

# Learning Climate Sensitivity from Future Observations, Fast and Slow

Adam Michael Bauer<sup>\*1,2</sup>, Cristian Proistosescu<sup>3,4</sup>, and Kelvin K Droegeemeier<sup>3,5</sup>

<sup>1</sup>*Department of Physics, University of Illinois Urbana-Champaign, Loomis Laboratory, Urbana, Illinois*

<sup>2</sup>*Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois*

<sup>3</sup>*Department of Climate, Meteorology, and Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, Illinois*

<sup>4</sup>*Department of Earth Sciences and Environmental Change, University of Illinois Urbana-Champaign, Urbana, Illinois*

<sup>5</sup>*National Center for Supercomputing Applications, Urbana, Illinois*

*Preprint – Comments Welcome*

July 21, 2025

## Abstract

Climate sensitivity has remained stubbornly uncertain since the Charney Report was published some 45 years ago. Two factors in future climate projections could alter this dilemma: (i) an increased ratio of CO<sub>2</sub> forcing relative to aerosol cooling, owing to both continued accumulation of CO<sub>2</sub> and declining aerosol emissions, and (ii) a warming world, whereby CO<sub>2</sub>-induced warming becomes more pronounced relative to climate variability. Here, we develop a novel modeling approach to explore the rates of learning about equilibrium climate sensitivity and the transient climate response (TCR) and identify the physical drivers underpinning these learning rates. Our approach has the advantage over past work by accounting for the full spectrum of parameter uncertainties and covariances, while also taking into account serially correlated internal climate variability. Moreover, we provide a physical explanation of how quickly we may hope to learn about climate sensitivity. We find that, although we are able to constrain future TCR regardless of the true underlying value, constraining ECS is more difficult, with low values of ECS being more easily ascertained than high values. This asymmetry can be explained by most of the warming this century being attributable to the fast climate mode, which is more useful for constraining TCR than it is for ECS. We further show that our inability to constrain the deep ocean response is what limits our ability to learn high values of ECS.

## <sup>1</sup> 1 Introduction

<sup>2</sup> Broadly speaking, climate sensitivity relates the amount of atmospheric carbon dioxide (CO<sub>2</sub>) to the

<sup>3</sup> amount of global average warming that will result on various timescales. Climate sensitivity can be

---

\*Corresponding author email: ambauer@uchicago.edu

4 more precisely defined depending on the timescale under consideration. On long timescales (e.g.,  
5 centuries or more), one often considers the equilibrium climate sensitivity (ECS), which quantifies the  
6 equilibrium amount of warming that would result from a doubling of atmospheric CO<sub>2</sub> relative to  
7 pre-industrial times. On near-term timescales (e.g., years to decades), the transient climate response  
8 (TCR) is more common, which quantifies the warming at the point where CO<sub>2</sub> has doubled in a  
9 simulation where CO<sub>2</sub> is increased by 1% each year and the physical climate has not yet reached  
10 equilibrium ([Intergovernmental Panel On Climate Change \(IPCC\), 2023](#)).

11 Though each of these measures of the climate response have qualitative differences, they share a  
12 common characteristic of being uncertain owing to the stochasticity and complexity of the climate  
13 system. At the most basic level, the uncertain nature of climate sensitivity, regardless of its precise  
14 definition, is the foundation of climate risk, making climate sensitivity a metric of substantive impor-  
15 tance in climate policy ([Broecker, 1987](#); [Barnett et al., 2020](#); [Bauer et al., 2024b](#)). It immediately  
16 follows that any reduction in uncertainty in each of these sensitivity measures will result improved  
17 decision-making and, hopefully, cost-savings for policymakers formulating carbon dioxide abatement  
18 policy ([Hope et al., 1993](#)).

19 Frustratingly, narrowing uncertainty in climate sensitivity, regardless of the timescale, has remained  
20 a major challenge, with the “likely” range for ECS not narrowing significantly since the Charney Report  
21 published in the 1970s ([National Research Council, 1979](#); [Roe and Baker, 2007](#); [Sherwood et al., 2020](#);  
22 [Intergovernmental Panel On Climate Change \(IPCC\), 2023](#)). One reason for this is the signal of climate  
23 change – that is, CO<sub>2</sub>-induced warming relative to preindustrial – has not been present for long enough  
24 in the observational record to place a strong constraint on ECS ([Proistosescu and Huybers, 2017](#)).  
25 Another is that this signal is confounded by a number of factors; the main factor is aerosols that have a  
26 very uncertain cooling effect on the global climate, offsetting some amount of CO<sub>2</sub> warming and making  
27 the task of ascertaining how much of the current climate signal is attributable to CO<sub>2</sub> difficult ([Knutti  
et al., 2002](#)).

29 The utility of the observational record in constraining the distribution of climate sensitivity has  
30 thus far been limited. In its sixth assessment report, the Intergovernmental Panel on Climate Change  
31 (IPCC) took a multi-pronged approach to constraining climate sensitivity, combining observations,  
32 coupled climate models, paleoclimate records, simplified climate emulators, and emergent constraints  
33 in a Bayesian framework ([Intergovernmental Panel On Climate Change \(IPCC\), 2023](#)). Within the  
34 context of this estimation framework, the observationally-inferred historical energy budget provides a

35 strong constraint on the low end of ECS, but did not meaningfully constrain the high end of the ECS  
36 distribution (Sherwood et al., 2020). This is because of both the state-dependence of climate feedbacks  
37 on the pattern of global warming (Zhou et al., 2017; Dong et al., 2019) and the confounding effects of  
38 aerosol emissions. The historical observational record can be used to indirectly constrain the high end  
39 of the ECS distribution via climate model weighting schemes that down-weight climate models that  
40 are ‘too hot’ in the historical record and generally have unrealistically high ECS values (Knutti, 2010;  
41 Liang et al., 2020; Tokarska, 2020; Ribes et al., 2021; Hausfather et al., 2022; McDonnell et al., 2024),  
42 but these approaches usually only constrain the extreme tail of the ECS distribution.

43 However, two socio-economic trends could break the current stalemate in narrowing climate sensi-  
44 tivity uncertainty using observations. Each of these trends would increase the CO<sub>2</sub>-induced warming  
45 signal-to-noise ratio, making climate sensitivity, in principle, easier to ascertain. First, a widespread  
46 trend of increasingly stringent air quality laws have caused steady declines in aerosol emissions over the  
47 last several decades (US Environmental Protection Agency Office of Air and Radiation, 2011); indeed,  
48 almost all of the shared socio-economic pathways (SSPs) project declining aerosol emissions into the  
49 future (Riahi et al., 2017). This implies that, if future air quality laws become more stringent, less  
50 CO<sub>2</sub>-induced warming will be masked by the cooling effect of aerosols, thus increasing the signal-to-  
51 noise ratio of CO<sub>2</sub>-attributable warming. The second factor is that as more CO<sub>2</sub> is emitted, the signal  
52 of CO<sub>2</sub> warming will become more pronounced relative to internal climate variability, making the sig-  
53 nal of climate change more easily discernible and bolstering our ability to constrain climate sensitivity  
54 using new observations.

55 Another important consideration relates to which climate sensitivity parameter one hopes to learn.  
56 On the one hand, because the characteristic timescales of the climate system’s thermal response can  
57 exceed two hundred years (Geoffroy et al., 2013a; Leach et al., 2021), near-term warming information  
58 is likely to be of limited use in improving our estimates of ECS (as the “slow” mode of warming will  
59 not be manifest for perhaps hundreds of years). On the other hand, improving our estimates of TCR  
60 may yet be possible, as the near-term scope of TCR makes near-term warming information potentially  
61 useful, especially in light of the two projected socio-economic trends mentioned above. This important  
62 dynamic trade-off between when the short- and long-term timescales of warming manifest was explored  
63 by Roe and Baker (2007) and Baker and Roe (2009).

64 Past studies have used synthetic future observations of the climate, either generated with an energy  
65 balance model (EBM) or general circulation models (GCM), to constrain estimates of climate sensitivity

66 using Bayesian frameworks (Kelly and Kolstad, 1999; Leach, 2007; Padilla et al., 2011; Ring and  
67 Schlesinger, 2012; Urban et al., 2014; Kelly and Tan, 2015; Mori and Shiogama, 2018). Although the  
68 specifics of each framework in the literature varies, a common thread is to jointly estimate a subset of  
69 the EBM parameter space (e.g., joint uncertainty in the climate feedback, ocean diffusivity, and the  
70 sensitivity of aerosol forcing, as done in Ring and Schlesinger (2012)) to estimate a “learning rate”  
71 of climate sensitivity once future observations are incorporated into the assimilation window. Most  
72 studies find that learning about climate sensitivity is possible by the end of the century, although  
73 some suggest that uncertainty may widen before it begins to narrow (Hannart et al., 2013). Another  
74 study focused on projecting the decline in anthropogenic radiative forcing uncertainty to constraint  
75 the TCR uncertainty, finding that TCR uncertainty could be reduced by as much as 50% in the  
76 near future (Myhre et al., 2015). Finally, the Allen, Stott and Kettleborough (ASK) method (Allen  
77 et al., 2000; Stott and Kettleborough, 2002; Gillett et al., 2012; Stott et al., 2013) has been used to  
78 constrain future temperature rise conditional on sequentially assimilated observations, but a recent  
79 application did not explicitly tie narrowing uncertainty in temperature rise to learning about the  
80 underlying distribution of ECS or TCR (Shiogama et al., 2016).

81 While the above studies have shed light on the rate at which we narrow uncertainty in climate  
82 sensitivity (which we coin “climate sensitivity learning rates” throughout), they share a number of  
83 limitations. First, current state-of-the-art simplified EBMs (such as FaIR (Leach et al., 2021) or  
84 MAGICC (Meinshausen et al., 2011) that are used in the IPCC’s estimation methodology mentioned  
85 above) are calibrated by treating every parameter in the EBM as uncertain, not just a few that  
86 are deemed to be the most vital (Cummins et al., 2020). This implies that different methodologies  
87 are being used by those assessing climate sensitivity using EBMs based on past observations (see,  
88 e.g., Chapter 4 in Intergovernmental Panel On Climate Change (IPCC), 2023) and those making  
89 predictions about how climate uncertainty will evolve in time. As a concrete example, we are not  
90 currently aware of any study on the learning rate of climate sensitivity that treats the deep ocean heat  
91 capacity as uncertain, despite the known difficulty in constraining it (see, for example, Geoffroy et al.,  
92 2013a). By neglecting certain EBM parameters in the parameter estimation problem, past work may  
93 have considerably underestimated the underlying covariance structure between many EBM parameters,  
94 which could significantly change climate learning rate estimates both quantitatively and qualitatively.

95 Second, all previous studies treat the sensitivity of radiative forcing to CO<sub>2</sub> concentrations as given,  
96 despite known uncertainty in the magnitude of the response of forcing to an injection of CO<sub>2</sub> into the

97 atmosphere (Zelinka et al., 2020; Intergovernmental Panel On Climate Change (IPCC), 2023; He et al.,  
98 2023). Past work therefore assumes all uncertainty in ECS comes from uncertainty in the climate  
99 feedbacks, despite radiative forcing sensitivity uncertainty strongly trading-off with climate feedback  
100 uncertainty in estimates of ECS uncertainty. Capturing this trade-off would likely alter the qualitative  
101 features of climate learning rates.

102 Lastly, though some studies have pointed to the different characteristic timescales of the climate  
103 response to warming as an important factor for climate learning rates (Urban et al., 2014), explicitly  
104 decomposing the learning rates of the different climate modes, and connecting them to the overall  
105 climate sensitivity learning rate, remains lacking. This analysis would provide physical insights into why  
106 and how we may expect to learn about the climate in the future rather than relying on computational  
107 models alone.

108 In light of the limitations and challenges described above, this paper presents a modeling framework  
109 that merges an IPCC-consistent climate emulator with ensemble variational data assimilation to assess  
110 the prospects and perils of learning ECS and TCR from hypothetical future observations of global  
111 average temperature and ocean heat content. Our framework utilizes a weak-constraint, ensemble of  
112 data assimilations (EDA) approach similar to those used in the weather forecasting literature (see, e.g.,  
113 Isaksen et al., 2010) to incorporate “pseudo-observations” of the future climate to estimate the posterior  
114 distribution of climate sensitivity and other climate parameters conditional on future observations (see  
115 *Materials and Methods*).

116 Our modeling framework has two distinct advantages over past work. The first is we are able to  
117 jointly estimate the conditional probability distribution of each parameter in our EBM, as opposed to  
118 a subset of our parameter space. This allows us to capture trade-offs between CO<sub>2</sub> radiative forcing  
119 sensitivity and climate feedbacks on the posterior distribution of ECS, while also accounting for the  
120 impact of difficult-to-learn parameters, such as the deep ocean heat capacity, on our estimates of  
121 learning rates. As we will see later, this has an especially substantive impact on the rate we are able  
122 to learn ECS from future observations. The second is that, by using a so-called “weak-constraint”  
123 variational framework (WC-VAR, see Evensen et al., 2022, and *Materials and Methods*), we are able  
124 to account for the trade-off between estimating the climate *state* – that is, what the true climatic  
125 temperature is in the absence of internal climate variability – and estimating the internal climate  
126 variability itself (Nicklas et al., 2024), a novelty in the climate learning rates literature.

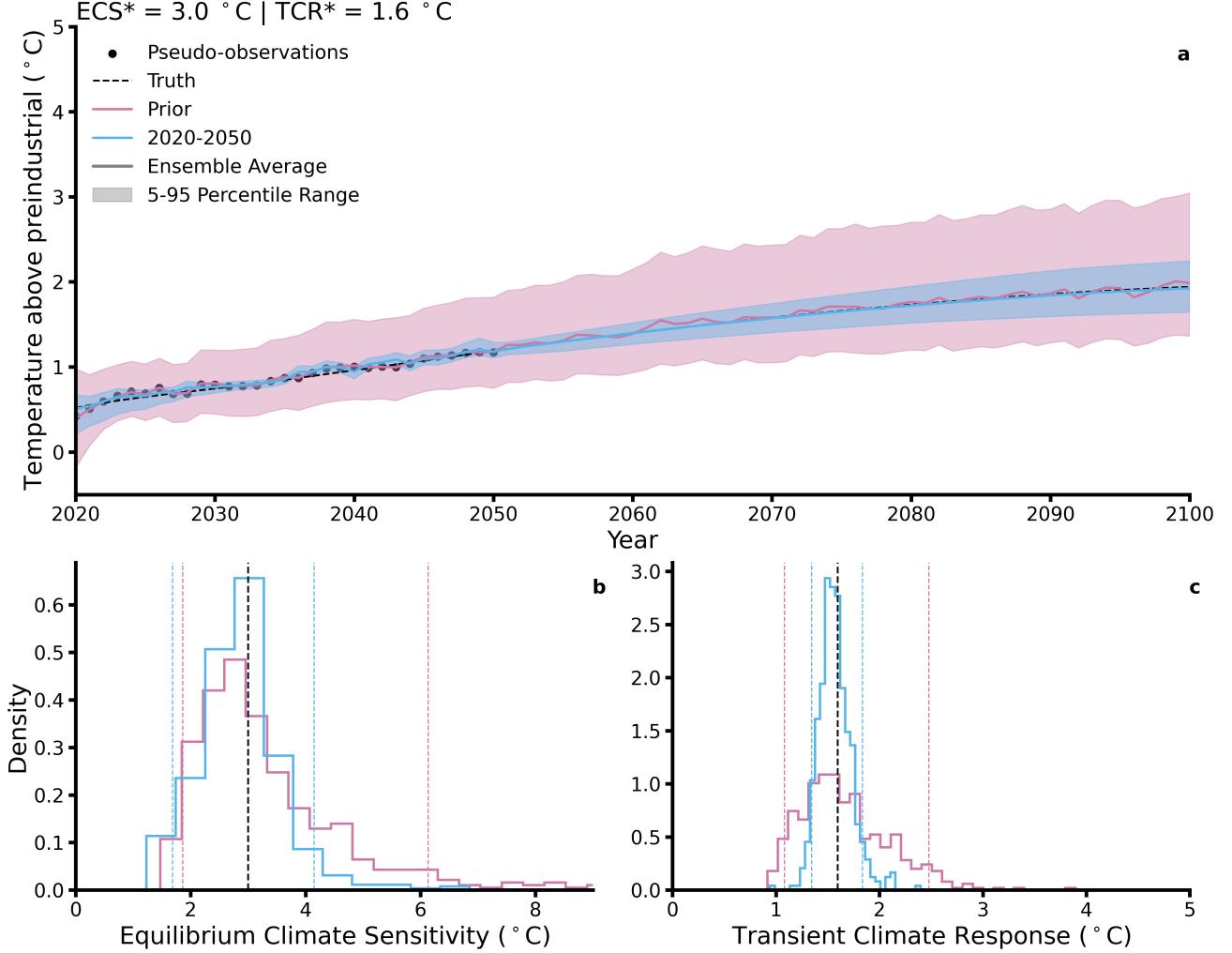
127 **2 Results**

128 We begin by demonstrating how the forecasted global surface temperature and the distributions of  
129 ECS and TCR change as we add observations of global temperature and ocean heat content into  
130 our assimilation window (see *Materials and Methods*). We generate pseudo-observations for 2020–  
131 2050 by forcing our two-layer model (Geoffroy et al., 2013a, and *Materials and Methods*) with CO<sub>2</sub>  
132 concentrations and sulfur dioxide (SO<sub>2</sub>) emissions from SSP2–4.5 (Riahi et al., 2017). We compute the  
133 effective radiative forcing using the parameterizations in Leach et al. (2021). We then treat the time  
134 period of 2020–2050 as our “analysis” period (Park and Zupanski, 2022), over which we minimize our  
135 WC-VAR cost function for each of our 500 ensemble members. Because each ensemble member cost  
136 function is independent, this procedure results in randomized maximum likelihood (RML) sampling of  
137 the posterior conditional probability distribution function (PDF) (Evensen et al., 2022). We then use  
138 the posterior distribution to forecast future temperature change.

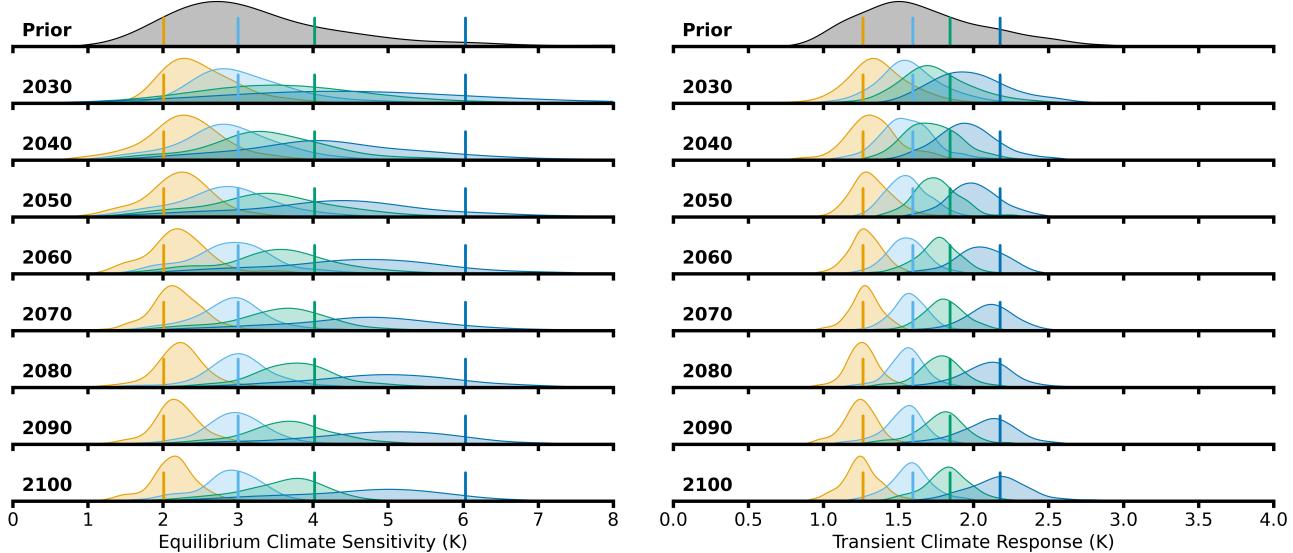
139 In Figure 1, we show the projected future temperature rise using the prior distribution (pink), and  
140 the forecasted future temperature rise using the posterior parameter distributions from the 2020–2050  
141 assimilation window (blue) for a true values of ECS of 3 °C and a true TCR of 1.6 °C (Fig. 1a). Note  
142 these values are equal to our ensemble prior central estimates; we will probe the dependence of our  
143 results on what the true underlying value is later. We also compute the posterior distributions of ECS  
144 and TCR for the prior distribution and the parameter posterior (Figs. 1b–c).

145 We find the estimated range of 2100 warming levels is significantly reduced after observations  
146 of 2020–2050 temperature and ocean heat content are assimilated. Using our parameter priors, the  
147 end-of-century warming 5–95 percentile range is 1.6 °C. After incorporating thirty additional years  
148 of observations into our assimilation window, this range reduces to about 0.5 °C (a 66% reduction  
149 relative to the prior) when the true value of ECS is 3 °C. This suggests that future observations may  
150 be able to place a relatively tight constraint on end-of-century temperature rise, consistent with past  
151 work (Shiogama et al., 2016).

152 Our ability to constrain end-of-century temperature rise may seem natural given that, in 2050,  
153 our proximity to the end of the century is increased, and so the range of possible outcomes has less  
154 time to “spread out”. However, our results would suggest that the narrowing of the end-of-century  
155 temperature rise 5–95 percentile range is also in large part due to an improved ability to estimate  
156 model parameters which, in turn, constrain ECS and TCR (see Figs. 1b–c and corner plots in the



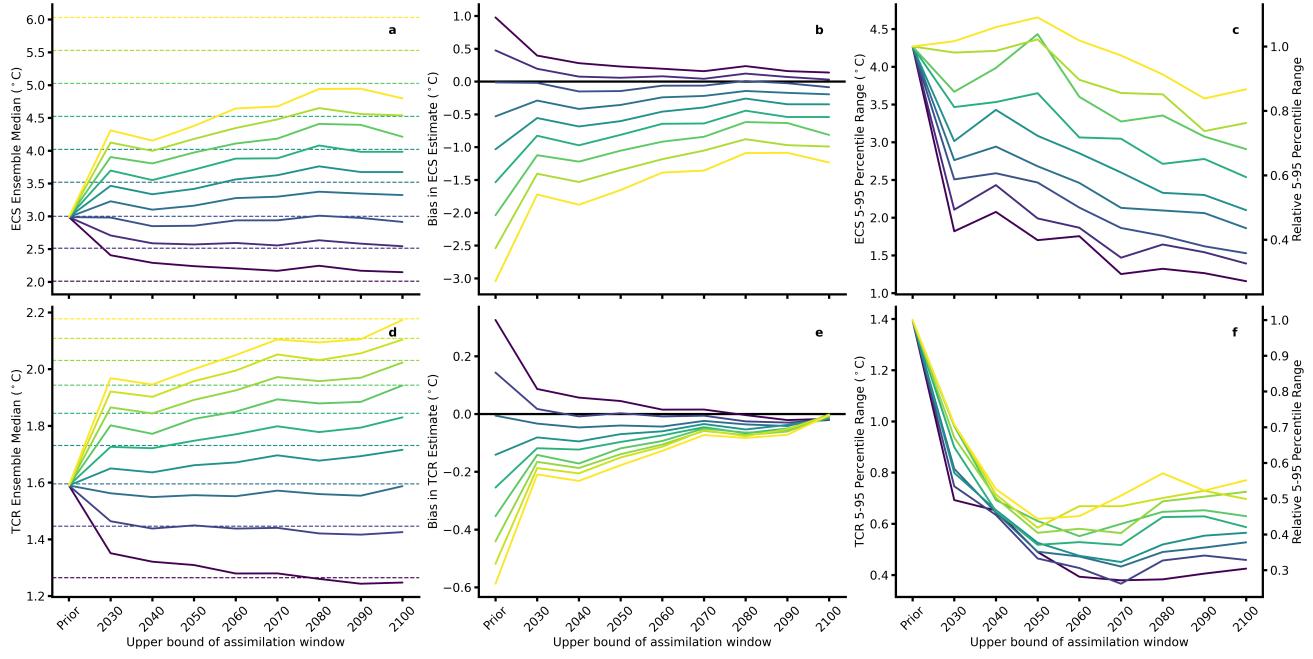
**Figure 1: Temperature forecasts, ECS, and TCR posteriors conditional on observations up to 2050.** **a** shows the forecast of future temperatures using parameter posteriors from the 2020-2050 assimilation window (blue) and using the prior distribution (pink) where the true values of ECS and TCR are 3  $^{\circ}\text{C}$  and 1.6  $^{\circ}\text{C}$ , respectively. The shaded regions are the 5-95 percentile range, and future temperature pseudo-observations are shown in the black dots. **b** shows the prior ECS histogram (pink) and the 2020-2050 assimilation window ECS posterior (blue), with the black vertical line representing the true ECS value. **c** is as **b**, but for TCR. Pseudo-observations are generated using SSP2-4.5, and in **b-c**, dashed lines represent the 5-95 percentile range.



**Figure 2: Evolution of ECS and TCR posteriors.** On the left, we show how the posterior distribution of ECS evolves for different true values of ECS (solid lines) as the upper bound of our assimilation window is pushed further into the future (and more observations are added to our assimilation window). The right is as the left, but for TCR.

157 *Supplementary Information).* We find that by mid-century, our approach provides a fairly robust  
 158 constraint on the central value of both ECS and TCR. In particular, we estimate the ECS ensemble  
 159 median to be  $2.86^{\circ}\text{C}$  (5% error) and estimate a TCR ensemble median of  $1.56^{\circ}\text{C}$  (3% error). Not  
 160 only are the ECS and TCR distribution medians estimated relatively well, but we also find nontrivial  
 161 reductions in uncertainty: the 5–95 percentile range the ECS distribution decreases by 42%, while the  
 162 5–95 percentile range of the TCR distributions decreases by 65%. These results – the narrowing of the  
 163 uncertainty band for end-of-century temperature rise, ECS, and TCR – would tentatively suggest that  
 164 there is some reason for hope in constraining uncertainty in these quantities in the future.

165 Are these results robust regardless of whether we live in an extremely high or low ECS or TCR  
 166 world? And how do these results change as more (or less) observations are incorporated into the  
 167 analysis window? To answer these questions, we carry out the same exercise demonstrated by Figure 1  
 168 for different true values of ECS and TCR and display the ECS and TCR posteriors in Figure 2 for  
 169 numerous analysis window upper bounds. We find that the degree to which one can ascertain the true  
 170 value of ECS by the end of the century varies significantly with what the true underlying value is.  
 171 For low values of ECS (yellow lines and distributions on the left in Figure 2), we find that the ECS is  
 172 learned efficiently by the end of the century, with very little difference in the posterior central estimate  
 173 and the true value (6.8% error) and a relatively tight 5–95 percentile range ( $1.16^{\circ}\text{C}$ ). However, as we



**Figure 3: ECS and TCR posterior central estimates and learning rates.** **a** shows the central estimate of the posterior ECS distribution (solid lines) and the true value of the ECS used to generate the observations (dashed lines) as the assimilation window is moved further into the future (and more observations are used for analysis). **b** shows the difference between the posterior central estimate and the true value, while **c** shows the raw posterior 5-95 percentile range and the 5-95 percentile range relative to the prior (left and right axes, respectively). **d–f** are as **a–c**, but for TCR.

increase the true value of ECS, our approach grows increasingly unable to learn the true value; for a high value of ECS ( $\sim 5$  °C, navy lines and distributions in Figure 2), the percent error between the end of century posterior central estimate and the true value is about 16.2%, while posterior distribution remains highly uncertain (5-95 percentile range of 2.9 °C).

However, we find that the persistence of uncertainty and bias in posterior central estimates is only present for ECS. For TCR, even extremely high values are able to be learned by the end-of-century, with uncertainty narrowing at a similar rate for each true values of TCR we consider (see the right side of Figure 2). The percent difference between the true value and the end of century posterior central estimate range from 0.19%-1.45%, a far smaller range than that of ECS. This would suggest that while the prospects for ascertaining *equilibrium* climate metrics, such as ECS, may depend on the underlying true value, *transient* climate metrics, like TCR, do not. Corner plots in the *Supplementary Information* show that high climate sensitivity experiments struggle to disentangle the trade-offs between the climate feedback strength and the CO<sub>2</sub> radiative forcing sensitivity (estimating a too-strong forcing and a too-strong feedback), while low climate sensitivity experiments estimate both parameters effectively.

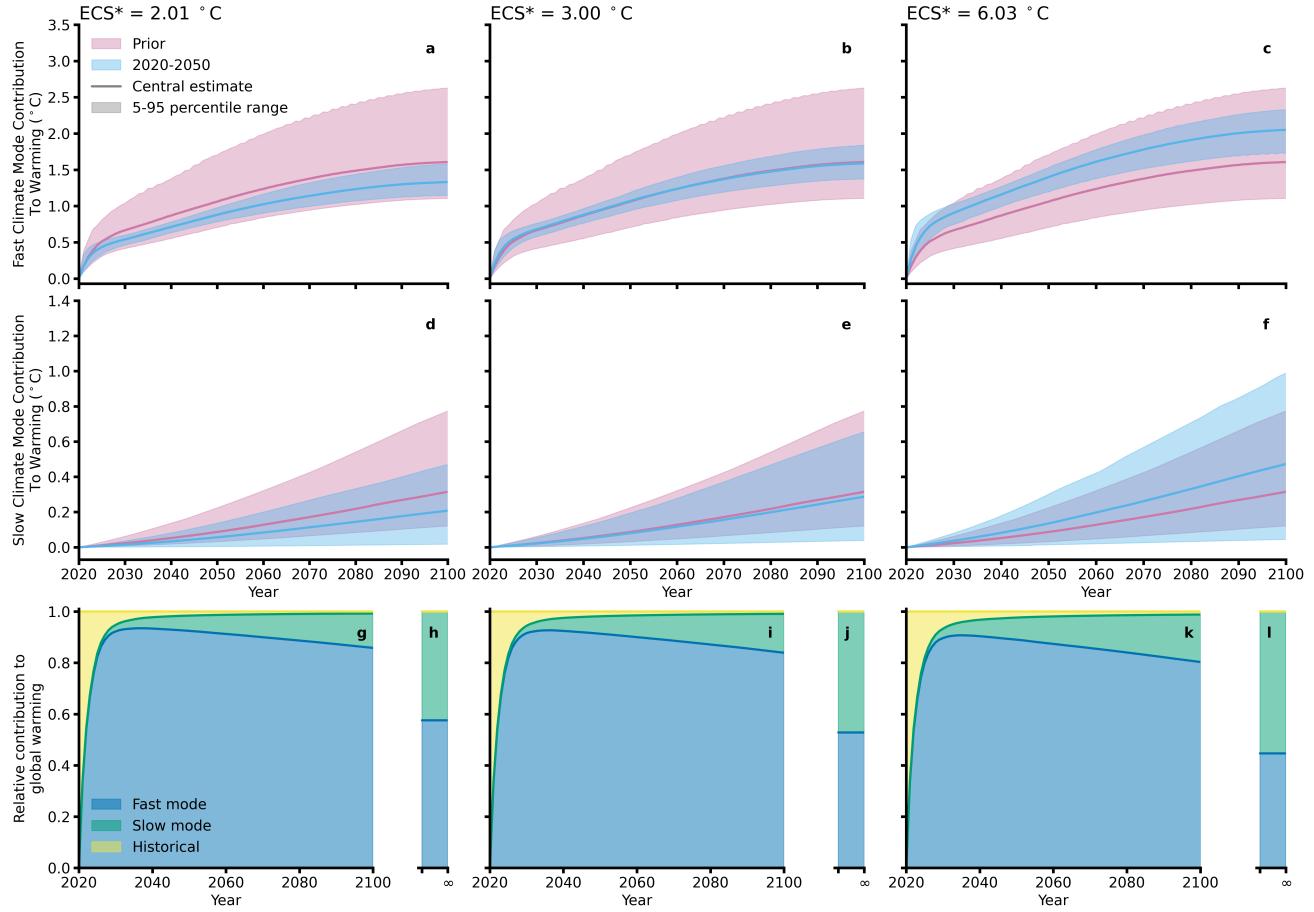
We can summarize the above findings by showing the evolution of the posterior central estimate

189 and posterior 5-95 percentile range as more observations are assimilated in Figure 3 for numerous true  
190 values of ECS and TCR. Figure 3 supports the findings of Figure 2: high values of ECS are unable to  
191 be robustly estimated even with new observations through the end of the century (see Figs. 3a, b),  
192 while the bias in the central estimate of TCR becomes virtually negligible by 2100 (Fig. 3e).

193 We further find that the ECS 5-95 percentile range decreases at a much slower rate (in relative  
194 terms) than the TCR 5-95 percentile range (Figs. 3c, f). Indeed, we find that the degree to which  
195 uncertainty decreases in the TCR is mostly independent of the true value; for each true value we  
196 consider, the 5-95 percentile range is 33%-44% that of the prior once observations from 2020-2050  
197 are assimilated, and is roughly constant thereafter. The ECS posterior, on the other hand, has a  
198 5-95 percentile range between 40%-109% of the prior 5-95 range once observations from 2020-2050  
199 are assimilated, and continues to decline somewhat thereafter. There is also significant stratification  
200 between the different true values of ECS; in other words, our ability to narrow uncertainty in ECS  
201 is more limited when ECS is high than when it is low. Hence, not only is our ability to match the  
202 ensemble central estimate and the true value different for when we are estimating ECS or TCR, but  
203 the rate at which we decrease uncertainty depends on if we are considering an equilibrium or transient  
204 climate metric as well.

205 Why can we learn TCR faster and more robustly than ECS, both in terms of the central estimate  
206 and the rate of uncertainty decline? We postulate that the physical intuition behind the differing  
207 learning rates displayed in Figure 3 manifests in the relative amounts of warming attributed to the  
208 fast and slow climate modes this century. Using our parameter prior and posterior after assimilating  
209 observations from 2020-2050 (*à la* Figure 1), we compute the contribution to temperature rise from  
210 the fast and slow climate modes in Figure 4a-f for low, average, and extreme values of ECS (see the  
211 *Materials and Methods* for details on our decomposition) when forced with SSP2-4.5. Contributions  
212 from historical warming are negligible past a decade or two, and have been neglected for our discussion  
213 here. We compute the relative fraction of warming from each climate mode in Figure 4g-i.

214 We find that most of the uncertainty reduction from thirty years of observations (comparing the  
215 pink and blue envelopes in Figure 4a-f) comes from constraining the fast climate mode. Comparatively,  
216 slow mode uncertainty is recalcitrant, experiencing very little reduction in uncertainty (we even find a  
217 minor increase in uncertainty in slow mode-contributed warming when ECS is large). These patterns  
218 are consistent regardless of whether or not the true underlying ECS is low or high. Additionally, the  
219 difference in uncertainty reduction between the fast and slow modes happens even while the central



**Figure 4: Fast and slow mode decomposition of 21<sup>st</sup> century warming and learning rates.** a–c show the contribution to global warming from the fast climate mode for three different true ECS values (see the titles). d–f is as a–c, but for the slow climate mode. g, i, k show the relative fraction of warming between the fast mode, slow mode, and historical warming for each true value of ECS. h, j, l show the same relative contribution to warming as g, i, k, but after the climate system has reached equilibrium.

220 estimate of warming attributable to each mode changes to reflect the true climate sensitivity.

221 The difference in uncertainty reduction for the fast and slow climate mode can be explained by how  
222 much near-term warming is attributable to the different modes of warming. We find that, in the 80  
223 years following the beginning of our simulations, most of the warming found in our pseudo-observations  
224 is attributable to the fast climate mode, regardless of the true value of ECS (panels 4g–i). Physically,  
225 this is reflective of the vastly different characteristic timescales of adjustment between the atmosphere  
226 and upper mixed ocean layer versus the deep ocean; the mixed ocean layer has an average characteristic  
227 timescale of  $\sim$ 10 years, whereas the deep ocean has a characteristic timescale of  $\sim$ 200 years (Geoffroy  
228 et al., 2013a). Because warming attributed to the deep ocean manifests over far longer timescales,  
229 near-term information on temperature rise is more useful for constraining the fast climate mode, which  
230 is the primary contributor to near-term temperature changes and, in turn, the estimated TCR and end  
231 of century temperature change (Baker and Roe, 2009).

232 This difference in characteristic timescales between the fast and slow mode of warming can further  
233 explain the bias we observed in our central estimates of high values of ECS in Figures 2 and 3. We  
234 found that if the true underlying value of ECS is high, our central estimate of ECS maintains a lingering  
235 bias, even when pseudo-observations are assimilated through the end of the century. This was not the  
236 case for low values of ECS. The difference between cases where ECS is high or low is related to how  
237 much warming is attributable to each climate mode *in equilibrium*, see Figure 4g–l. For low values  
238 of ECS, the amount of warming attributable to the slow mode in equilibrium is about 42%, whereas  
239 for high values of ECS, the relative amount of warming is about 56%. This means that for low values  
240 of ECS, the deep ocean accounts for less overall warming than the fast mode even after the climate  
241 system has equilibrated, whereas the when ECS is high, the deep ocean accounts for the majority of  
242 the equilibrium warming. (We prove this explicitly in the *Materials and Methods*.) Hence, our ability  
243 to constrain the fast mode of warming when ECS is low contributes to a tighter constraint on *both*  
244 ECS and TCR, while for high values of ECS, the fast mode can only partially constrain ECS while still  
245 tightly constraining TCR.

246 These findings imply that no matter what the true value of ECS is, the information from the next  
247 50 years will be primarily useful for constraining the fast climate mode. In turn, information on the  
248 fast climate mode will be primarily useful for constraining transient climate metrics, such as the TCR,  
249 and only partially useful for constraining ECS. The remaining information to constrain ECS – that  
250 is, the warming attributable to the slow mode – will not arrive until hundreds of years hence when

251 the climate system approaches equilibrium, implying that prospects for constraining ECS based on  
252 future observations remain limited from a purely physical perspective, and is corroborated by our data  
253 assimilation approach.

### 254 3 Discussion

255 Our work highlights the key underlying factors important for constraining climate sensitivity from new  
256 observations of global temperature and ocean heat content. Our key takeaway is that the prospects for  
257 constraining climate sensitivity critically depends on if one is aiming to constrain an equilibrium or a  
258 transient climate metric.

259 For equilibrium climate sensitivity, we find that our ability to constrain ECS varies depending on  
260 the true underlying value; low values of ECS may be constrained by century's end, while high values  
261 of ECS are unable to be efficiently estimated even after an additional 80 years of pseudo-observations.  
262 Likewise, uncertainty in the distribution of ECS will remain wide for high values of ECS, while for low  
263 values of ECS uncertainty may be narrowed by mid-century. The core reason for this asymmetry in  
264 the learning rate of ECS for different true values lies in two physical characteristics of the climate: the  
265 slow ocean response timescale, which limits the utility of near-term information for constraining ECS,  
266 and the fact that the slow mode accounts for a large (small, resp.) share of equilibrium warming when  
267 ECS is larger (smaller, resp.) than average. In this way, our work would suggest that if the true value  
268 of ECS is large, not only will climate damages be more acute (because temperatures are higher than if  
269 ECS is low), but climate risk will also be persistent through the end of the century.

270 On the other hand, our work would suggest that transient climate metrics, such as end-of-century  
271 temperature rise and the transient climate response, may be estimated with some confidence by end  
272 of century. We find that the central estimate of TCR has very little bias by the end of the century  
273 regardless of the true underlying value, and uncertainty narrows at a relatively rapid rate. This can be  
274 explained by end-of-century temperature rise and TCR being primarily controlled by the fast climate  
275 mode, which can be tightly constrained in the near-term as CO<sub>2</sub>-induced warming rises and aerosol  
276 emissions decline. One may therefore be optimistic about narrowing TCR uncertainty in the coming  
277 years, despite the prospects for constraining ECS being less easy to judge.

278 One limitation of our analysis is that we treat our climate feedback parameter and the sensitivity  
279 of radiative forcing on CO<sub>2</sub> concentrations as independent from the underlying climate state. However,

it has been shown that both of these processes are dependent on the underlying pattern of warming (Zhou et al. (2017), Armour (2017), and Dong et al. (2019) explore the spatial dependence of climate feedbacks, while He et al. (2023) explores the state dependence of radiative forcing). In general, it is difficult to capture these effects in global, simplified climate models given their spatial dependence, though we do include an efficacy parameter in our energy balance model that modulates the top of atmosphere energy imbalance to approximate them (Held et al., 2010; Cummins et al., 2020; Leach et al., 2021). Including state dependent feedbacks would likely lower our learning rates presented here, given that the transient patterns of warming would obfuscate or ability to learn the “true” underlying climate sensitivity during the adjustment period. We note that our framework is amenable with a dynamic, time-dependent climate feedback or radiative forcing sensitivity, and future work could incorporate these effects to evaluate their impact on climate learning rates. We expect that while the our quantitative results would be impacted by including these effects, our core physical insight (that transient metrics can be learned more effectively than equilibrium metrics) will still hold in the modified framework.

Our findings have direct implications for climate risk assessment and policymaking. At a technical level, our modeling framework puts forth a methodology to update uncertainty as more pseudo-observations are incorporated into climate model projections. One can image merging our framework with a scenario-based approach to climate risk management in the insurance or financial sector, which would allow climate risk can evolve endogenously with the emissions scenario under consideration, enabling for a more holistic assessment of climate risk in decision-making (see, e.g., Neubersch et al., 2014).

For policymakers, the rate at which one can learn about the climate system – often coined “active learning” – has been shown to be important for estimates of the social cost of carbon (SCC) (Kelly and Kolstad, 1999; Leach, 2007; Kelly and Tan, 2015; Lemoine and Rudik, 2017). Our results suggest that the degree to which active learning plays a role in policymaking depends on whether one is considering long- or short-run climate policies. Long-run climate policies often rely on ECS as the key source of climate risk (Weitzman, 2012), though recently some models utilize the transient climate response to emissions to map cumulative CO<sub>2</sub> emissions to temperature rise (Allen et al., 2009; Campiglio et al., 2022). Short-run policies, on the other hand, often rely on carbon budgets as their main source of climate uncertainty (e.g., Bauer et al., 2024a). Given that the remaining carbon budget is a transient climate metric, it is likely that the quantitative implications of active learning will play a larger role

311 for carbon budget-based policies than ECS-based policies.

312 Overall, our analysis suggests that while uncertainty will remain persistent despite new information  
313 on the climate over the coming years, we can expect some decline in uncertainty as more observations  
314 of the CO<sub>2</sub>-warming-dominated climate are assimilated into climate models. Our results provide a  
315 physical underpinning for why transient climate metrics are learned at a disproportionately fast rate  
316 compared to equilibrium climate metrics. However, our results are meant to be a stylized demonstration  
317 of this fact, and many climatic processes likely will obscure learning from happening as quickly as we  
318 have demonstrated here. For example, our quantitative results would need to be reassessed if, say,  
319 climate change triggers fundamental process-level changes that are not included in our two-layer box  
320 model (e.g., accelerated permafrost melt). Despite of these limitations, our results suggests that there  
321 are fundamental physical reasons to expect that we may be able to constrain climate sensitivity from  
322 new observations in the future, and that, depending on the climate metric of interest, these constraints  
323 may be substantial.

## 324 4 Materials and Methods

### 325 4.1 Summary of our modeling approach

326 We utilize pseudo-observations and an ensemble of data assimilations (EDA) approach to estimate  
327 the rate at which climate sensitivity, and other climate metrics, are learned conditional on future  
328 observations of the global climate. Using a large ensemble (in our case, 500 members), we minimize  
329 an ensemble of cost functions that allows us to sample the posterior distribution of model states,  
330 parameters, and errors conditional on observations of the climate. Throughout, we use a two-layer  
331 climate emulator ([Geoffroy et al., 2013a](#)) to compute the global surface temperature and ocean heat  
332 content, and use the forcing representation of FaIR v2.0.0 ([Leach et al., 2021](#)) to compute the radiative  
333 forcing of each of our forcing agents (e.g., aerosols and greenhouse gases). Our model is forced with CO<sub>2</sub>  
334 concentrations time series and SO<sub>2</sub> emissions time series from SSP2–4.5 (and SSP5–8.5 in a sensitivity  
335 analysis) ([Riahi et al., 2017](#)).

336 We apply our modeling approach in a so-called “perfect model” context ([Urban et al., 2014](#)), where  
337 we generate synthetic observations of the past and future using the same model we attempt to fit the  
338 data with. The only difference in the “perfect model” case is that the observations are forced with a

339 known stochastic component, whereas the retrieved model fit is forced with an unknown forcing that  
 340 we treat as a control variable (as is standard in WC-VAR). We vary the true underlying value of ECS  
 341 and TCR by altering the climate feedback parameter,  $\lambda$ , which allows us to probe how the learning  
 342 rate about climate sensitivity changes for different true values of ECS or TCR.

343 **4.2 2-layer energy balance model**

344 We consider a stochastic two-layer energy balance model (EBM) to emulate the global climate ([Ge-](#)  
 345 [offroy et al., 2013a,b](#)). We force the EBM with an exogenous path of CO<sub>2</sub> concentrations and sulfur  
 346 dioxide (SO<sub>2</sub>) emissions, each of which are mapped to an effective radiative forcing through the pa-  
 347 rameterization of [Leach et al. \(2021\)](#). In general, if the species of molecules whose radiative forcing  
 348 are specified by concentrations (emissions, resp.) are given by the set  $\mathcal{C}$  ( $\mathcal{E}$ , resp.) we can write the  
 349 cumulative radiative forcing at a time  $t$  as

$$\begin{aligned}\mathcal{F}_t = & \sum_{c \in \mathcal{C}} f_1^{(c)} \ln \left( \frac{C_t^{(c)}}{C_0^{(c)}} \right) + f_2^{(c)} \left( C_t^{(c)} - C_0^{(c)} \right) + f_3^{(c)} \left( \sqrt{C_t^{(c)}} - \sqrt{C_0^{(c)}} \right) \\ & + \sum_{e \in \mathcal{E}} f_1^{(e)} \ln \left( 1 + \frac{E_t^{(e)}}{C_0^{(e)}} \right) + f_2^{(e)} E_t^{(e)}\end{aligned}\quad (4.1)$$

350 The radiative forcing given by (4.1) forces the two layer model given by

$$C_1 \frac{dT_t^{(1)}}{dt} = \mathcal{F}_t - \lambda T_t^{(1)} + \varepsilon \gamma \left( T_t^{(2)} - T_t^{(1)} \right) + q_t \quad (4.2)$$

$$C_2 \frac{dT_t^{(2)}}{dt} = \gamma \left( T_t^{(1)} - T_t^{(2)} \right) \quad (4.3)$$

351 where layer  $i$  has a heat capacity  $C_i > 0$  with  $i \in \{1, 2\}$ ,  $\lambda > 0$  is the climate feedback,  $\gamma > 0$  is the  
 352 transfer coefficient between layers,  $\varepsilon > 0$  is the ocean heat uptake efficacy, and  $q_t$  are model errors.  
 353 When we use (4.2)-(4.3) to generate noisy synthetic observations,  $q_t = \varepsilon_t$ , where  $\varepsilon_t$  is a stochastic  
 354 forcing term that represents internal climate variability. We draw vectors of internal variability such  
 355 that  $\vec{\varepsilon} \sim \mathcal{N}(\vec{0}, \mathbf{I})$ , where  $\mathbf{I}$  is the temporal covariance matrix of internal climate variability. In general,  
 356 the time average of  $q_t$  represents model bias, and the temporal covariance represents the covariance  
 357 structure of internal climate variability.

358 Importantly, from the integrated paths of (4.2)-(4.3), we can compute the ocean heat content,  $Q_t$ ,

359 as

$$Q_t = C_1 T_t^{(1)} + C_2 T_t^{(2)}. \quad (4.4)$$

360 The ocean heat content is a commonly assimilated observation in estimating climate parameters owing  
 361 to its dependence on the deep ocean temperature,  $T_t^{(2)}$ . We can then write an amended two layer model  
 362 where the ocean heat content is also integrated, as

$$C_1 \frac{dT_t^{(1)}}{dt} = \mathcal{F}_t - \lambda T_t^{(1)} + \varepsilon \gamma \left( T_t^{(2)} - T_t^{(1)} \right) + q_t \quad (4.5)$$

$$C_2 \frac{dT_t^{(2)}}{dt} = \gamma \left( T_t^{(1)} - T_t^{(2)} \right) \quad (4.6)$$

$$\frac{dQ}{dt} = \mathcal{F}_t - \lambda T_t^{(1)} + (\varepsilon - 1) \gamma \left( T_t^{(2)} - T_t^{(1)} \right) + q_t. \quad (4.7)$$

363 We discretize (4.5)-(4.7) using a standard forward Euler scheme.

#### 364 4.2.1 Eigenmode decomposition

365 Following Geoffroy et al. (2013a), we can use eigenmode decomposition to write the generalized ana-  
 366 lytical solution to (4.2)-(4.3). Notice we may rewrite the system as

$$\frac{d\mathbf{T}_t}{dt} = \mathbf{D}\mathbf{T}_t + \mathbf{F}_t, \quad (4.8)$$

367 with

$$\mathbf{T}(t) = \begin{pmatrix} T^{(1)}(t) \\ T^{(2)}(t) \end{pmatrix}, \quad (4.9)$$

$$\mathbf{D} = \begin{pmatrix} -(\gamma\varepsilon + \lambda)/C_1 & \gamma\varepsilon/C_1 \\ \gamma/C_2 & -\gamma/C_2 \end{pmatrix}, \quad (4.10)$$

369 and

$$\mathbf{F}(t) = \begin{pmatrix} \mathcal{F}(t)/C_1 \\ 0 \end{pmatrix}. \quad (4.11)$$

**Table 1: Summary of eigenmode decomposition parameters and relationships.** Abbreviated version of Table 1 from [Geoffroy et al. \(2013a\)](#) including an efficacy term,  $\varepsilon$ .

General parameters	
$b = \frac{\lambda + \varepsilon\gamma}{C_1} + \frac{\gamma}{C_2}$	
$b^* = \frac{\lambda + \varepsilon\gamma}{C_1} - \frac{\gamma}{C_2}$	
$\delta = b^2 - 4\frac{\lambda\gamma}{C_1 C_2}$	
Mode parameters	
Fast	Slow
$\phi_f = \frac{C_1}{2\gamma} (b^* - \sqrt{\delta})$	$\phi_s = \frac{C_1}{2\gamma} (b^* + \sqrt{\delta})$
$\tau_f = \frac{C_1 C_2}{2\lambda\gamma} (b - \sqrt{\delta})$	$\tau_s = \frac{C_1 C_2}{2\lambda\gamma} (b + \sqrt{\delta})$

370 Eqn. (4.8) is well-known to admit a two timescale solution ([Proistosescu and Huybers, 2017; Strogatz,](#)  
371 [2018](#)). Using the variation of parameters method, one can show the analytic solution for the upper  
372 layer temperature,  $T_t^{(1)}$ , is given by

$$T^{(1)}(t) = T_{hist}^{(1)}(t) + \underbrace{\frac{\phi_s}{C_1(\phi_s - \phi_f)} \int_0^t \mathcal{F}(\zeta) e^{-(t-\zeta)/\tau_f} d\zeta}_{=: T_{fast}^{(1)}(t)} - \underbrace{\frac{\phi_f}{C_1(\phi_s - \phi_f)} \int_0^t \mathcal{F}(\zeta) e^{-(t-\zeta)/\tau_s} d\zeta}_{=: T_{slow}^{(1)}(t)}, \quad (4.12)$$

373 where  $\tau_f$ ,  $\tau_s$ ,  $\phi_s$ , and  $\phi_f$  can be found using the quantities given in Table 1,

$$T_{hist}^{(1)}(t) = \frac{1}{\phi_s - \phi_f} \left( [\phi_s T_0^{(1)} - T_0^{(2)}] e^{-t/\tau_f} + [T_0^{(2)} - \phi_f T_0^{(1)}] e^{-t/\tau_s} \right), \quad (4.13)$$

374 and  $T_0^{(1)}$  and  $T_0^{(2)}$  are the initial conditions of each box ([Geoffroy et al., 2013a](#)). We plot each of these  
375 components individually in Figure 4 using our parameter posteriors.

#### 376 4.2.2 Fast and slow mode contributions to equilibrium warming

377 In the main text, we asserted that higher ECS worlds have more warming that is attributable to the  
378 slow climate mode in equilibrium. We now make this notion theoretically concrete. Suppose the fast  
379 mode contribution to warming can be well-approximated by the transient climate response, given by

$$\Delta T_{fast} \approx TCR = \frac{F_{2\times}}{\lambda + \gamma}. \quad (4.14)$$

380 Then the slow mode contribution can be written as the remaining warming in equilibrium, i.e.,

$$\Delta T_{slow} \approx ECS - TCR = \frac{F_{2\times}}{\lambda} - \frac{F_{2\times}}{\lambda + \gamma} \quad (4.15)$$

$$= \frac{\gamma F_{2\times}}{\lambda(\gamma + \lambda)}. \quad (4.16)$$

381 Then the ratio of warming attributed to each mode can be written as

$$\frac{\Delta T_{slow}}{\Delta T_{fast}} = \frac{\gamma}{\lambda}. \quad (4.17)$$

382 Therefore, assuming minimal correlation between  $\gamma$  and  $\lambda$ , in worlds with low  $\lambda$  (high ECS), the  
383 contribution from the slow mode increases nonlinearly, as supported by the calculations in Figure 4.

### 384 4.3 Ensemble of data assimilations

385 The EDA has two primary components: a tangent linear model (TLM) and a cost function, which we  
386 develop in turn below. Of particular importance is how a TLM, which can easily be translated into an  
387 adjoint model (ADJM), allows for efficient computation of the gradient of the cost function ([Errico, 1997](#)). Throughout we closely follow [Park and Zupanski \(2022\)](#) (see their Chapter 4), and refer readers  
388 to their exposition for additional technical details.  
389

#### 390 4.3.1 Tangent linear model

391 In this section, we introduce the general concept of TLMs, and how they can be verified. In general,  
392 any dynamical operator (or “propagator”),  $\mathbf{M}_t : \mathbb{R}^N \rightarrow \mathbb{R}^N$ , that maps a vector of inputs  $\mathbf{X}_t \in \mathbb{R}^N$  at  
393 a time  $t$  to the derivative of  $\mathbf{X}$  can be written as

$$\frac{d\mathbf{X}_t}{dt} = \mathbf{M}_t(\mathbf{X}_t). \quad (4.18)$$

394 Note that  $\mathbf{X}_t$  is a collection of model states, and if one is interested in estimating model parameters  
395 or model errors, these are included in  $\mathbf{X}_t$  and are assumed to be stationary; i.e.,  $d\xi/dt = 0$  for an

396 arbitrary model parameter or error  $\xi$ . For a small perturbation  $\delta\mathbf{X}_t$ , we can Taylor expand (4.18) as

$$\frac{d}{dt}(\mathbf{X}_t + \delta\mathbf{X}_t) = \mathbf{M}_t(\mathbf{X}_t) + \frac{\partial\mathbf{M}_t}{\partial\mathbf{X}_t}\Big|_{\mathbf{X}_t} \delta\mathbf{X}_t + \mathcal{O}(|\delta\mathbf{X}_t|^2) \quad (4.19)$$

$$\approx \mathbf{M}_t(\mathbf{X}_t) + \mathbf{L}_t \delta\mathbf{X}_t \quad (4.20)$$

397 where

$$\mathbf{L}_t := \frac{\partial\mathbf{M}_t}{\partial\mathbf{X}_t}\Big|_{\mathbf{X}_t} \quad (4.21)$$

398 is the TLM at time  $t$  and we have neglected higher order terms. In general, the TLM is a square matrix  
399 with shape  $N \times N$ , where  $N$  is the number of control variables one has in their VAR problem. As we  
400 will see later, the TLM is of practical importance because of its relationship to the ADJM, and via  
401 the ADJM, the gradient of the cost function. To verify that the TLM has been coded correctly and is  
402 operating as expected, we can introduce an arbitrary parameter  $\alpha > 0$  and use (4.20) to define

$$R_\alpha := \frac{|\mathbf{M}_t(\mathbf{X}_t + \alpha\delta\mathbf{X}_t) - \mathbf{M}_t(\mathbf{X}_t)|_2}{\alpha|\mathbf{L}_t\delta\mathbf{X}_t|_2} \quad (4.22)$$

403 where  $|\cdot|_2$  is the  $L_2$ -norm. If the TLM is coded correctly, one will find that  $\lim_{\alpha \rightarrow 0} R_\alpha \rightarrow 1$ .

#### 404 4.3.2 Cost function

405 The cost function,  $J$ , can be written in terms of two components:  $J_{guess}$ , which quantifies the cost  
406 of control variables far away from their “first-guess” estimates, and  $J_{obs}$ , that penalizes model misfit  
407 with observations. Both parts of the cost function are assumed to be derived from Gaussian likeli-  
408 hoods (Evensen et al., 2022).

409 We begin with  $J_{guess}$ . Assume that each control variable listed in Table 2 is written as  $\vec{x}_0$ , and that  
410 each control variable as a first-guess of  $\vec{\chi}$  with covariance of the joint prior distribution given by  $\mathbf{P}$ .  
411 Then we can write  $J_{guess}$  as

$$J_{guess} = \frac{1}{2} (\vec{x}_0 - \vec{\chi})^T \mathbf{P}^{-1} (\vec{x}_0 - \vec{\chi}). \quad (4.23)$$

412 As for  $J_{obs}$ , we write the superposition of two cost functions that quantify deviations in the surface  
413 temperature observations and ocean heat content. If  $\vec{T}^{(obs)}$  is a vector of observations of surface

**Table 2: Control variables in our EDA approach.**

Parameter name	Symbol	Initial condition, parameter, or model error
Initial state of surface temperature	$T_0^{(1)}$	Initial condition
Initial state of deep ocean temperature	$T_0^{(2)}$	Initial condition
Climate sensitivity	$\lambda$	Parameter
Heat transfer coefficient	$\gamma$	Parameter
Efficacy factor	$\varepsilon$	Parameter
Surface layer heat capacity	$C_1$	Parameter
Deep layer heat capacity	$C_2$	Parameter
Logarithmic sensitivity of concentrations-driven radiative forcing	$f_1^{(c)}$	Parameter
Linear sensitivity of concentrations-driven radiative forcing	$f_2^{(c)}$	Parameter
Square-root sensitivity of concentrations-driven radiative forcing	$f_3^{(c)}$	Parameter
Logarithmic sensitivity of aerosol emissions-driven radiative forcing	$f_1^{(e)}$	Parameter
Aerosol forcing shape parameter	$C_0^{(e)}$	Parameter
Linear sensitivity of aerosol emissions-driven radiative forcing	$f_2^{(e)}$	Parameter
Model errors	$q_t$	Model error

414 temperature and  $\vec{Q}^{(obs)}$  is a vector of observations of ocean heat content, we can write  $J_{obs}$  as

$$J_{obs} = \frac{1}{2} \left( \vec{T}^{(1)} - \vec{T}^{(obs)} \right)^T \Sigma_T^{-1} \left( \vec{T}^{(1)} - \vec{T}^{(obs)} \right) + \frac{1}{2} \left( \vec{Q} - \vec{Q}^{(obs)} \right)^T \Sigma_Q^{-1} \left( \vec{Q} - \vec{Q}^{(obs)} \right) \quad (4.24)$$

415 where  $\Sigma_T$  ( $\Sigma_Q$ , resp.) is the error covariance matrix between observations of surface temperature  
416 (ocean heat content, resp.).

417 Combining (4.23) and (4.24), we can write our cost function as

$$\begin{aligned} J &= J_{guess} + J_{obs} \\ &= \frac{1}{2} (\vec{x}_0 - \vec{\chi})^T \mathbf{P}^{-1} (\vec{x}_0 - \vec{\chi}) \\ &\quad + \frac{1}{2} \left( \vec{T}^{(1)} - \vec{T}^{(obs)} \right)^T \Sigma_T^{-1} \left( \vec{T}^{(1)} - \vec{T}^{(obs)} \right) \\ &\quad + \frac{1}{2} \left( \vec{Q} - \vec{Q}^{(obs)} \right)^T \Sigma_Q^{-1} \left( \vec{Q} - \vec{Q}^{(obs)} \right). \end{aligned} \quad (4.25)$$

418 The goal of VAR is to minimize this cost function; if we are able to find a global minimum of (4.25),  
419 then our EDA approach is guaranteed to be a randomized maximum likelihood sampling (RML) of the  
420 posterior probability distribution function (PDF). Finding the global minimum requires the gradient  
421 of the cost function, and we develop a formalism for computing the cost function gradient next.

#### 422 4.3.3 Connecting the TLM and the cost function gradient

423 The final piece of the VAR approach is to compute the gradient of the cost function (4.25). Throughout,  
424 we will define a vector of control variables at a time  $t$  as  $\vec{x}_t$ , which is simply a vector containing all  
425 of the listed quantities in Table 2. We will work in two parts, focusing on computing the gradient of  
426  $J_{guess}$  and  $J_{obs}$  separately; for  $J_{obs}$ , we will follow the chain rule-based approach of [Talagrand \(1991\)](#)  
427 and [Errico and Vukicevic \(1992\)](#) to compute the gradient of the cost function with respect to the initial  
428 condition of the  $n^{\text{th}}$  component of  $\vec{x}_t$  inductively.

429 Starting with  $J_{guess}$ , the gradient is relatively trivial. Let  $x_0^{(n)}$  be the  $n^{\text{th}}$  component of  $\vec{x}_0$ , suppose  
430 there are  $N$  control variables, and let  $p_{i,j}$  be the  $(i,j)^{\text{th}}$  element of  $\mathbf{P}^{-1}$ . Then directly differentiat-

<sup>431</sup> ing (4.23), we can write

$$\frac{\partial J_{\text{guess}}}{\partial x_0^{(n)}} = p_{n,n} \left( x_0^{(n)} - \chi^{(n)} \right) + \sum_{j=0}^N p_{n,j} \left( x_0^{(j)} - \chi^{(j)} \right) \quad (4.26)$$

<sup>432</sup> where we have exploited the fact that covariance matrices are symmetric by definition (and therefore  
<sup>433</sup> so are their inverses).

<sup>434</sup> For  $J_{\text{obs}}$ , we derive the gradient with respect to  $x_0^{(n)}$  inductively. Let  $\sigma_{i,j}^{(T)}$  ( $\sigma_{i,j}^{(Q)}$ , resp.) be the  
<sup>435</sup>  $(i,j)^{\text{th}}$  component of  $\Sigma_T^{-1}$  ( $\Sigma_Q^{-1}$ , resp.) and define

$$\Delta_t^{(T)} := T_t^{(1)} - T_t^{\text{obs}}, \quad (4.27)$$

<sup>436</sup> as well as

$$\Delta_t^{(Q)} := Q_t - Q_t^{\text{obs}}. \quad (4.28)$$

<sup>437</sup> Starting with the final timestep for which we have model states,  $T$ , we can write the gradient of  
<sup>438</sup>  $J_{\text{obs}}$  with respect to  $x_T^{(n)}$  as

$$\frac{\partial J_{\text{obs}}}{\partial x_T^{(n)}} = \left( \sigma_{T,T}^{(T)} \Delta_T^{(T)} + \sum_{\substack{t'=0 \\ t' \neq T}} \sigma_{t',T}^{(T)} \Delta_{t'}^{(T)} \right) \frac{\partial \Delta_T^{(T)}}{\partial x_T^{(n)}} + \left( \sigma_{T,T}^{(Q)} \Delta_T^{(Q)} + \sum_{\substack{t'=0 \\ t' \neq T}} \sigma_{t',T}^{(Q)} \Delta_{t'}^{(Q)} \right) \frac{\partial \Delta_T^{(Q)}}{\partial x_T^{(n)}} \quad (4.29)$$

<sup>439</sup> or more succinctly,

$$\frac{\partial J_{\text{obs}}}{\partial x_T^{(n)}} = \sum_{k \in \{T,Q\}} \left( \sigma_{T,T}^{(k)} \Delta_T^{(k)} + \sum_{\substack{t'=0 \\ t' \neq T}} \sigma_{t',T}^{(k)} \Delta_{t'}^{(k)} \right) \frac{\partial \Delta_T^{(k)}}{\partial x_T^{(n)}} =: F_T^{(n)} \quad (4.30)$$

<sup>440</sup> where we have again exploited the symmetric property of covariance matrices.

<sup>441</sup> We can carry out the same process for  $t = T - 1$  and show that

$$\frac{\partial J_{\text{obs}}}{\partial x_{T-1}^{(n)}} = F_{T-1}^{(n)} + \sum_{a=0}^N \frac{\partial x_T^{(a)}}{\partial x_{T-1}^{(n)}} F_T^{(a)}. \quad (4.31)$$

442 Comparing our general formulation of the TLM in (4.21) and (4.31), we can see that

$$\sum_{a=0}^N \frac{\partial x_T^{(a)}}{\partial x_{T-1}^{(n)}} F_T^{(a)} \equiv (\mathbf{L}_T)^* \vec{F}_T \quad (4.32)$$

443 where  $(\mathbf{L}_t)^*$  is the adjoint of the TLM. In this way, misfits in the gradient at time  $t = T$  are propagated  
 444 backward in time by the ADJM to  $t = T - 1$ . This approach can be carried out inductively to form  
 445 a dynamical system, where at every time  $t$  there is a “forcing” term  $\vec{F}_t$  (given by (4.30) where  $T$  is  
 446 replaced by  $t$ ) and a propagated term involving the ADJM and all other forcings for  $t' > t$ . Integrating  
 447 this dynamical system back to the initial condition (i.e., from  $t = T \rightarrow t = 0$ ) results in the gradient of  
 448  $J_{obs}$ , as desired.

449 We can check the validity of our ADJM and the gradient of the cost function by introducing two  
 450 quantities. The first tests the validity of the ADJM; using the definition of the adjoint (Park and  
 451 Zupanski, 2022), we can write

$$\Lambda_t := \frac{\langle \mathbf{L}\delta\mathbf{X}, \mathbf{L}\delta\mathbf{X} \rangle}{\langle \delta\mathbf{X}, \mathbf{L}^*(\mathbf{L}\delta\mathbf{X}) \rangle} \quad (4.33)$$

452 where  $\delta\mathbf{X}$  is a perturbed state vector. This identity essentially compares the forward propagated path  
 453 of the TLM and the backward propagated path of the ADJM. If the program is running correctly, for  
 454 small numbers of timesteps,  $\Lambda_t \approx 1$ , and  $\Lambda_t$  will diverge from one the more timesteps are taken.

455 The second quantity we can introduce serves to check our computation of the cost function gradient.

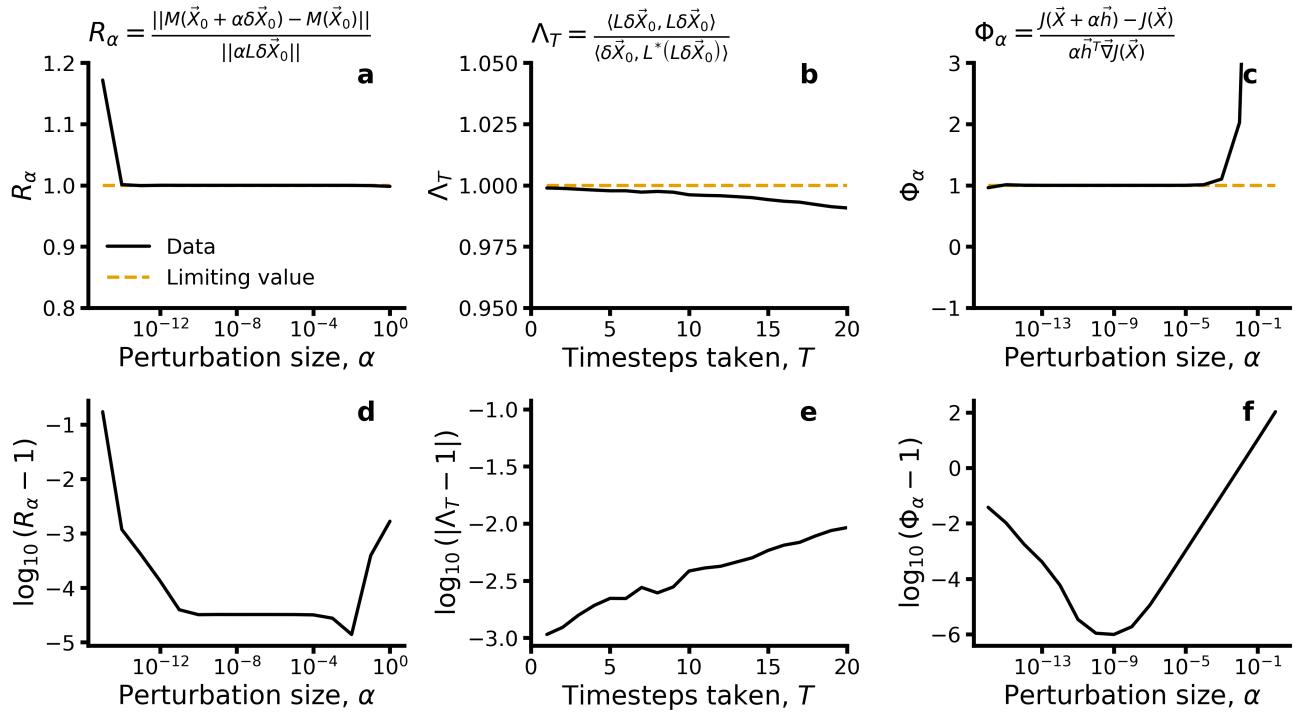
456 For an arbitrary parameter  $\alpha > 0$  and vector  $\mathbf{h}$ , we can define

$$\Phi_\alpha := \frac{J(\mathbf{X} + \alpha\mathbf{h}) - J(\mathbf{X})}{\alpha\mathbf{h}^T \nabla J(\mathbf{X})}, \quad (4.34)$$

457 using the definition of a Taylor expansion similar to our derivation of (4.22). It is common practice to  
 458 set  $\mathbf{h} = \nabla J(\mathbf{X})$ , but (4.34) is obviously general for any  $\mathbf{h}$ . As was the case with (4.22), we expect a  
 459 correctly formulated cost function gradient to have the limiting behavior of  $\lim_{\alpha \rightarrow 0} \Phi_\alpha \rightarrow 1$ .

#### 460 4.3.4 Verification of TLM, ADJM, and cost function gradient

461 To verify the validity of our formulation, we compute (4.22), (4.33), and (4.34) for a number of pertur-  
 462 bation sizes and timesteps in Figure 5. We find that each quantity displays the expected behavior for  
 463 a correctly programmed TLM, ADJM, and cost function gradient. In particular,  $R_\alpha$  and  $\Phi_\alpha$  approach  
 464 unity for decreasingly small perturbation sizes, and  $\Lambda_t$  starts close to unity for a small number of

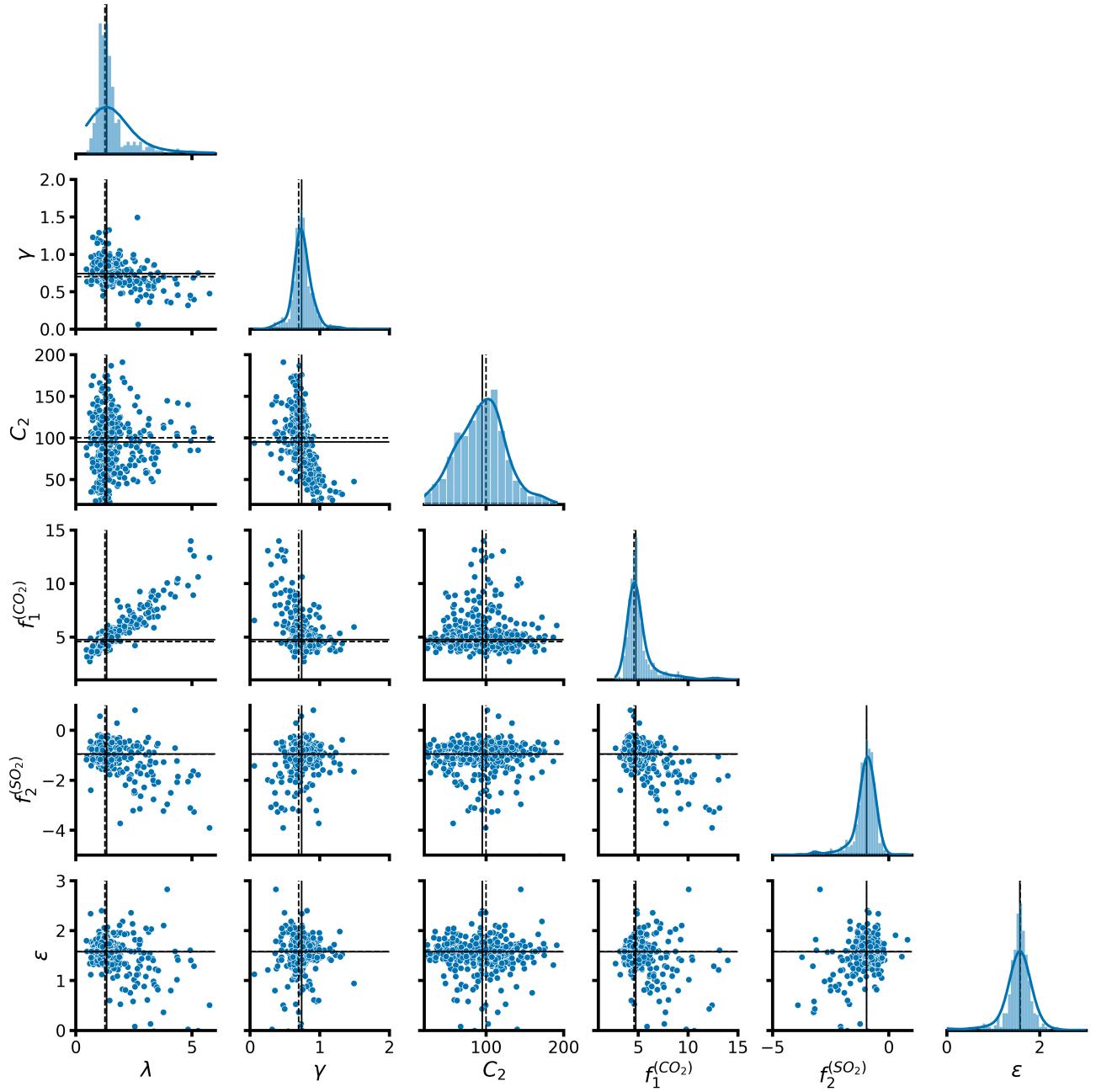


**Figure 5: Check plots for TLM, ADJM, and cost function gradient.** Panels a–c show (4.22), (4.33), and (4.34) for a number of values of  $\alpha$  (or timesteps taken, in the case of panel b). d–f are as a–c, but on a logarithmic scale.

465 timesteps before diverging as more are taken. Moreover, on a logarithmic scale,  $R_\alpha$  and  $\Phi_\alpha$  display the  
 466 normal “V”-shaped patterns as is usually found in such computations (Park and Zupanski, 2022).

#### 467 4.3.5 Synthesis

468 We now combine each of the prior subsections to summarize our EDA approach. Assuming the control  
 469 variables have a prior distribution  $\vec{\chi} \sim \mathcal{N}(\vec{\chi}^p, \mathbf{P})$ , we sample this prior  $N_{ens}$  times to form an ensemble  
 470 of “particles”. Each particle’s cost function is given by (4.25), where  $\vec{\chi}$  is the prior draw for that particle.  
 471 This prior draw serves as the first “first-guess” in the VAR inner- and outer-loop procedure (Park  
 472 and Zupanski, 2022); it is then updated in the outer-loop of the VAR gradient descent algorithm  
 473 as the algorithm searches for better solutions to the minimization. We use sequential least-squares  
 474 programming to solve the inner-loop minimization problem (Noceda and Wright, 2006). An example  
 475 of an abbreviated corner plot when the true ECS is average is shown in Figure 6 where we show  
 476 the abbreviated parameter posteriors for the 2020–2050 assimilation window; see the *Supplementary*  
 477 *Information* for more corner plots when ECS is low and high, as well as the full parameter corner  
 478 plots. See Rabier and Liu (2003) and Bannister (2017) for a review of the technical aspects of the VAR



**Figure 6: Abbreviated parameter posteriors for average true value of ECS.** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window. This is a truncated corner plot where we only show a subset parameters that we estimate in our data assimilation approach:  $\lambda$ ,  $\gamma$ ,  $C_2$ ,  $f_1^{(CO_2)}$ ,  $f_2^{(SO_2)}$ ,  $\varepsilon$ .

**Table 3: Cost function summary statistics.** Average median, 95<sup>th</sup> percentile, 97<sup>th</sup> percentile, 98<sup>th</sup> percentile, and 99<sup>th</sup> percentile for each finite posterior cost function distribution, computed for each true value of ECS and assimilation window upper bound we consider.

Quantity	$\log_{10}$ (Value)
Median	0.506
95 <sup>th</sup> percentile	1.433
97 <sup>th</sup> percentile	1.900
98 <sup>th</sup> percentile	2.485
99 <sup>th</sup> percentile	13.137

479 approach.

480 We minimize each cost function using VAR, utilizing the cost function gradient derived using (4.26)  
481 and by integrating (4.31) recursively backwards in time. The result is an optimized ensemble of state  
482 estimates that allow us to sample the posterior of our model states, model parameters, and model  
483 errors conditional on our data in the assimilation window; indeed, insofar as we find global minima  
484 of (4.25), we are guaranteed a randomized maximum likelihood sampling of the posterior since each  
485 particle's cost function is independent (Evensen et al., 2022). We establish a convergence criterion  
486 that, after 100 iterations of VAR, the cost function must be below  $10^2$ , which places the cut-off just  
487 over the average 97<sup>th</sup> percentile in our summary statistics, see Table 3. See Figure 7 for an example of  
488 our screening methodology.

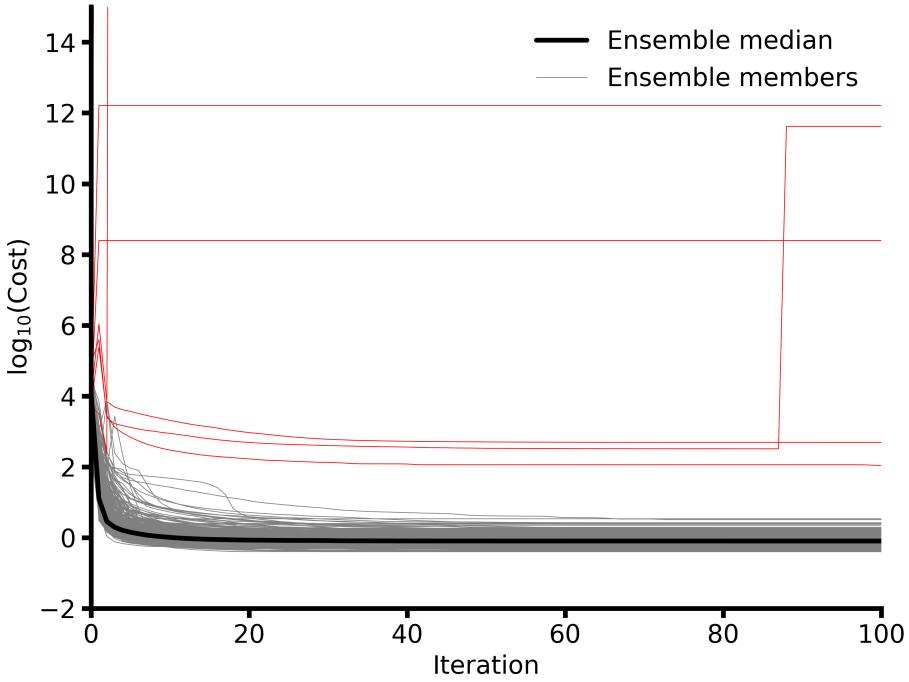
489 **4.4 Details on prior**

490 It remains to describe how we chose our prior for each parameter we estimate in our approach. For  
491 each box's initial condition,  $T_0^{(1)}$  and  $T_0^{(2)}$ , we warm start our model starting in 1850, and integrate  
492 the model until 2020 using the true values of each other parameter in our model. We then assign a 0.2  
493 °C standard deviation to the prior distribution to account for internal variability and the possibility of  
494 errors in measuring the current global average temperature.

495 We pull the central estimates for our thermal response parameters ( $\lambda$ ,  $\gamma$ ,  $C_1$ , and  $C_2$ ) from Geoffroy  
496 et al. (2013a). Note we slightly revise upwards the estimate of  $\lambda$  (1.13 → 1.258) to achieve an average  
497 ECS of 3 °C, in line with recent estimates (Sherwood et al., 2020). However, Geoffroy et al. (2013a)  
498 calibrates their two-layer model on only 16 climate models, which calls into question their reported

Table 4: Information on prior calibration.

Parameter and units	Prior distribution	Source	Notes
$T_0^{(1)}$ [K]	$\mathcal{N}(0.46, 0.2)$	Model warm start	—
$T_0^{(2)}$ [K]	$\mathcal{N}(0.06, 0.2)$	Model warm start	—
$\lambda$ [W K <sup>-1</sup> ]	$\mathcal{N}(1.258, 0.38)$	Geoffroy et al. (2013a)	—
$\gamma$ [W K <sup>-1</sup> ]	$\mathcal{N}(0.7, 0.21)$	Geoffroy et al. (2013a)	—
$C_1$ [J K <sup>-1</sup> ]	$\mathcal{N}(8, 2.4)$	Geoffroy et al. (2013a)	—
$C_2$ [J K <sup>-1</sup> ]	$\mathcal{N}(100, 30)$	Geoffroy et al. (2013a)	—
$f_1^{(\text{CO}_2)}$ [W m <sup>-2</sup> ]	$\mathcal{N}(4.58, 0.519)$	Leach et al. (2021)	Var. from Zelinka et al. (2020)
$f_3^{(\text{CO}_2)}$ [W m <sup>-2</sup> (ppm CO <sub>2</sub> ) <sup>-1/2</sup> ]	$\mathcal{N}(0.086, 0.026)$	Leach et al. (2021)	—
$f_1^{(\text{SO}_2)}$ [W]	$\mathcal{N}(-0.0047, 0.0014)$	Leach et al. (2021)	—
$f_2^{(\text{SO}_2)}$ [W (MtSO <sub>2</sub> / yr) <sup>-1</sup> ]	$\mathcal{N}(-0.96, 0.29)$	Leach et al. (2021)	—
$C_0^{(\text{SO}_2)}$ [MtSO <sub>2</sub> / yr]	$\mathcal{N}(170.6, 51.2)$	Leach et al. (2021)	—
$\varepsilon$ [-]	$\mathcal{N}(1.58, 0.128)$	Cummins et al. (2020)	—
$q_t$ [W]	$\mathcal{N}(0, 0.27)$	Proistescu and Huybers (2017)	AR(1), $\varphi = 0.2$



**Figure 7: Cost function versus iteration.** We plot the cost function for each ensemble member against the VAR iteration number, where valid ensemble members are shown in gray, the ensemble median is in black, and rejected ensemble members are in red. These data are generated for our two-layer model forced with SSP2–4.5 with a true ECS of 6 °C. Note the upper bound of the assimilation window is 2090.

499 standard deviations for each of these parameter distributions, especially given the correlation structure  
 500 they report between parameters in their posteriors (see their Tables 2 and 3). These factors complicate  
 501 the prior for the thermal response parameters. We therefore take a more naïve approach: to assign  
 502 a standard deviation of 30% error for each parameter. This procedure results in an approximately  
 503 similar distribution for  $\lambda$ , a wider distribution for  $\gamma$  and  $C_1$ , and a slightly narrower distribution for  
 504  $C_2$ . Sensitivity tests with a 40% spread in the prior yield similar results as what was reported in the  
 505 main text (see the *Supplementary Information*).

506 We pull the central estimate and standard deviation of  $f_1^{(CO_2)}$  from [Zelinka et al. \(2020\)](#), their Table  
 507 S1. Note that [Zelinka et al. \(2020\)](#) reports the central estimate and variance of effective radiative forcing  
 508 from a doubling of CO<sub>2</sub> in CMIP6 climate models; their values were converted into a central estimate  
 509 and standard deviation for the  $f_1^{(CO_2)}$  parameter by dividing by ln 2.

510 For the remaining radiative forcing parameters ( $f_3^{(CO_2)}$ ,  $f_1^{(SO_2)}$ ,  $f_2^{(SO_2)}$ , and  $C_0^{(SO_2)}$ ), we pull the  
 511 central estimates from [Leach et al. \(2021\)](#). Unfortunately, [Leach et al. \(2021\)](#) does not report individual  
 512 parameter uncertainties, and only the uncertainty in overall radiative forcing from different forcing  
 513 agents. We therefore take a similar approach to the above and assign a 30% error for each parameter.

514 See the below for sensitivity tests; we find that a wider prior results in similar findings. Note that  
515  $f_2^{(CO_2)} = 0$ , following [Leach et al. \(2021\)](#).

516 Finally, we calibrate the internal climate variability using the AR(1) process following [Proistosescu](#)  
517 [and Huybers \(2017\)](#). We follow their formulation to write variability in the radiative forcing,  $q_t$ , as an  
518 AR(1) process such that

$$q_{t+1} = \varphi q_t + \varepsilon_{t+1} \quad (4.35)$$

519 where  $\varepsilon_t \sim \mathcal{N}(0, \sigma)$  and  $\varphi \geq 0$ . We take the central estimate for  $\sigma$  from [Proistosescu and Huybers](#)  
520 ([2017](#)), and perform sensitivity tests for higher value of  $\sigma$  ( $0.4 \text{ W m}^{-2}$ ) below. We take a slightly  
521 higher value of  $\varphi$  (0.2, compared to 0.03) in our calibration compared to [Proistosescu and Huybers](#)  
522 ([2017](#)) to account for the higher autocorrelation in global temperature than top of atmosphere radiative  
523 imbalance. See Table 4 for a summary.

## 524 Acknowledgments

525 The authors would like to firstly thank Jidong Gao for his help in setting up, verifying, and utilizing  
526 our data assimilation approach; this manuscript would not have been possible without his patience and  
527 insight. The authors would also like to thank Dan Amrhein, Alfonso Ladino, Joseph Neid, Massimo  
528 Pascale, Chris Smith, Ryan Sriver, as well as participants at the American Geophysical Union's 2024  
529 fall meeting for helpful discussions related to this work. AMB acknowledges support from a National  
530 Science Foundation Graduate Research Fellowship grant No. DGE 21-46756. Computations were  
531 performed on the Keeling computing cluster, a computing resource operated by the School of Earth,  
532 Society and Environment (SESE) at the University of Illinois Urbana-Champaign.

## 533 Data Availability Statement

534 All the code needed to reproduce the analysis of this paper is located on the author's Github, see  
535 <https://www.github.com/adam-bauer-34/BPD-Climate-Learning-Rates-Reprod>.

## 536 References

537 M. R. Allen, P. A. Stott, J. F. B. Mitchell, R. Schnur, and T. L. Delworth. Quantifying the uncertainty  
538 in forecasts of anthropogenic climate change. *Nature*, 407(6804):617–620, Oct. 2000. ISSN 0028-0836,

- 539 1476-4687. doi: 10.1038/35036559. URL <https://www.nature.com/articles/35036559>.
- 540 M. R. Allen, D. J. Frame, C. Huntingford, C. D. Jones, J. A. Lowe, M. Meinshausen, and N. Mein-  
541 shausen. Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature*,  
542 458(7242):1163–1166, Apr. 2009. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature08019. URL  
543 <https://www.nature.com/articles/nature08019>.
- 544 K. C. Armour. Energy budget constraints on climate sensitivity in light of inconstant climate  
545 feedbacks. *Nature Climate Change*, 7(5):331–335, May 2017. ISSN 1758-678X, 1758-6798. doi:  
546 10.1038/nclimate3278. URL <https://www.nature.com/articles/nclimate3278>.
- 547 M. B. Baker and G. H. Roe. The Shape of Things to Come: Why Is Climate Change So  
548 Predictable? *Journal of Climate*, 22(17):4574–4589, Sept. 2009. ISSN 0894-8755, 1520-0442.  
549 doi: 10.1175/2009JCLI2647.1. URL <https://journals.ametsoc.org/view/journals/clim/22/17/2009jcli2647.1.xml>.
- 551 R. N. Bannister. A review of operational methods of variational and ensemble-variational data assim-  
552 ilation. *Quarterly Journal of the Royal Meteorological Society*, 143(703):607–633, Jan. 2017. ISSN  
553 0035-9009, 1477-870X. doi: 10.1002/qj.2982. URL <https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.2982>.
- 555 M. Barnett, W. Brock, and L. P. Hansen. Pricing Uncertainty Induced by Climate Change. *The  
556 Review of Financial Studies*, 33(3):1024–1066, Mar. 2020. ISSN 0893-9454, 1465-7368. doi: 10.1093/  
557 rfs/hhz144. URL <https://academic.oup.com/rfs/article/33/3/1024/5735312>.
- 558 A. M. Bauer, F. McIsaac, and S. Hallegatte. How Delayed Learning about Climate Uncertainty Impacts  
559 Decarbonization Investment Strategies. Working Paper WPS10743, The World Bank, Wash-  
560 ington DC, 2024a. URL <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099829103282438373/1du1f2d86d77127091490d1a6df1dc342f15d10b>.
- 562 A. M. Bauer, C. Proistosescu, and G. Wagner. Carbon Dioxide as a Risky Asset. *Climatic Change*,  
563 177(5):72, May 2024b. ISSN 0165-0009, 1573-1480. doi: 10.1007/s10584-024-03724-3. URL <https://link.springer.com/10.1007/s10584-024-03724-3>.
- 565 W. S. Broecker. Unpleasant surprises in the greenhouse? *Nature*, 328(6126):123–126, July 1987.  
566 ISSN 0028-0836, 1476-4687. doi: 10.1038/328123a0. URL <https://www.nature.com/articles/328123a0>.

- 568 E. Campiglio, S. Dietz, and F. Venmans. Optimal Climate Policy as If the Transition Matters. Working  
569 Paper 10139, CESifo, Munich, 2022. URL <https://www.cesifo.org/en/publications/2022/working-paper/optimal-climate-policy-if-transition-matters>.
- 571 D. P. Cummins, D. B. Stephenson, and P. A. Stott. Optimal Estimation of Stochastic Energy Balance  
572 Model Parameters. *Journal of Climate*, 33(18):7909–7926, Sept. 2020. ISSN 0894-8755, 1520-0442.  
573 doi: 10.1175/JCLI-D-19-0589.1. URL <https://journals.ametsoc.org/view/journals/clim/33/18/jcliD190589.xml>.
- 575 Y. Dong, C. Proistosescu, K. C. Armour, and D. S. Battisti. Attributing Historical and Future Evolution  
576 of Radiative Feedbacks to Regional Warming Patterns using a Green’s Function Approach: The  
577 Preeminence of the Western Pacific. *Journal of Climate*, 32(17):5471–5491, Sept. 2019. ISSN 0894-  
578 8755, 1520-0442. doi: 10.1175/JCLI-D-18-0843.1. URL <https://journals.ametsoc.org/view/journals/clim/32/17/jcli-d-18-0843.1.xml>.
- 580 R. M. Errico. What Is an Adjoint Model? *Bulletin of the American Meteorological Society*, 78(11):2577–  
581 2591, Nov. 1997. ISSN 0003-0007, 1520-0477. doi: 10.1175/1520-0477(1997)078<2577:WIAAM>2.0.  
582 CO;2. URL [http://journals.ametsoc.org/doi/10.1175/1520-0477\(1997\)078<2577:WIAAM>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2).
- 583 R. M. Errico and T. Vukicevic. Sensitivity Analysis Using an Adjoint of the PSU-NCAR Mesoseale  
584 Model. *Monthly Weather Review*, 120(8):1644–1660, Aug. 1992. ISSN 0027-0644, 1520-0493. doi: 10.  
585 1175/1520-0493(1992)120<1644:SAUAAO>2.0.CO;2. URL [http://journals.ametsoc.org/doi/10.1175/1520-0493\(1992\)120<1644:SAUAAO>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0493(1992)120<1644:SAUAAO>2.0.CO;2).
- 588 G. Evensen, F. C. Vossepoel, and P. J. Van Leeuwen. *Data Assimilation Fundamentals: A Unified  
589 Formulation of the State and Parameter Estimation Problem*. Springer Textbooks in Earth Sciences,  
590 Geography and Environment. Springer International Publishing, Cham, 2022. ISBN 9783030967086  
591 9783030967093. doi: 10.1007/978-3-030-96709-3. URL <https://link.springer.com/10.1007/978-3-030-96709-3>.
- 593 O. Geoffroy, D. Saint-Martin, G. Bellon, A. Volodire, D. J. L. Olivié, and S. Tytéca. Transient Climate  
594 Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-  
595 Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, 26(6):1859–1876, Mar.

- 596 2013a. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-12-00196.1. URL <https://journals.ametsoc.org/view/journals/clim/26/6/jcli-d-12-00196.1.xml>.
- 597  
598 O. Geoffroy, D. Saint-Martin, D. J. L. Olivie, A. Voldoire, G. Bellon, and S. Tytéca. Transient Cli-  
599 mate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter  
600 Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, 26(6):1841–1857, Mar. 2013b.  
601 ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-12-00195.1. URL <http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00195.1>.
- 602  
603 N. P. Gillett, V. K. Arora, G. M. Flato, J. F. Scinocca, and K. Von Salzen. Improved constraints on  
604 21st-century warming derived using 160 years of temperature observations. *Geophysical Research  
Letters*, 39(1):2011GL050226, Jan. 2012. ISSN 0094-8276, 1944-8007. doi: 10.1029/2011GL050226.  
605 URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2011GL050226>.
- 606  
607 A. Hannart, M. Ghil, J.-L. Dufresne, and P. Naveau. Disconcerting learning on climate sensitivity  
608 and the uncertain future of uncertainty. *Climatic Change*, 119(3-4):585–601, Aug. 2013. ISSN 0165-  
609 0009, 1573-1480. doi: 10.1007/s10584-013-0770-z. URL <http://link.springer.com/10.1007/s10584-013-0770-z>.
- 610  
611 Z. Hausfather, K. Marvel, G. A. Schmidt, J. W. Nielsen-Gammon, and M. Zelinka. Climate sim-  
612 ulations: recognize the ‘hot model’ problem. *Nature*, 605(7908):26–29, May 2022. ISSN 0028-  
613 0836, 1476-4687. doi: 10.1038/d41586-022-01192-2. URL <https://www.nature.com/articles/d41586-022-01192-2>.
- 614  
615 H. He, R. J. Kramer, B. J. Soden, and N. Jeevanjee. State dependence of CO<sub>2</sub> forcing and its implica-  
616 tions for climate sensitivity. *Science*, 382(6674):1051–1056, Dec. 2023. ISSN 0036-8075, 1095-9203.  
617 doi: 10.1126/science.abq6872. URL <https://www.science.org/doi/10.1126/science.abq6872>.
- 618  
619 I. M. Held, M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis. Probing the Fast and  
620 Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of  
Climate*, 23(9):2418–2427, May 2010. ISSN 1520-0442, 0894-8755. doi: 10.1175/2009JCLI3466.1.  
621 URL <http://journals.ametsoc.org/doi/10.1175/2009JCLI3466.1>.
- 622  
623 C. Hope, J. Anderson, and P. Wenman. Policy analysis of the greenhouse effect. *Energy Policy*,  
624 21(3):327–338, Mar. 1993. ISSN 03014215. doi: 10.1016/0301-4215(93)90253-C. URL <https://linkinghub.elsevier.com/retrieve/pii/030142159390253C>.

- 625 Intergovernmental Panel On Climate Change (IPCC), editor. *Climate Change 2022 - Mitigation of*  
626 *Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergov-*  
627 *ernmental Panel on Climate Change.* Cambridge University Press, 1 edition, Aug. 2023. ISBN  
628 9781009157926. doi: 10.1017/9781009157926. URL <https://www.cambridge.org/core/product/identifier/9781009157926/type/book>.
- 630 L. Isaksen, M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud. Ensemble  
631 of data assimilations at ECMWF. 2010. doi: 10.21957/OBKE4K60. URL <https://www.ecmwf.int/node/10125>.
- 633 D. L. Kelly and C. D. Kolstad. Bayesian learning, growth, and pollution. *Journal of Economic Dynamics*  
634 and Control, 23(4):491–518, Feb. 1999. ISSN 01651889. doi: 10.1016/S0165-1889(98)00034-7. URL  
635 <https://linkinghub.elsevier.com/retrieve/pii/S0165188998000347>.
- 636 D. L. Kelly and Z. Tan. Learning and climate feedbacks: Optimal climate insurance and fat  
637 tails. *Journal of Environmental Economics and Management*, 72:98–122, July 2015. ISSN  
638 00950696. doi: 10.1016/j.jeem.2015.05.001. URL <https://linkinghub.elsevier.com/retrieve/pii/S0095069615000431>.
- 640 R. Knutti. The end of model democracy?: An editorial comment. *Climatic Change*, 102(3-4):395–  
641 404, Oct. 2010. ISSN 0165-0009, 1573-1480. doi: 10.1007/s10584-010-9800-2. URL <http://link.springer.com/10.1007/s10584-010-9800-2>.
- 643 R. Knutti, T. F. Stocker, F. Joos, and G.-K. Plattner. Constraints on radiative forcing and future cli-  
644 mate change from observations and climate model ensembles. *Nature*, 416(6882):719–723, Apr. 2002.  
645 ISSN 0028-0836, 1476-4687. doi: 10.1038/416719a. URL <https://www.nature.com/articles/416719a>.
- 647 A. J. Leach. The climate change learning curve. *Journal of Economic Dynamics and Control*, 31(5):  
648 1728–1752, May 2007. ISSN 01651889. doi: 10.1016/j.jedc.2006.06.001. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165188906001266>.
- 650 N. J. Leach, S. Jenkins, Z. Nicholls, C. J. Smith, J. Lynch, M. Cain, T. Walsh, B. Wu, J. Tsutsui, and  
651 M. R. Allen. FaIRv2.0.0: a generalized impulse response model for climate uncertainty and future  
652 scenario exploration. *Geoscientific Model Development*, 14(5):3007–3036, May 2021. ISSN 1991-9603.  
653 doi: 10.5194/gmd-14-3007-2021. URL <https://gmd.copernicus.org/articles/14/3007/2021/>.

- 654 D. Lemoine and I. Rudik. Managing Climate Change Under Uncertainty: Recursive Integrated  
655 Assessment at an Inflection Point. *Annual Review of Resource Economics*, 9(1):117–142, Oct.  
656 2017. ISSN 1941-1340, 1941-1359. doi: 10.1146/annurev-resource-100516-053516. URL <https://www.annualreviews.org/doi/10.1146/annurev-resource-100516-053516>.
- 658 Y. Liang, N. P. Gillett, and A. H. Monahan. Climate Model Projections of 21st Century Global Warm-  
659 ing Constrained Using the Observed Warming Trend. *Geophysical Research Letters*, 47(12), June  
660 2020. ISSN 0094-8276, 1944-8007. doi: 10.1029/2019GL086757. URL <https://onlinelibrary.wiley.com/doi/10.1029/2019GL086757>.
- 662 A. McDonnell, A. M. Bauer, and C. Proistosescu. To What Extent Does Discounting ‘Hot’ Cli-  
663 mate Models Improve the Predictive Skill of Climate Model Ensembles? *Earth’s Future*, 12  
664 (10):e2024EF004844, Oct. 2024. ISSN 2328-4277, 2328-4277. doi: 10.1029/2024EF004844. URL  
665 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024EF004844>.
- 666 M. Meinshausen, S. C. B. Raper, and T. M. L. Wigley. Emulating coupled atmosphere-ocean and  
667 carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration.  
668 *Atmospheric Chemistry and Physics*, 11(4):1417–1456, Feb. 2011. ISSN 1680-7324. doi: 10.5194/  
669 acp-11-1417-2011. URL <https://acp.copernicus.org/articles/11/1417/2011/>.
- 670 S. Mori and H. Shiogama. The value of knowledge accumulation on climate sensitivity uncertainty:  
671 comparison between perfect information, single stage and act then learn decisions. *Sustainability  
Science*, 13(2):351–368, Mar. 2018. ISSN 1862-4065, 1862-4057. doi: 10.1007/s11625-018-0528-7.  
673 URL <http://link.springer.com/10.1007/s11625-018-0528-7>.
- 674 G. Myhre, O. Boucher, F.-M. Bréon, P. Forster, and D. Shindell. Declining uncertainty in transient  
675 climate response as CO<sub>2</sub> forcing dominates future climate change. *Nature Geoscience*, 8(3):181–185,  
676 Mar. 2015. ISSN 1752-0894, 1752-0908. doi: 10.1038/ngeo2371. URL <https://www.nature.com/articles/ngeo2371>.
- 678 National Research Council. *Carbon Dioxide and Climate: A Scientific Assessment*. National Academies  
679 Press, Washington, D.C., Jan. 1979. ISBN 9780309333146. doi: 10.17226/19856. URL <https://www.nap.edu/catalog/19856>.
- 681 D. Neubersch, H. Held, and A. Otto. Operationalizing climate targets under learning: An application

- 682 of cost-risk analysis. *Climatic Change*, 126(3-4):305–318, Oct. 2014. ISSN 0165-0009, 1573-1480. doi:  
683 10.1007/s10584-014-1223-z. URL <http://link.springer.com/10.1007/s10584-014-1223-z>.
- 684 J. M. Nicklas, B. Fox-Kemper, and C. Lawrence. Efficient Estimation of Climate State and Its Un-  
685 certainty Using Kalman Filtering with Application to Policy Thresholds and Volcanism. *Journal of*  
686 *Climate*, Oct. 2024. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-23-0580.1. URL <https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-23-0580.1/JCLI-D-23-0580.1.xml>.
- 688 J. Noceda and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Finan-  
689 cial Engineering. Springer New York, 2006. ISBN 9780387303031. doi: 10.1007/978-0-387-40065-5.  
690 URL <http://link.springer.com/10.1007/978-0-387-40065-5>.
- 691 L. E. Padilla, G. K. Vallis, and C. W. Rowley. Probabilistic Estimates of Transient Climate Sensitivity  
692 Subject to Uncertainty in Forcing and Natural Variability. *Journal of Climate*, 24(21):5521–5537,  
693 Nov. 2011. ISSN 0894-8755, 1520-0442. doi: 10.1175/2011JCLI3989.1. URL <http://journals.ametsoc.org/doi/10.1175/2011JCLI3989.1>.
- 695 S. K. Park and M. Zupanski. *Principles of Data Assimilation*. Cambridge University Press, 1 edition,  
696 Sept. 2022. ISBN 9781108924238 9781108831765. doi: 10.1017/9781108924238. URL <https://www.cambridge.org/core/product/identifier/9781108924238/type/book>.
- 698 C. Proistosescu and P. J. Huybers. Slow climate mode reconciles historical and model-based estimates  
699 of climate sensitivity. *Science Advances*, 3(7):e1602821, July 2017. ISSN 2375-2548. doi: 10.1126/  
700 sciadv.1602821. URL <https://www.science.org/doi/10.1126/sciadv.1602821>.
- 701 F. Rabier and Z. Liu. Variational data assimilation : Theory and overview. 2003. URL <https://api.semanticscholar.org/CorpusID:160032700>.
- 703 K. Riahi, D. P. van Vuuren, E. Kriegler, J. Edmonds, B. C. O'Neill, S. Fujimori, N. Bauer, K. Calvin,  
704 R. Dellink, O. Fricko, W. Lutz, A. Popp, J. C. Cuaresma, S. Kc, M. Leimbach, L. Jiang, T. Kram,  
705 S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. A. Da Silva, S. Smith,  
706 E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Strefler, L. Drouet, V. Krey,  
707 G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J. C. Doelman, M. Kainuma, Z. Klimont,  
708 G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni. The Shared Socioeco-  
709 nomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview.

- 710      *Global Environmental Change*, 42:153–168, Jan. 2017. ISSN 09593780. doi: 10.1016/j.gloenvcha.2016.  
711      05.009. URL <https://linkinghub.elsevier.com/retrieve/pii/S0959378016300681>.
- 712      A. Ribes, S. Qasmi, and N. P. Gillett. Making climate projections conditional on historical observations.  
713      *Science Advances*, 7(4):eabc0671, Jan. 2021. ISSN 2375-2548. doi: 10.1126/sciadv.abc0671. URL  
714      <https://www.science.org/doi/10.1126/sciadv.abc0671>.
- 715      M. J. Ring and M. E. Schlesinger. Bayesian Learning of Climate Sensitivity I: Synthetic Observations.  
716      *Atmospheric and Climate Sciences*, 02(04):464–473, 2012. ISSN 2160-0414, 2160-0422. doi: 10.4236/  
717      acs.2012.24040. URL <http://www.scirp.org/journal/doi.aspx?DOI=10.4236/acs.2012.24040>.
- 718      G. H. Roe and M. B. Baker. Why Is Climate Sensitivity So Unpredictable? *Science*, 318(5850):  
719      629–632, Oct. 2007. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1144735. URL <https://www.science.org/doi/10.1126/science.1144735>.
- 721      S. C. Sherwood, M. J. Webb, J. D. Annan, K. C. Armour, P. M. Forster, J. C. Hargreaves, G. Hegerl,  
722      S. A. Klein, K. D. Marvel, E. J. Rohling, M. Watanabe, T. Andrews, P. Braconnot, C. S. Bretherton,  
723      G. L. Foster, Z. Hausfather, A. S. Heydt, R. Knutti, T. Mauritsen, J. R. Norris, C. Proistosescu,  
724      M. Rugenstein, G. A. Schmidt, K. B. Tokarska, and M. D. Zelinka. An Assessment of Earth's  
725      Climate Sensitivity Using Multiple Lines of Evidence. *Reviews of Geophysics*, 58(4), Dec. 2020.  
726      ISSN 8755-1209, 1944-9208. doi: 10.1029/2019RG000678. URL <https://onlinelibrary.wiley.com/doi/10.1029/2019RG000678>.
- 728      H. Shiogama, D. Stone, S. Emori, K. Takahashi, S. Mori, A. Maeda, Y. Ishizaki, and M. R. Allen. Pre-  
729      dicting future uncertainty constraints on global warming projections. *Scientific Reports*, 6(1):18903,  
730      Jan. 2016. ISSN 2045-2322. doi: 10.1038/srep18903. URL <https://www.nature.com/articles/srep18903>.
- 732      P. Stott, P. Good, G. Jones, N. Gillett, and E. Hawkins. The upper end of climate model temperature  
733      projections is inconsistent with past warming. *Environmental Research Letters*, 8(1):014024, Mar.  
734      2013. ISSN 1748-9326. doi: 10.1088/1748-9326/8/1/014024. URL <https://iopscience.iop.org/article/10.1088/1748-9326/8/1/014024>.
- 736      P. A. Stott and J. A. Kettleborough. Origins and estimates of uncertainty in predictions of twenty-first  
737      century temperature rise. *Nature*, 416(6882):723–726, Apr. 2002. ISSN 0028-0836, 1476-4687. doi:  
738      10.1038/416723a. URL <https://www.nature.com/articles/416723a>.

- 739 S. H. Strogatz. *Nonlinear Dynamics and Chaos*, volume 1. Westview Press, Boulder, CO, 2 edition,  
740 2018. ISBN 9780813349107.
- 741 O. Talagrand. The use of adjoint equations in numerical modelling of the atmospheric circulation. In  
742 A. Griewank and G. F. Corliss, editors, *Automatic Differentiation of Algorithms: Theory, Imple-*  
743 *mentation, and Application*, pages 169–180. SIAM, Philadelphia, PA, 1991. ISBN 0-89871-284-X.
- 744 K. Tokarska. Past warming trend constrains future warming in CMIP6 models. *Science Advances*,  
745 page 14, 2020.
- 746 N. M. Urban, P. B. Holden, N. R. Edwards, R. L. Srivat, and K. Keller. Historical and future learning  
747 about climate sensitivity. *Geophysical Research Letters*, 41(7):2543–2552, Apr. 2014. ISSN 00948276.  
748 doi: 10.1002/2014GL059484. URL <http://doi.wiley.com/10.1002/2014GL059484>.
- 749 US Environmental Protection Agency Office of Air and Radiation. *The Benefits and Costs of the Clean  
750 Air Act from 1990 to 2020*. US Environmental Protection Agency, Washington DC, Apr. 2011. URL  
751 [https://www.epa.gov/sites/default/files/2015-07/documents/fullreport\\_rev\\_a.pdf](https://www.epa.gov/sites/default/files/2015-07/documents/fullreport_rev_a.pdf).
- 752 M. L. Weitzman. GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of  
753 Public Economic Theory*, 14(2):221–244, Mar. 2012. ISSN 1097-3923, 1467-9779. doi: 10.1111/  
754 j.1467-9779.2011.01539.x. URL [https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9779.2011.01539.x](https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9779.<br/>755 2011.01539.x).
- 756 M. D. Zelinka, T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and  
757 K. E. Taylor. Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*,  
758 47(1):e2019GL085782, Jan. 2020. ISSN 0094-8276, 1944-8007. doi: 10.1029/2019GL085782. URL  
759 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019GL085782>.
- 760 C. Zhou, M. D. Zelinka, and S. A. Klein. Analyzing the dependence of global cloud feedback on  
761 the spatial pattern of sea surface temperature change with a green's function approach. *Journal  
762 of Advances in Modeling Earth Systems*, 9(5):2174–2189, Sept. 2017. ISSN 1942-2466, 1942-2466.  
763 doi: 10.1002/2017MS001096. URL [https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017MS001096](https://agupubs.onlinelibrary.wiley.com/doi/10.1002/<br/>764 2017MS001096).

# Supplementary Information: Learning Climate Sensitivity from Future Observations, Fast and Slow

Adam Michael Bauer<sup>\*1,2</sup>, Cristian Proistosescu<sup>3,4</sup>, and Kelvin K Droegeemeier<sup>3,5</sup>

<sup>1</sup>*Department of Physics, University of Illinois Urbana-Champaign, Loomis Laboratory, Urbana, Illinois*

<sup>2</sup>*Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois*

<sup>3</sup>*Department of Climate, Meteorology, and Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, Illinois*

<sup>4</sup>*Department of Earth Sciences and Environmental Change, University of Illinois Urbana-Champaign, Urbana, Illinois*

<sup>5</sup>*National Center for Supercomputing Applications, Urbana, Illinois*

July 21, 2025

## <sup>1</sup> 1 Sensitivity tests of main results

### <sup>2</sup> 1.1 Other values of ECS

<sup>3</sup> In Figure 1, we show Figure 1 from the main text with two values of ECS to illustrate how our  
<sup>4</sup> temperature projections change with different true values of ECS.

### <sup>5</sup> 1.2 Emissions scenario

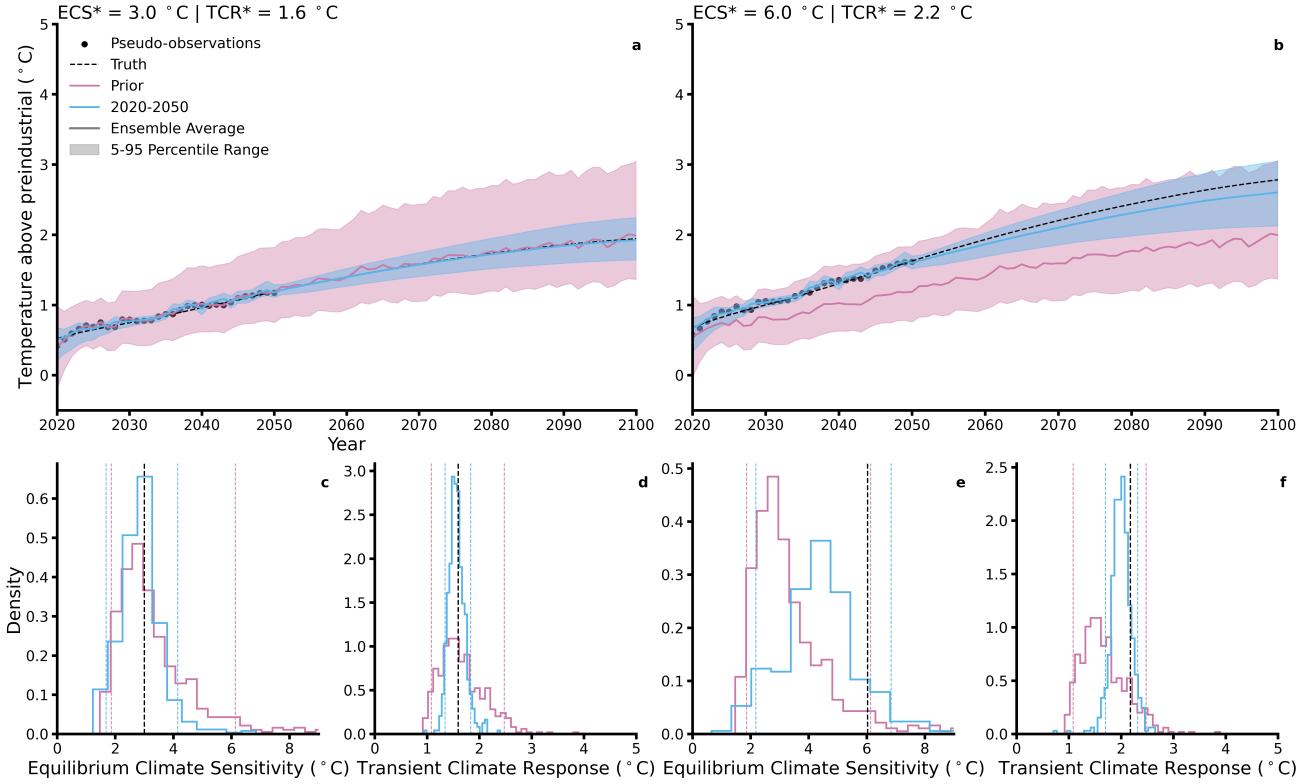
<sup>6</sup> In Figures 2 and 3, we show Figures 1 and 3 from the main text where we have forced our model (with  
<sup>7</sup> default specification on the prior and internal variability statistics) with concentrations and aerosol  
<sup>8</sup> emissions from SSP5-8.5. Our results are robust to this specification.

### <sup>9</sup> 1.3 Wider prior

<sup>10</sup> In Figures 4 and 5, we show Figures 1 and 3 from the main text where we have increased the error  
<sup>11</sup> that sets the spread of our prior from 30% → 40%. Insights are robust to this specification.

---

\*Corresponding author email: ambauer@uchicago.edu



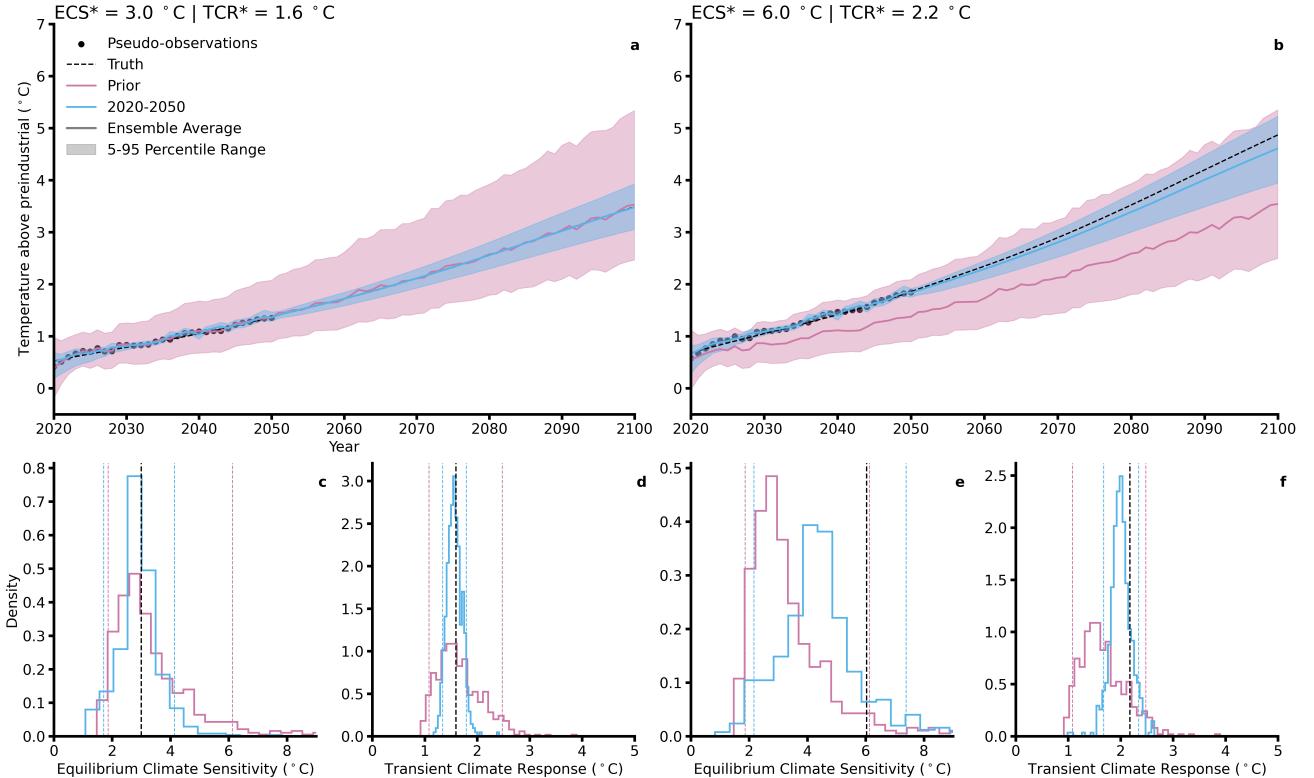
**Figure 1: Temperature forecasts, ECS, and TCR posteriors conditional on observations up to 2050, two values of ECS.** **a** shows the forecast of future temperatures using parameter posteriors from the 2020-2050 assimilation window (blue) and using the prior distribution (pink) where the true value of ECS is  $3 \text{ }^{\circ}\text{C}$  and the TCR is  $1.6 \text{ }^{\circ}\text{C}$ . **b** is as **a**, but with a true value of ECS of  $6 \text{ }^{\circ}\text{C}$  and TCR of  $2.2 \text{ }^{\circ}\text{C}$ . The shaded regions are the 5-95 percentile range, and future temperature pseudo-observations are shown in the black dots. **c** shows the prior ECS histogram (pink) and the 2020-2050 assimilation window ECS posterior (blue), with the black vertical line representing the true ECS value. **d** is as **c**, but for TCR. **e–f** are as **c–d**, but for a true value of ECS of  $6 \text{ }^{\circ}\text{C}$  and TCR of  $2.2 \text{ }^{\circ}\text{C}$ . Pseudo-observations are generated using SSP2–4.5, and in **b–c**, dashed lines represent the 5-95 percentile range.

#### 12 1.4 Higher magnitude of internal variability

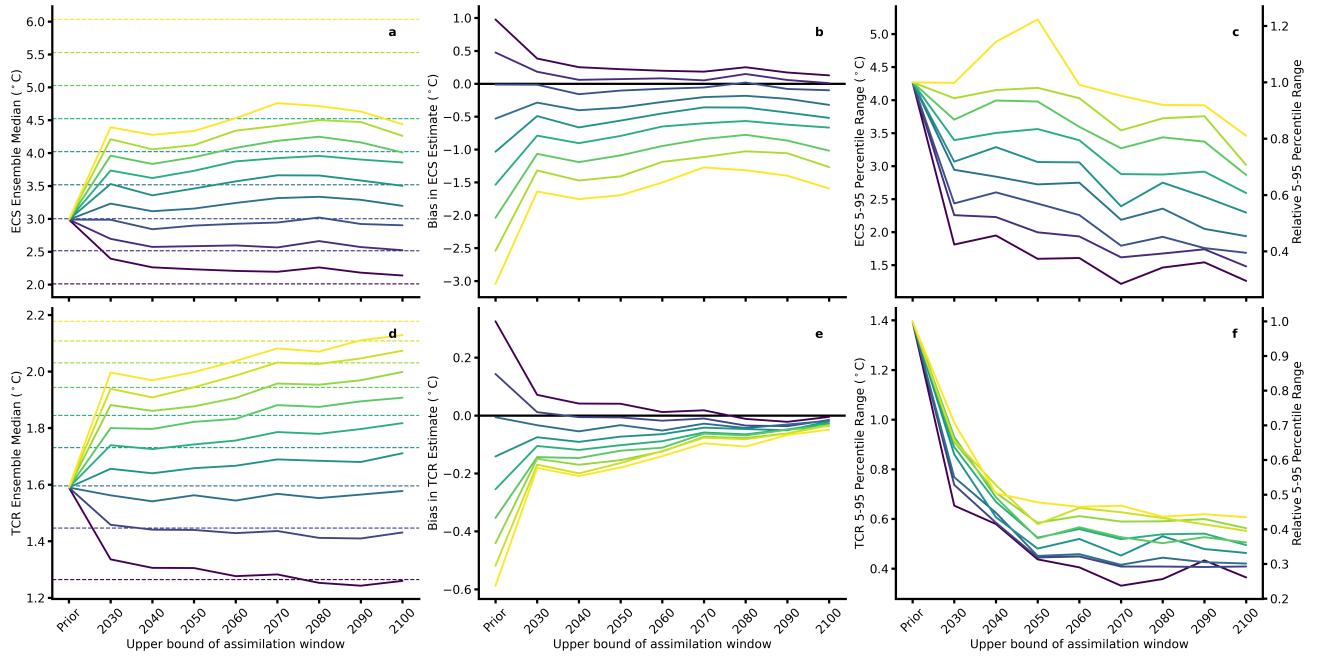
13 In Figures 6 and 7, we show Figures 1 and 3 from the main text where we have set  $\sigma = 0.4 \text{ W m}^{-2}$ .

14 Our results are robust to this specification.

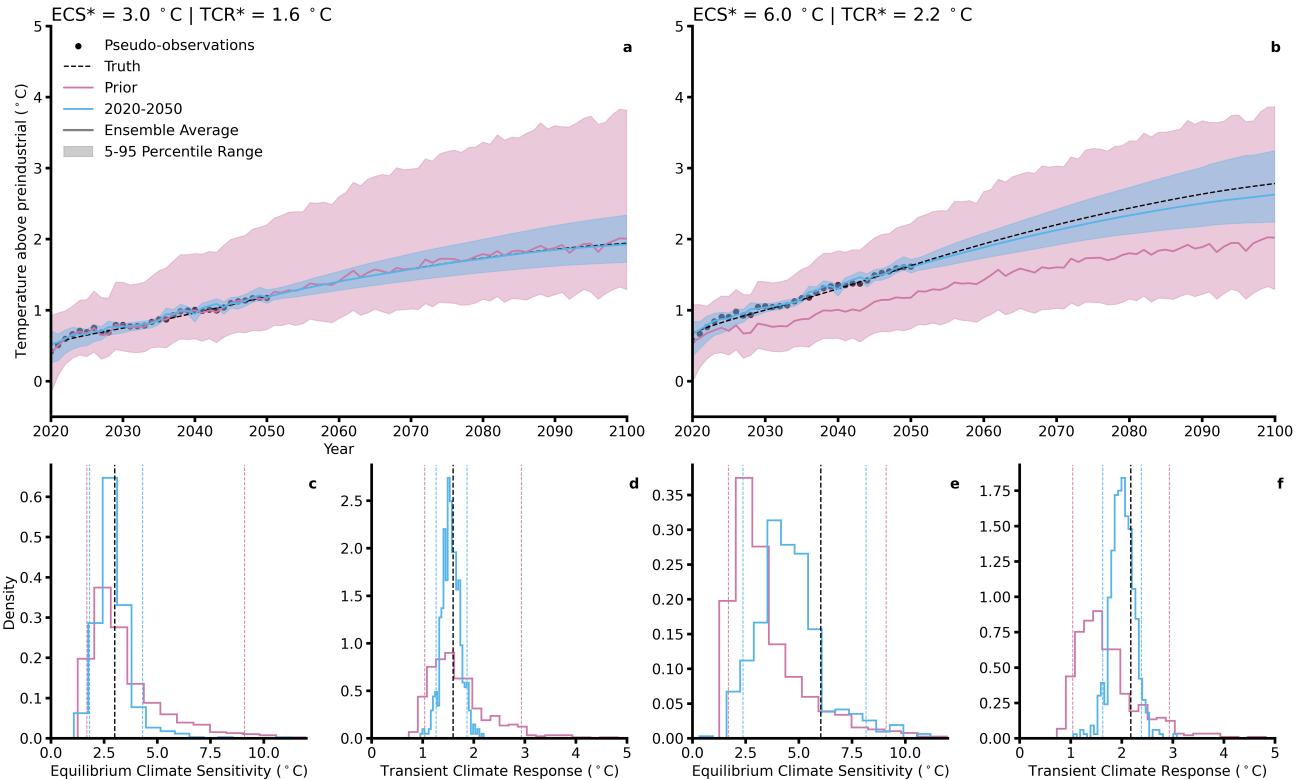
## 15 2 Additional figures



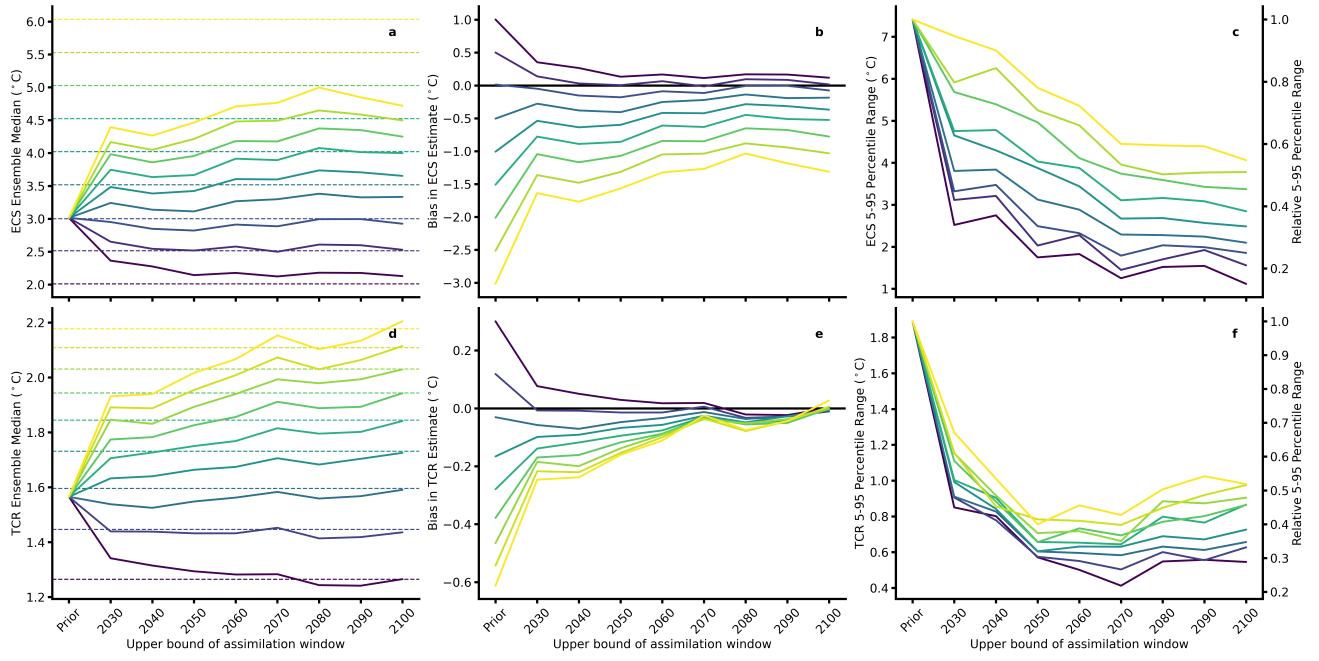
**Figure 2: Temperature forecasts, ECS, and TCR posteriors conditional on observations up to 2050, SSP5-8.5 sensitivity test.** **a** shows the forecast of future temperatures using parameter posteriors from the 2020-2050 assimilation window (blue) and using the prior distribution (pink) where the true value of ECS is 3 °C. **b** is as **a**, but with a true value of ECS of 6 °C. The shaded regions are the 5-95 percentile range, and future temperature pseudo-observations are shown in the black dots. **c** shows the prior ECS histogram (pink) and the 2020-2050 assimilation window ECS posterior (blue), with the black vertical line representing the true ECS value. **d** is as **c**, but for TCR. **e-f** are as **c-d**, but for a true value of ECS of 6 °C. Pseudo-observations are generated using SSP5-8.5, and in **b-c**, dashed lines represent the 5-95 percentile range.



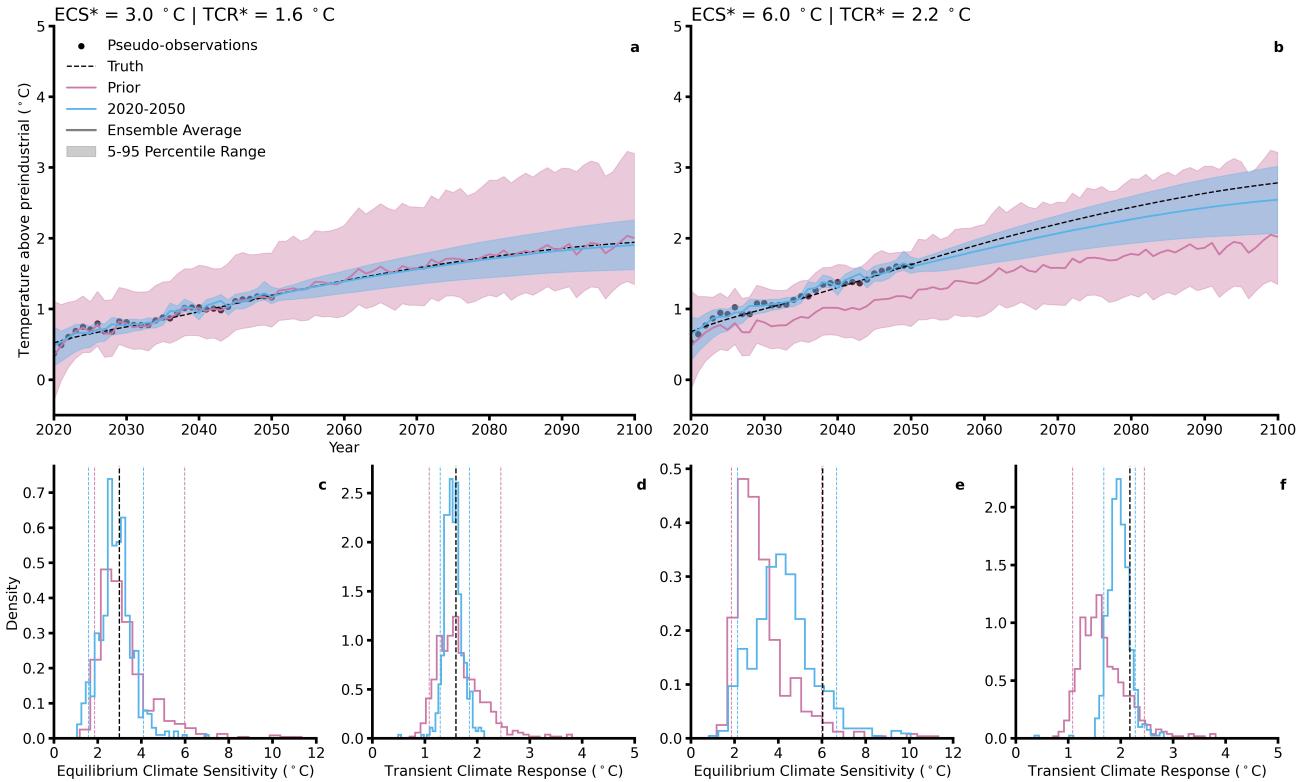
**Figure 3: ECS and TCR posterior central estimates and learning rates, SSP5-8.5 sensitivity test.** **a** shows the central estimate of the posterior ECS distribution (solid lines) and the true value of the ECS used to generate the observations (dashed lines) as the assimilation window is moved further into the future (and more observations are used for analysis). **b** shows the difference between the posterior central estimate and the true value, while **c** shows the raw posterior 5-95 percentile range and the relative 5-95 percentile range. **d–f** are as **a–c**, but for TCR.



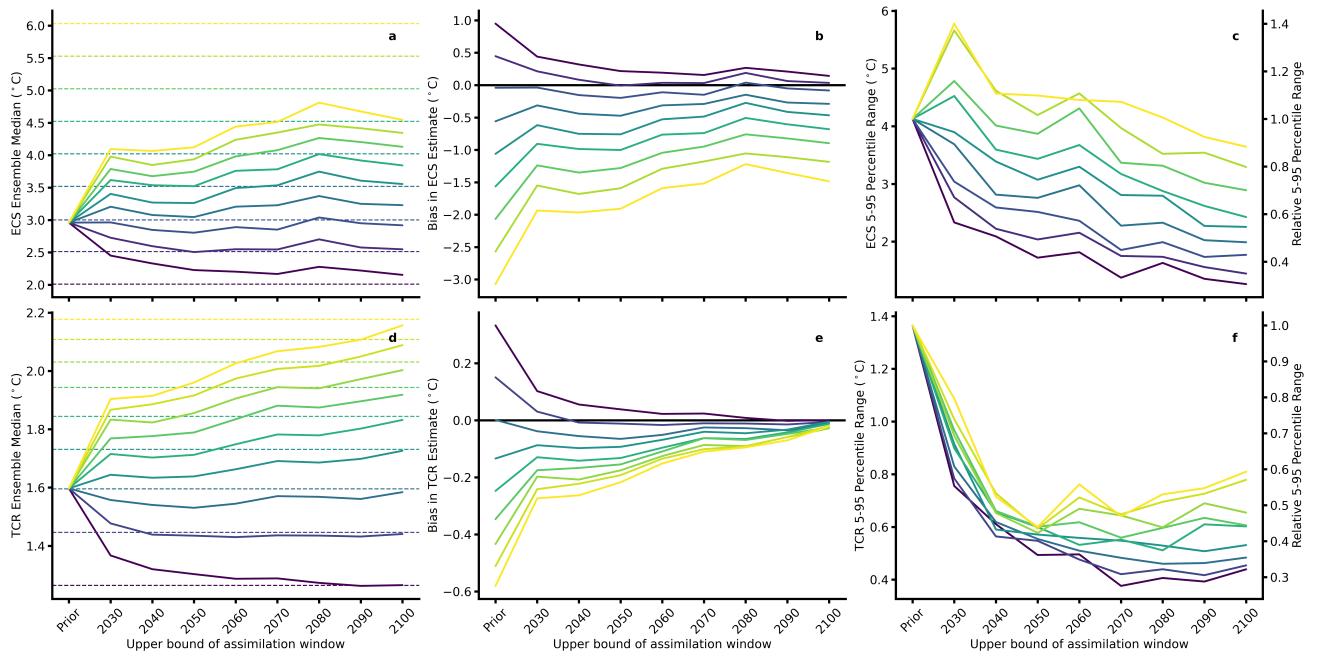
**Figure 4: Temperature forecasts, ECS, and TCR posteriors conditional on observations up to 2050, wide prior sensitivity test.** **a** shows the forecast of future temperatures using parameter posteriors from the 2020-2050 assimilation window (blue) and using the prior distribution (pink) where the true value of ECS is  $3^{\circ}\text{C}$ . **b** is as **a**, but with a true value of ECS of  $6^{\circ}\text{C}$ . The shaded regions are the 5-95 percentile range, and future temperature pseudo-observations are shown in the black dots. **c** shows the prior ECS histogram (pink) and the 2020-2050 assimilation window ECS posterior (blue), with the black vertical line representing the true ECS value. **d** is as **c**, but for TCR. **e-f** are as **c-d**, but for a true value of ECS of  $6^{\circ}\text{C}$ . Pseudo-observations are generated using SSP2-4.5, and in **b-c**, dashed lines represent the 5-95 percentile range.



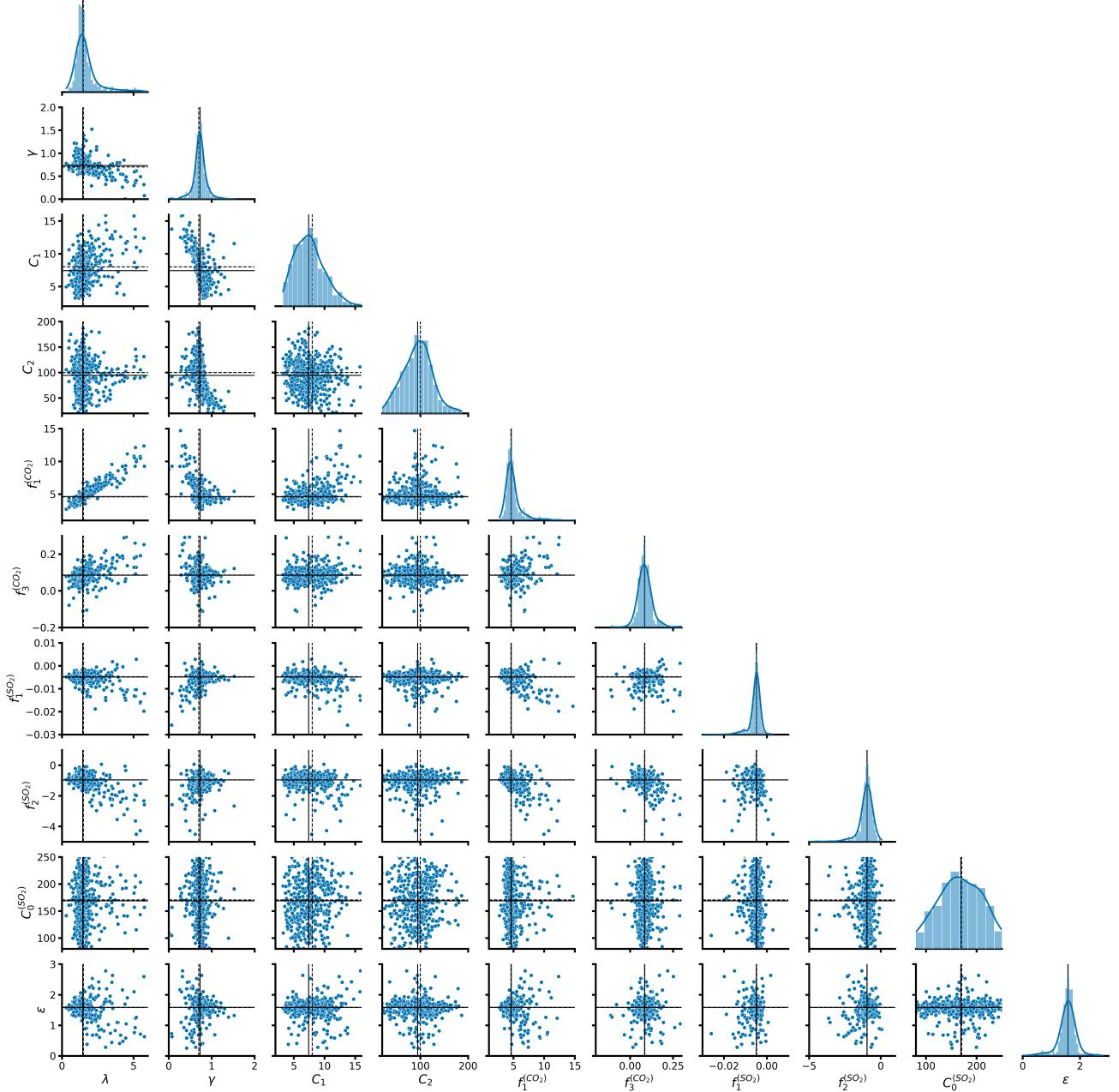
**Figure 5: ECS and TCR posterior central estimates and learning rates, wide prior sensitivity test.** **a** shows the central estimate of the posterior ECS distribution (solid lines) and the true value of the ECS used to generate the observations (dashed lines) as the assimilation window is moved further into the future (and more observations are used for analysis). **b** shows the difference between the posterior central estimate and the true value, while **c** shows the raw posterior 5-95 percentile range and the relative 5-95 percentile range. **d–f** are as **a–c**, but for TCR.



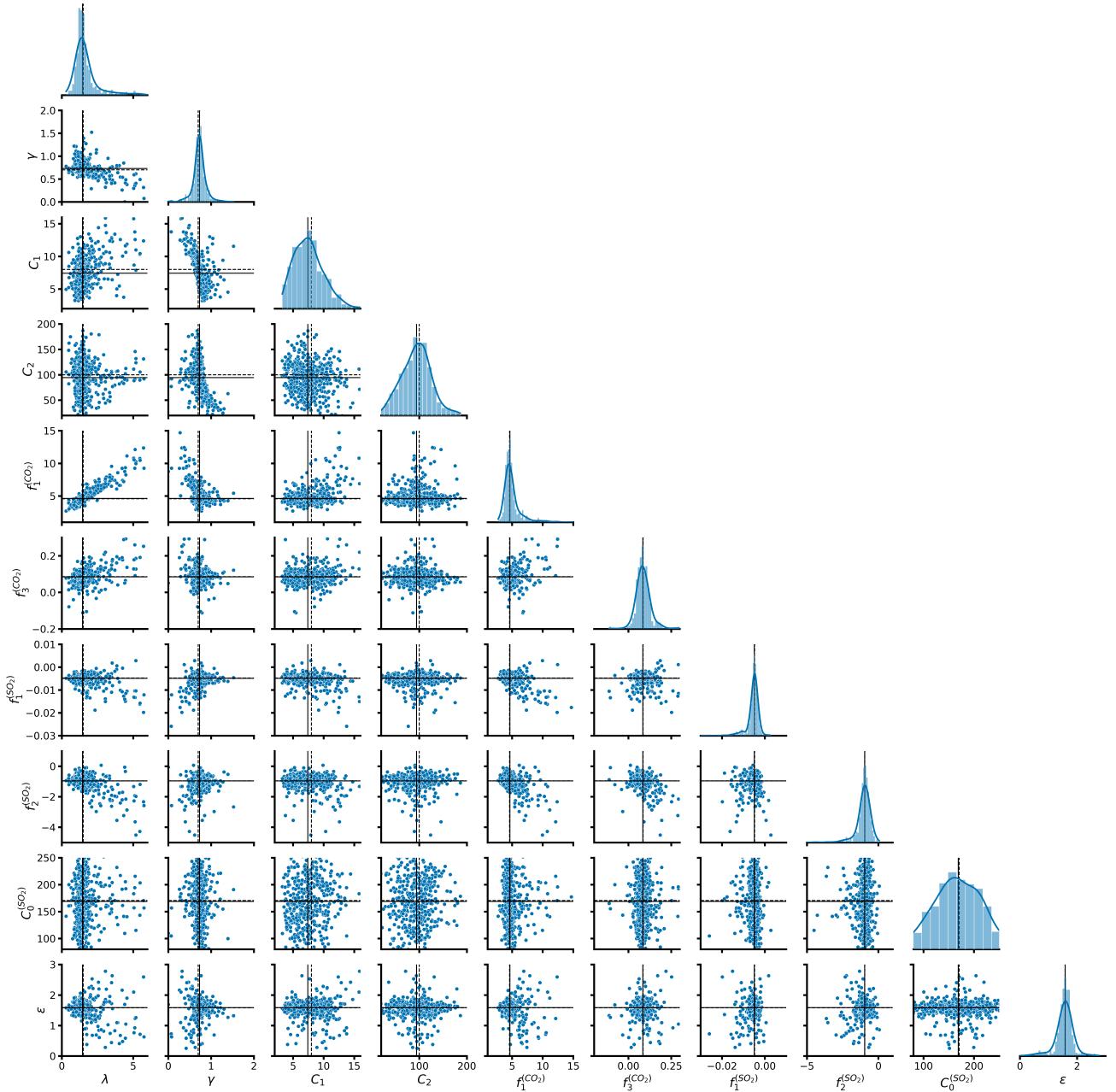
**Figure 6: Temperature forecasts, ECS, and TCR posteriors conditional on observations up to 2050, high internal variability sensitivity test.** **a** shows the forecast of future temperatures using parameter posteriors from the 2020-2050 assimilation window (blue) and using the prior distribution (pink) where the true value of ECS is  $3 \text{ } ^\circ\text{C}$ . **b** is as **a**, but with a true value of ECS of  $6 \text{ } ^\circ\text{C}$ . The shaded regions are the 5-95 percentile range, and future temperature pseudo-observations are shown in the black dots. **c** shows the prior ECS histogram (pink) and the 2020-2050 assimilation window ECS posterior (blue), with the black vertical line representing the true ECS value. **d** is as **c**, but for TCR. **e-f** are as **c-d**, but for a true value of ECS of  $6 \text{ } ^\circ\text{C}$ . Pseudo-observations are generated using SSP2-4.5, and in **b-c**, dashed lines represent the 5-95 percentile range.



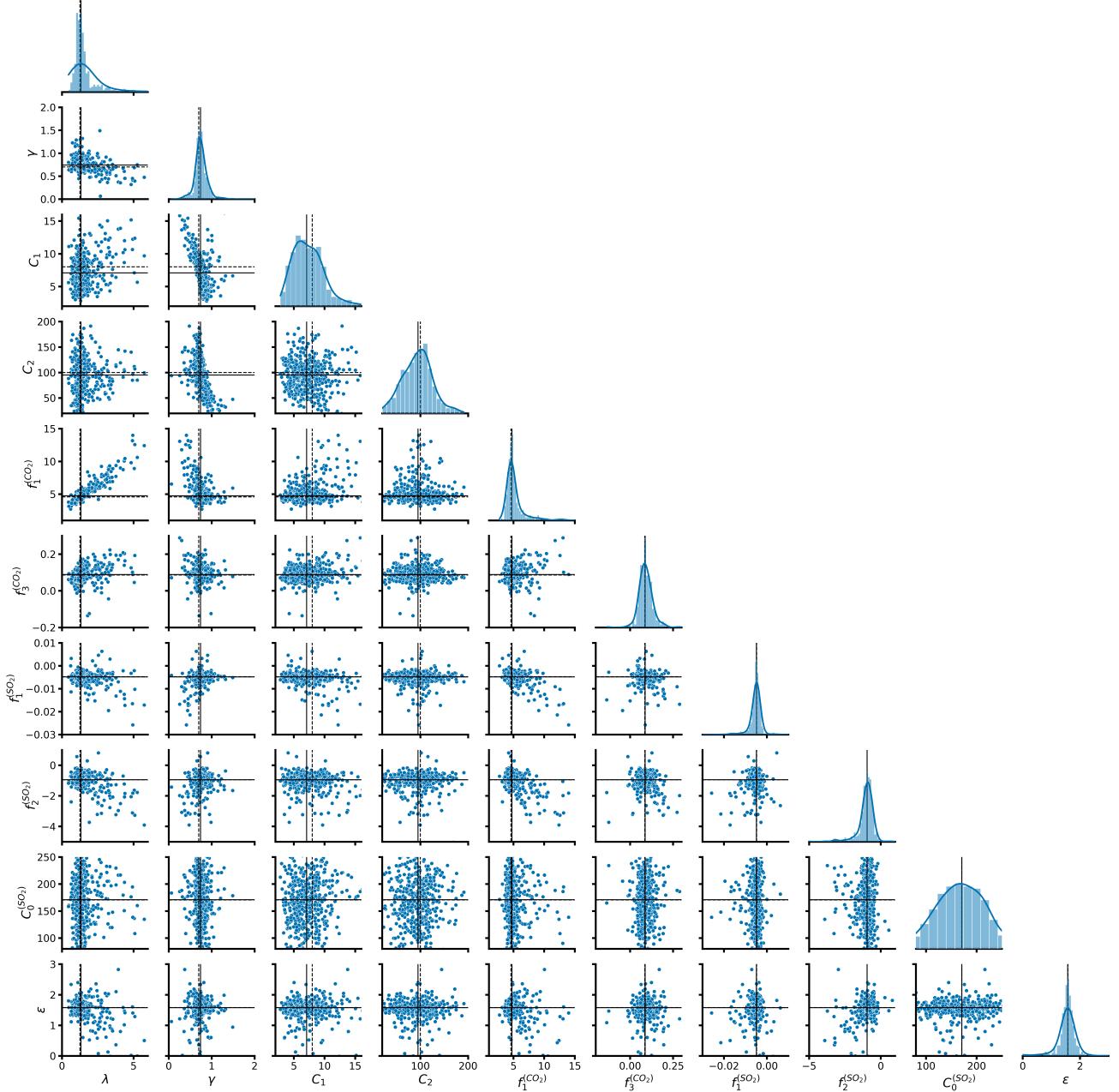
**Figure 7: ECS and TCR posterior central estimates and learning rates, high internal variability sensitivity test.** **a** shows the central estimate of the posterior ECS distribution (solid lines) and the true value of the ECS used to generate the observations (dashed lines) as the assimilation window is moved further into the future (and more observations are used for analysis). **b** shows the difference between the posterior central estimate and the true value, while **c** shows the raw posterior 5-95 percentile range and the relative 5-95 percentile range. **d–f** are as **a–c**, but for TCR.



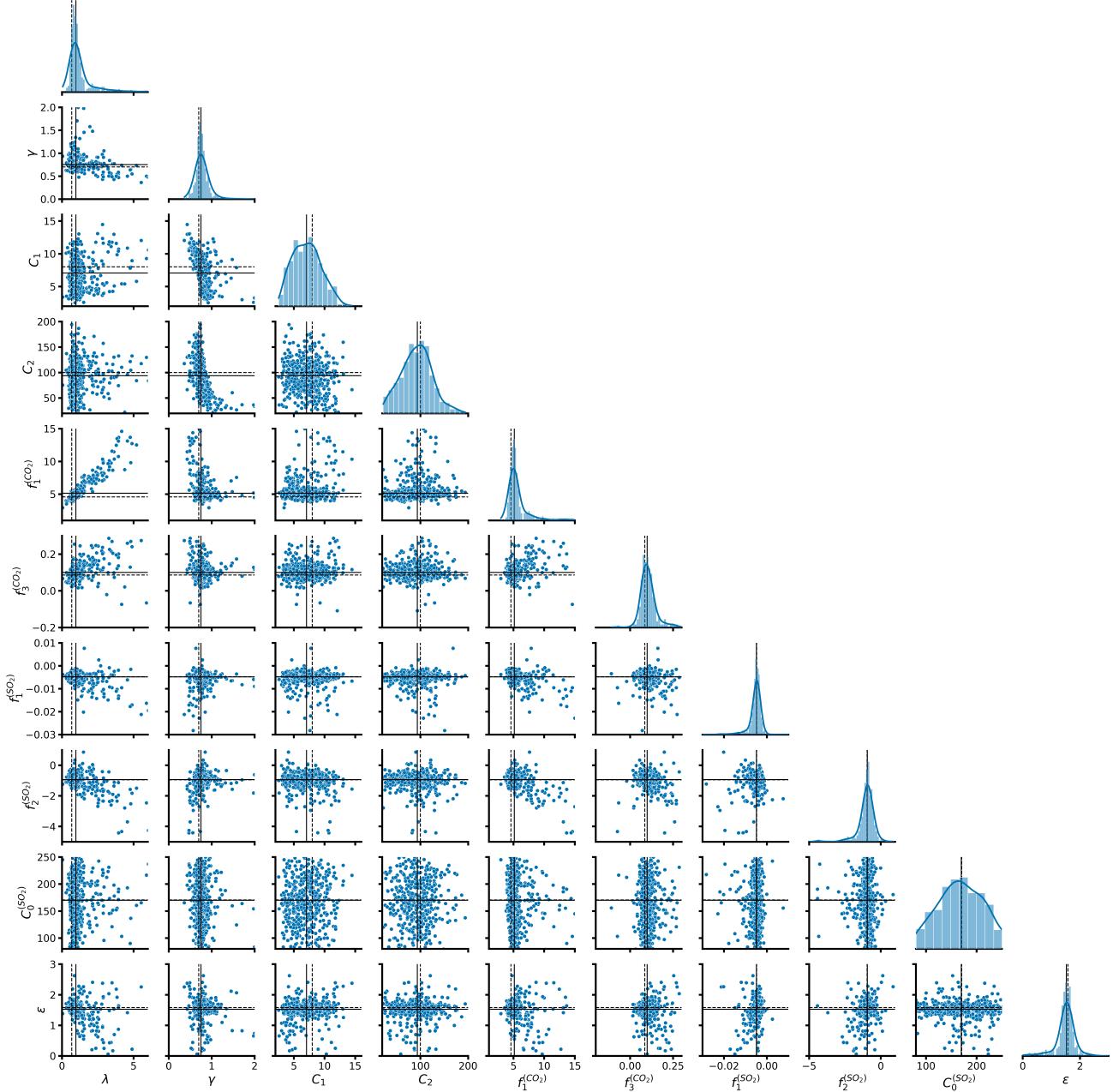
**Figure 8: Parameter posteriors for low true value of ECS (2 °C).** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window.



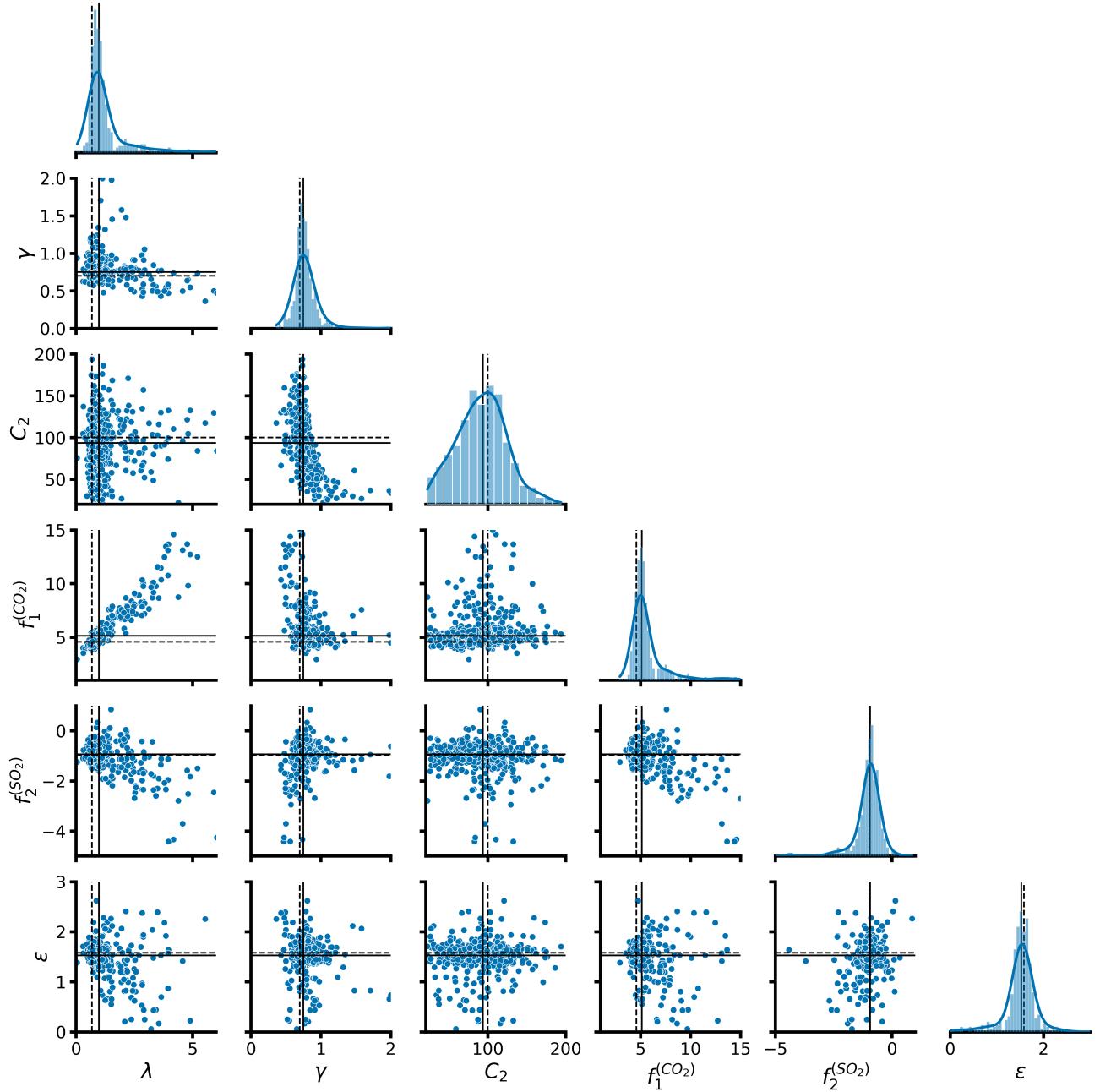
**Figure 9: Abbreviated parameter posteriors for low true value of ECS (2 °C).** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window. This is a truncated corner plot where we only show a subset parameters that we estimate in our data assimilation approach:  $\lambda$ ,  $\gamma$ ,  $C_2$ ,  $f_1^{(CO_2)}$ ,  $f_2^{(SO_2)}$ ,  $\varepsilon$ .



**Figure 10: Parameter posteriors for average true value of ECS (3 °C).** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window.



**Figure 11: Parameter posteriors for high true value of ECS (6 °C).** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window.



**Figure 12: Abbreviated parameter posteriors for high true value of ECS (6 °C).** Black solid lines represent the ensemble median estimate, while the dashed lines represent the true value. Parameter posteriors shown are for the 2020-2060 assimilation window. This is a truncated corner plot where we only show a subset parameters that we estimate in our data assimilation approach:  $\lambda$ ,  $\gamma$ ,  $C_2$ ,  $f_1^{(CO_2)}$ ,  $f_2^{(SO_2)}$ ,  $\varepsilon$ .