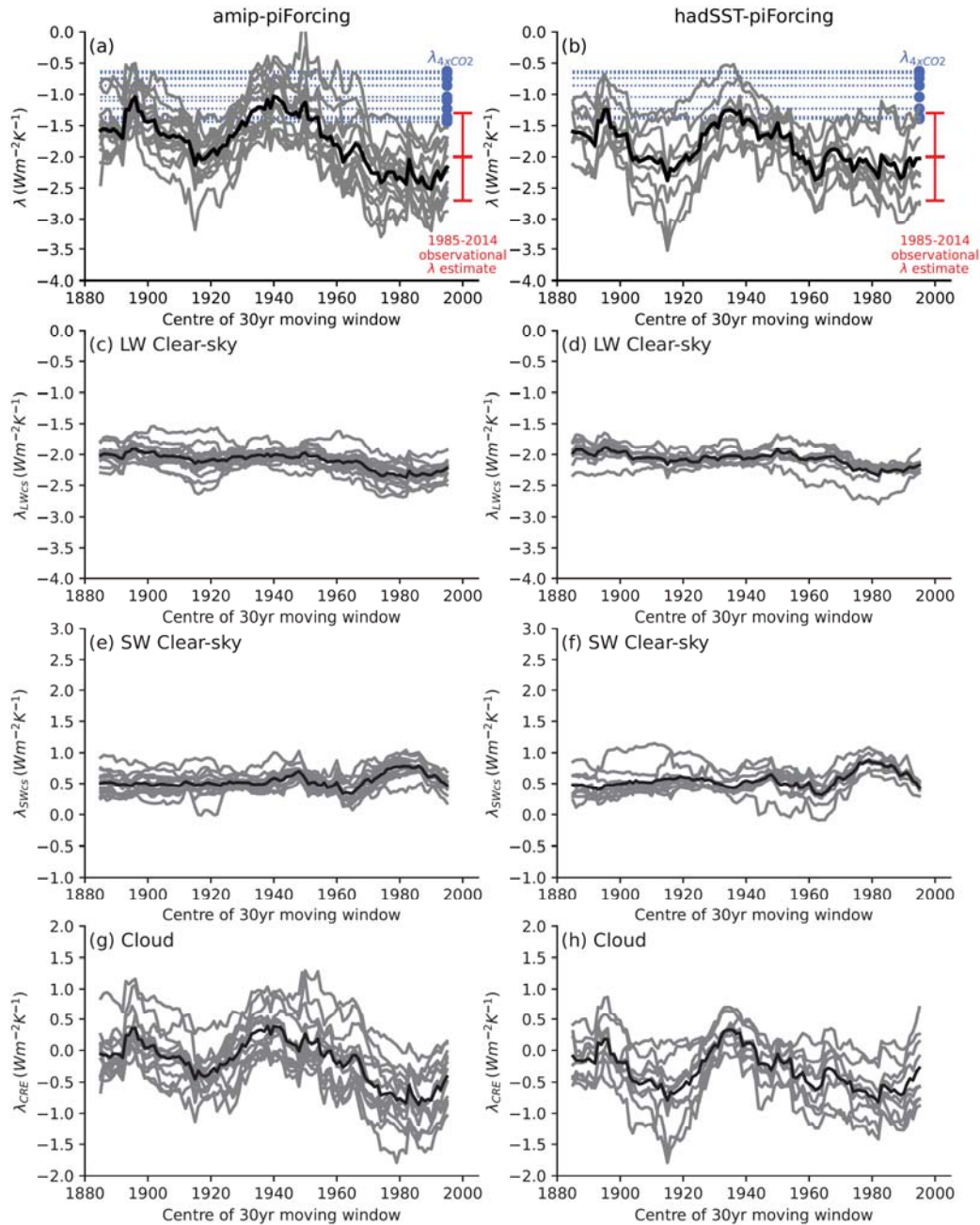
1463

1465

1467

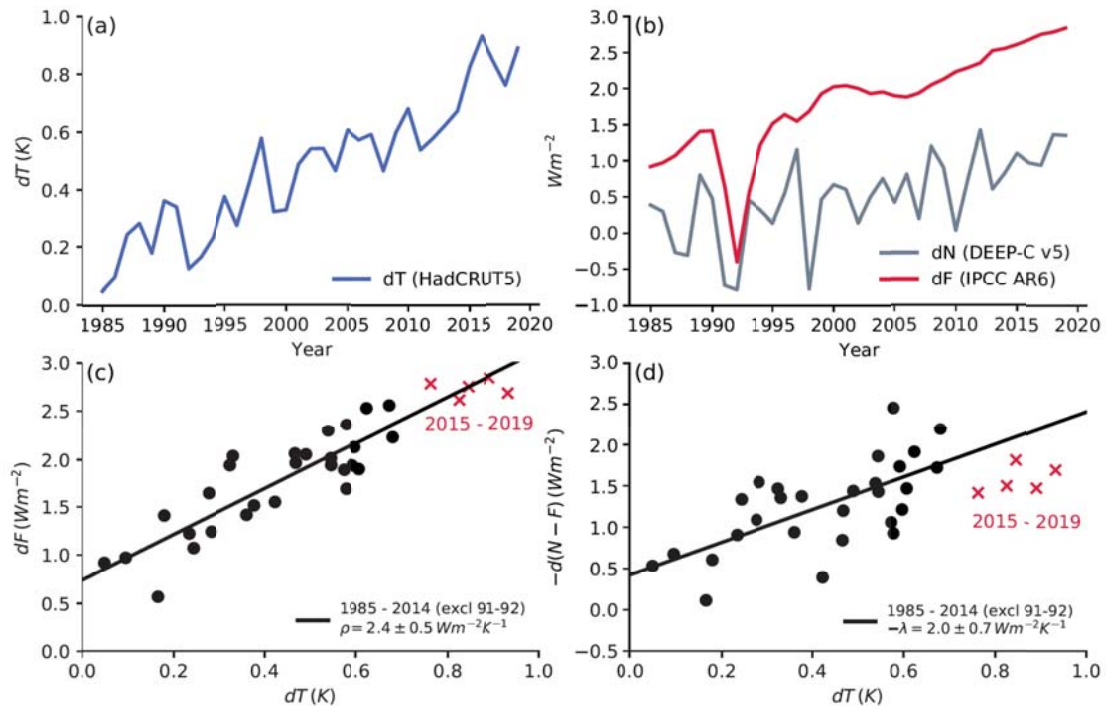
1468

Figure 5: Relationships between model simulated feedbacks in *amip-piForcing* over years 1871-1980 (blue) or 1981-2010 (grey) and (a) λ_{4xCO_2} from *abrupt-4xCO_2*, (b) λ_{hist} over the entire historical record (1871-2010), (c) λ_{4xCO_2} from *abrupt-4xCO_2* over years 1-20 and (d) years 21-150.



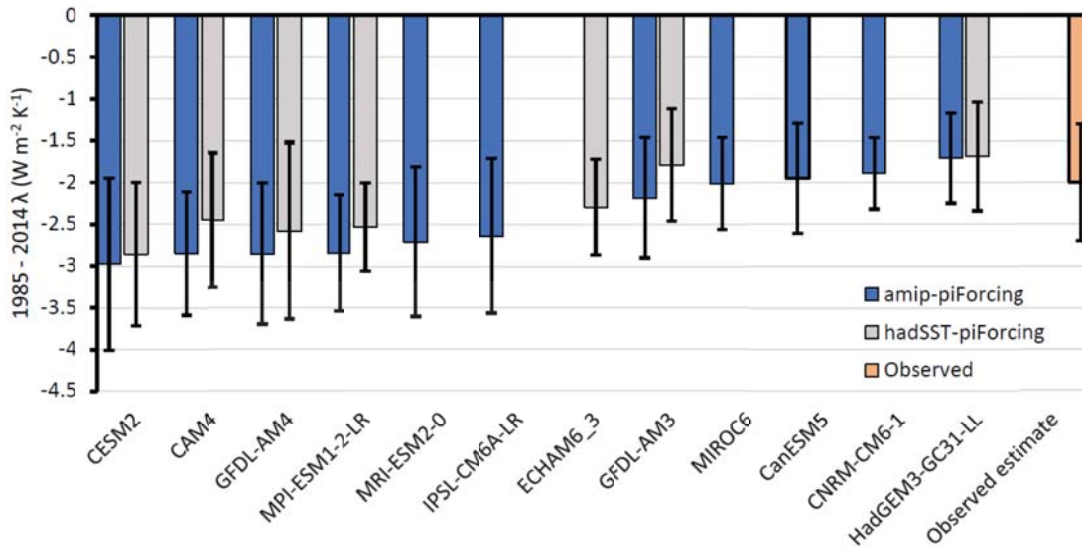
1466

1467 **Figure 6: Decadal variation in the feedback parameter λ from 1871 to 2010. Left column shows**
 1468 **results from *amip-piForcing* and right column shows results from *hadSST-piForcing*. Each grey line**
 1469 **represents a single AGCM (see Table 1). Thick black is the ensemble-mean of the results. X-axis**
 1470 **represents the centre of a 30 year moving window in which $\lambda = dN/dT$ is calculated from OLS**
 1471 **regression on annual-mean data, i.e. λ at 1980.5 represents the feedback parameter over years**
 1472 **1966 to 1995. Shown in (a) and (b) is the net feedback parameter. Blue dots and lines represent**
 1473 **the corresponding λ_{4xCO_2} values from AOGCM *abrupt-4xCO_2* simulations (Table 2). Red shows an**
 1474 **observational estimate and 5-95% uncertainty of $\lambda = d(N - F)/dT \sim -2.0 \pm 0.7 \text{ W m}^{-2} \text{ K}^{-1}$ over years**
 1475 **1985-2014 (see Section 4). (c) – (h) shows the corresponding LW clear-sky, SW clear-sky and cloud**
 1476 **radiative effect (CRE) components of λ .**



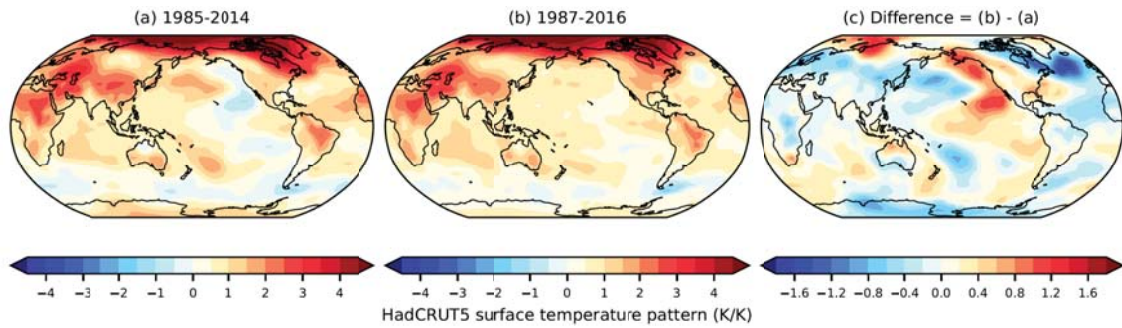
1473

1476 **Figure 7: Observational estimate of the Earth's 1985-2019 energy balance. All points are global-**
 1477 **annual-means. (a) dT (HadCRUT5 analysis dataset; Morice et al., 2021), (b) dN (DEEP-C v5; Allan et**
 1488 **al., 2014; Liu and Allan, 2022) and dF (IPCC AR6; Forster et al., 2021; Smith et al., 2021). (c) $\rho =$**
 1489 **dF/dT relationship and (d) $-\lambda_{\text{hist}} = -d(N - F)/dT$ relationship over years 1985-2014. Black dots are**
 1480 **global-annual means over years 1985-2014 excluding years 1991-2 which are strongly influenced**
 1481 **by the Pinatubo explosive volcanic eruption (see red line panel b). Red points in (c) and (d) are**
 1482 **years 2015-2019. The stated 5-95% uncertainties are $\pm 1.645\sigma$ from the standard error of the linear**
 1483 **fit.**



1487

1484 **Figure 8: Comparison of the 1985-2014 feedback parameter, $\lambda_{\text{hist}} = d(N - F)/dT$, in *amip-piForcing***
 1485 **and *hadSST-piForcing* simulations to an observed estimate based on DEEP-C V5 dN (Allan et al.,**
 1486 **2014; Liu and Allan, 2022), HadCRUT5 analysis dT (Morice et al. 2021) and IPCC AR6 dF (Forster et**
 1497 **al., 2021; Smith et al., 2021). The 5-95% uncertainty is simply 1.645σ from the standard error of**
 1498 **the linear fit, with no allowance for systematic uncertainties. Note also that years 1991-2 are**
 1499 **excluded from the calculation as these years are identified as being strongly impacted by the**
 1500 **volcanic forcing from the Pinatubo eruption (Figure 7b).**



1495

1490 **Figure 9: Pattern of near-surface temperature change (local dT per global-mean dT) for the time-**
 1491 **periods (a) 1985-2014 and (b) 1987-2016, and (c) shows the difference (b minus a). Data is the**
 1492 **HadCRUT5 analysis dataset (Morice et al. 2021). Patterns are calculated from the slope of the**
 1493 **linear regression of local temperature change against global-mean temperature change using**
 1494 **annual-mean data points. Note that by definition the global-means of panels (a) and (b) are unity.**

Figure 1.

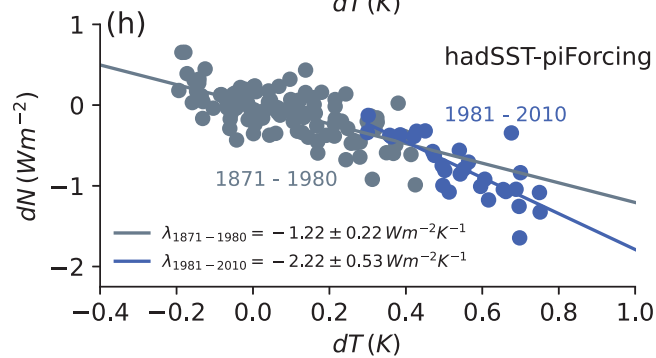
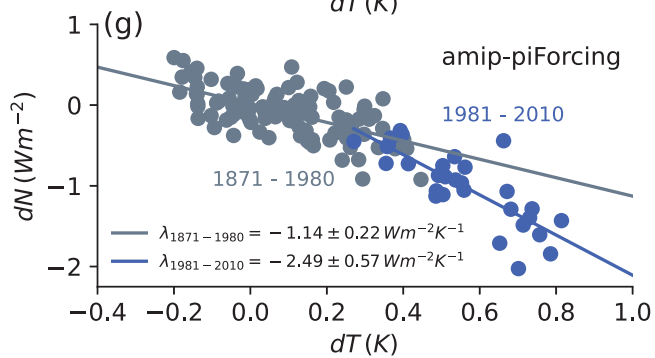
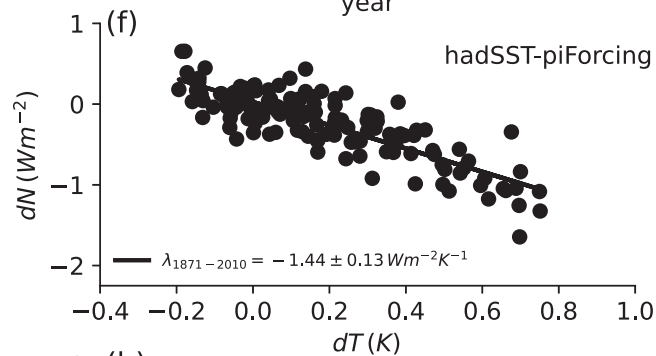
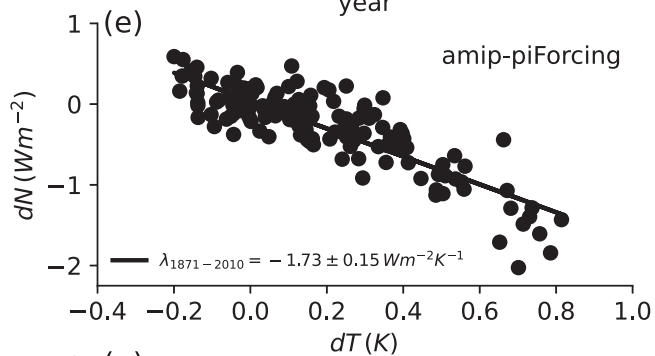
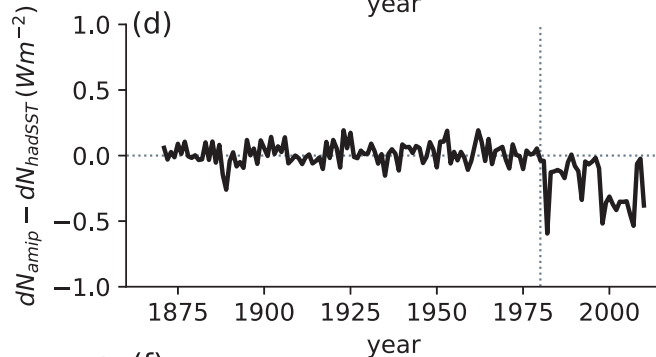
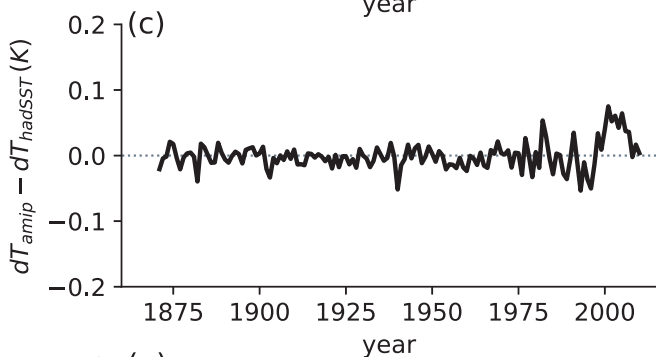
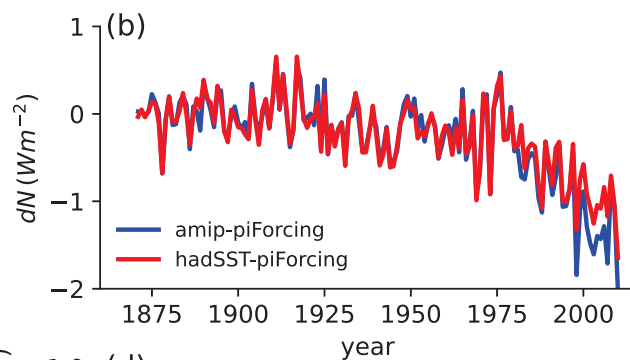
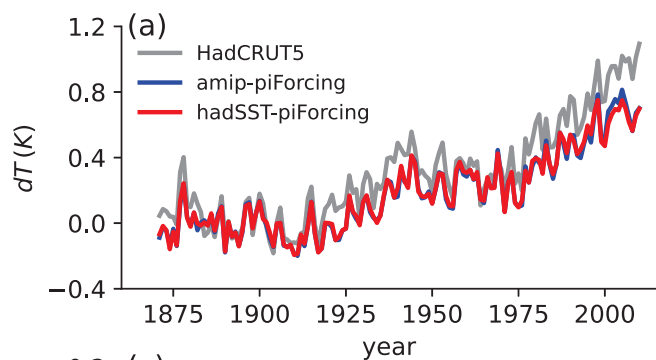


Figure 2.

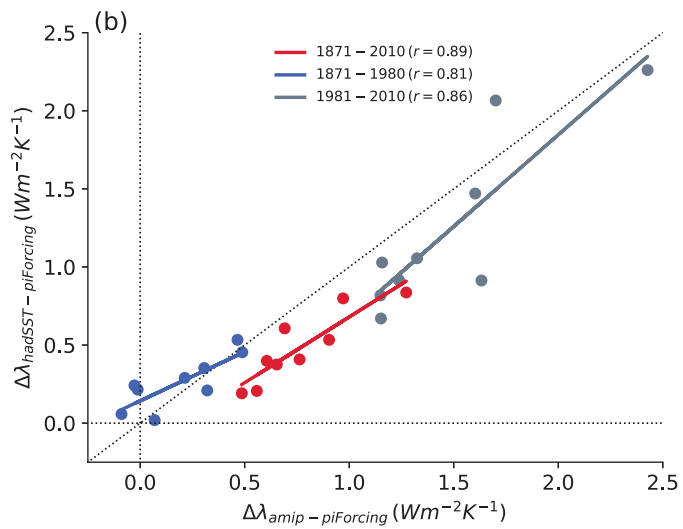
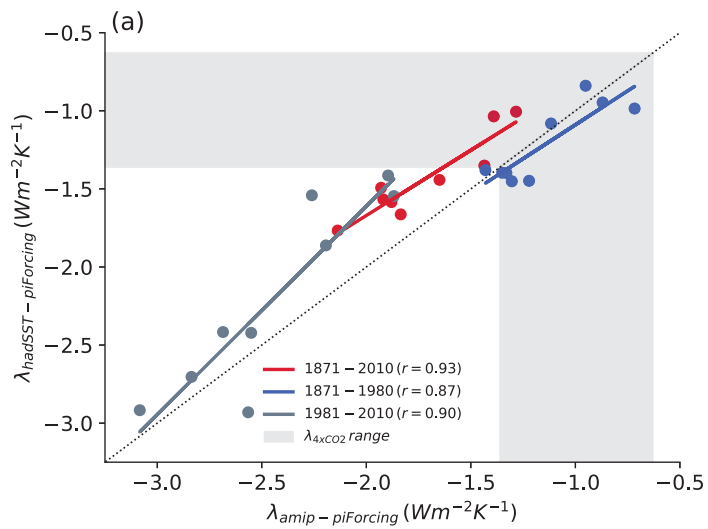


Figure 3.

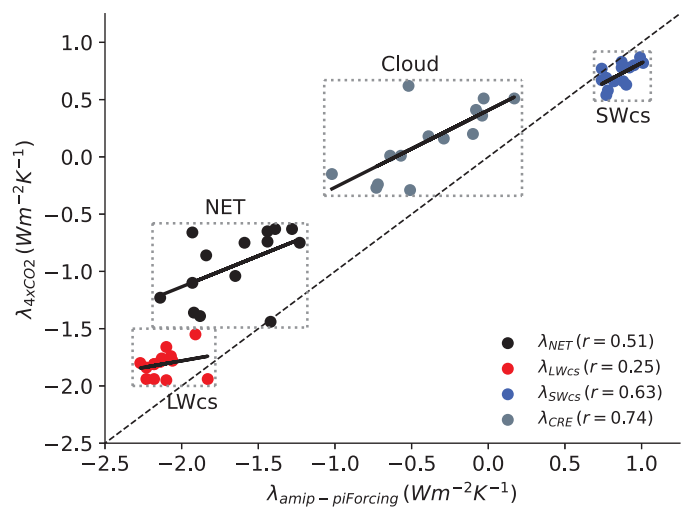
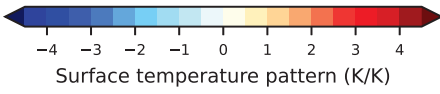
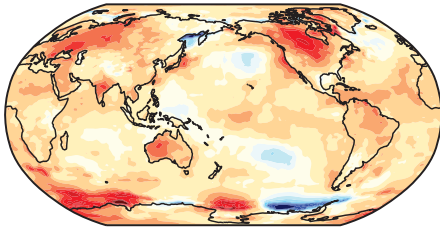
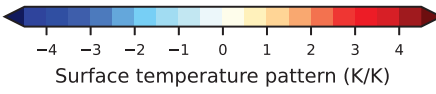
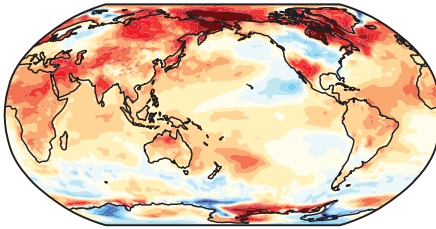


Figure 4.

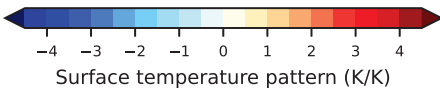
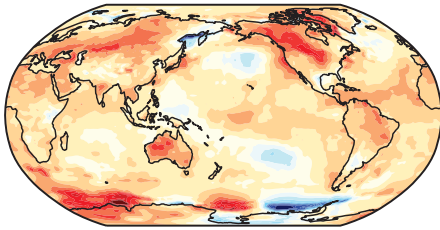
(a) amip-piForcing 1871-1980



(b) amip-piForcing 1981-2010



(c) hadSST-piForcing 1871-1980



(d) hadSST-piForcing 1981-2010

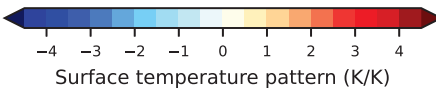
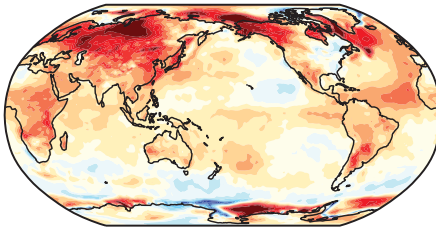


Figure 5.

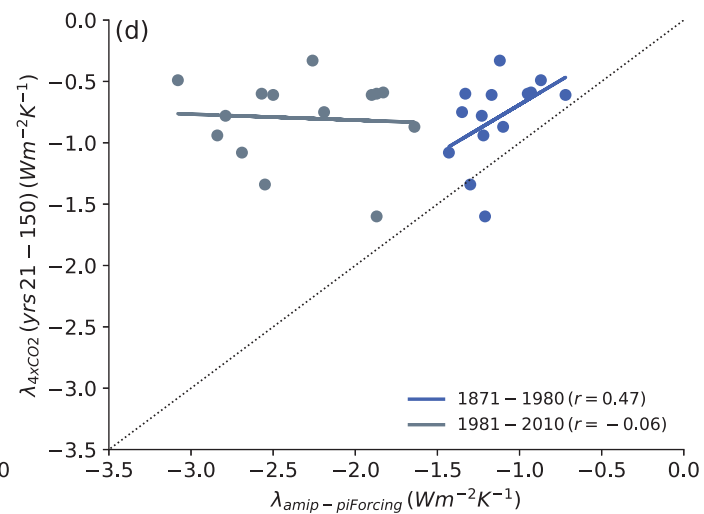
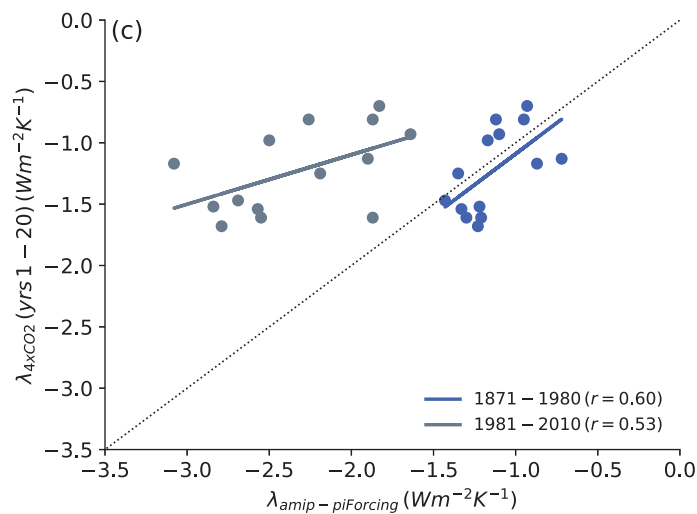
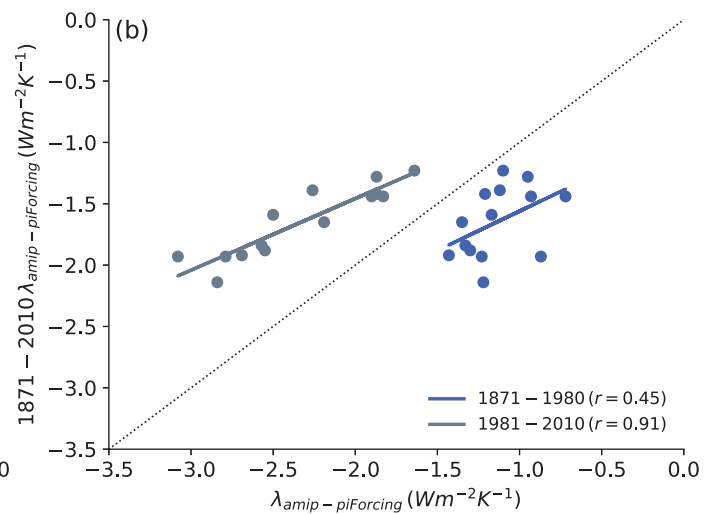
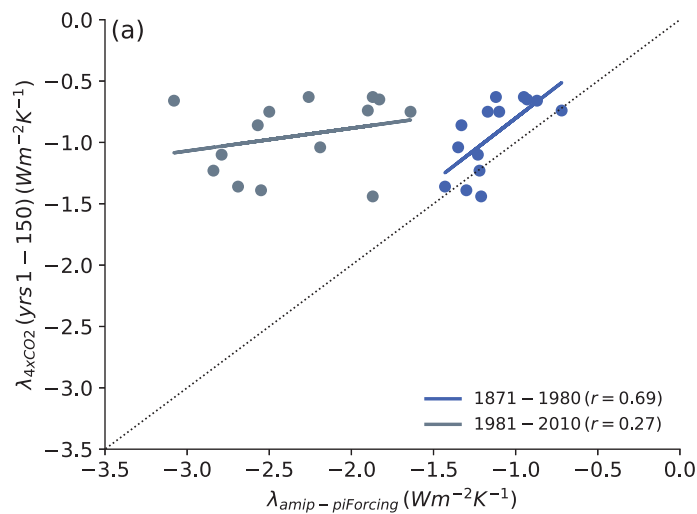


Figure 6.

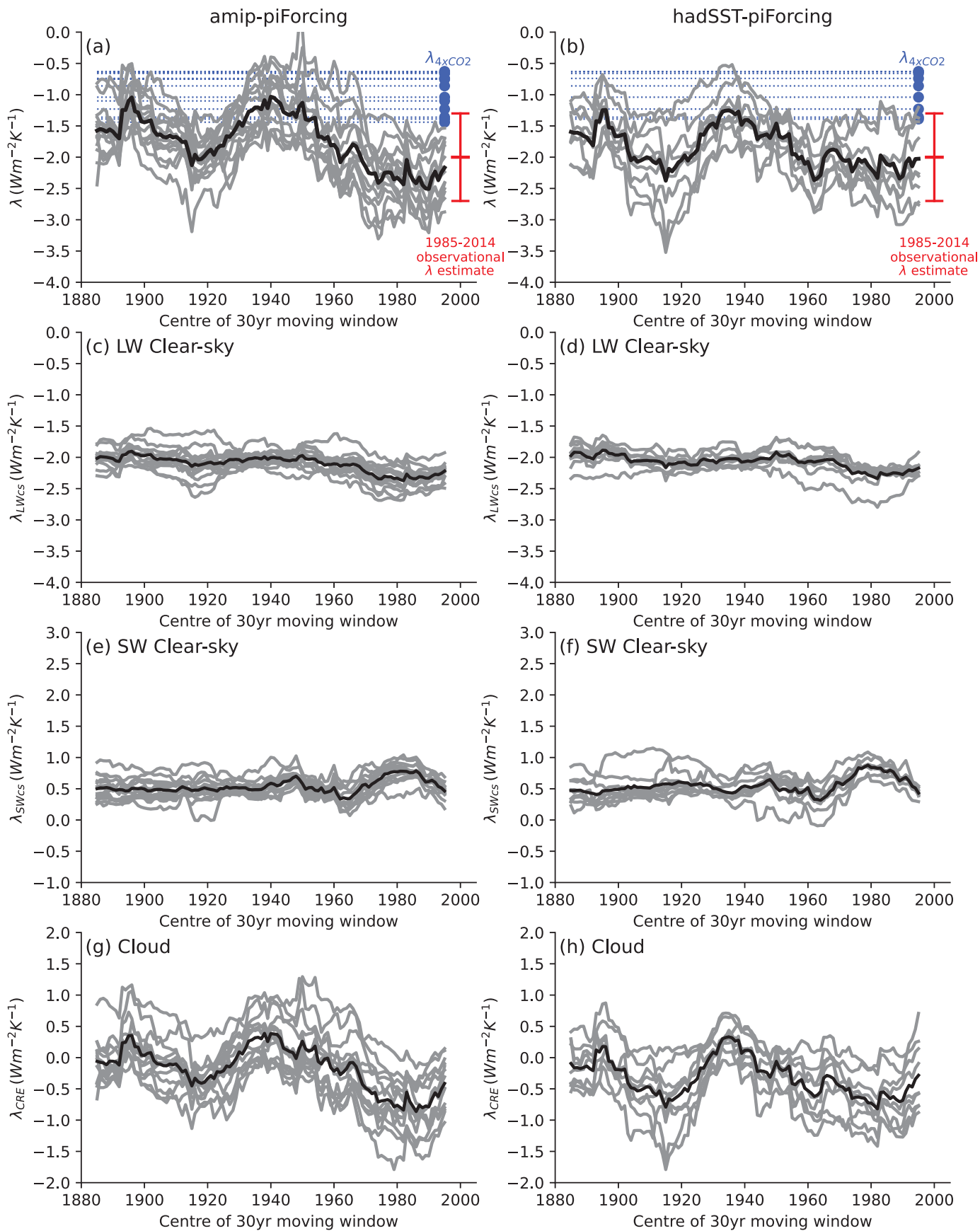


Figure 7.

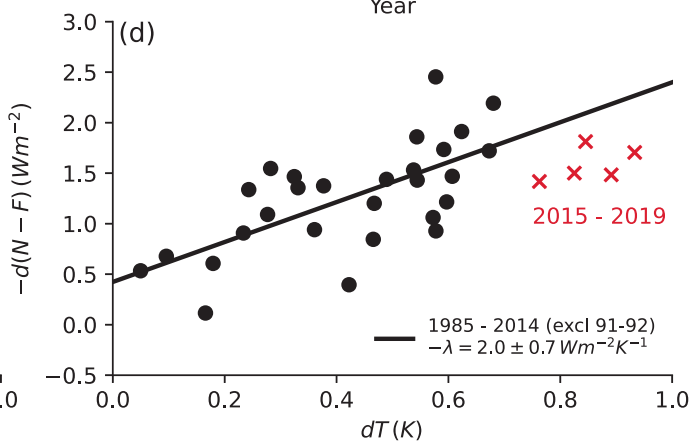
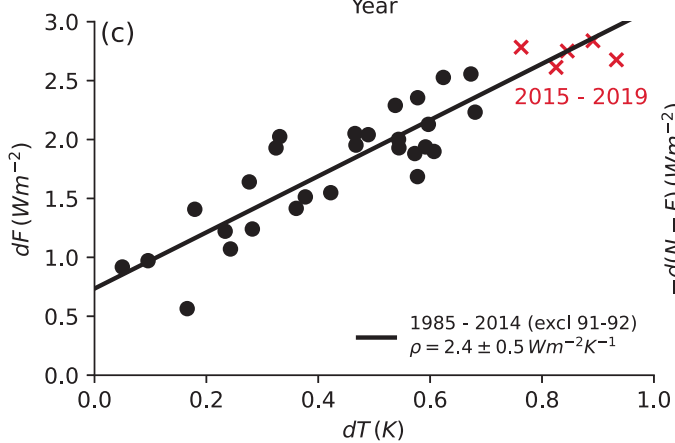
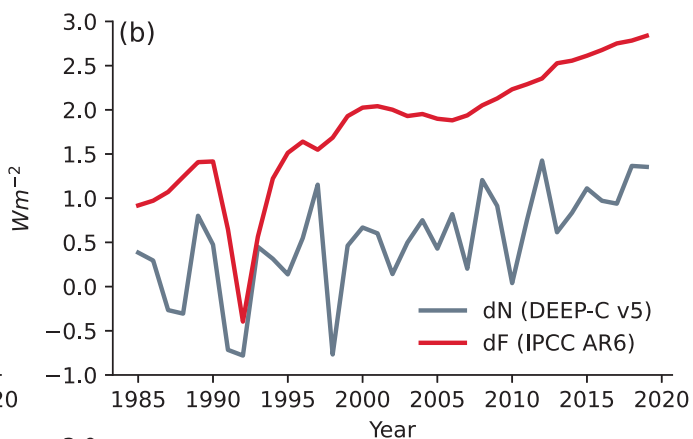
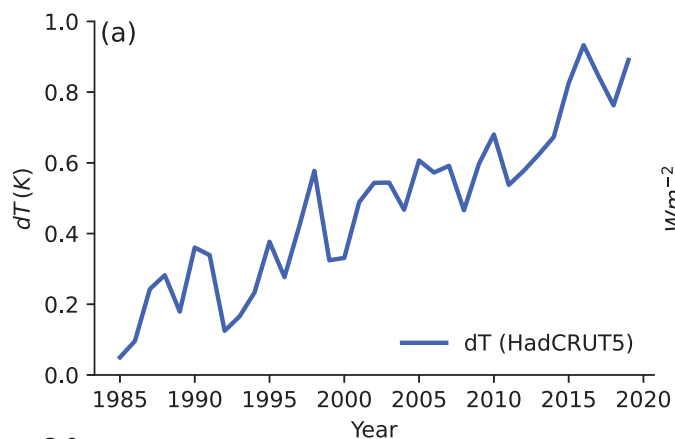


Figure 8.

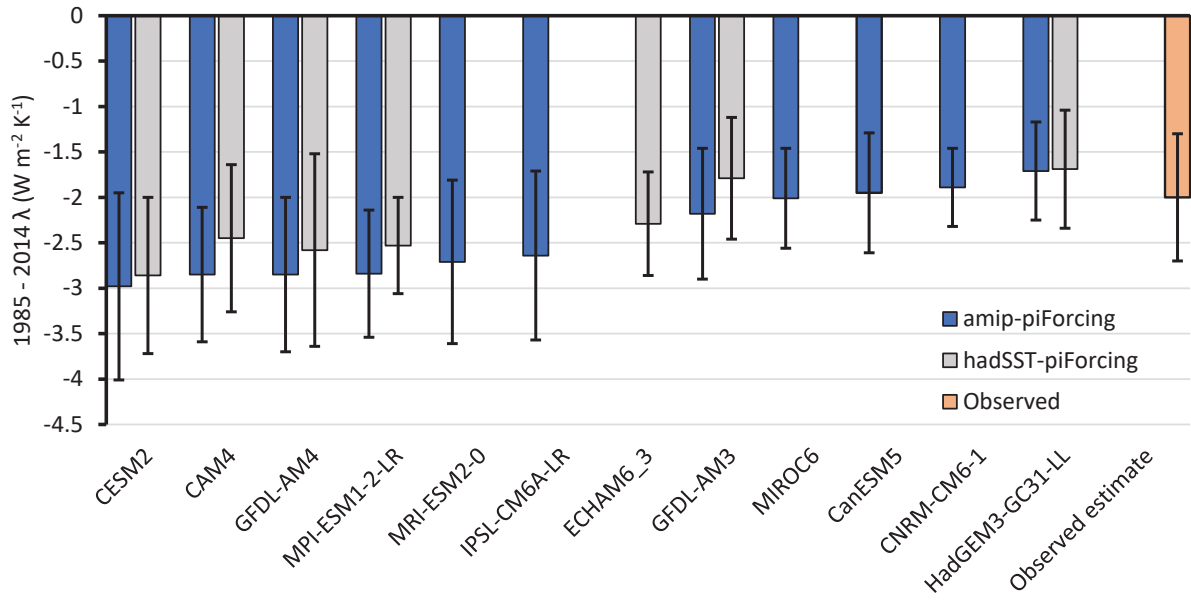
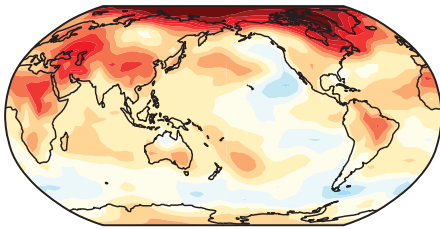
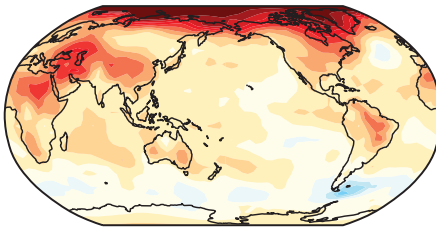


Figure 9.

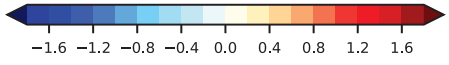
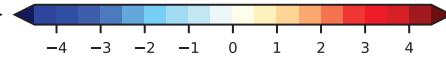
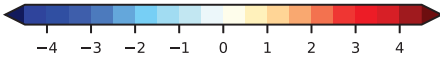
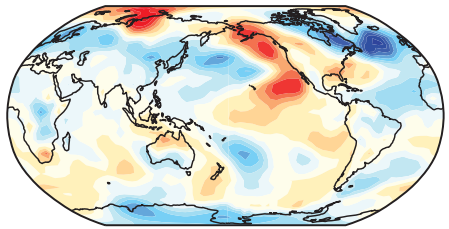
(a) 1985-2014



(b) 1987-2016



(c) Difference = (b) - (a)



HadCRUT5 surface temperature pattern (K/K)