Regional mixed layer depth as a climate diagnostic and emergent constraint

G. Hall[1], B. Fox-Kemper[1,2]

[1]Institute at Brown for Environment and Society, Brown University, Providence, Rhode Island 02912, USA
[2]Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, Rhode Island 02912, USA

Key Points:

- Regional oceanic mixed layer depth predicts the rate and extent of warming in CMIP6 models: across a 25-model ensemble and validated with 9 out-of-sample models.
- North and South Ocean surface mixing correlate strongly with different patterns of ocean ventilation.
- Mixed layer depth observations allow us to constrain the range of model uncertainty in future warming by nearly 40%.

Corresponding author: Baylor Fox-Kemper, `baylor@brown.edu`

Abstract

The global ocean modulates Earth's warming. Most research has nevertheless ignored using ocean processes to improve warming projections in general circulation models (GCMs). We show that regional mixed layer depth (MLD) constrains climate sensitivity through its relation to ocean heat uptake. We correlate the parameters of two-layer energy balance models (EBMs) to pre-forcing MLD in the north, tropical and south ocean for a 25-member ensemble, and use the results to explain 47% of the variance in the effective climate sensitivity (S) of a 9-member validation ensemble. Using a climatology of observations from the Argo float network, we then constrain the EBMs which alters the range of S for the whole 34-member EBM ensemble from 4.42 $(3.09-5.65)$°C to 4.51 $(3.81-5.21)$°C – a 45% reduction in the span of the 66% (likely) range. This result indicates new potential mechanisms for ensemble spread, a new use for Argo measurements, and path to improving GCMs.

Plain Language Summary

Climate or earth system models predict the rate and scale of global warming as the result of land, ocean, and atmospheric processes in response to greenhouse gas concentration change, also called climate sensitivity. Here a set of ocean mixing measures are shown to correlate with climate models' sensitivity, so ocean processes are linked to climate change. Prior studies have focused on atmospheric mechanisms such as cloud feedbacks. We show that the regional depth of the oceanic mixed layer is strongly correlated to warming in an ensemble of 34 climate models including both atmospheric and oceanic processes. This relation together with observed mixed layer depths constrains the uncertainty range for the earth's climate sensitivity, reducing the range of sensitivities that are consistent with observations by 45%.

1 Introduction

Many climate change impacts scale with warming (Flato et al., 2013), so process-level understanding of the amount of global warming (e.g., Jones & Friedlingstein, 2020; Ehlert & Zickfeld, 2017)) and the speed at which it arrives (e.g., Fyfe et al., 2016; Solomon et al., 2010) are central tasks for climate models. Constraining the real-world equilibrium climate sensitivity (ECS) has proven difficult, with small improvements over decades of research. Recent studies of *emergent constraints* on climate sensitivity have improved these predictions using observed climatic quantities. However, emergent constraints research has focused mostly on atmospheric and cloud processes, despite the central role that the global ocean is known to play in modulating the climate's temperature response to forcing by absorbing energy. We propose a new set of constraints, based on the regional mixed layer depths (MLDs) of the modelled ocean. Rather than link these depths directly to climate model outcomes, instead we investigate their relationships to parameters in a two-layer energy balance model (EBM) that emulates the climate system and connects MLD to both the ECS and the near-term speed of the temperature response to radiative forcing. Thereby, interpretation of what the parameters in that two-layer EBM relate to in the real ocean is also improved.

The ECS of the earth system refers to the steady-state warming resulting from a doubling of atmospheric CO_2. In a simple formulation, the heat uptake N following a forcing perturbation F is determined by the change in average global surface temperature ΔT modified by a feedback parameter λ:

$$N = F - \lambda \Delta T, \tag{1}$$

such that when $N = 0$, warming ceases and the corresponding ΔT is the ECS. (Quantities without subscripts refer to global averages.)

ECS is to be distinguished from effective climate sensitivity, S, which is the equilibrium sensitivity extrapolated from only those feedbacks active in the 150 years after a modelled doubling of CO_2, and which is frequently estimated by finding the ΔT-intercept of the linear relationship between N and ΔT over that time period (Gregory, 2004).

Two recent developments have influenced this area of research. First, some GCMs in the sixth iteration of the Coupled Model Intercomparison Project (CMIP6, see Eyring et al. (2016)) have produced ECS estimates higher than the models preceding them (P. M. Forster et al., 2020). Much present research focuses on whether, or to what extent, these new models' projections are accurate when compared to older ones, and why (Zelinka et al., 2020; Zhu et al., 2020; Meehl et al., 2020; Roberts et al., 2020; Notz & Community, 2020; Hermans et al., 2021). The second development concerns emergent constraints on climate sensitivity (Eyring et al., 2019). Emergent constraints are empirical relationships between present and future climate model variables (call them A and B), motivated by a physical mechanism. An emergent constraint allows an update to probability distributions of future relationships based on how present observations relate to modelled variables. If the modelled relationship between A_{model} and B_{model} is consistent across models and justified *by a physical mechanism*, then $A_{observed}$ should also constrain $B_{observed}$ (Nijsse & Dijkstra, 2018; Williamson et al., 2018). A recent review synthesizes 17 emergent constraints to inform estimates of S (Sherwood et al., 2020), and a few recent papers have put forward processes that might explain high ECS in nature and models (Proistosescu & Huybers, 2017; Bjordal et al., 2020; Zelinka et al., 2020; Gjermundsen et al., 2021). The review concludes that both low ($< 2.7°$C) and very high ($> 4.7°$C) ECS values can likely be ruled out. Other authors prefer other metrics over ECS (Knutti et al., 2017), affected by a different collection of process considerations (Bronselaer & Zanna, 2020).

In addition to ECS or S, the rate of warming is key. Recent progress on transient warming has focused on ocean heat uptake (Yoshimori et al., 2016; Von Schuckmann et al., 2020). The deep ocean is frequently assumed to set the warming rate over long timescales (Rosenthal et al., 2017) and affects decadal variability (Liu & Xie, 2018). The surface mixed layer of the ocean, our focus, modulates seasonal and diurnal atmospheric cycles (Frankignoul & Hasselmann, 1977), and its peak wintertime conditions set water properties of subduction to the deeper ocean through Stommel's "Demon" (Williams et al., 1995; Stommel, 1979). Here the mixed layer depth in different regions is explored as an observable proxy for a variety of upper ocean processes. The upper ocean stratification and mixed layer depth relate both to climate change and its impacts (Sallée et al., 2021) and are sensitive indicators of the representation of upper ocean processes in models (Fox-Kemper et al., 2011; Belcher et al., 2012; Li et al., 2019). Regional surface mixing also ventilates deeper waters to affect the warming rate (Marzocchi et al., 2021), and similar GCMs differ in rates of transient warming due to differing surface mixing strength (Semmler et al., 2021).

1.1 Theory

Hasselmann (1976) proposed a model of the ocean response to weather and climate variability in which the ocean's large heat capacity reservoir integrates over transient atmospheric perturbations. Frankignoul and Hasselmann (1977) propose the ocean mixed layer as the reservoir, with stochastic fluxes from weather and a negative feedback restoring conditions toward climatology. With a uniform mixed layer of depth h, ocean surface temperature change is given by:

$$\frac{d\Delta T_o}{dt} = \frac{F_o}{c_s h} - \lambda_o \Delta T_o, \tag{2}$$

where c_s is the heat capacity per unit depth of the mixed layer, and the subscript o refers to ocean surface averages (rather than global). If the Hasselmann hypothesis motivating equation 2 is correct, and a time-invariant mixed layer depth holds approximately, then intermodel differences in h might explain some of the differences in the transient and perhaps even equilibrium warming, thus S and ECS.

Gregory (2000) proposed a two-layer ocean model of the climate system to explain the difference between near and long-term rates of warming in climate models. In this model, the heat capacity of the climate system comprises two constant-volume ocean layers termed the upper and lower ocean. The upper ocean (or *active* layer) responds to radiative forcing and exchanges heat with the lower ocean at a rate proportional to the temperature difference between the two,

$$\text{exchange rate} = \gamma(\Delta T - \Delta T_D)$$

where γ is the (constant) ocean heat uptake *efficiency*. Although the two layers in this model are global in extent, their depth, temperature, and other properties are homogeneous. Furthermore, ΔT encodes the *global* average surface temperature, not just the ocean surface temperature, because the ocean is the dominant heat sink for global heat imbalances. Winton et al. (2010) proposed an ocean heat uptake *efficacy* parameter, ε, which modifies the effect of heat exchange between the upper and lower ocean versus upper ocean heat anomaly remediation through atmospheric and radiative processes. This quantity distills vertical mixing and eddy processes that vary across models with resolution and parameterization choices (Raper et al., 2002; Griffies et al., 2015). The ratio between the effect of radiative feedbacks and ocean heat uptake feedbacks on global warming is encoded by ε (Winton et al., 2010; Held et al., 2010). With constant ε, Geoffroy et al. (2013a) give the resulting system:

$$\begin{aligned} C_S \frac{d\Delta T}{dt} &= F - \lambda \Delta T - \varepsilon\gamma(\Delta T - \Delta T_D), && (3) \\ C_D \frac{d\Delta T_D}{dt} &= \gamma(\Delta T - \Delta T_D); && (4) \\ \text{where } N &= C_S \frac{d\Delta T}{dt} + C_D \frac{d\Delta T_D}{dt} && (5) \end{aligned}$$

where C_S and C_D are the surface and deep-ocean heat capacities (which depend on layer volume or average layer depth). Note that while equations 3-5 are written with reference to global surface air temperature (GSAT), they might instead reference sea surface temperature (SST); primary differences between the two approaches result from sea ice freezing and insulation and continental climate variability. From an energy balance perspective the ocean is the dominant reservoir on decadal to centennial timescales suggesting SST or mean ocean temperature may be the more direct measure (Hansen et al., 2011; Trenberth & Fasullo, 2012; Trenberth et al., 2016), but we follow convention by using GSAT. The depth D_S of a uniform global ocean surface layer of heat capacity C_S is $D_S = C_S/(\rho c_p A_0)$, where $c_p = 4180$ J kg^{-1} K^{-1} is the heat capacity of seawater, $\rho = 1030$ kg m^{-3} is the density, and $A_0 = 3.6 \cdot 10^{14}$ m^2 is the area of Earth's surface covered by oceans. Geoffroy et al. (2013a) find that fitting this two-layer system, called EBM-ε, to estimate S gives better fits to modelled temperature changes than the linear model in Equation 1, and we typically use this approach to find S.

Recent studies have shown that the 2-layer EBM-ε is valuable as an emulator of metrics of climate sensitivity (Geoffroy et al., 2013a; Soldatenko & Colman, 2019; Winton et al., 2020) and related metrics such as thermosteric sea level rise (Palmer et al., 2018). For this reason, there is value in better understanding the variability of the parameters of these models and the processes they approximate in more complete earth system models. To this end, we follow Geoffroy et al. (2013b) and Kostov et al. (2014) in linking the process variability across a CMIP ensemble to the variations in 2-layer model parameters. However, our study uses a larger ensemble, CMIP6 models rather than CMIP5

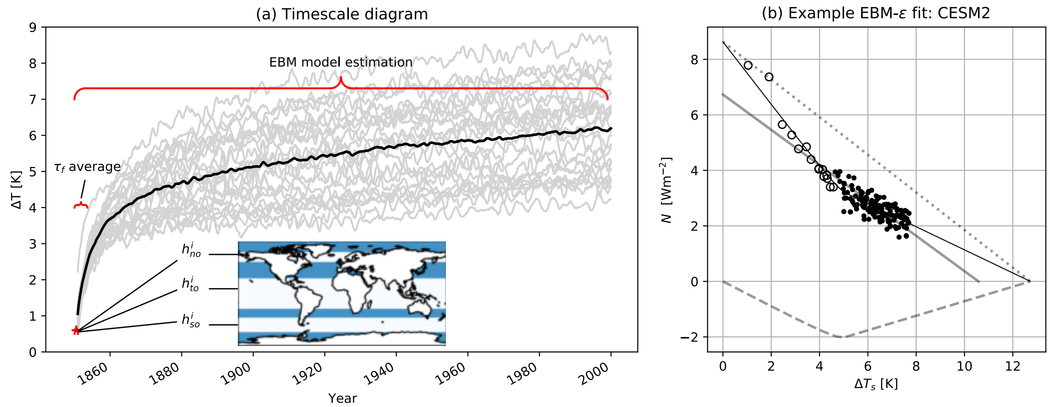

Figure 1. (a) Time ranges for calculated variables: Grey lines show each model's warming trend under the $4{\times}CO_2$ experiment; the black line shows the ensemble mean warming. Inset: shown in white, the latitudinal ranges for each zonal-mean ML quantity. (b) Example results of the EBM-ε model for ΔT data from a $4{\times}CO_2$ run with `CESM2`. TOA radiative imbalance N is plotted against the yearly average ΔT; empty circles represent the first 15 years and dots represent the remaining years. The thin black line shows the fitted relationship between N and ΔT from the EBM-ε. The solid gray line shows the same fit using Gregory (2004)'s one-layer relation. The dotted and dashed gray lines show the linear contribution $F - \lambda\Delta T$ and the $(1-\epsilon)H$ components of N, respectively. This figure mimics the layout in Geoffroy et al. (2013b).

models, and emphasizes the *observable* metric of mixed layer depth rather than the EBM-ε fitted parameters.

Note that the two-layer EBM-ε model described here does not admit of an immediate real-world interpretation of the layers or other parameters. Their depth or variability cannot be measured directly; only through simulations or extended observations and fitting of parameters are they estimable. The effective heat capacities C_S and C_D can be related to approximate depths of each ocean layer, D_S and D_D, but the corresponding surface layer of the global ocean may not follow the warming in Equation 3. Geoffroy et al. (2013b) find that D_S in an ensemble of CMIP5 models varies near 64 meters, which is near the mixed layer depth over much of the globe; however, Gregory (2000) argues that the active layer in 80 year simulations does not share geographical similarity to the mixed layer, despite the assumed relationship in earth system models with "slab" oceans and the short timescale model of Hasselmann (1976).

2 Methods

This study establishes diagnostic tests relating mixed layer depth $h(t, \text{longitude}, \text{latitude})$ to measures of global climate sensitivity. When *observed* regional h^i (where i stands for "initial", i.e. before forcing is applied) are used to estimate the constrained parameters, *observed* h^i is a 5-95% p-box range (highest high estimate, lowest low estimate) taken from the two different estimation methods in Holte et al. (2017). The first, combining the spatial average of profile-by-profile density threshold (with a criterion of 0.03 kg m^{-3}) following de Boyer Montégut et al. (2004), and the second, a profile-by-profile density algorithm method following Holte and Talley (2009) are combined into a p-box range. This climatology of mixed layer depths is based on the January 2000 to December 2019 Argo observations (Holte et al., 2017). The CMIP6 *mlotst* variable, representing h, uses the "sigma-t" density threshold method with a criterion of 0.03 kg m^{-3} in most models (Levitus, 1982; Griffies et al., 2016). Some inaccuracy in *mlotst* stems from different interpretations of this metric: some models measure a density difference from the surface grid cell and some from the 10 m depth, and one (excluded) model uses a different threshold value.

The CMIP6 variables used to determine the two-layer model parameters, correlations, and sensitivity include h, GSAT, ocean potential temperature θ, and the Eulerian mean meridional overturning streamfunction, which was used to calculate the AMOC strength and depth to aid interpretation of the results. The first GCM outputs in the CMIP6 ensemble were selected based on the 25 models with all of the MLD, GSAT, and TOA radiative imbalance variables used in this study available for download by March 2021. At that time, only 15 models' full θ profiles were available, and only 11 overturning streamfunctions; however, neither variable is central to our analysis. The correlations shown in the figures and tables in the main text text are based on these 25 models only. By February 2022 an additional 9 models had become available and were used as an out-of-sample validation of the relationships found with the first 25 models. All datasets were restricted to the first 150 years of data, which was the maximum available from all sets. The pre-forcing mixed layer depths were taken from averages over preindustrial control runs, and in cases where that data was not available, they were averaged from the first two years of a linear 1% per year forcing run because the adjustment in h over that timespan is negligible. When taking zonal mean and latitudinal band averages, variables were grouped in 2 degree bands and, where necessary, regridded using a bilinear algorithm onto a common grid prior to averaging. Datasets of all global and zonal mean variables used in this study were assembled and can be viewed in the supplemental materials.

2.1 TCR and effective ECS estimation

Transient climate response (TCR) was assessed as the average change in temperature (ΔT) from the initial value at the time of CO_2 doubling in the 1%/year scenarios, averaged over years 65-75. Effective climate sensitivity S was calculated using equations 1, 3, and 4 applied to model runs with an abrupt quadrupling of CO_2 (henceforth $4 \times CO_2$). To estimate the solution to equations 3 and 4, the multilinear approximation described in (Geoffroy et al., 2013b) was reproduced and applied to the $4 \times CO_2$ data (Fig. 1b). S therefore represents the effective equilibrium climate sensitivity estimated from the two-layer EBM-ε. In addition to providing S, this method estimates the model parameters C_S, C_D, λ, γ, F and ϵ. The detailed estimation procedure and all fitted parameters are provided in the supplemental material.

2.2 AMOC depth and strength calculation

AMOC depth was calculated using the Sun et al. (2020) method: the average depth at which the Atlantic meridional overturning streamfunction equals zero between 0 and 30 degrees latitude. However, some models had highly variable depth of the zero streamline contour in the meridional direction *after* warming, so the 0.1 Sv line was used as the point of reference instead. The 0 Sv and 0.1 Sv lines were similar in all cases before warming. Following Liu et al. (2020), AMOC strength was calculated as the maximum annual mean meridional overturning streamfunction below 500m in the North Atlantic Cell.

2.3 Modeling the two-layer model dependence on regional h^i

Our goal is to constrain our estimate of the real-world climate sensitivity S, as approximated by a two-layer model with parameters adjusted to fit constraints from the real-world zonal-mean northern, tropical, and southern MLD, $\mathbf{h}^{obs} = (h^{obs}_{no}, h^{obs}_{to}, h^{obs}_{so})$. The precise latitude ranges for each regional h were chosen to maximize the predictive power of the linear constraint: these were, respectively, $h_{no} \in [55°N, 75°N]$, $h_{to} \in [25°S, 25°N]$, and $h_{so} \in [65°S, 45°S]$. Furthermore, h_{no} is averaged over the northern winter (DJF),

h_{so} is averaged over the austral summer (DJF), and h_{to} is the annual mean. Because of the complex nature of the constraint and the presence of significant cross-correlations between the dependent variables correlated with mixed layer depths, we examine four methods to estimate the constrained parameter distributions and find they provide similar constrained projections. Each of these methods and their parameter distributions are described in the Supplemental Information. The first method, integration over a normal uncertainty range (Cox et al., 2018), is the most conceptually straightforward and common in prior literature, so unless otherwise stated we report constraints derived using this method.

3 Results

Fitting the EBM-ε model results in increased estimates of effective climate sensitivity, S, compared to the (Gregory, 2004) method, with an average increase of 0.39°C. This increase is almost entirely due to increases in the estimated initial forcing F, rather than changes to λ. In fact, λ tended to increase slightly (by 0.02 W/m^{2}K on average) in the EBM-ε model, so that this effect alone decreased the S estimates. F increased by 0.35 W/m^2 on average. The ranges of EBM-ε parameters obtained here deviate slightly from those found in Geoffroy et al. (2013b) (see Table S.1 in the supplemental material). In what follows, S refers to the ECS estimate from the two-layer models rather than the Gregory method. The range of ensemble active layer depths D_S (47-79m) overlaps heavily with the range of global average *initial* mixed layer depth (h^i_{ga}, 38-77m), calculated from the pre-forcing control runs. However variation in h^i_{ga} only explains approximately 20% of the variation in C_S. Thus, as Gregory (2000) notes, the active surface layer in EBM models is *not identical* to the oceanic mixed layer, as Equation 2 and Frankignoul and Hasselmann (1977) hypothesize on shorter timescales, but nonetheless mixed layer dynamics remain involved in modulating the active-layer response to forcing.

Figure 2 shows the correlated scatter between variation in pre-forcing regional MLD and EBM-ε variables in the right column panels, with shaded regions indicating the 5-95% range of present-day observed MLD measures. The left-hand panels show correlations between the same MLD variables and the zonal-mean change in ocean potential temperature by the last ten years of the $4 \times CO_2$ experiments. Panel 2a shows that a deeper northern initial MLD (h^i_{no}, 55°N to 75°N latitude) correlates with greater North Atlantic Deep Water warming under forcing. Panel 2b shows that upper pycnocline warming is related to a deeper tropical initial MLD (h^i_{to}). Panel 2c shows that a deeper southern initial MLD (h^i_{so}) strongly predicts warming in the top 1200m of the ocean structure, the Antarctic Intermediate Water, and cooling of the North Atlantic Deep Water. Only two models have h^i_{to} within the observed range.

Variation in h^i_{ga} explains a small-but-significant variation in efficacy, ε. Efficacy quantifies the ratio between equilibrium radiative climate feedbacks and deep ocean climate change, and therefore distinguishes between the feedbacks affecting transient versus equilibrium warming. Here the correlation arises primarily from variation in tropical h^i. We therefore define the tropical ocean mixed layer depth (h^i_{to}) as the average mixed layer depth between 26°S and 26°N latitude, and find that h^i_{to} explains 35% of the variation in ε and is correlated primarily with tropical upper ocean heating (Figure 2c, d). Note, however, that h^i_{to} is also positively correlated with a factor in each term in equation 3, i.e., C_S, S and ΔT through S, and insignificantly in γ. The Hasselmann (1976) model for a deeper active layer would predict larger ε to reflect an overall magnitude increase of the active layer budget equation 3 rather than a change in air-sea coupling relative to layer depth. Note that larger ε alone would indicate stronger atmospheric feedbacks that tend to reduce S (Winton et al., 2010; Rose et al., 2014).

Variation in the relative depth of the northern wintertime mixed layer (h^i_{no}) explains 45% of the intermodel variation in γ. This correlation points to the importance

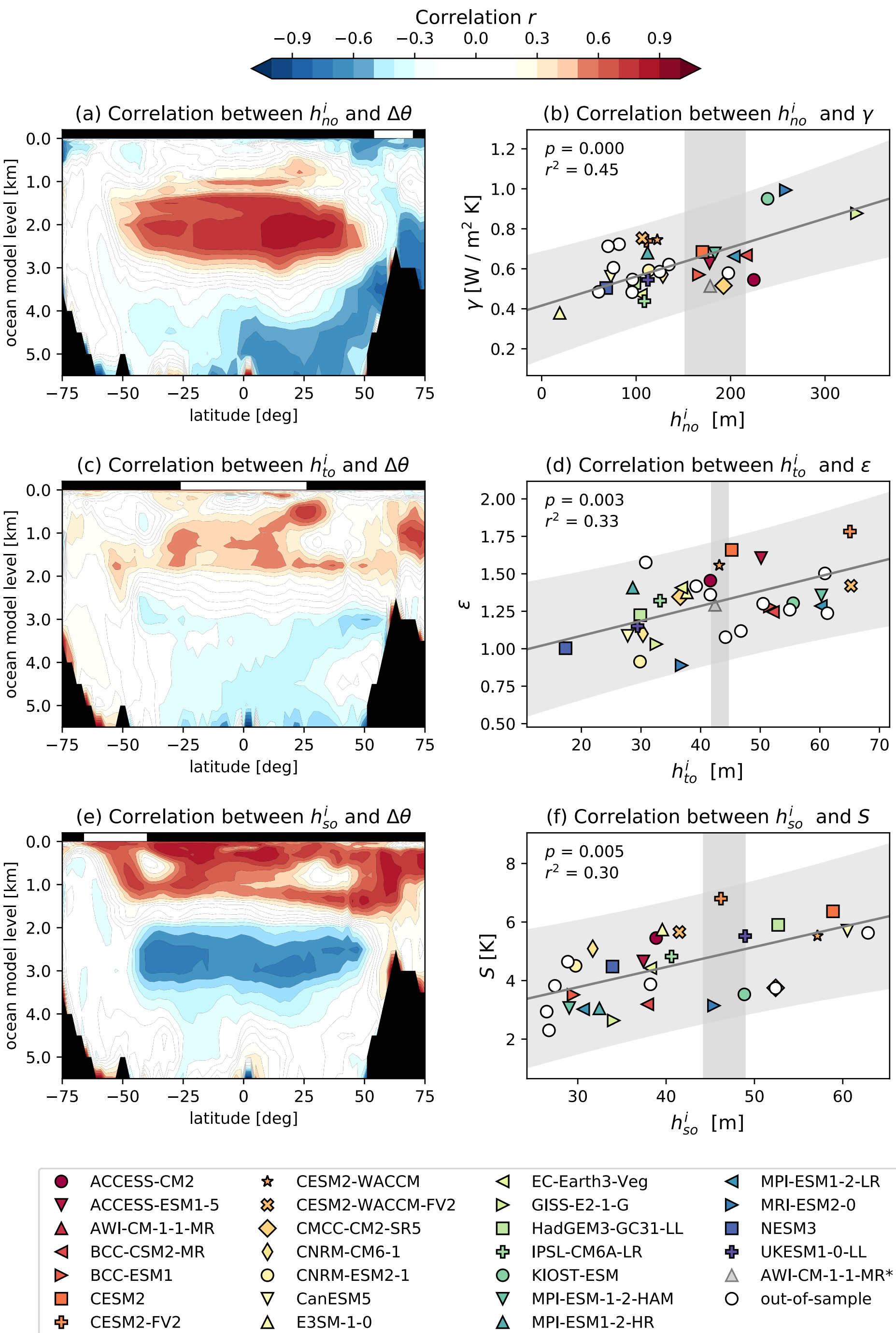

Figure 2. (left) Correlations between zonal-mean pre-forcing MLD h^i and $\Delta\theta$ averaged over years 140-150. The ranges of these zonal averages are indicated by the white bars at the top of figures (a)-(c). h^i_{no} is averaged over the northern winter (DJF), h^i_{so} is averaged over the austral summertime (DJF), and h^i_{to} is the yearly average. Intermodel variation in the different regions of h^i correlates with variation in different segments of the ocean warming structure. Colors emphasize significant correlations: $p < 0.05$ when $|r| > 0.513$ and $p < 0.01$ when $|r| > 0.641$. (Right) The right-hand panels demonstrate significant ($p < 0.01$) positive correlations between h^i_{no} and the ocean heat uptake efficiency γ, h^i_{to} and the ocean heat uptake efficacy ε, and h^i_{so} and the effective climate sensitivity S. The AWI model is excluded as an outlier, as they used a different density criterion that is more targeted at Arctic mixed layers (0.125 instead of 0.03 kg m^{-3}, Q. Wang pers. comm.). Shaded regions around the regressions show 90% confidence intervals, and vertical shaded regions indicate $5-95\%$ estimates for the present day values of h_{no}, h_{to}, and h_{so}. White circles indicate models added after regression as an out-of-sample validation.

of winter ocean mixing in determining deep ocean heat uptake, as hypothesized by Stommel (1979) and many others since. Furthermore, models with a deeper wintertime mixed layer have a larger-volume deep layer C_D and a larger equilibrium feedback parameter λ. The impact of h^i_{no} on deep water mixing can be seen in Figure 2b, which shows that models with greater h^i_{no} experience significantly larger γ, and Figure 2a shows that the warming is primarily in the Deep Water (particularly in North Atlantic Deep Water, but shown in zonal mean between 1000 m and 3000 m and 50°S to 50°N). Six models have h^i_{no} within the observed range.

Variation in the relative depth of the southern wintertime mixed layer (h^i_{so}, 45°S to 65°S latitude) explains 31% of the intermodel variation in S. Figure 2e shows that warming correlated with (h^i_{so}) opposes the Deep Water changes induced by (h^i_{no}), and warming of surface and intermediate waters is positively correlated. Four models have h^i_{so} within the observed range.

The right column of Figure 2 also includes the 9 out-of-sample models shown as white circles. These models fall within the expected rang of the correlations with h^i_{no}, h^i_{to}, and h^i_{so}. If these 9 out-of-sample models are included in the regression, then the EBM-ε parameters (Table 1) vary only slightly (Supplemental Information Table S.3).

The ocean heat uptake efficacy ε predicts the ratio between transient and equilibrium warming (Winton et al., 2010). Knowledge of ε, which is correlated to inter-model variation in h^i_{to}, should therefore improve predictions of the remaining warming alongside knowledge of h_{no} and h_{so}, which predict equilibrium warming. We define the heat uptake temperature T_H as the difference between the equilibrium warming and the realized warming at a given time t and forcing $F(t)$:

$$T_H(t) \equiv \Delta T_{eq}(F(t)) - \Delta T(t). \tag{6}$$

In 1%/year forcing experiments ΔT_{eq} varies with the forcing $F = \alpha t$ where α is the rate of increase, while in $4\times CO_2$ experiments F does not change so $\Delta T_{eq} = 2\times\text{ECS}$ at all times.

Figure 3c-d shows the correlations between T_H and ΔT, and between T_H and $\{h^i_{no}, h^i_{to}, h^i_{so}\}$. In both the $4\times\text{CO}_2$ and 1%/year experiments, the three regional initial mixed layer depths predict T_H significantly better than ΔT predicts T_H (note that r is adjusted for the greater number of independent variables using MLD). In the 1%/year ensemble, knowledge of just the pre-forcing MLDs explains over 60% of the inter-model spread in T_H.

Because the three regional mixed layer depths constrain each of the EBM-ε parameters (Table 1), we can estimate a constrained range of EBM-ε temperature predictions. Panel (a) of Figure 3 shows the CMIP6 ensemble spread ΔT for both experiments in light shading (with values after the year 2000 extrapolated using the EBM-ε prediction without MLD constraints), and the range of MLD-constrained EBM-ε predictions in darker shading. Panel (b) shows the EBM-ε estimation of T_H at all times.

4 Discussion

Our results show that the pre-forcing, initial mixed layer depth provides information about future warming of the upper ocean and climate system. Indeed, Figures 3c, d show that initial mixed layer depth is a much more effective predictor of remaining warming than the observed warming before equilibrium is reached is. Improving the climate sensitivity of models (Chapman et al., 2020; Small et al., 2020) deserves significant effort. The results here show that initial, or in these idealized forcing experiments the equivalent of present-day (year 0-10), mixed layer depth biases affect the active layer depth over the following century. However, the simple interpretation of Hasselmann (1976) that the mixed layer is essentially the active layer–i.e., governing the heat capacity of the ocean

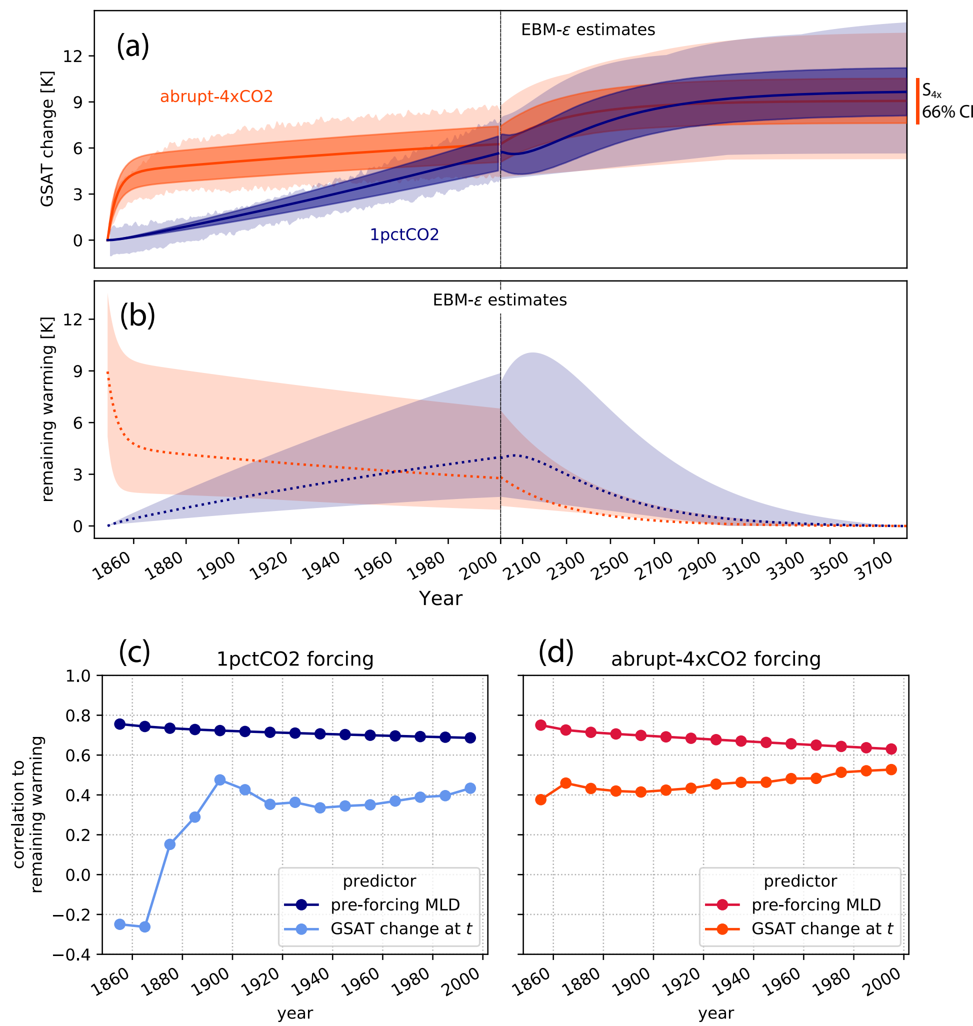

Figure 3. Regional mixed layer depth predicts equilibrium warming. (a) the ensemble mean surface warming ΔT for both experiments. Lighter shaded areas indicate the model spread, extrapolated past year 2000 using the EBM-ε prediction. Dark shaded regions show the 5-95% range of the surface warming under each scenario using EBM-ε predictions with parameters constrained by $h^i_{n,t,s}$. Note the time scale discontinuity at year 2000. (b) The ensemble difference between equilibrium warming and realized warming for forcing $F(t)$, as predicted by the EBM-ε fits. (c-d) Scatter plots showing the correlation between decadal-mean realized and remaining warming which is a standard metric to predict remaining warming (heat uptake temperature), and between the proposed emergent constraints of initial regional mixed layer depths ($h^i_{n,t,s}$) and remaining warming, for c) `1pctCO2` and d) `abrupt-4xCO2` simulations.

Table 1. Ensemble correlations between EBM-ε parameters as well as multiple measures of ocean and climate variables. Overbars indicate variables calculated prior to forcing. S is calculated using the EBM-ε model. The AWI GCM was excluded from all correlations because of its outlier density criterion leading to a different southern ocean mixed layer depth. Statistical significance is indicated in **bold** $\rightarrow p < 0.01$ and with $* \rightarrow p < 0.05$.

	C_D	ε	λ	γ	F	S	TCR	h^i_{ga}	h^i_{no}	h^i_{to}	h^i_{so}	AMOC S_i	AMOC S_f	AMOC D_i	AMOC D_f
	heat	efficacy	rate	exch.	forcing	equil.	trans.	global	55N-75N	26S-26N	65S-45S	strength	strength	depth	depth
C_S	-0.24	**0.52**	0.22	-0.29	0.49*	-0.02	-0.15	0.41*	-0.11	0.42*	-0.05	-0.23	-0.09	**-0.55**	-0.05
C_D		-0.28	0.44*	**0.61**	0.05	-0.38	-0.43*	0.27	**0.55**	0.18	-0.38	0.1	-0.13	0.33	0.22
ε			-0.18	-0.06	0.41*	0.43*	0.37	0.43*	-0.13	**0.58**	0.25	-0.44*	-0.26	-0.4	-0.06
λ				0.31	0.46*	**-0.87**	**-0.57**	0.25	**0.58**	0.13	-0.46*	0.23	0.21	0.06	**0.78**
γ					-0.18	-0.36	-0.34	**0.53**	**0.67**	0.39	0.1	-0.05	0.15	0.21	0.26
F						-0.01	-0.12	0.01	0	0.08	0.08	-0.37	-0.28	-0.4	-0.08
S							**0.59**	-0.24	**-0.62**	-0.04	**0.55**	-0.47*	-0.36	-0.32	**-0.77**
TCR								-0.23	-0.4	-0.04	0.22	**-0.76**	-0.18	-0.03	-0.45*
h^i_{ga}									**0.57**	**0.87**	0	-0.27	0.08	-0.45*	-0.13
h^i_{no}										0.28	-0.13	0.16	0.28	0.11	0.15
h^i_{to}											-0.08	-0.32	-0.11	-0.18	-0.15
h^i_{so}												**-0.57**	-0.01	-0.5*	**-0.55**
AMOC S_i													**0.65**	0.17	**0.74**
AMOC S_f														-0.09	0.43*
AMOC D_i															0.17

surface energy reservoir–is demonstrated insufficient to capture the different roles that mixed layers play in different regions. In the tropical oceans on decadal to centennial timescales the Hasselmann theory seems to apply, but the role of mixed layers and surface temperature response in polar heat uptake is quite different from tropical as revealed when patterns of deep ocean heat uptake are varied (Rose et al., 2014). The different roles of the mixed layer in different regions, as well as the different connectivity by region of the surface ocean to the deep ocean, is a key aspect of the ocean warming pattern effect on regional and global climate (Xie, 2020).

The northern ocean mixed layer is correlated with the deep layer heat capacity (C_D), the surface to deep exchange coefficient (γ), and the equilibrium feedback parameter (λ). The equilibrium feedback parameter (λ) is also highly correlated with AMOC depth at the end of the simulation (Table 1). Thus, the mixed layer and AMOC both contribute to watermass transformation and the ventilation of North Atlantic Deep Water, and thereby climate sensitivity consistent with past studies (Marshall & Zanna, 2014; Petit et al., 2020; Pickart & Spall, 2007; Jackson & Wood, 2020; Kostov et al., 2014; Heuzé, 2017, 2021; Sun et al., 2020). Unlike Kostov et al. (2014) find in an 8 member CMIP5 ensemble, we do not find statistically significant correlations between AMOC strength and γ. This may be related to the cancellation in NADW change between northern overturning and southern overturning, both of which contribute to γ. Here the direct correlations between initial northern ocean mixed layer depth and final AMOC strength or depth across models are not strong, indicating that other effects participate as well.

The Southern Ocean mixed layer affects both the Antarctic Intermediate Water, through mixing and ventilation of the southern meridional overturning, and North Atlantic Deep Water. Many studies emphasize the role of the Southern Ocean mixed layer in heat uptake in near decades (e.g., Morrison et al., 2016), and a recent study directly relates convective mixing shutdown to the potential for long-term warming in CMIP6 models (Gjermundsen et al., 2021). They argue that models with high climate sensitivity result from Southern Ocean positive low cloud feedbacks, which in turn result from warmer SSTs and weaker atmospheric boundary layer inversions (see also Gettelman & Sherwood, 2016). Studies comparing ocean mixed layer depths reveal that SST biases on seasonal and longer timescales are not connected to mixed layer depth biases, but by different responses to winds, waves, and convection (Belcher et al., 2012; Li et al., 2019). Weaker atmospheric inversions tend to make momentum and moisture transfer easier from the free atmosphere to the atmospheric boundary layer, raising winds and lowering humidity at the sea surface, thus deepening the ocean mixed layer through both wind-driven and evaporative (latent cooling) forcing. As a correlation between S and h_{so} does not imply causation but potentially a confounding agent that causes both to occur, note that too weak atmospheric boundary layer inversions in the Southern Ocean would cause both excessive positive cloud feedbacks and too deep h_{so}. Parameterization biases in both fluids will complicate this effect further. Deep Southern Ocean mixed layers have the opposite effect on NADW as deep northern ocean mixed layers (Gnanadesikan, 1999; Marshall & Zanna, 2014; Heuzé, 2021), which is one reason why examining the *global* mean mixed layer depth does not reveal the processes controlling NADW temperature: h_{so} deepening correlates with NADW cooling, while h_{no} deepening correlates with NADW warming (Figure 2a and 2c). Southern Ocean warming biases have been noted as a critical failing in models for some time (Belcher et al., 2012; Sallée et al., 2013; Durack et al., 2014), but the relationship of these biases to mixed layer processes (Belcher et al., 2012; Li et al., 2019), changing stratification (Sallee et al., 2020; Sallée et al., 2021), and unresolved fronts and eddies (Bachman & Klocker, 2020) is not fully understood.

5 Conclusions: Constrained Projections and Outlook

The opposing correlations of southern and northern ocean mixing (Fig. 2a, e) and the distinct tropical effects require us to constrain S using variation in h^i_{no} and h^i_{so} rather than h^i_{ga}. Geoffroy et al. (2013b, 2013a) provide solutions for the effective ECS and $\Delta T_S(t)$ given the two-layer model parameters, and Table 1 provides the proportionality of those parameters on mixed layer depth in each region. We can therefore constrain the EBM-ε parameters based on Argo-observed mixed layer depths and use the constrained EBM-ε emulator to estimate future warming by the CMIP6 models were they to have had mixed layer depths equal to the observed depths in the real ocean. In particular, since S has opposite-sign dependencies on h^i_{no} and h^i_{so}, which are in practice independent within the ensemble, using both MLDs at once provides a relatively strong constraint on S.

Using a variety of multiple linear regression techniques described in the supplemental material with the original 25-member ensemble, the 66% confidence range of effective climate sensitivity S changes from (3.13–5.71)°C to (3.88–5.43)°C under adjustment to the observed northern and southern mixed layer depths. This change is a 40% reduction in the uncertainty range. Under regional h constraints the mean S value warms from 4.51°C to 4.66°C. Testing the relationship between (h_{no}, h_{so}) and S derived from the original 25-member ensemble on the 9 out-of-sample models explained 47% of the variation in S (Figure S.2). Finally, using constraints on all 34 models to find S, we arrive at a reduction in the 66% uncertainty range for S from 4.42 (3.09−5.65)°C to 4.51 (3.81−5.21)°C.

Note that the pre-constraint range is from the selected CMIP6 model ensemble not an assessed range (P. Forster et al., 2021). The mean values of S and most of the other EBM-ε parameters change insignificantly, because observed MLDs are near the middle of the ensemble ranges, but the MLD emergent constraints shrink the uncertainty ranges on each variable considerably (Supplemental Figure 1). These adjustments are as large as several of the emergent constraint adjustments suggested in (Sherwood et al., 2020). The large scatter in Figure 3a ensemble timeseries, and the much narrower ensemble range for the constrained timeseries, illustrate how much of the uncertainty in warming is explained by initial mixed layer depth biases across the ensemble. Thus, the MLD constraints revealed here are not trivial adjustments but constitute a large potential for using mixing of the upper ocean to constrain climate sensitivity.

However, a study based on diagnosis of simulations of this kind cannot distinguish causality directly, nor follow all of the consequences of altering the GCMs so as to arrive at more realistic MLDs. It is not clear what the direct mechanism of these correlations between mixed layer depth and two-layer model parameters is. For example, the mixed layers affect temperature through entrainment of colder water to the surface and through ventilation of deeper water, but they also affect clouds and cloud feedbacks, tend to be deeper when winds are stronger but stronger winds have many other effects, change the intensity of seasonal and diurnal cycles, and other important consequences, correlates, and confounding variables abound. Thus, it is unclear if the mixed layer depth biases are the cause or merely a symptom of the model biases leading to spread in their sensitivity. So long as the mixed layer correlations with sensitivity remain valid, using regional h^i as an emergent contraint is valid even if the causal links are not fully clear. Recall that these changes are predicted by the *initial* mixed layer depths, $h^i_{so}, h^i_{to}, h^i_{no}$, not the evolving mixed layer depths, which does suggest causality based on biases in mixed layer depth preceding other consequences in time. However, the mixed layer biases in each model tend to persist throughout each run so this sequencing is not dispositive. Thus, important next steps are to show that altering processes or parameterizations that change the regional mixed layer depths also change climate sensitivity of the magnitude noted here, and furthermore to note the processes and mechanisms triggered by and leading to these mixed layer depth changes.

5.1 Open Research

CMIP6 data was provided and accessed by Pangeo and ESGF (Abernathey et al., 2021; Eyring et al., 2016). All processed datasets (including zonal and global-mean variables) and the Jupyter notebook necessary to computationally reproduce these results are available at https://doi.org/10.7910/DVN/NYFZJJ.

References

Abernathey, R. P., Augspurger, T., Banihirwe, A., Blackmon-Luca, C. C., Crone, T. J., Gentemann, C. L., … Signell, R. P. (2021, mar). Cloud-native repositories for big scientific data. *Computing in Science & Engineering*, *23*(02), 26-35. doi: 10.1109/MCSE.2021.3059437

Bachman, S. D., & Klocker, A. (2020). Interaction of jets and submesoscale dynamics leads to rapid ocean ventilation. *Journal of Physical Oceanography*, *50*(10), 2873–2883.

Belcher, S. E., Grant, A. L. M., Hanley, K. E., Fox-Kemper, B., Van Roekel, L., Sullivan, P. P., … Polton, J. A. (2012, September). A global perspective on Langmuir turbulence in the ocean surface boundary layer. *Geophysical Research Letters*, *39*(18). Retrieved 2020-05-15, from `http://doi.wiley.com/10.1029/2012GL052932` doi: 10.1029/2012GL052932

Bjordal, J., Storelvmo, T., Alterskjær, K., & Carlsen, T. (2020). Equilibrium climate sensitivity above 5°C plausible due to state-dependent cloud feedback. *Nature Geoscience*. Retrieved from `https://doi.org/10.1038/s41561-020-00649-1` doi: 10.1038/s41561-020-00649-1

Bronselaer, B., & Zanna, L. (2020). Heat and carbon coupling reveals ocean warming due to circulation changes. *Nature*, *584*(7820), 227–233.

Chapman, C. C., Lea, M.-A., Meyer, A., Sallée, J.-B., & Hindell, M. (2020). Defining southern ocean fronts and their influence on biological and physical processes in a changing climate. *Nature Climate Change*, 1–11.

Cox, P. M., Huntingford, C., & Williamson, M. S. (2018, January). Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature*, *553*(7688), 319–322. Retrieved 2020-08-05, from `https://www.nature.com/articles/nature25450` (Number: 7688 Publisher: Nature Publishing Group) doi: 10.1038/nature25450

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, *109*(C12).

Durack, P. J., Gleckler, P. J., Landerer, F. W., & Taylor, K. E. (2014). Quantifying underestimates of long-term upper-ocean warming. *Nature Climate Change*, *4*(11), 999–1005.

Ehlert, D., & Zickfeld, K. (2017). What determines the warming commitment after cessation of co_2 emissions? *Environmental Research Letters*, *12*(1), 015002.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016, May). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. Retrieved 2020-04-26, from `https://www.geosci-model-dev.net/9/1937/2016/` (Publisher: Copernicus GmbH) doi: https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., … Williamson, M. S. (2019, February). Taking climate model evaluation to the next level. *Nature Climate Change*, *9*(2), 102–110. Retrieved 2020-05-27, from `https://www.nature.com/articles/s41558-018-0355-y` (Number: 2 Publisher: Nature Publishing Group) doi: 10.1038/s41558-018-0355-y

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., ... Rummukainen, M. (2013). Evaluation of climate models. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the IPCC* (pp. 741–866). Cambridge University Press. Retrieved from http://bit.ly/2wNMqXD

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., ... Zhang, H. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In V. Masson-Delmotte et al. (Eds.), (chap. The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity). Cambridge University Press. (In press)

Forster, P. M., Maycock, A. C., McKenna, C. M., & Smith, C. J. (2020, January). Latest climate models confirm need for urgent mitigation. *Nature Climate Change*, *10*(1), 7–10. Retrieved 2020-08-05, from https://www.nature.com/articles/s41558-019-0660-0 (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41558-019-0660-0

Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Griffies, S. M., Hallberg, R. W., Holland, M. M., ... Samuels, B. L. (2011). Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Modelling*, *39*, 61–78. Retrieved from http://dx.doi.org/10.1016/j.ocemod.2010.09.002 doi: 10.1016/j.ocemod.2010.09.002

Frankignoul, C., & Hasselmann, K. (1977, January). Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, *29*(4), 289–305. Retrieved 2020-03-26, from https://www.tandfonline.com/doi/full/10.3402/tellusa.v29i4.11362 doi: 10.3402/tellusa.v29i4.11362

Fyfe, J. C., Meehl, G. A., England, M. H., Mann, M. E., Santer, B. D., Flato, G. M., ... others (2016). Making sense of the early-2000s warming slowdown. *Nature Climate Change*, *6*(3), 224–228.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., & Tytéca, S. (2013a, March). Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, *26*(6), 1841–1857. Retrieved 2020-05-01, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00195.1 doi: 10.1175/JCLI-D-12-00195.1

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., & Tytéca, S. (2013b, March). Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, *26*(6), 1859–1876. Retrieved 2020-04-30, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00196.1 doi: 10.1175/JCLI-D-12-00196.1

Gettelman, A., & Sherwood, S. (2016). Processes responsible for cloud feedback. *Current climate change reports*, *2*(4), 179–189.

Gjermundsen, A., Nummelin, A., Olivié, D., Bentsen, M., Seland, Ø., & Schulz, M. (2021). Southern ocean convection shutdown controls potential for long-term climate warming by greenhouse gases.
(preprint)

Gnanadesikan, A. (1999). A simple predictive model for the structure of the oceanic pycnocline. *Science*, *283*(5410), 2077–2079.

Gregory, J. M. (2000). Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, *16*(7), 501–515. (Place: Berlin/Heidelberg Publisher: Springer-Verlag) doi: 10.1007/s003820000059

Gregory, J. M. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3), L03205. Retrieved 2020-04-28, from http://doi.wiley.com/10.1029/2003GL018747 doi: 10.1029/

2003GL018747
Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., … others (2016). Omip contribution to cmip6: Experimental and diagnostic protocol for the physical component of the ocean model intercomparison project. *Geoscientific Model Development*, *9*(9), 3231–3296.
Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., … others (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, *28*(3), 952–977.
Hansen, J., Sato, M., Kharecha, P., & von Schuckmann, K. (2011, December). Earth's energy imbalance and implications. *Atmospheric Chemistry and Physics*, *11*(24), 13421–13449. Retrieved 2020-04-19, from `https://www.atmos-chem-phys.net/11/13421/2011/` doi: 10.5194/acp-11-13421-2011
Hasselmann, K. (1976, January). Stochastic climate models Part I. Theory. *Tellus*, *28*(6), 473–485. Retrieved 2020-05-01, from `https://www.tandfonline.com/doi/full/10.3402/tellusa.v28i6.11316` doi: 10.3402/tellusa.v28i6.11316
Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., & Vallis, G. K. (2010, May). Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of Climate*, *23*(9), 2418–2427. Retrieved 2020-04-18, from `http://journals.ametsoc.org/doi/10.1175/2009JCLI3466.1` doi: 10.1175/2009JCLI3466.1
Hermans, T. H., Gregory, J. M., Palmer, M. D., Ringer, M. A., Katsman, C. A., & Slangen, A. B. (2021). Projecting global mean sea-level change using cmip6 models. *Geophysical Research Letters*, *48*(5), e2020GL092064.
Heuzé, C. (2017). North atlantic deep water formation and amoc in cmip5 models. *Ocean Science*, *13*(4), 609–622.
Heuzé, C. (2021). Antarctic bottom water and north atlantic deep water in cmip6 models. *Ocean Science*, *17*(1), 59–90.
Holte, J., & Talley, L. (2009). A new algorithm for finding mixed layer depths with applications to argo data and subantarctic mode water formation. *Journal of Atmospheric and Oceanic Technology*, *26*(9), 1920–1939.
Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An argo mixed layer climatology and database. *Geophysical Research Letters*, *44*(11), 5618–5626.
Jackson, L., & Wood, R. (2020). Fingerprints for early detection of changes in the amoc. *Journal of Climate*, *33*(16), 7027–7044.
Jones, C. D., & Friedlingstein, P. (2020). Quantifying process-level uncertainty contributions to tcre and carbon budgets for meeting paris agreement climate targets. *Environmental Research Letters*, *15*(7), 074019.
Knutti, R., Rugenstein, M. A., & Hegerl, G. C. (2017). Beyond equilibrium climate sensitivity. *Nature Geoscience*, *10*(10), 727–736.
Kostov, Y., Armour, K. C., & Marshall, J. (2014). Impact of the atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*, *41*(6), 2108–2116.
Levitus, S. (1982). *Climatological atlas of the world ocean* (Vol. 13). US Department of Commerce, National Oceanic and Atmospheric Administration.
Li, Q., Reichl, B. G., Fox-Kemper, B., Adcroft, A. J., Belcher, S. E., Danabasoglu, G., … Zheng, Z. (2019). Comparing Ocean Surface Boundary Vertical Mixing Schemes Including Langmuir Turbulence. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3545–3592. Retrieved 2020-05-14, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001810` doi: 10.1029/2019MS001810
Liu, W., Fedorov, A. V., Xie, S.-P., & Hu, S. (2020, June). Climate impacts of a weakened Atlantic Meridional Overturning Circulation in a warming climate. *Science Advances*, *6*(26), eaaz4876. Retrieved 2020-11-04, from `https://advances.sciencemag.org/content/6/26/eaaz4876` (Publisher: American Association for the Advancement of Science Section: Research Arti-

cle) doi: 10.1126/sciadv.aaz4876

Liu, W., & Xie, S.-P. (2018). An ocean view of the global surface warming hiatus. *Oceanography*, *31*(2), 72–79.

Marshall, D. P., & Zanna, L. (2014, November). A Conceptual Model of Ocean Heat Uptake under Climate Change. *Journal of Climate*, *27*(22), 8444–8465. Retrieved 2020-05-23, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-13-00344.1 doi: 10.1175/JCLI-D-13-00344.1

Marzocchi, A., Nurser, A. J. G., Clément, L., & McDonagh, E. L. (2021, July). Surface atmospheric forcing as the driver of long-term pathways and timescales of ocean ventilation. *Ocean Science*, *17*(4), 935–952. Retrieved 2021-08-31, from https://os.copernicus.org/articles/17/935/2021/ (Publisher: Copernicus GmbH) doi: 10.5194/os-17-935-2021

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., ... Schlund, M. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the cmip6 earth system models. *Science Advances*, *6*(26), eaba1981.

Morrison, A. K., Griffies, S. M., Winton, M., Anderson, W. G., & Sarmiento, J. L. (2016). Mechanisms of southern ocean heat uptake and transport in a global eddying climate model. *Journal of Climate*, *29*(6), 2059–2075.

Nijsse, F. J. M. M., & Dijkstra, H. A. (2018, August). A mathematical approach to understanding emergent constraints. *Earth System Dynamics*, *9*(3), 999–1012. Retrieved 2020-07-30, from https://esd.copernicus.org/articles/9/999/2018/ doi: 10.5194/esd-9-999-2018

Notz, D., & Community, S. (2020). Arctic sea ice in cmip6. *Geophysical Research Letters*, *47*(10), e2019GL086749.

Palmer, M. D., Harris, G. R., & Gregory, J. M. (2018). Extending cmip5 projections of global mean temperature change and sea level rise due to thermal expansion using a physically-based emulator. *Environmental Research Letters*, *13*(8), 084003.

Petit, T., Lozier, M. S., Josey, S. A., & Cunningham, S. A. (2020). Atlantic deep water formation occurs primarily in the iceland basin and irminger sea by local buoyancy forcing. *Geophysical Research Letters*, *47*(22), e2020GL091028.

Pickart, R. S., & Spall, M. A. (2007). Impact of labrador sea convection on the north atlantic meridional overturning circulation. *Journal of Physical Oceanography*, *37*(9), 2207–2227.

Proistosescu, C., & Huybers, P. J. (2017). Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Science advances*, *3*(7), e1602821.

Raper, S. C., Gregory, J. M., & Stouffer, R. J. (2002). The role of climate sensitivity and ocean heat uptake on aogcm transient temperature response. *Journal of Climate*, *15*(1), 124–130.

Roberts, M. J., Jackson, L. C., Roberts, C. D., Meccia, V., Docquier, D., Koenigk, T., ... others (2020). Sensitivity of the atlantic meridional overturning circulation to model resolution in cmip6 highresmip simulations and implications for future changes. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS002014.

Rose, B. E. J., Armour, K. C., Battisti, D. S., Feldl, N., & Koll, D. D. B. (2014). The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophysical Research Letters*, *41*(3), 1071–1078. Retrieved 2020-11-19, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013GL058955 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013GL058955) doi: https://doi.org/10.1002/2013GL058955

Rosenthal, Y., Kalansky, J., Morley, A., & Linsley, B. (2017). A paleo-perspective on ocean heat content: Lessons from the holocene and common era. *Quaternary*

Science Reviews, *155*, 1–12.

Sallee, J.-B., Pellichero, V., Akhoudas, C., Pauthenet, E., Vignes, L., Schmidtko, S., … Sutherland, P. (2020). Fifty-year changes of the world ocean's surface layer in response to climate change. *Nature*. (Submitted)

Sallée, J.-B., Pellichero, V., Akhoudas, C., Pauthenet, E., Vignes, L., Schmidtko, S., … Kuusela, M. (2021). Summertime increases in upper-ocean stratification and mixed-layer depth. *Nature*, *591*(7851), 592–598.

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., & Wang, Z. (2013). Assessment of Southern Ocean mixed-layer depths in CMIP5 models: Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, *118*(4), 1845–1862. Retrieved 2020-05-28, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrc.20157 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrc.20157) doi: 10.1002/jgrc.20157

Semmler, T., Jungclaus, J. H., Danek, C., Goessling, H., Koldunov, N. V., Rackow, T., & Sidorenko, D. (2021, June). *Ocean model formulation influences transient climate response* (preprint). Climatology (Global Change). Retrieved 2021-08-31, from http://www.essoar.org/doi/10.1002/essoar.10507255.1 doi: 10.1002/essoar.10507255.1

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., … Zelinka, M. D. (2020, July). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *n/a*(n/a), e2019RG000678. Retrieved 2020-08-05, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000678 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000678) doi: 10.1029/2019RG000678

Small, R. J., DuVivier, A. K., Whitt, D. B., Long, M. C., Grooms, I., & Large, W. G. (2020). On the control of subantarctic stratification by the ocean circulation. *Climate Dynamics*, 1–29.

Soldatenko, S., & Colman, R. (2019). Climate variability from annual to multidecadal timescales in a two-layer stochastic energy balance model: analytic solutions and implications for general circulation models. *Tellus A: Dynamic Meteorology and Oceanography*, *71*(1), 1554421.

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., & Plattner, G.-K. (2010). Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science*, *327*(5970), 1219–1223.

Stommel, H. (1979). Determination of water mass properties of water pumped down from the ekman layer to the geostrophic flow below. *Proceedings of the National Academy of Sciences*, *76*(7), 3051–3055.

Sun, S., Eisenman, I., Zanna, L., & Stewart, A. L. (2020). Surface constraints on the depth of the atlantic meridional overturning circulation: Southern ocean versus north atlantic. *Journal of Climate*, *33*(8), 3125–3149.

Trenberth, K. E., & Fasullo, J. T. (2012). Tracking earth's energy: From el niño to global warming. *Surveys in geophysics*, *33*(3-4), 413–426.

Trenberth, K. E., Fasullo, J. T., Von Schuckmann, K., & Cheng, L. (2016). Insights into earth's energy imbalance from multiple sources. *Journal of Climate*, *29*(20), 7495–7505.

Von Schuckmann, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., Aich, V., … others (2020). Heat stored in the earth system: where does the energy go? *Earth System Science Data*, *12*(3), 2013–2041.

Williams, R. G., Marshall, J. C., & Spall, M. A. (1995). Does Stommel's Mixed Layer "Demon" Work? *Journal of physical oceanography*, *25*(12), 3089–3102. doi: 10.1175/1520-0485(1995)025⟨3089:DSMLW⟩2.0.CO

Williamson, M. S., Cox, P. M., & Nijsse, F. J. M. M. (2018, January). Theoretical foundations of emergent constraints: relationships between climate sensitivity and global temperature variability in conceptual models. *Dynamics and Statistics of the Climate System*, *3*(1). Retrieved 2020-07-30, from https://academic.oup.com/climatesystem/article/3/1/dzy006/5382541 (Publisher: Oxford Academic) doi: 10.1093/climsys/dzy006

Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., ... Zhang, R. (2020, January). Climate Sensitivity of GFDL's CM4.0. *Journal of Advances in Modeling Earth Systems*, *12*(1). Retrieved 2020-04-16, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001838 doi: 10.1029/2019MS001838

Winton, M., Takahashi, K., & Held, I. M. (2010, May). Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, *23*(9), 2333–2344. Retrieved 2020-04-28, from http://journals.ametsoc.org/doi/10.1175/2009JCLI3139.1 doi: 10.1175/2009JCLI3139.1

Xie, S.-P. (2020). Ocean warming pattern effect on global and regional climate change. *AGU advances*, *1*(1), e2019AV000130.

Yoshimori, M., Watanabe, M., Shiogama, H., Oka, A., Abe-Ouchi, A., Ohgaito, R., & Kamae, Y. (2016, December). A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. *Progress in Earth and Planetary Science*, *3*(1), 21. Retrieved 2020-04-26, from http://progearthplanetsci.springeropen.com/articles/10.1186/s40645-016-0096-3 doi: 10.1186/s40645-016-0096-3

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., ... Taylor, K. E. (2020). Causes of higher climate sensitivity in cmip6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782.

Zhu, J., Poulsen, C. J., & Otto-Bliesner, B. L. (2020). High climate sensitivity in cmip6 model not supported by paleoclimate. *Nature Climate Change*, *10*(5), 378–379.

Acknowledgments

BFK and GH designed the project goals and diagnostics for evaluation. GH wrote the code to evaluate the metrics and create figures. Both authors wrote and edited the text. BFK was supported by NOAA NA19OAR4310366, NSF 1655221, and ONR N00014-17-1-2393. GH recieved support from Brown University through a SPRINT award. Conversations with Piers Forster, Jochem Marotzke, and Anne-Marie Treguier and two thoughtful anonymous reviews helped improve this work.

Supplemental Information for "Regional mixed layer depth as a climate diagnostic and emergent constraint"

G. Hall[1], B. Fox-Kemper[1,2]

[1]Institute at Brown for Environment and Society, Brown University, Providence, Rhode Island 02912, USA

[2]Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, Rhode Island 02912, USA

Contents of this file

Text S1. Methods

1. Normal uncertainty integration

Cox, Huntingford, and Williamson (2018) demonstrate a method for constraining S based on a linear regression of S against modelled temperature variability. Taking instead (for example) the North Ocean MLD h_{no}, once the regression parameters are estimated, the conditional probability distribution of S on h_n is given by

$$P(S|h_{no}) = \mathcal{N}(f(h_{no}), \sigma_f(h_{no})), \tag{1}$$

where σ_f is the prediction error of the regression (see Cox et al. (2018) for a full description). We estimate conditional distributions for each parameter on each regional MLD with a significant ($p < 0.05$) explanatory relationship, and take the integrated product of the conditional probabilities as the constrained distribution. For example, S depends on h_{no} and h_{so}, giving:

$$P(S) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(S|h_{no}) P(S|h_{so}) dh_{no} dh_{so}. \tag{2}$$

2. Linear regression with fixed errors

We first estimate a multiple linear regression between $\mathbf{h}$ and the parameters of the two-layer model $\mathbf{P} = (C_S, C_D, \lambda, S, \gamma, \varepsilon)$ within the GCM ensemble, linearizing about the ensemble mean:

$$(\mathbf{P}^{gcm} - \overline{\mathbf{P}}^{gcm}) = \mathbf{M} \cdot (\mathbf{h}^{gcm} - \overline{\mathbf{h}}^{gcm}) + \epsilon^{gcm} \tag{3}$$

where the GCM subscripts indicate that the values are calculated from the ensemble, and ϵ^{gcm} is the residual term. To keep only meaningful relationships we set $M_{ij} = 0$ for all parameters with $p > 0.05$ and recalculate the coefficients. We then enter in the observed values $\mathbf{h}^{obs}$ to estimate the parameters of the "real" two-layer model:

$$\mathbf{P}^{pred} = \mathbf{M} \cdot (\mathbf{h}^{obs} - \overline{\mathbf{h}}^{gcm}) + \overline{\mathbf{P}}^{gcm} + \epsilon^{gcm}. \tag{4}$$

Solving for the residual term in equation 3 we have

$$\epsilon^{gcm} = (\mathbf{P}^{gcm} - \overline{\mathbf{P}}^{gcm}) + \mathbf{M} \cdot (\overline{\mathbf{h}}^{gcm} - \mathbf{h}^{gcm}).$$

Assuming that the residual remains the same for the constrained model parameters, this leaves

$$\mathbf{P}^{pred} = \mathbf{M} \cdot (\mathbf{h}^{obs} - \mathbf{h}^{gcm}) + \mathbf{P}^{gcm}, \tag{5}$$

which gives a *discrete* 24-member dataset of adjusted the EBM-ε parameters using the observed mixed layer depths ("constant residuals" in figure S1).

3. Bootstrapped linear regression

We additionally bootstrap uncertainty intervals using 10,000 random draws with replacement of the 24 models (excluding AWI-CM-1-1-MR for the Southern Ocean constraints, as described in the text). For each draw we conduct the same emergent constraint calculation described above using random noise, and take the distribution of predicted values as our uncertainty range.

4. Monte Carlo constraint

Ordinary least squares linear regressions risk misrepresenting the constrained parameter ranges by ignoring cross-correlations between the dependent values (for instance between γ and C_D). It may be more accurate to treat MLD as constraining the joint probability distribution of all the parameters rather than the independent probabilities of each parameter; instead of finding

$$P(S|\mathbf{h}), P(\lambda|\mathbf{h}), P(\varepsilon|\mathbf{h})...$$

the joint probability is preferred,

$$P(S, \lambda, C_S, C_D, \varepsilon, \gamma|\mathbf{h}).$$

We estimate a continuous joint distribution of $(S, \lambda, C_S, C_D, \varepsilon, \gamma, h_n, h_s, h_t)$ matching the cross-correlations in our ensemble, and take random samples from it weighted by the probabilities of observing each sampled h. We generate the distribution using independent Gaussian kernel density estimations (KDEs) of each parameter and the three MLD regions from the GCMs, then take a large number of random draws from each and correlate the random draws using the Cholesky matrix of the GCM distribution. This transformation gives each continuous random variable the variance and cross-correlations found in the original dataset, including all of the correlations to h. We then weight each draw by the combined probability of observing each of the regional MLD values in the draw: $P(h_{n,t,s}) = P(h_n)P(h_t)P(h_s)$. (The pdfs of each regional MLD's observed values are assumed to be normal and independent.) Then we take a much smaller probability-weighted sample, so that we are approximately again taking a weighted sample from a continous distribution. We repeat this process in batches of 50,000 draws, taking a weighted sub-sample of 500 each time, until we have constructed a dataset of 50,000 subsamples which represents the MLD-constrained joint parameter distribution.

This method relies on the fewest assumptions about the shape of the constrained distribution and consequently reports larger spreads in some constrained ranges and the largest changes under constraint in the mean value (Figure S1), although all methods show smaller uncertainty ranges than the unconstrained GCM ensemble. We therefore use this method to report headline numbers.

5. Constrained time-series predictions

The h regression gives us a constraint on S directly. Further, the GSAT change under `abrupt-4xCO2` and `1pctCO2` forcings can be determined analytically in the two-layer model as a function of S, λ and the other parameters without ever dividing S by λ to obtain F:

$$\Delta T_{4\times CO2}(t) = S \times \left\{1 - a_s \exp(\frac{-t}{\tau_f}) - a_d \exp(\frac{-t}{\tau_s})\right\} \tag{6}$$

$$\begin{aligned}\Delta T_{linear}(t) = S \times \Big\{ & t_0(t) \\ & - \tau_f a_s [1 - \exp(\frac{-t}{\tau_f})] \exp(\frac{-t + t_0(t)}{\tau_f}) \\ & - \tau_s a_d [1 - \exp(\frac{-t}{\tau_s})] \exp(\frac{-t + t_0(t)}{\tau_s}) \Big\}\end{aligned} \tag{7}$$

where the τ and a variables are calculated from $(C_s, C_d, \lambda, \gamma)$ using the formulas in Geoffroy et al. (2013a). In the linear forcing case, where forcing stops at year 150,

$$t_0(t) = \begin{cases} t, & t < 150\text{yr} \\ 150\text{yr}, & t > 150\text{yr} \end{cases}$$

.

References

Cox, P. M., Huntingford, C., & Williamson, M. S. (2018, January). Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature*, *553*(7688), 319–322. Retrieved 2020-08-05, from https://www.nature.com/

articles/nature25450 (Number: 7688 Publisher: Nature Publishing Group) doi: 10.1038/nature25450

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., & Tytéca, S. (2013, March). Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, *26*(6), 1859–1876. Retrieved 2020-04-30, from http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00196.1 doi: 10.1175/JCLI-D-12-00196.1

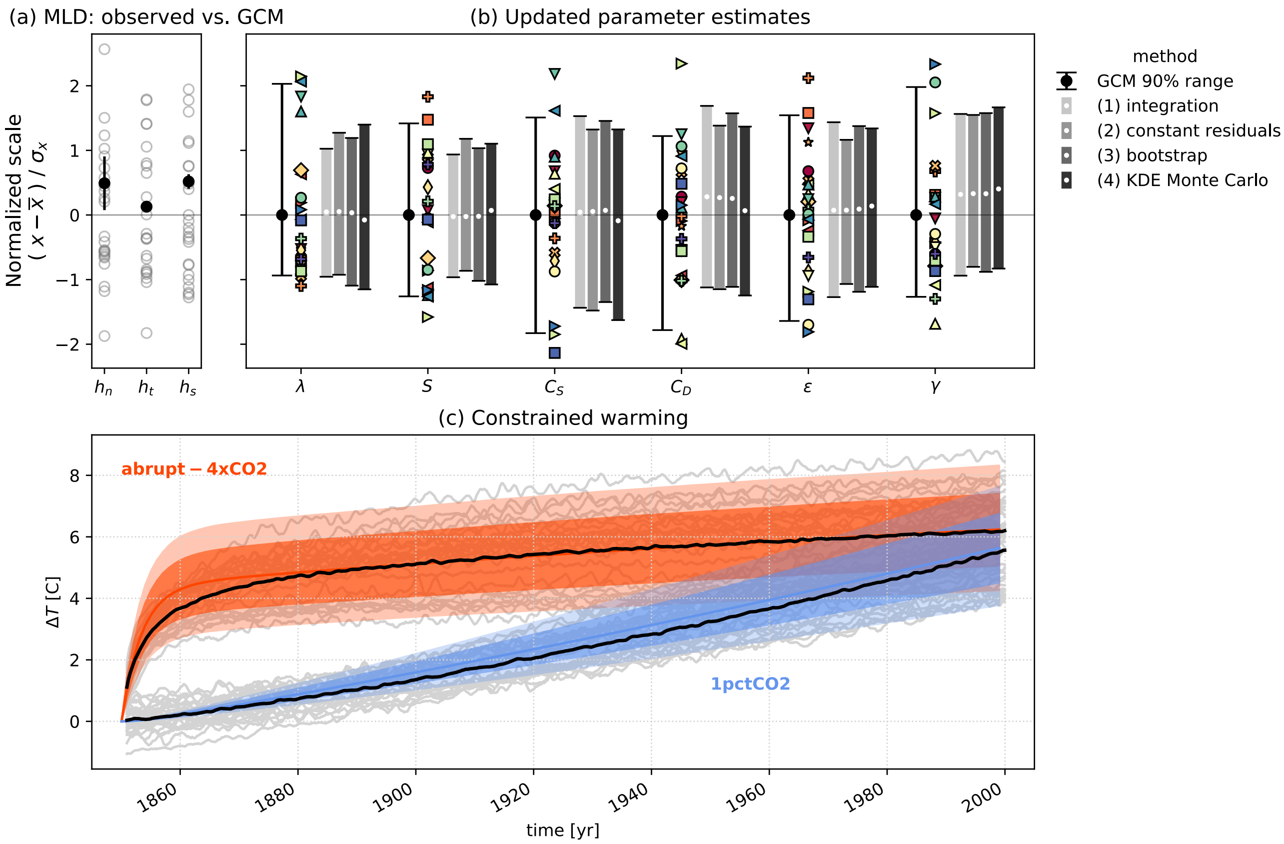

Figure S1. (a) the simulated h ranges from 25 GCMs (grey circles) and the observed 5-95% confidence ranges (black lines). (b) The ensemble values of the 1-layer EBM λ and S and each of the EBM-ε parameters (grey circles), and the constrained values using observed h values (black and orange circles and lines). Values in (a) and (b) are normalized about the ensemble means for easier comparison. (c) Predicted $\Delta T_S(t)$ for 1%/year (orange) and $4\times CO_2$ (blue) experiments, estimated using the constrained EBM-ε parameters. The lighter and darker shaded regions show 90% and 66% confidence intervals respectively. Grey lines show individual GCM temperatures and black lines the ensemble means.

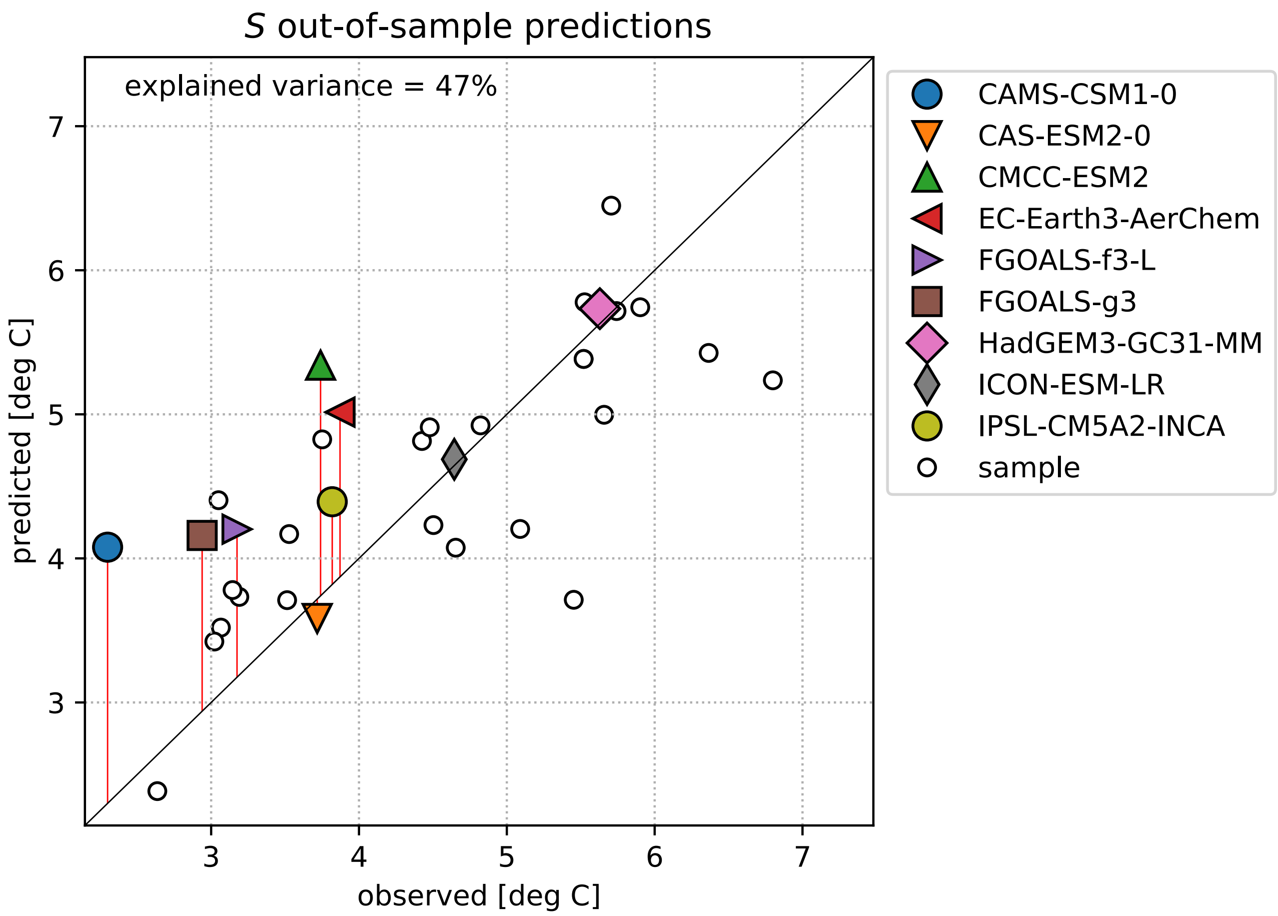

Figure S2. Comparison between the 2-layber EBM estimates of S for our out-of-sample ensemble, and the S values predicted for those models using only their pre-forcing mixed layer depths and the emergent constraints.

Table S1. EBM-ε parameters calculated from the 25 member CMIP6 ensemble. The rightmost two columns mean and standard deviation are taken from Geoffroy et al. (2013), which uses an ensemble of 16 CMIP5 models. Our results give lower λ estimates, in line with the higher reported values of S in the CMIP6 ensemble. The set of GCMs in each ensemble are not the same. The institutions providing CMIP6 data differ from those providing CMIP5 data, so these columns are not just an update of the same model ensemble for direct comparison.

Variable	CMIP6 Mean	CMIP6 Std. dev.	CMIP5 Mean	CMIP5 Std. dev.
D_S (m)	63	7	86	9
D_D (m)	976	269	1141	544
ε	1.3	0.23	1.3	0.25
λ (W m^{-2} K^{-1})	0.89	0.30	1.18	0.37
γ (W m^{-2} K^{-1})	0.64	0.14	0.67	0.15
$F_{4\times CO_2}$ (W m^{-2})	7.5	0.82	7.6	1.0

Table S2. All CMIP6 models used in the study. Ensemble 1 members were used to calculate the emergent constraint, and ensemble 2 members were downloaded later as an out-of-sample test.

Model	Ensemble
ACCESS-CM2	1
ACCESS-ESM1-5	1
AWI-CM-1-1-MR	1
BCC-CSM2-MR	1
BCC-ESM1	1
CESM2	1
CESM2-FV2	1
CESM2-WACCM	1
CESM2-WACCM-FV2	1
CMCC-CM2-SR5	1
CNRM-CM6-1	1
CNRM-ESM2-1	1
CanESM5	1
E3SM-1-0	1
EC-Earth3-Veg	1
GISS-E2-1-G	1
HadGEM3-GC31-LL	1
IPSL-CM6A-LR	1
KIOST-ESM	1
MPI-ESM-1-2-HAM	1
MPI-ESM1-2-HR	1
MPI-ESM1-2-LR	1
MRI-ESM2-0	1
NESM3	1
UKESM1-0-LL	1
CAMS-CSM1-0	2
CAS-ESM2-0	2
CMCC-ESM2	2
EC-Earth3-AerChem	2
FGOALS-f3-L	2
FGOALS-g3	2
HadGEM3-GC31-MM	2
ICON-ESM-LR	2
IPSL-CM5A2-INCA	2

Table S3. Linear correlations between all variables in the extended 34-model ensemble.

	C_D	ε	λ	γ	F	S	h_g	h_n	h_t	h_s
C_S	-0.19	**0.46**	**0.43**	-0.31	**0.53**	-0.22	0.27	-0.24	0.44*	-0.24
C_D		-0.04	0.13	**0.51**	0.13	-0.13	-0.06	0.23	-0.07	-0.26
ε			-0.05	-0.04	**0.49**	0.31	0.20	-0.17	0.43*	0.05
λ				0.20	**0.54**	**-0.87**	0.15	0.26	0.25	**-0.51**
γ					-0.10	-0.24	**0.45**	**0.63**	0.29	0.14
F						-0.14	-0.07	-0.13	0.11	-0.08
S							-0.14	-0.36*	-0.17	**0.59**
h_g								**0.56**	**0.76**	0.23
h_n									0.15	0.13
h_t										-0.14