

The Climate Response to Multiple Volcanic Eruptions Mediated by Ocean Heat Uptake: Damping Processes and Accumulation Potential

MUKUND GUPTA AND JOHN MARSHALL

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 17 October 2017, in final form 31 July 2018)

ABSTRACT

A hierarchy of models is used to explore the role of the ocean in mediating the response of the climate to a single volcanic eruption and to a series of eruptions by drawing cold temperature anomalies into its interior, as measured by the ocean heat exchange parameter q ($\text{W m}^{-2} \text{K}^{-1}$). The response to a single (Pinatubo-like) eruption comprises two primary time scales: one fast (year) and one slow (decadal). Over the fast time scale, the ocean sequesters cooling anomalies induced by the eruption into its depth, enhancing the damping rate of sea surface temperature (SST) relative to that which would be expected if the atmosphere acted alone. This compromises the ability to constrain atmospheric feedback rates measured by λ ($\sim 1 \text{ W m}^{-2} \text{K}^{-1}$) from study of the relaxation of SST back toward equilibrium, but yields information about the transient climate sensitivity proportional to $\lambda + q$. Our study suggests that q can significantly exceed λ in the immediate aftermath of an eruption. Shielded from damping to the atmosphere, the effect of the volcanic eruption persists on longer decadal time scales. We contrast the response to an “impulse” from that of a “step” in which the forcing is kept constant in time. Finally, we assess the “accumulation potential” of a succession of volcanic eruptions over time, a process that may in part explain the prolongation of cold surface temperatures experienced during, for example, the Little Ice Age.

1. Introduction

Large volcanic eruptions are a natural, impulse-like perturbation to the climate system. The sulfur particles ejected into the stratosphere during these eruptions are rapidly converted to sulfate aerosols that diminish the net incoming solar flux at the top of the atmosphere, resulting in a cooling of the surface climate. These sulfate aerosols have a residence time of about 1–2 years in the stratosphere (Robock 2000) but can cause surface cooling for many more years after the eruption.

The response of the climate to volcanic eruptions is of interest for at least two reasons. First, it can teach us about how robust the climate is to a perturbation and the rate at which it relaxes back to equilibrium (e.g., Wigley et al. 2005; Bender et al. 2010; Merlis et al. 2014). Second, because of its large effective heat capacity, the

ocean can perhaps remember the effect of successive eruptions, enabling an accumulation larger than any single event (e.g., Free and Robock 1999; Crowley et al. 2008; Stenchikov et al. 2009). Some of the issues are illustrated in Fig. 1, which shows the hypothetical response of the climate to a volcanic eruption in two limit cases. In the first, the atmosphere is imagined to be coupled to a slab ocean. The relaxation of the system here depends simply on the climatic feedback parameter λ ($\text{W m}^{-2} \text{K}^{-1}$) and the slab’s heat capacity. The larger the value of λ , the smaller the equilibrium climate sensitivity and the faster the system relaxes back to equilibrium. In the second, the slab lies atop an interior ocean that can sequester heat away from the surface at a rate proportional to the ocean heat exchange parameter q ($\text{W m}^{-2} \text{K}^{-1}$), which reduces the peak cooling and increases the rate of SST recovery in the initial stages. However, on longer time scales, the sequestered heat anomaly is shielded from damping to space, leading to a prolongation of the signal. Thus, interaction with the interior ocean changes the response from that of a simple exponential decay on one time scale to a two-time-scale process, as evidenced by the kinked profiles

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-17-0703.s1>.

Corresponding author: Mukund Gupta, guptam@mit.edu

DOI: 10.1175/JCLI-D-17-0703.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

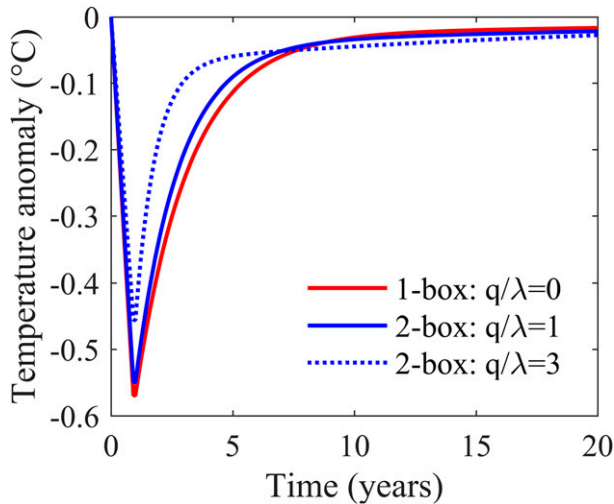


FIG. 1. Surface temperature response of the box model to an idealized Pinatubo eruption (-4 W m^{-2} for a year) in the 1- (red) and 2-box cases (blue) in terms of the ratio of ocean mixing strength to the climatic feedback parameter $\mu = q/\lambda$ with $\lambda = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$. All other parameters are as in Table 1. The “area under the curve” is the same in all cases when integrating to infinity, but with a smaller peak and a longer “tail” as q (or μ) increases.

in Fig. 1, which become more prominent as the ratio $\mu = q/\lambda$ increases.

Many studies have explored the role of the subsurface ocean in the climatic response to external forcings (e.g., Hansen et al. 1985; Gregory 2000; Stouffer 2004; Winton et al. 2010; Held et al. 2010; Geoffroy et al. 2013). Volcanic responses have been explored in simple box models (e.g., Lindzen and Giannitsis 1998) as well as in state-of-the-art global climate models and observations (e.g., Church et al. 2005; Glecker et al. 2006; Stenchikov et al. 2009; Merlis et al. 2014; Schurer et al. 2015). Here, we investigate the role of the ocean in sequestering thermal anomalies to depth and enhancing initial surface temperature relaxation rates, while temporarily shielding those anomalies from damping processes and thereby extending the response time scale. As we will see, this mechanism can promote accumulation of the cooling signal from successive eruptions and cause the response to span multidecadal time scales. While previous studies (e.g., Geoffroy et al. 2013; Kostov et al. 2014) have reported $\mu \sim 1$, we argue that over relatively short time scales (years to a decade or so), μ can be considerably larger than 1. We explore the consequences for estimating λ in the immediate aftermath of a volcanic eruption from the relaxation time scale of SST. We also quantify the role of the interior ocean in prolonging the response from volcanic eruptions and contrast the response of the system to an impulse from that of a step.

Our study employs a hierarchy of idealized models, ranging from box models to a coupled general circulation

model (GCM). Section 2 explores results from idealized volcanic eruptions in a GCM. In section 3, we interpret those results using 1.5-, 2-, and 3-box models of the ocean and investigate the role of μ . In section 4, we apply the resulting insights to study the climate response to a series of volcanic eruptions during the last millennium. In section 5, we conclude.

2. Experiments with an idealized coupled aquaplanet model

a. Experiment description

This study uses the Massachusetts Institute of Technology GCM (MITgcm), which simulates the physics of an ocean-covered planet coupled to an atmosphere, with no land, sea ice, or clouds (see appendix). Geometrical constraints are imposed on the ocean circulation through the effect of two narrow barriers extending from the North Pole to 35°S and set 90° apart. These barriers extend from the seafloor (assumed flat) to the surface and separate the ocean into a large and a small basin that are connected in a circumpolar region to the south. Despite the simplicity of the geometry, this “double drake” configuration captures aspects of the present climate, including plausible energy transport by the oceans and atmosphere and a deep meridional overturning circulation that is dominated by the small basin (Ferreira et al. 2010).

The atmospheric component of the model employs a simplified radiation scheme where the shortwave flux does not interact with the atmosphere, and hence, the planetary albedo is equivalent to the surface albedo, as described in Frierson et al. (2006). Idealized volcanic eruptions are simulated by reducing the net incoming shortwave radiative flux by a uniform amount over the globe, while ensuring that it does not become negative anywhere. The forcing is applied as a 1-yr square pulse in time starting 1 January. Both single and multiple pulses (separated by a specified time interval) are considered. To isolate the role of the interior ocean, numerical experiments are run using the “full ocean” (dynamic and 3D) configuration of the MITgcm, as well as a “slab ocean” configuration that has a single spatially varying vertical layer representing the annual-mean mixed layer depth of the model (with a globally averaged depth of 43 m). In the “slab ocean,” a prescribed lateral flux of heat in the mixed layer helps to maintain a climatological SST close to that of the coupled system.

b. Idealized volcanic responses

Figure 2 shows the globally averaged SST response of the MITgcm to a forcing of -4 W m^{-2} for 1 year, which

MITgcm response to idealized single volcanic eruptions

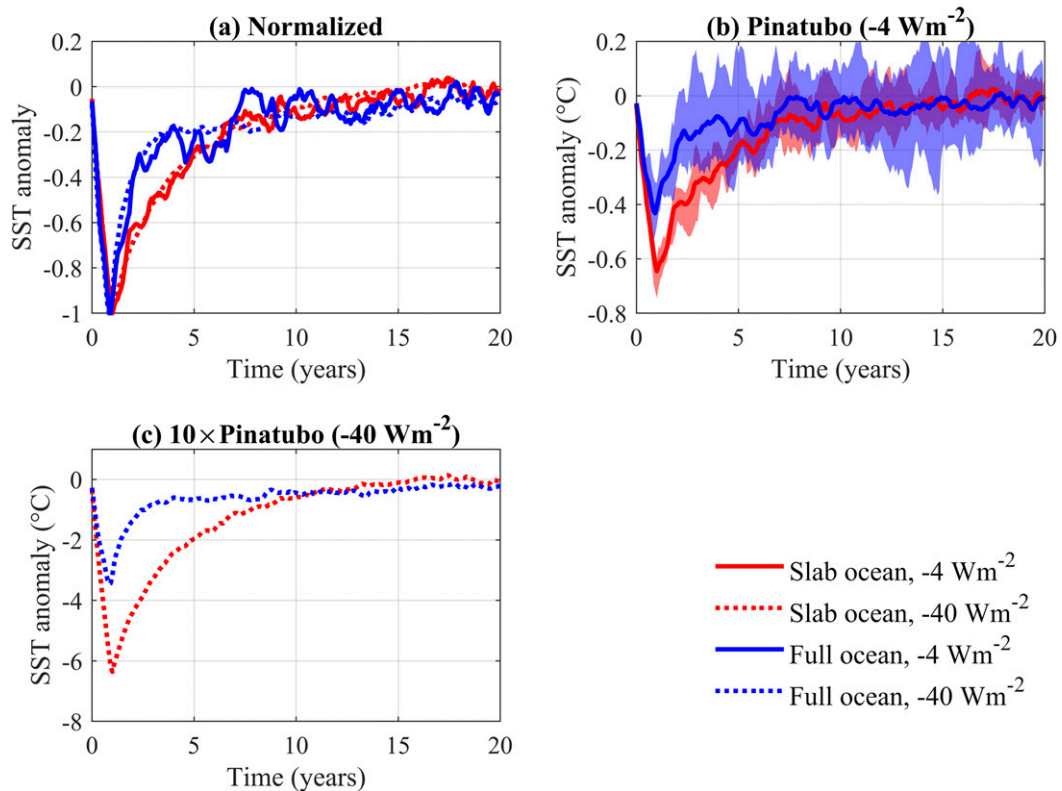


FIG. 2. MITgcm responses (with monthly mean data) to a Pinatubo-like forcing (-4 W m^{-2} for a year) and a $10 \times$ Pinatubo forcing (-40 W m^{-2} for a year) for the slab ocean (red) and the full ocean configuration (blue). (a) Ensemble-mean responses normalized with respect to their peak cooling temperature. (b) Non-normalized responses for the Pinatubo forcing with the shaded envelopes of five ensemble members for the slab ocean (red) and 10 ensemble members for the full ocean (blue). The solid lines are the corresponding ensemble mean. (c) Non-normalized responses for the $10 \times$ Pinatubo forcing with one ensemble member for the slab and full ocean, respectively.

crudely emulates the radiative effect of the 1991 Mount Pinatubo eruption. A theoretical $10 \times$ Pinatubo eruption was also simulated using a forcing of -40 W m^{-2} for a year. Ensemble members (five for the Pinatubo forcing and one for the $10 \times$ Pinatubo forcing) were initialized from a long control integration of the model separated by 10-yr intervals. Anomalies were calculated by subtracting the response of the forced run from the control run. Figure 2a shows all model responses normalized with respect to their peak cooling value. The slab ocean curves decay over a single e -folding time scale of about 4 years, whereas the full ocean curve displays an initial fast relaxation rate and a long-lasting tail (5%–10% of the signal present after 20 years). The shape of these response functions is interpreted using box models of the climate in section 3.

Figure 2b shows that in the Pinatubo-like simulations, the SST anomaly reaches a minimum value of -0.62°C

for the slab and -0.41°C for the full ocean. This difference is the result of some of the cooling being sequestered into the subsurface ocean in the case of a dynamic ocean, as argued by Held et al. (2010). Soden et al. (2002) report an observed globally averaged tropospheric temperature anomaly of -0.5°C the year after the Pinatubo eruption, broadly in accord with our calculations. The shading in Fig. 2b is the envelope corresponding to the response of the various ensemble members, whereas the solid lines are the ensemble means. The standard deviation in the SST anomaly is constrained to be zero at $t = 0$, but eventually settles to 0.11°C for the full ocean and 0.06°C for the slab, characteristic of the noise levels in these respective configurations. Figure 2c shows that for a $10 \times$ Pinatubo forcing, the slab displays a maximum cooling of -6.1°C , compared to only -3.7°C for the full ocean. This peak cooling scales linearly with the forcing amplitude in the slab, but is 10% smaller than linear scaling

Temperature anomaly ($^{\circ}\text{C}$) with latitude and depth for idealized Pinatubo (top) and $10\times$ Pinatubo (bottom) forcings

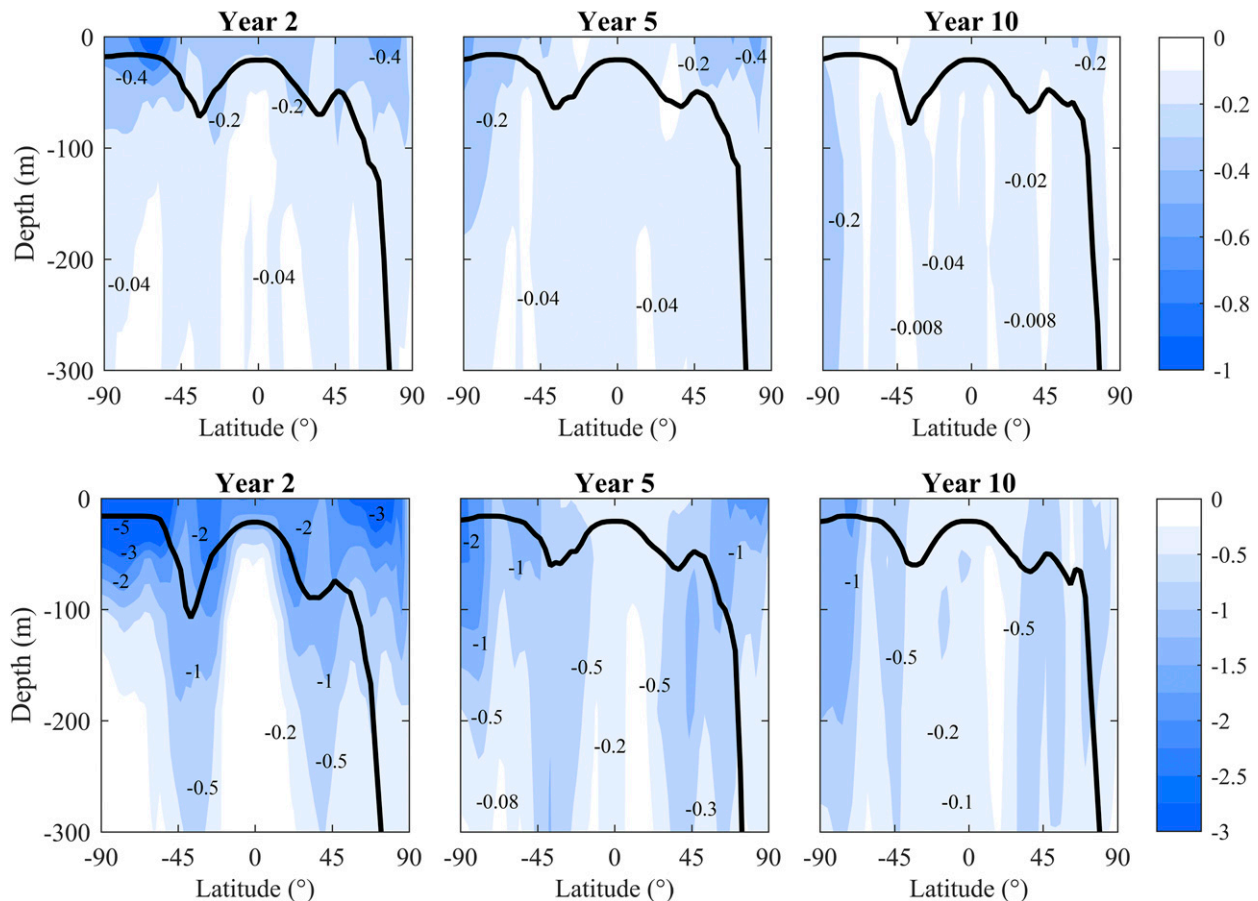


FIG. 3. MITgcm zonally averaged temperature anomaly in the ocean with depth and latitude in the full ocean configuration. The temperature evolution is shown for (left) 2, (middle) 5, and (right) 10 years after the eruption initiation. (top) Mean responses of 10 ensemble members for the Pinatubo-like forcing (-4 W m^{-2} for a year) and (bottom) the responses for a single ensemble member of the $10 \times$ Pinatubo forcing (-40 W m^{-2} for a year). The thick black line represents the model-diagnosed zonally averaged mixed layer depth.

when the ocean is active. This nonlinearity can be explained by the fact that the larger forcing causes the mixed layer to deepen, which allows the cooling signal to penetrate farther down into the ocean.

As evidenced in Fig. 2b, unforced variability can readily obscure the response to volcanic eruptions. To explore this issue, we conduct a statistical analysis of the globally and annually averaged SST in long control simulations of the slab and full ocean configurations. The full ocean simulations show more variability than those corresponding to the slab, due to the many additional degrees of freedom imparted by the presence of a dynamic ocean. Based on a single-sided Student's t test, we find that the slab ocean response in the $10 \times$ Pinatubo

simulation is significant for 15 years at -0.13°C , whereas the full ocean response remains significant for 22 years at -0.18°C (both at a 95% confidence level). However, for Pinatubo-like events, we require a large number of ensembles (~ 10) to tease out a significant response for 10–20 years. While noise levels may differ in the real ocean, this analysis suggests that unforced variability poses severe limitations on the ability to detect SST signals resulting from volcanic eruptions, except, perhaps, for the most significant events, such as Santa María, Mount Agung, El Chichón, and Mount Pinatubo, during the recent historical past.

Figure 3 shows the evolution of the ocean temperature anomaly as a function of latitude and depth for the

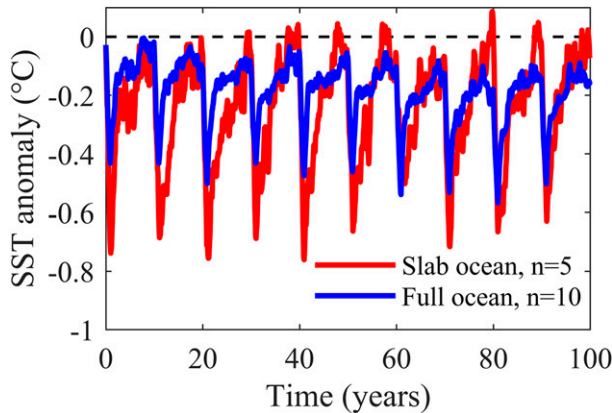


FIG. 4. MITgcm ensemble-mean response (with monthly mean data) to a Pinatubo-like eruption (-4 W m^{-2} for a year) every 10 years in the slab ocean (red) and full ocean (blue) configurations. The slab and full ocean configuration were run for 5 and 10 ensemble members, respectively.

Pinatubo and $10 \times$ Pinatubo forcings (full ocean configuration). Within 2 years of the eruption, a significant amount of cooling is transported below the mixed layer. Temperature anomalies on the order of 5%–10% of the peak surface cooling exist at 200-m depth and persist for more than 10 years after the cooling pulse. A combination of processes may be acting to spread the anomaly vertically, such as turbulent diffusion, Ekman pumping, seasonal convection, mixing in the wind-driven gyres, and large-scale overturning circulation (e.g., Gregory 2000; Stouffer 2004; Stenchikov et al. 2009). Figure 3 reveals signatures of Ekman pumping within the subtropical gyres, particularly visible for the $10 \times$ Pinatubo forcing. At the poles, the penetration of the anomaly to depth happens over a longer time scale than in midlatitudes. Several studies (e.g., Stenchikov et al. 2009; Otterå et al. 2010; Mignot et al. 2011) discuss a strengthening of the meridional overturning circulation in response to volcanic eruptions, which can also contribute to the vertical exchange. For the $10 \times$ Pinatubo eruption, the globally averaged mixed layer depth increases by 37% (63 m) in the year of the eruption and relaxes back to its base value (43 m) within 3 years. This increase occurs principally in the midlatitudes, where most of the anomalous subduction of cooling occurs in the first few years after the eruption. This might explain the slight nonlinearity in the $10 \times$ Pinatubo response visible in Fig. 2 and mentioned above.

Figure 4 shows simulations of a series of Pinatubo-like eruptions occurring every 10 years in the slab and full ocean configurations. In the slab, the response falls back to zero after each eruption. On the other hand, the

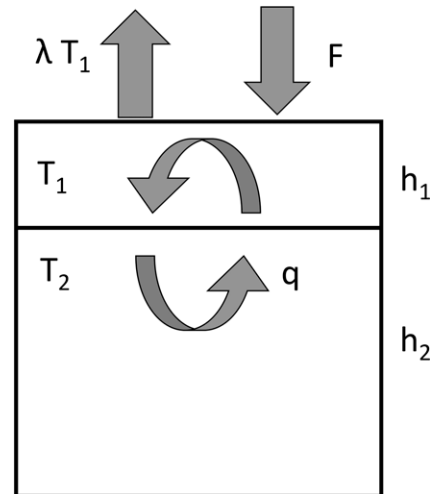


FIG. 5. A 2-box model comprising a mixed layer of depth h_1 and a deeper ocean of depth h_2 with temperature anomalies T_1 and T_2 , respectively. The model is driven from the top by an external forcing F and damped by the climate feedback λT_1 . The two boxes exchange heat through the parameter q .

full ocean response slowly builds over time, as evidenced by the 20% increase in peak cooling achieved approximately 60 years after the first eruption. This suggests that the presence of a deeper ocean can facilitate the buildup of a cooling signal from successive eruptions. In section 3, we discuss the conditions that can lead to signal accumulation using box models as a guide.

3. Interpretation using box models

a. Response to an impulse forcing

The globally averaged SST responses of the MITgcm aquaplanet to an idealized volcanic forcing can be interpreted using simple analytical models. We find that the shapes of the temperature response functions are most readily recovered and interpreted through the use of a 2-box model (shown in Fig. 5). The model was introduced by Gregory (2000) and was subsequently employed by Held et al. (2010), Geoffroy et al. (2013), Kostov et al. (2014), and others. It consists of a mixed layer and a deeper ocean box of temperatures T_1 and T_2 , respectively, driven from the top by an external forcing F and damped by the climate feedback λT_1 . The governing equations can be written as follows:

$$\rho c_w h_1 \frac{dT_1}{dt} = -\lambda T_1 + q(T_2 - T_1) + F(t), \quad (1)$$

and

$$\rho c_w h_2 \frac{dT_2}{dt} = q(T_1 - T_2), \quad (2) \quad \mu = \frac{q}{\lambda}, \quad (6a)$$

where h_1 and h_2 are the thicknesses of the mixed layer and deeper ocean boxes, respectively. The density and heat capacity of water are ρ and c_w , respectively. The parameter q controls vertical ocean heat exchange; it is positive for an active deeper ocean and zero for a slab ocean. We represent an idealized volcanic eruption by imposing a delta function forcing $F(t) = V\delta(t)$ in Eq. (1), where V is the integrated amount of energy instantaneously extracted from the system. The impulse (or Green's function) response provides information on the first-order climate response to a volcanic eruption and lends itself to convolution with a more realistic time series of forcing (see section 4). The analytical solution to Eqs. (1) and (2) is presented in the online supplemental information (SI) and is consistent with the work of Geoffroy et al. (2013), Kostov et al. (2014), and Tsutsui (2017), who presented the solution to a step in the forcing. The solution for T_1 is the sum of two decaying exponentials:

$$T_1(t) = T_f e^{-t/\tau_f} + T_s e^{-t/\tau_s}, \quad (3a)$$

and

$$T_f + T_s = T_c, \quad (3b)$$

where T_c , T_f , T_s , τ_f , and τ_s are written out in the SI, together with the solution for T_2 . Equation (3a) describes the relaxation of T_1 back to equilibrium after the forcing F has ceased to act. The relaxation occurs over a fast and a slow time scale with e -folding values τ_f and τ_s , respectively. In the case of a delta function forcing, the peak cooling T_c occurs instantaneously at $t = 0$ and is given by

$$T_c = \frac{V}{\rho c_w h_1}, \quad (4)$$

where V is written as follows:

$$V = \int_0^{\infty} F(t) dt. \quad (5)$$

Equation (4) suggests that the peak cooling T_c does not depend on the climatic feedback λ and oceanic damping q , but this is only valid for an idealized instantaneous forcing, as will be seen in section 3c.

The full analytical solutions are unwieldy and not very insightful, but may be simplified by introducing the parameters μ and r :

and

$$r = \frac{h_1}{h_2}. \quad (6b)$$

The quantity μ represents the ratio of the ocean damping strength versus climatic damping, and r is the ratio of heat capacities between the two boxes. We consider three limiting cases: (i) r is small, (ii) μ is small, and (iii) μ is large and r is small. A fit to the MITgcm simulations with monthly mean data gives $r \sim 1/3$ and $\mu \sim 2$, suggesting that the limit of small r and large μ is perhaps the most relevant. A detailed discussion of these fits in subsequent paragraphs shows that these parameter values may differ when considering annual mean data, but here we focus on fitting the monthly resolved signal.

In the SI, we show that in the limit of $r \ll 1$, the parameters τ_f , τ_s , T_f , and T_s , are given by

$$\tau_f \approx \frac{\rho c_w h_1}{\lambda(1 + \mu)}, \quad (7a)$$

$$\tau_s \approx \rho c_w h_1 \frac{(1 + \mu)}{qr}, \quad (7b)$$

$$T_f \approx \frac{(1 + \mu)^2}{(1 + \mu)^2 + r\mu^2}, \quad (8a)$$

and

$$T_s \approx \frac{r\mu^2}{(1 + \mu)^2 + r\mu^2} T_c. \quad (8b)$$

When $\mu \ll 1$, a 1-box model is retrieved. In this case, the transfer of heat to the deeper ocean is limited, and the atmosphere is the only significant medium responsible for damping the surface anomaly. The solution reduces to a single exponential decay controlled by damping to the atmosphere:

$$T_1(t) = T_c e^{-t/\tau_m}, \quad (9a)$$

with

$$\tau_m = \frac{\rho c_w h_1}{\lambda}. \quad (9b)$$

When $\mu \gg 1$ and $r \ll 1$ (i.e., when ocean heat transport is large relative to damping to the atmosphere), the two-time-scale solution becomes

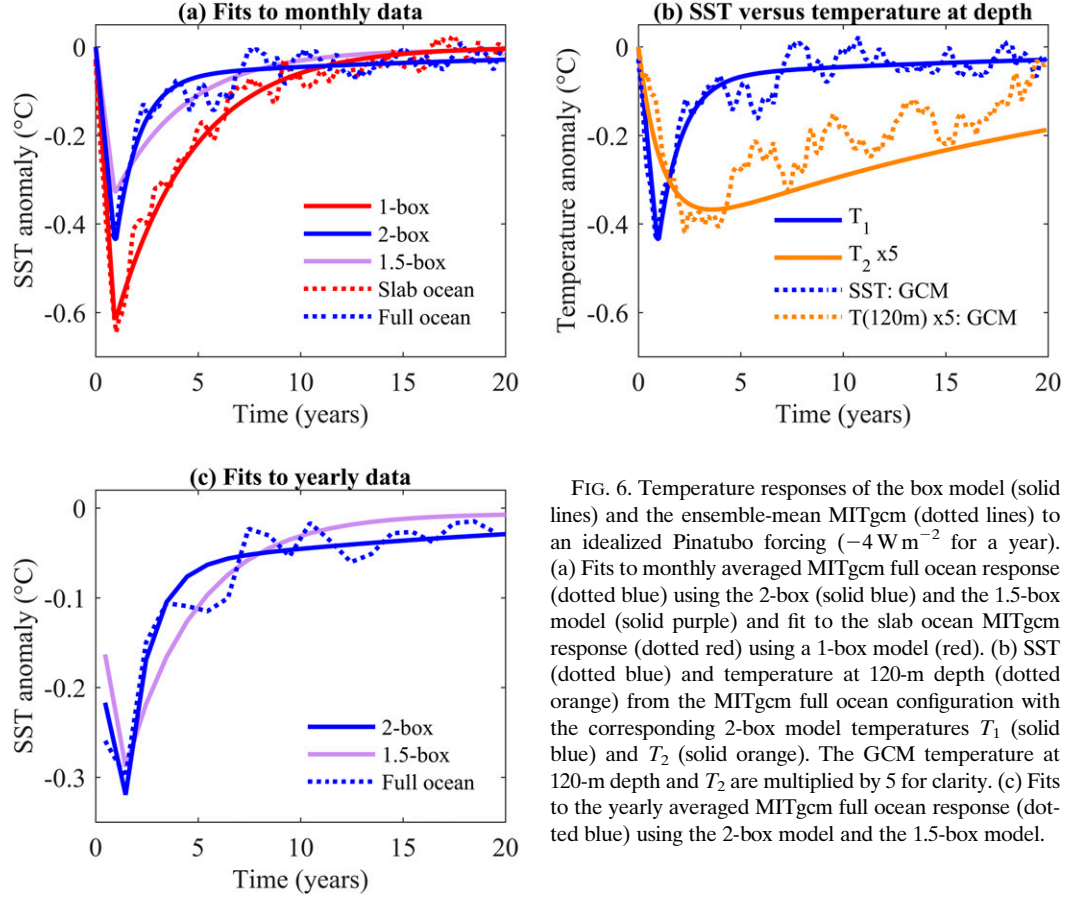


FIG. 6. Temperature responses of the box model (solid lines) and the ensemble-mean MITgcm (dotted lines) to an idealized Pinatubo forcing (-4 W m^{-2} for a year). (a) Fits to monthly averaged MITgcm full ocean response (dotted blue) using the 2-box (solid blue) and the 1.5-box model (solid purple) and fit to the slab ocean MITgcm response (dotted red) using a 1-box model (red). (b) SST (dotted blue) and temperature at 120-m depth (dotted orange) from the MITgcm full ocean configuration with the corresponding 2-box model temperatures T_1 (solid blue) and T_2 (solid orange). The GCM temperature at 120-m depth and T_2 are multiplied by 5 for clarity. (c) Fits to the yearly averaged MITgcm full ocean response (dotted blue) using the 2-box model and the 1.5-box model.

$$\tau_f \approx \frac{\rho c_w h_1}{q}, \quad (10a)$$

$$\tau_s \approx \frac{\rho c_w h_2}{\lambda} = \frac{\mu \tau_f}{r}, \quad (10b)$$

$$T_f \approx \frac{T_c}{1+r}, \quad (11a)$$

and

$$T_s \approx \frac{r}{1+r} T_c. \quad (11b)$$

In this limit, it is interesting (and curious) to note that the fast time scale is controlled by oceanic damping q , whereas the slow time scale is controlled by the climatic feedback λ . The two time scales lead to the kinked profiles evident in Fig. 1, which become more prominent as the ratio μ increases. Physically, we can understand this as a rapid initial response during which the temperature anomaly is sequestered in the deeper ocean, followed by a slower evolution during which the anomaly is damped by climatic feedbacks. In this limit, the coefficients T_f

and T_s only depend on T_c and r . We now assess the magnitudes of r and μ by fitting the analytical solutions to curves obtained from the GCM.

Figure 6a shows the 1-box and 2-box model fits to the MITgcm slab and full ocean responses, respectively. The value of h_1 is set to 43 m, the globally and annually averaged mixed layer depth diagnosed from a long control simulation of the coupled model. To estimate λ , we use the equilibrium response of the slab ocean configuration to a constant, spatially uniform forcing F_s . Setting $F(t) = F_s$ in Eq. (1) produces a response that asymptotes to the equilibrium climate sensitivity (ECS):

$$\text{ECS} = \frac{F_s}{\lambda}. \quad (12)$$

While Eq. (12) is usually applied to a doubling of CO_2 , here we consider a negative forcing $F_s = -4 \text{ W m}^{-2}$ that produces an equilibrium SST response of $\text{ECS} = -2.67^\circ\text{C}$, from which we infer $\lambda = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$. Least squares minimization with respect to the full ocean Pinatubo response in Fig. 2b then gives $q = 3.5 \text{ W m}^{-2} \text{ K}^{-1}$ and $h_2 = 150 \text{ m}$ (and thus, $r = 0.29$ and $\mu = 2.3$) with a fitting

TABLE 1. The 2-box model parameters obtained by curve fitting the monthly SST response of the full ocean MITgcm to an idealized Pinatubo eruption (-4 W m^{-2} for a year).

Parameter	Physical interpretation	Exact fit	Approximation ($r \ll 1$)	Approximation ($r \ll 1, \mu \gg 1$)
h_1	Mixed layer depth	43 m	43 m	43 m
h_2	Deeper ocean depth	150 m	150 m	150 m
λ	Climatic feedback	$1.5 \text{ W m}^{-2} \text{ K}^{-1}$	$1.5 \text{ W m}^{-2} \text{ K}^{-1}$	$1.5 \text{ W m}^{-2} \text{ K}^{-1}$
q	Oceanic mixing	$3.5 \text{ W m}^{-2} \text{ K}^{-1}$	$3.5 \text{ W m}^{-2} \text{ K}^{-1}$	$3.5 \text{ W m}^{-2} \text{ K}^{-1}$
μ	Ratio of ocean to climatic damping (q/λ)	2.3	2.3	2.3
r	Heat capacity ratio (h_1/h_2)	0.29	0.29	0.29
τ_f	Fast time scale	1.0 yr	1.2 yr	1.6 yr
τ_s	Slow time scale	22.0 yr	19.2 yr	13.4 yr
T_f/T_c	Fast amplitude	0.86	0.88	0.78
T_s/T_c	Slow amplitude	0.14	0.12	0.22

accuracy of $R^2 = 0.87$. The single-exponential fit to the slab ocean configuration ($R^2 = 0.97$) is obtained by setting $q = 0$. The relaxation time of the slab ocean curve is $\tau_m = 4$ years, whereas the fast and slow time scales corresponding to the full ocean simulations are $\tau_f = 1$ year and $\tau_s = 22$ years, respectively. Parameter fits are summarized in Table 1, where the goodness of the approximate expressions Eqs. (7)–(11) is assessed by comparison with the full analytical expression. The limit solutions for $r \ll 1$ given by Eqs. (7) and (8) are very good approximations to the exact GCM fits. When we additionally assume $\mu \gg 1$, the fast time scale prediction remains relatively accurate, but the slow time scale reduces to 13 years and is hence underestimated by a factor of 2. We conclude that the $r \ll 1$ and $\mu \gg 1$ limit [Eqs. (10) and (11)] provide useful insight and have some limited quantitative skill.

Figure 6b plots the evolution of T_1 and T_2 for the best-fit solution along with the globally averaged SST and temperature at 120-m depth from the MITgcm, enabling the analytical solution to be compared to the GCM. Immediately after the eruption, the large temperature difference between the mixed layer and the deeper ocean results in significant vertical heat exchange, with surface cooling being sequestered into the thermocline. In this first phase of relaxation, T_2 decreases, and ocean heat exchange works in the same sense as climate feedbacks to damp the SST response. The fast (order 1 year) time scale τ_f given by Eq. (7a) is controlled by $\lambda(1 + \mu)$. Since $\mu = 2.3$, vertical ocean heat exchange dominates over climatic feedbacks in setting the time scale of the fast response. Held et al. (2010) and Gregory et al. (2016) also note that the ocean plays a significant role on these short time scales but assume $\mu \leq 1$. Over time, the temperature anomalies T_1 and T_2 approach one another, and the system behaves like a single layer of thickness $h_1 + h_2$ relaxing on a much longer (20 year) time scale set by climate feedbacks. The sequestration of the temperature

anomaly into the interior ocean acts to temporarily shield it from surface damping, resulting in a lingering of the cold signal.

Previous studies (e.g., Held et al. 2010; Gregory et al. 2016) have argued that a model with a deep h_2 (~ 1000 m), which we refer to as a 1.5-box model, along with $q \sim 1 \text{ W m}^{-2} \text{ K}^{-1}$, can fit the response to a volcanic eruption. It is clear that this type of model cannot fit the fast (~ 1 year) time scale observed in the MITgcm response when the signal is resolved with monthly mean data (see Fig. 6a). However, as shown in Fig. 6c, it gives a relatively good fit to annually resolved data ($R^2 = 0.90$) when using a thicker upper ocean ($h_1 = 80$ m), $q = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$, and $h_2 = 1000$ m. While the 1.5-box model underestimates the first-year peak cooling because of its thicker h_1 (see Fig. 6c), it matches the subsequent decay with its dominant e -folding time scale of 3.7 years. One might reasonably interpret the h_1 of the 1.5-box model as the winter mixed layer depth from which fluid “escapes” into the main thermocline (see Williams et al. 1995), as opposed to the annual-mean value used in the 2-box model. The large q of the 2-box model represents fast (subannual) heat exchange processes between the mixed layer and the seasonal thermocline (e.g., winter deepening of the mixed layer, Ekman pumping, etc.), whereas the lower q of the 1.5-box model represents heat exchange with the deeper ocean (~ 1000 m). The absence of a subannual structure in the response of other models (e.g., Held et al. 2010; Merlis et al. 2014; Gregory et al. 2016) is likely due to a more realistic forcing spread over several years, which can mask the fast (~ 1 year) response time scale. Our results, however, are consistent with the work of Wigley et al. (2005), who report a sharp (2–3 years) decay time scale after the eruption followed by a long “tail” in the signal.

b. Impulse versus step responses

In Fig. 7, we explore the difference between the impulse (volcanic) and the step responses of the MITgcm

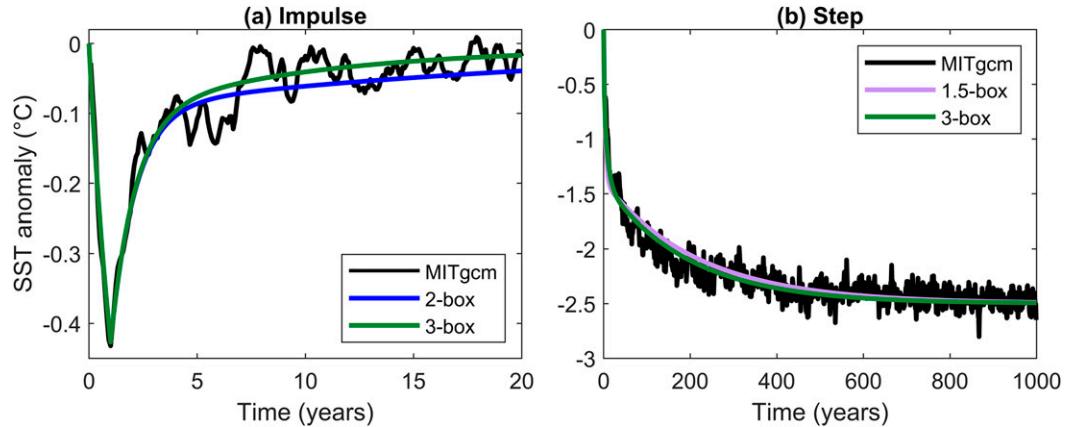


FIG. 7. Fits to the MITgcm full ocean response. (a) Fits to the ensemble-mean idealized Pinatubo response (-4 W m^{-2} for a year) over monthly averaged data (black) using a 2-box (blue) and a 3-box (green) model. (b) Fits to the MITgcm step response with -4 W m^{-2} forcing (black) using the 1.5- (purple) and 3-box models.

full ocean configuration. The step response is of interest because it mimics the effect of, for example, impulsively changing the level of greenhouse gases in the atmosphere. The step forcing (-4 W m^{-2}) is simulated in exactly the same manner as the impulse, but held constant over time until surface equilibrium is reached. As shown in Fig. 7b, the step response is well approximated by the 1.5-box model introduced in the previous section, but cannot be captured by the 2-box model, which only resolves the top $\sim 200 \text{ m}$ of the ocean. The 1.5-box model can thus emulate both the step and impulse responses for yearly averaged data, but not the monthly resolved impulse response. A box model that can simultaneously capture yearly, decadal, and centennial time scales must have three separate layers. Hence, we introduce a 3-box model consisting of the 2-box model with an additional deep layer ($h_3 = 2000 \text{ m}$) underneath h_2 and a second heat exchange parameter $q_2 = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$ linking those two boxes. These parameters are found by simultaneously fitting to the step and impulse responses with monthly mean data (see Fig. 7). The model has time scales of 1, 11, and 273 years, which make up 85%, 14%, and 0.05% of the impulse response, respectively, and 22%, 41%, and 37% of the step response, respectively. The 3-box model is thus useful in linking the impulse and step responses of the GCM, but the 2-box model is better suited for analytical interpretation due to its simplicity.

c. Inferring climate sensitivity from volcanic eruptions

A number of studies (e.g., Lindzen and Giannitsis 1998; Wigley et al. 2005; Yokohata et al. 2005; Hegerl et al. 2006; Bender et al. 2010; Merlis et al. 2014) have attempted to relate the response of SST following a volcanic eruption to some measure of the climate sensitivity.

The methods can be grouped as follows: (i) inferring the ECS from the peak cooling in SST after an eruption, (ii) inferring the ECS from the integrated SST response, and (iii) inferring the transient climate response (TCR) from the integrated SST signal. We now critically review these methods, guided by our simulations and the simple models discussed previously.

1) PEAK COOLING AND ECS

Figure 8 shows idealized Pinatubo-like eruptions in the 2-box model for increasing values of λ , with all other parameters constant as in Table 1. Past studies (e.g., Wigley et al. 2005; Bender et al. 2010) have attempted to connect this peak cooling to λ (and hence, the ECS), but did not find a strong link between the two quantities.

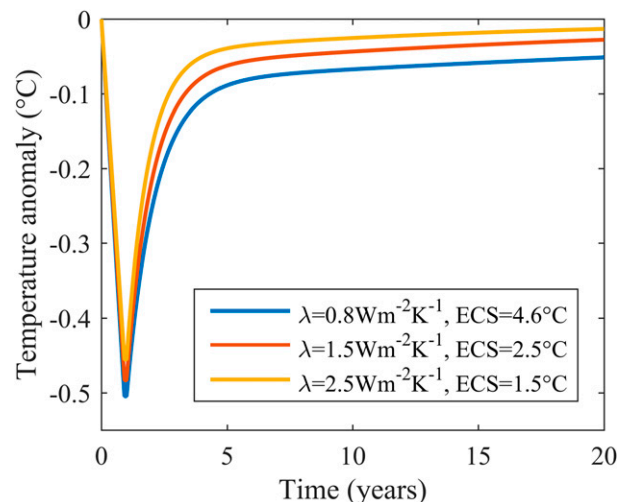


FIG. 8. The 2-box model responses to an idealized Pinatubo forcing (-4 W m^{-2} for a year) for a range of λ (or ECS) values. All other parameters are fixed to those in Table 1.

While the effect of unforced variability was invoked to explain the lack of correlation, the 2-box model can be used to explore further. Equation (1) can be used to obtain an approximate expression for the peak cooling T_c after a pulse forcing that lasts a small but finite time Δt , assumed to be approximately 1 year (details in the SI):

$$T_c \approx \frac{V}{\rho c_w h_1 + \frac{\lambda \Delta t (1 + \mu)}{2}}, \quad (13)$$

which reduces to Eq. (4) as Δt tends to zero. When the forcing time is finite, the peak cooling T_c depends on both λ and q (through the parameter μ). In the limit of small μ , the ocean does not play a significant role, and, in principle, the value of λ could be inferred from knowledge of V , T_c , Δt , and h_1 . However, if $\mu \geq 1$, oceanic damping becomes as important as λ in reducing T_c , and, hence, any correlation between the two can be confounded by the influence of ocean heat sequestration. Moreover, Fig. 8 shows that the influence of λ on the peak cooling is relatively small (especially for small Δt) and can easily be obscured by noise, as argued by Wigley et al. (2005), Bender et al. (2010), and Merlis et al. (2014).

2) INTEGRATED RESPONSE AND ECS

To mitigate the effect of noise, previous studies (e.g., Yokohata et al. 2005; Bender et al. 2010; Wigley et al. 2005) have attempted to link the ECS to the time-integrated SST response, rather than just the peak cooling value. Integrating Eq. (1) in time from $t = 0$ to ∞ gives

$$\lambda \int_0^{\infty} T_1(t) dt = V. \quad (14)$$

Equation (14) is a statement of conservation of energy: the energy extracted from the system by the volcanic eruption (rhs) must be balanced by the total energy recovered through climatic feedbacks (lhs) over the entire duration of the process. We note here that since the time-integrated response does not depend on ocean damping, the presence of an active deeper ocean underneath the mixed layer does not change the value of the integrated temperature response. A larger value of q shifts the weight of the response toward longer time scales, without affecting the total “area under the curve” (see Fig. 1).

The absence of the parameter q in Eq. (14) also means that the time-integrated response can in theory be used to infer λ (or the ECS), without the conflating influence of ocean damping. A common problem, however, is that in complex GCMs and observations, the response typically becomes indistinguishable from noise 5–10 years

after Pinatubo-like eruptions. If μ is small, the time scale of the response is dominated by the mixed layer, and in that case, an integration time of 5–10 years may be enough to obtain a reliable estimate of λ . However, if μ is large, a significant part of the cooling energy is stored in the subsurface ocean, and using Eq. (14) overestimates λ . For example, applying this method to our 2-box model fit with an integration time of 15 years gives $\lambda = 2.9 \text{ W m}^{-2} \text{ K}^{-1}$, which is much larger than the value of $1.5 \text{ W m}^{-2} \text{ K}^{-1}$ obtained from our GCM’s ECS. Moreover, Fig. 8 shows that the response curves (corresponding to different λ values) are tightly packed in the initial fast decay stage (0–3 years) but later become more distinct from each other. This overall behavior is reflective of the conclusion we drew from Eq. (10): in the limit of large μ (and small r), the fast time scale is controlled by q , whereas the slower time scale is set by λ . Since in practice we are constrained to integrate over short periods of time due to noise, this further limits the usefulness of Eq. (14) for estimating the ECS.

It is likely that the sensitivity of the short time evolution of SST to q , as well as λ , accounts for the large range of estimates of ECS inferred from volcanic eruptions that have been reported in the literature. Lindzen and Giannitsis (1998) simulated volcanic eruptions representing the ocean as a 1D diffusion model to argue that a high ECS is not realistic, because it implies a much longer decay time scale than seen in observations. However, Wigley et al. (2005) argued that the slow decays simulated by Lindzen and Giannitsis (1998) were likely too long and find that an ECS as high as 4.5°C per doubling of CO_2 ($F_s = 3.7 \text{ W m}^{-2}$) cannot be discarded. Yokohata et al. (2005) rule out very high sensitivities (6.3°C) but find in their model that an ECS of 4.0°C produces results consistent with observations. For the same values of h_1 , λ , and diffusivity used by Lindzen and Giannitsis (1998) in a 1D diffusion model, we obtain (not shown) significantly shorter decay time scales than they reported. Our own results are more consistent with the time scales found by Santer et al. (2001) and Wigley et al. (2005).

3) INTEGRATED RESPONSE AND TCS

Since the volcanic SST signal rapidly fades to noise for typical modern-era volcanic eruptions, Merlis et al. (2014) suggested that the SST response could provide a more reliable constraint on the transient climate sensitivity (TCS) instead of the long-term ECS. The TCS is a measure of the response of the system while the deep ocean temperature has not been significantly affected by the forcing. This is perhaps a more relevant characterization of the evolution of the climate under anthropogenic CO_2 forcing than the ECS (e.g., Held et al. 2010).

The TCS can be derived by setting $T_2 \ll T_1$ in Eq. (1) to yield the 1.5-layer model:

$$\rho c_w h_1 \frac{dT_1}{dt} \approx -(\lambda + q)T_1 + F(t), \quad (15)$$

which gives

$$\text{TCS} = \frac{F_s}{\lambda + q} \quad (16)$$

in the steady state, where $F(t) = F_s$.

The TCS is equivalent to the commonly used TCR, if F_s is the $2 \times \text{CO}_2$ forcing (e.g., Gregory and Forster 2008). It is inversely proportional to the sum $\lambda + q$, as is the approximate fast time scale τ_f given in Eq. (7a). Merlis et al. (2014) use the above 1.5-box model to estimate the TCS by integrating Eq. (15) up to a time t_I short enough that $T_2 \ll T_1$, but long enough that the lhs of Eq. (15) becomes negligible. It is also assumed that the forcing has ceased to act before time t_I . The energy balance then becomes

$$(\lambda + q) \int_0^{t_I} T_1(t) dt \approx V. \quad (17)$$

Equation (17) states that the energy extracted by the forcing is approximately balanced by the energy dissipated by both atmospheric and oceanic damping up to time t_I [in contrast to Eq. (14), where the integral is carried to infinity]. Merlis et al. (2014) use $t_I = 15$ years and find values of q that are on the order of $1 \text{ W m}^{-2} \text{ K}^{-1}$. Using this method on the MITgcm signal expressed with annual-mean data (as appropriate for the 1.5-box model fit) and assuming that λ is known gives $q \sim 1.6 \text{ W m}^{-2} \text{ K}^{-1}$, which is a good approximation for the least squares fit value of $q = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$.

When considering monthly averaged data and the corresponding 2-box model fit, we argue based on Fig. 6b that an integration time of $t_I = 15$ years is too long to satisfy the condition $T_2 \ll T_1$ because beyond years 2–3, the deeper layer temperature anomaly (at around 120-m depth) is of the same order of magnitude as the SST anomaly. For such short times, the lhs of Eq. (15) can no longer be neglected, and one must integrate Eq. (15) and take account of the transient term on the lhs to give

$$(\lambda + q) \int_0^{t_I} T_1(t) dt \approx V - \rho c_w h_1 T_1(t_I). \quad (18)$$

Equation (18) can be used to estimate q in the 2-box model but requires knowledge of the mixed layer depth h_1 in addition to V and $T_1(t)$. Using Eq. (18) with $t_I = 3$

years, we find $\lambda + q = 4.4 \text{ W m}^{-2} \text{ K}^{-1}$, and hence, $q = 2.9 \text{ W m}^{-2} \text{ K}^{-1}$. This is an underestimate of the value obtained from curve fitting ($3.5 \text{ W m}^{-2} \text{ K}^{-1}$), but more accurate than the one obtained using Eq. (17) with $t_I = 15$ years for monthly averaged data ($1.4 \text{ W m}^{-2} \text{ K}^{-1}$). Further tests find that short integration times give better results than longer ones, despite still underestimating q . The method is also less accurate for large q values because this leads to a rapid increase in the magnitude of T_2 , causing the approximation $T_2 \ll T_1$ to break down after only a short time.

We conclude that using annual-mean data and the method outlined in Merlis et al. (2014) can provide a good estimate of the TCS pertaining to the 1.5-box model and hence to the step forcing. When considering monthly mean data, using Eq. (18) with a short integration time can give a reasonable approximation for $\lambda + q$, where q represents the fast heat exchange between the mixed layer and subsurface ocean.

d. Series of impulses

In Fig. 9, we employ a box model approach to assess the potential for a volcanic SST response to build upon a history of eruptions and display a larger cooling than it would have done in isolation. We develop a metric for accumulation potential by considering a series of uniform eruptions spaced at a regular interval τ and assessing how the peak cooling after each eruption grows over time. Each eruption is modeled to extract 1 year's worth of energy from the system described by Eqs. (1) and (2). As was seen in the aquaplanet simulations in Fig. 4, the peak magnitude of the response may increase over time if the response decay time scale is large relative to the interval between each eruption. We obtain an analytical expression for the curve that passes through the peak temperature response following each eruption, which we term the envelope of the signal T_{en} (see SI for a detailed derivation):

$$T_{\text{en}}(t) = T_f \frac{1 - e^{-(t+\tau)/\tau_f}}{1 - e^{-\tau/\tau_f}} + T_s \frac{1 - e^{-(t+\tau)/\tau_s}}{1 - e^{-\tau/\tau_s}}. \quad (19)$$

Equation (19) can easily be extended to a 3-box model by adding a term with the appropriate time constant. In the limit that the repeated eruptions occur for all time ($t \rightarrow \infty$), the envelope asymptotes to a finite value T_∞ given by (see SI)

$$T_\infty = \frac{T_f}{1 - e^{-\tau/\tau_f}} + \frac{T_s}{1 - e^{-\tau/\tau_s}}. \quad (20)$$

This limit is reached when the rate at which cooling is supplied by the eruptions equals the rate at which it is

Normalized accumulation potential

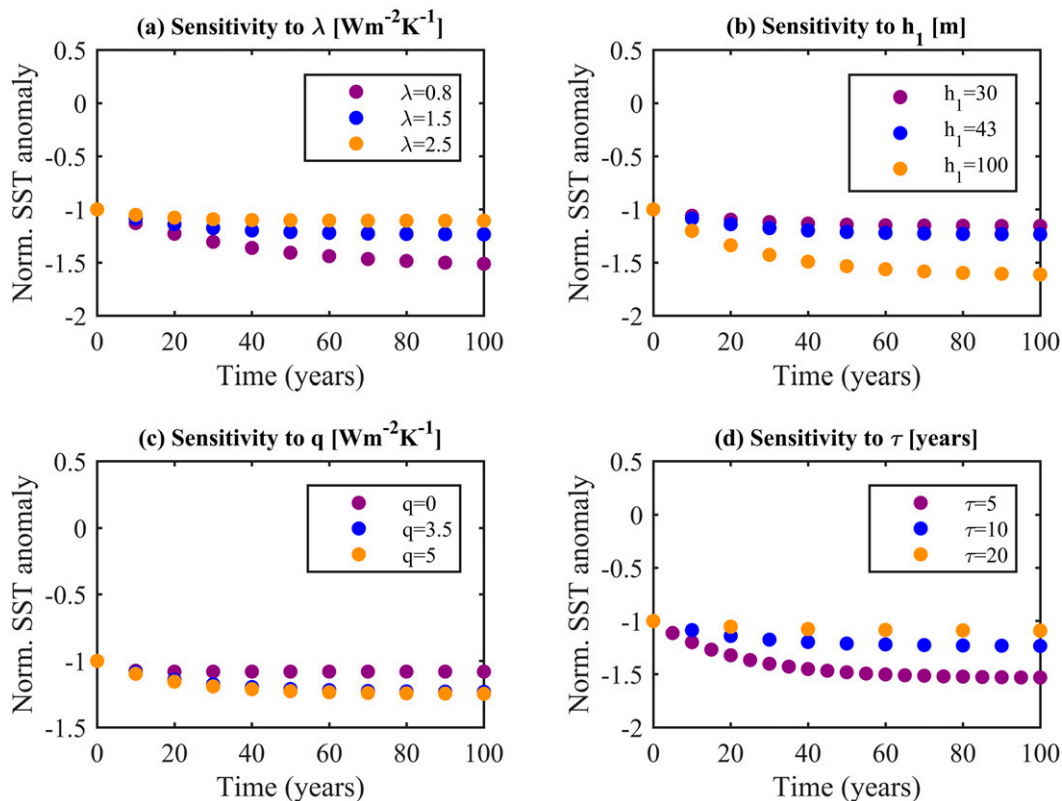


FIG. 9. Normalized temperature envelope T_{en} for a series of uniform and regularly spaced eruptions in the 2-box model. Each dot represents the peak cooling temperature after a new eruption. Parameter sensitivity is explored for (a) the climate feedback parameter λ , (b) the mixed layer depth h_1 , (c) the ocean exchange parameter q , and (d) the time interval between eruptions τ . All fixed parameters are as in Table 1, and the default τ value is 10 years.

lost through climate feedbacks. Equation (20) thus provides a theoretical maximum cooling resulting from successive uniform eruptions. The analytical expression for T_{∞} explicitly reveals how the potential for accumulation increases when the ratios τ/τ_f and τ/τ_s decrease. The first and second terms of Eq. (20) represent the contribution of the fast and the slow responses to the asymptotic peak temperature, respectively. For the parameter values in Table 1 and $\tau = 10$ years, we find that while the fast mode is negligibly enhanced, the slow mode grows over the slow time scale by a factor of 1.7 on moving from the initial to the equilibrium state. Moreover, the cycle-average temperature at equilibrium is given by (see SI)

$$\bar{T}_{\infty} = \frac{V}{\lambda\tau}, \quad (21)$$

where V is the energy extracted by a volcanic eruption persisting for 1 year. It is useful to note that \bar{T}_{∞} tends to

the ECS given by Eq. (12) when τ tends to 1 year, and the forcing becomes essentially continuous.

In Fig. 9, we explore the sensitivity of the temperature envelope T_{en} to λ , q , h_1 , and τ . The buildup amount is expressed relative to the peak cooling after the first eruption in the series. The blue points in each panel describe the accumulation curve obtained with the parameters from the 2-box fit to the MITgcm response (see Table 1). Figures 9a and 9b show that a smaller climate feedback parameter λ and a larger mixed layer depth h_1 elicit a larger accumulation potential T_{en} . Both these parameters directly affect the relaxation of the mixed layer temperature and hence are of primary importance in setting the amount of accumulation. Figure 9c shows the effect of q in enhancing the accumulation potential T_{en} . A comparison with the aquaplanet results from Fig. 4 shows that the 2-box model ($q = 3.5 \text{ W m}^{-2} \text{ K}^{-1}$) quantitatively captures the 20% increase in peak cooling seen in the full ocean configuration. Conversely, the 1-box model ($q = 0$) displays the same absence of

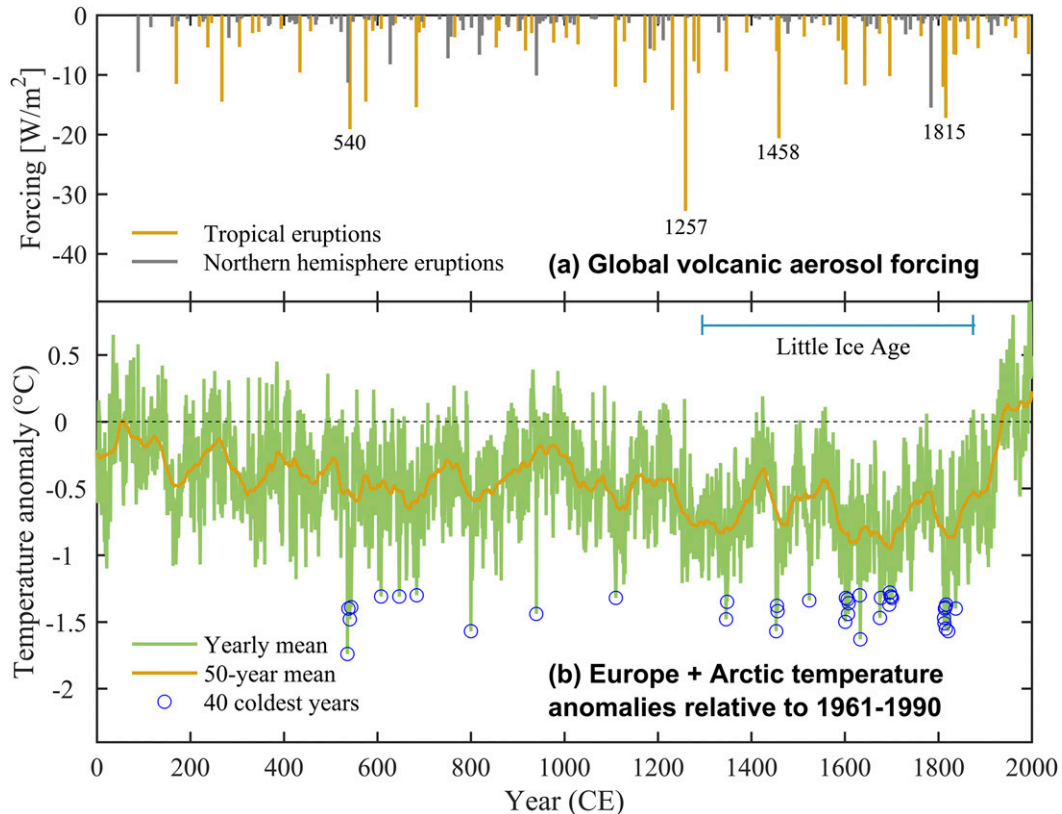


FIG. 10. (a) A 2000-yr reconstruction of global volcanic aerosol forcing from sulfate composite records from tropical (orange) and Northern Hemisphere (gray) eruptions. (b) A 2000-yr record of reconstructed summer temperature anomalies for Europe and the Arctic relative to 1961–90 shown at yearly resolution (green) and as a 50-yr running mean (orange). The 40 coldest single years are indicated with blue circles, and the approximate duration of the LIA is shown. [Data from Sigl et al. (2015).]

accumulation as was observed in the slab configuration. Figure 9d shows the increase in response buildup as the interval between eruptions τ is narrowed. Eruptions spaced by more than 20 years have a very low accumulation potential. Overall, this analysis shows that for the range of parameters considered, a regular series of uniform eruptions can yield a maximum accumulation of approximately 10%–50%. Moreover, ocean heat sequestration promotes accumulation, as indicated by the behavior of T_{en} with increasing q and h_1 .

4. Response to last millennium forcing

The role of the ocean in prolonging climate signals can be seen at work in the context of the volcanic forcing over the last millennium. A growing number of studies (e.g., Crowley 2000; Hegerl et al. 2003; Schurer et al. 2015; Atwood et al. 2016) have highlighted the importance of volcanic eruptions in instigating the coldest period of the Holocene, commonly referred to as the Little Ice Age (LIA; ~ 1250 –1850 CE). They point to

cooling induced by volcanoes as a major contributor to the LIA, beyond the effects of reduced insolation, changes in greenhouse gases, and land-use evolution. Several authors (e.g., Free and Robock 1999; Crowley et al. 2008; Stenchikov et al. 2009) have suggested that the ocean's long response time scales could help explain how eruptions that typically last only 1–2 years could engender cooling over multiple decades. Here, we make use of the MITgcm and box models to investigate the magnitude of the signal due to interaction with the ocean and the relative importance of small versus large eruptions.

Figure 10 [adapted from Sigl et al. (2015)] shows a reconstruction of Europe and Arctic temperatures along with global volcanic activity over the past 2000 years. The two panels show that 20 of the 40 coldest years in the series occurred during the LIA and that those cold years often coincided with the largest eruptions of that period. The LIA was characterized by the occurrence of cold spells during the mid-fifteenth, seventeenth, and early nineteenth centuries. The spatial extent of the cooling is

uncertain, as proxy records originate largely from land in the Northern Hemisphere. Nevertheless, Neukom et al. (2014) suggest that sustained periods of cooling could also have occurred in the Southern Hemisphere, particularly in the seventeenth century. Here, we focus on large tropical volcanic eruptions, because the stratospheric transport of particles toward the poles results in a considerable global impact (Robock 2000). Moreover, the volcanic forcing reconstruction in Fig. 10a indicates that tropical eruptions dominated over high-latitude events over the past 2000 years.

Figure 11a shows an estimate of the volcanic forcing over the last millennium (A. LeGrande 2016, personal communication). It is based on the reconstruction by Crowley and Unterman (2013) and represents the top-of-atmosphere shortwave flux anomaly in the GISS-E2-R model simulations. The forcing reveals the large volcanic eruptions ($\leq -4 \text{ W m}^{-2}$) of the thirteenth, fifteenth, seventeenth, and nineteenth centuries, as well as the smaller ($> -4 \text{ W m}^{-2}$) eruptions that occurred more regularly throughout the time series. In Figs. 11b and 11c, we plot the MITgcm response of globally averaged SSTs to this forcing, for the slab and full ocean configurations. These panels show that SSTs in the full ocean scenario tend to be colder than in the slab for the decades following clusters of large volcanic eruptions (thirteenth, fifteenth, seventeenth, and nineteenth centuries). It should be noted that since the model does not contain ice, it does not capture the positive sea ice feedback proposed by Miller et al. (2012) that links volcanism and the LIA. The differences in responses between the slab and full ocean configurations can thus be attributed to the sequestration of cold anomalies by the interior ocean.

Figure 11c shows that the strong volcanic activity of the thirteenth century, which has been related to the onset of the LIA (e.g., Miller et al. 2012; Cole-Dai et al. 2013), has an effect that spans multiple decades. At the end of this sequence of eruptions, the temperature anomaly in the full ocean configuration remains mostly colder than the slab until the middle of the fifteenth century. Similarly, after the large 1450s eruptions (Cole-Dai et al. 2013), the full ocean configuration displays a temperature anomaly of around -0.1°C that lasts for about 100 years, in contrast to the slab, whose response decays to noise after only 20 years. This prolonged cooling is significantly outside the range of internal variability seen in 100-yr segments of the control simulation. There is also some signal prolongation after the seventeenth-century eruptions that persists for around 20 years at the beginning of the 1800s. Finally, as reported by Crowley et al. (2008), the close packing of four eruptions between 1809 and 1835 (including the

Tambora eruption in 1815) leads to an accumulated climate response in the nineteenth century, because of the long time scales imparted by the global ocean.

Figure 11d shows the response of the 1-box, 2-box, and 3-box models to the historical forcing. Each additional box provides more signal prolongation, as deeper levels of the ocean participate in the response. The 1-box model matches the GCM slab solution well, with its single decay time scale of around 4 years. The temperature anomaly of the 2-box model is colder than the 1-box model for 70% of the time series, showing the effect of decadal-scale (negative) heat storage within the seasonal thermocline. The 3-box model displays the longest “tail” in the signal, with temperature anomalies of approximately -0.1°C lasting for several decades after large clusters of eruptions, as in the fully coupled GCM. It should be noted that these simple box models have been calibrated for a Pinatubo-size eruption, and thus, they do not capture the deepening of the mixed layer following eruptions of very large magnitude. As discussed in section 2b, a $10 \times$ Pinatubo eruption in the model causes a 10% reduction in peak cooling relative to linear scaling due to the transient expansion of the mixed layer depth. Nevertheless, the box model analysis shows that temperature anomalies stored in the deep ocean ($\sim 1000 \text{ m}$) for centennial time scales likely play an important part in these long periods of anomalously cold SSTs seen in the GCM.

In Fig. 11e, we use the 3-box model to estimate the contribution of the response from small eruptions ($> -4 \text{ W m}^{-2}$) versus large eruptions ($\leq -4 \text{ W m}^{-2}$). We find that small eruptions are frequent enough that their responses accumulate and cool the climate almost continuously throughout the entire time series by about 0.05°C . Large eruptions occur more rarely but can still lead to accumulation (e.g., thirteenth and nineteenth centuries). Thus, both small and large eruptions seem to play an important part in the cooling of the climate during the last millennium. Moreover, the large volcanic eruptions from 1250 to 1850, coupled with the heat sequestration from the deeper ocean, could have been a significant driver of the extended periods of cooling observed during the LIA.

5. Discussion and conclusions

We have explored the role of the ocean in modulating the globally averaged SST response of the climate to volcanic cooling, using a hierarchy of idealized models. We find that the presence of the deeper ocean beneath the mixed layer ($h_1 = 43 \text{ m}$) introduces a kink in the response, characterized by two time scales. This effect strengthens with the parameter μ , which represents the

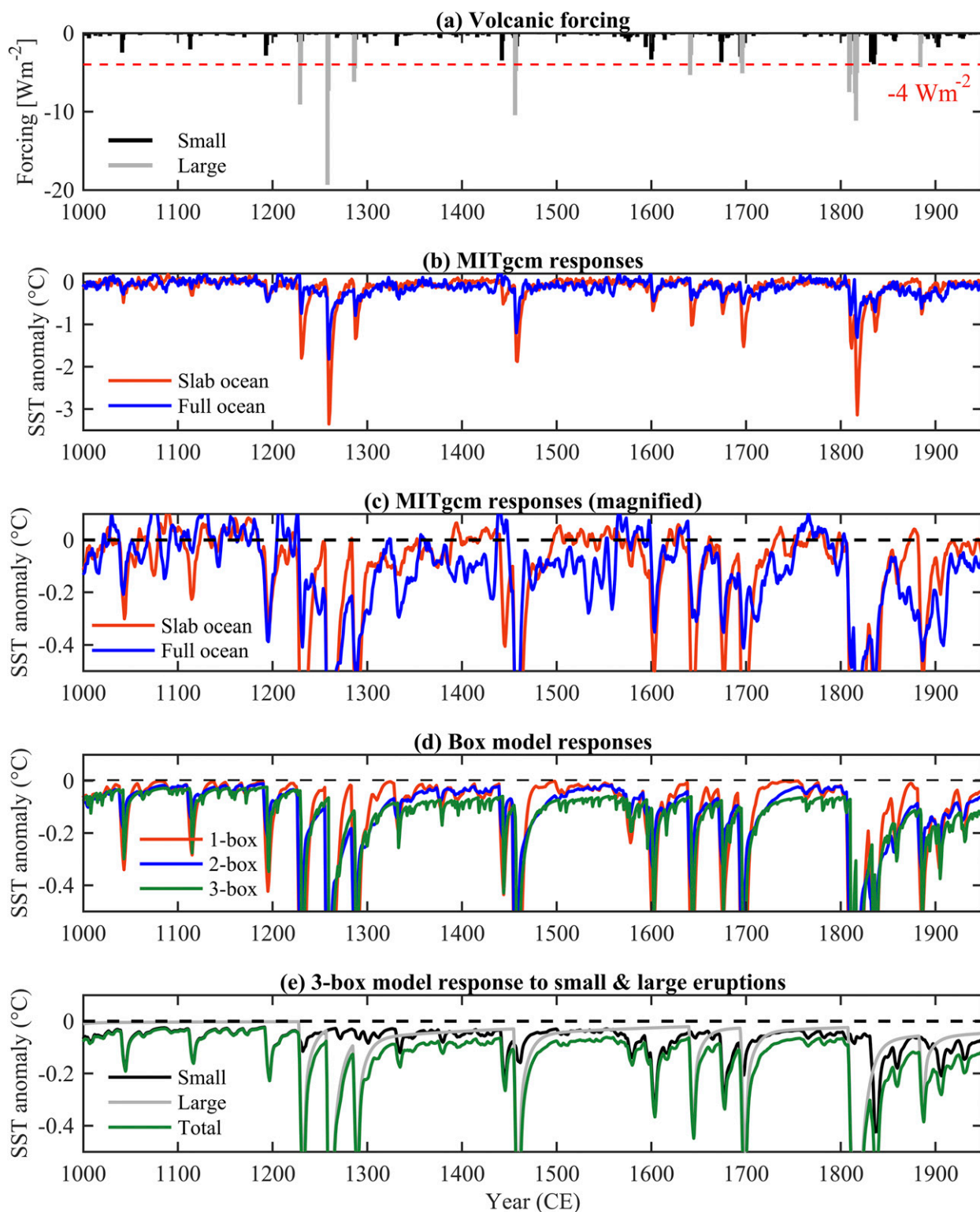


FIG. 11. (a) Tropical volcanic forcing of the last millennium (A. LeGrande, NASA GISS, 2016, personal communication) divided into small ($> -4 \text{ W m}^{-2}$) and large eruptions ($\leq -4 \text{ W m}^{-2}$). (b) Responses of the MITgcm coupled model with a full ocean (blue) and a slab ocean (red) to the volcanic forcing shown in (a). (c) The 5-yr running mean of (b) on a magnified scale. (d) The 1- (red), 2- (blue), and 3-box model (green) responses to the forcing. (e) The 5-yr running mean of the 3-box model response to the small (black), large (gray), and total (green) volcanic forcing.

ratio of the ocean heat exchange parameter q to the climatic feedback parameter λ . Fitting the (monthly mean) MITgcm response to a 2-box model gives $q = 3.5 \text{ W m}^{-2} \text{ K}^{-1}$, $\lambda = 1.5 \text{ W m}^{-2} \text{ K}^{-1}$, and $\mu = q/\lambda = 2.3$. This large value of μ leads to a pronounced kink in the response profile, with fast and slow time scales of 1 and 20 years, respectively. In the limit of large μ , the fast time scale is (perhaps counterintuitively) dominated by ocean damping, whereas the slow time scale is controlled by atmospheric feedbacks. Thus, in the first few years following the eruption, heat exchange with the subsurface ocean dominates over the climatic feedbacks in relaxing the SST response, sequestering the (negative) heat into the seasonal thermocline, and reducing the magnitude of the peak anomaly. Subsequently, the cooling stored in the subsurface ocean is delivered back to the surface over decadal periods, extending the response beyond the time scale implied by a slab ocean configuration.

We interpret the large q of the 2-box model as representing vertical mixing processes occurring in the seasonal thermocline over monthly time scales. It is likely that previous studies (e.g., Held et al. 2010; Merlis et al. 2014; Gregory et al. 2016) did not observe the sharp kink in the response because they simulated more realistic eruptions for which the forcing was spread over several years, thus masking the subannual structure. Indeed, when considering the MITgcm results with yearly mean SST anomalies instead of monthly means, the signal can be captured by a 1.5-box model (large $h_2 \sim 1000 \text{ m}$), with a thicker upper layer ($h_1 = 80 \text{ m}$) and a lower value of q ($1.5 \text{ W m}^{-2} \text{ K}^{-1}$). One might reasonably interpret the h_1 of the 1.5-box model as the winter mixed layer depth from which fluid “escapes” into the main thermocline (see Williams et al. 1995), as opposed to the annual-mean value used in the 2-box model. The 1.5-box model underestimates the first-year cooling because of its larger h_1 , but captures the subsequent decay occurring over an e -folding time scale of approximately 3.7 years. The same 1.5-box model can fit the step response of the GCM, and hence, we interpret its associated q value as a parameter representing heat exchange with the deeper levels of the ocean ($\sim 1000 \text{ m}$). However, a three-layer model with e -folding time scales of 1, 11, and 273 years is required to simultaneously capture the subannual and centennial scales relevant to the impulse and step responses, respectively.

We went on to review methods that have been suggested for constraining climate sensitivity using the global-mean SST response to a volcanic eruption: (i) peak cooling, (ii) integrated response to estimate the ECS, and (iii) integrated response to estimate the TCS. We find that unforced variability masking the volcanic

signal is a strong limiting factor in all such approaches. For methods (i) and (ii), we find that results can additionally be confounded by the effects of ocean heat uptake. When considering monthly mean data and the corresponding 2-box model fit, using method (iii) with a short integration time (~ 3 years) can yield reasonable estimates for $\lambda + q$, where q is relevant for heat exchange within the seasonal thermocline. With yearly mean data and the 1.5-box model fit, using method (iii) with a longer integration time (~ 15 years) may provide an estimate of the TCS relevant to the step response if enough ensemble members are available to resolve the decadal time scale.

When $\mu > 1$, the large kink in the 2-box SST anomaly profile implies a longer prolongation of the response beyond that expected from a slab ocean, which favors accumulation from successive eruptions. When forced by Pinatubo-like eruptions every 10 years, the peak anomaly response grows by 20% over 100 years in the full ocean simulations, but does not grow in the slab ocean case. The accumulation behaves rather linearly in the GCM and can thus be studied with simple box models. We find that there is a limit to the theoretical maximum amount of accumulation that can occur for a series of regularly spaced uniform eruptions, which decreases with the climatic feedback λ and increases with the mixed layer depth h_1 . For typical parameter values, this maximum accumulation potential is around 10%–50% of the initial peak cooling, and it decreases significantly as the interval between eruptions becomes larger than the slow decay time scale (~ 20 years).

Finally, we demonstrate how signal prolongation and accumulation due to the presence of the subsurface ocean reservoir could help explain the extended periods of cooling observed during the Little Ice Age (LIA; ~ 1250 – 1850 CE). After the large clusters of eruptions of the thirteenth, fifteenth, seventeenth, and nineteenth centuries, we find that the presence of an active deeper ocean prolongs the surface cooling over multiple decades. In particular, after the large eruptions of the 1450s, the simulation with an active ocean shows an SST anomaly of -0.1°C lasting for 100 years versus only 20 years in the slab ocean configuration. The box model analysis shows that both decadal- and centennial-scale storage of heat anomalies in the ocean is important in explaining the prolongation behavior, with the 3-box model most accurately emulating the GCM. The response functions obtained from box models also suggest that frequent small-scale eruptions ($> -4 \text{ W m}^{-2}$) were responsible for a continuous cooling of about 0.05°C throughout the last millennium, whereas larger eruptions were rarer, but could have played an important part in the extended periods of cooling during the LIA

when aided by ocean heat sequestration. These results are in line with the conclusions from [Crowley et al. \(2008\)](#), [Miller et al. \(2012\)](#), [Cole-Dai et al. \(2013\)](#), [Atwood et al. \(2016\)](#), and others. We thereby conclude that the mechanisms responsible for storing volcanic cooling in the subsurface ocean are relevant for questions pertaining to climatic cooling over decadal to centennial time scales.

Acknowledgments. MG acknowledges support from the John H. Carlson fellowship and JM from the NSF FESD Ozone project. We thank Allegra LeGrande from NASA GISS for her support in this work and for the data on volcanic forcing during the last millennium. We are most grateful for informative discussions with Susan Solomon, Jean-Michel Campin, Brian Green, and Paul O’Gorman. Finally, we thank the three anonymous reviewers who helped improve this paper.

APPENDIX

MITgcm Coupled Model

This study employs the MITgcm ([Marshall et al. 1997a,b](#)) in a coupled atmosphere–ocean configuration. The atmosphere and ocean fluids both use the same C32 cubed–sphere grid (32×32 cells per face), yielding a nominal horizontal resolution of 2.8° ([Adcroft et al. 2004](#); [Adcroft and Campin 2004](#)). The ocean is uniformly 3.4 km deep and has 15 vertical levels, with grid spacing increasing from 30 m at the surface to 400 m at depth. Effects of mesoscale eddies are parameterized as an advective process ([Gent and McWilliams 1990](#)) and isopycnal diffusion ([Redi 1982](#)). Convective adjustment is implemented as an enhanced vertical mixing of temperature, and salinity is used to represent ocean convection ([Klinger and Marshall 1995](#)). The background vertical diffusion is uniform and set to $3 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$.

The atmospheric component of the model has 26 pressure levels and employs a gray radiation scheme with parameterized convection and precipitation, as in [Frierson et al. \(2006\)](#). Surface fluxes are computed using standard drag laws, based on Monin–Obukhov similarity theory, given the atmospheric model’s lowest-level wind, temperature, and humidity and the surface roughness lengths, temperature, and humidity. The longwave optical thickness is modified by the distribution of water vapor, following [Byrne and O’Gorman \(2013\)](#). In this simplified radiation scheme, the shortwave flux does not interact with the atmosphere, and hence, the planetary albedo is the same as the surface albedo. There are no clouds in the model. A seasonal cycle of insolation

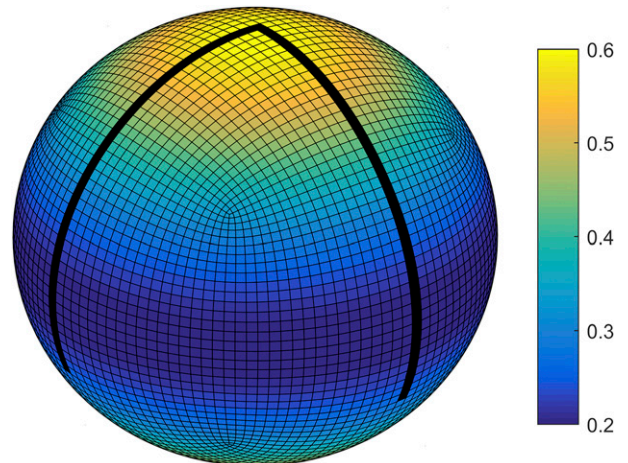


FIG. A1. Linear albedo gradient imposed at the surface of the MITgcm. The grid is in a cubed sphere configuration with 32×32 points per face, with a nominal horizontal resolution of 2.8° . The thick black lines indicate the solid ridges of the “double drake” setup extending from the North Pole to 35°S and set 90° apart.

at the top of the atmosphere is specified for a solar constant of 1360 W m^{-2} . The meridional albedo contrast is represented by a pole-to-equator albedo gradient varying linearly from 0.6 to 0.2 (see [Fig. A1](#)), in line with the observations presented in [Donohoe and Battisti \(2011\)](#).

We also make use of a “slab ocean” configuration of the MITgcm that has a single layer in the vertical whose thickness is fixed in time but varies spatially according to the annual-mean mixed layer depth, diagnosed from a long control simulation according to the method outlined in [Kara et al. \(2000\)](#). Surface heat fluxes are imposed as a stationary boundary condition to the slab ocean model. These heat fluxes are also diagnosed from the control simulation and represent ocean energy transport convergence into a given grid box.

REFERENCES

- Adcroft, A., and J.-M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284, <https://doi.org/10.1016/j.ocemod.2003.09.003>.
- , —, C. Hill, and J. Marshall, 2004: Implementation of an atmosphere–ocean general circulation model on the expanded spherical cube. *Mon. Wea. Rev.*, **132**, 2845–2863, <https://doi.org/10.1175/MWR2823.1>.
- Atwood, A. R., E. Wu, D. M. W. Frierson, D. S. Battisti, and J. P. Sachs, 2016: Quantifying climate forcings and feedbacks over the last millennium in the CMIP5–PMIP3 models. *J. Climate*, **29**, 1161–1178, <https://doi.org/10.1175/JCLI-D-15-0063.1>.
- Bender, F. A.-M., A. M. Ekman, and H. Rodhe, 2010: Response to the eruption of Mount Pinatubo in relation to climate sensitivity in the CMIP3 models. *Climate Dyn.*, **35**, 875–886, <https://doi.org/10.1007/s00382-010-0777-3>.

- Byrne, M. P., and P. A. O’Gorman, 2013: Land–ocean warming contrast over a wide range of climates: Convective quasi-equilibrium theory and idealized simulations. *J. Climate*, **26**, 4000–4016, <https://doi.org/10.1175/JCLI-D-12-00262.1>.
- Church, J. A., N. J. White, and J. M. Arblaster, 2005: Significant decadal-scale impact of volcanic eruptions on sea level and ocean heat content. *Nature*, **438**, 74–77, <https://doi.org/10.1038/nature04237>.
- Cole-Dai, J., D. G. Ferris, A. L. Lanciki, J. Savarino, M. H. Thiemens, and J. R. McConnell, 2013: Two likely stratospheric volcanic eruptions in the 1450s C.E. found in a bipolar, sub-annually dated 800 year ice core record. *J. Geophys. Res. Atmos.*, **118**, 7459–7466, <https://doi.org/10.1002/jgrd.50587>.
- Crowley, T. J., 2000: Causes of climate change over the past 1000 years. *Science*, **289**, 270–277, <https://doi.org/10.1126/science.289.5477.270>.
- , and M. B. Unterman, 2013: Technical details concerning development of a 1200 yr proxy index for global volcanism. *Earth Syst. Sci. Data*, **5**, 187–197, <https://doi.org/10.5194/essd-5-187-2013>.
- , G. Zielinski, B. Vinther, R. Udisti, K. Kreutz, J. Cole-Dai, and E. Castellano, 2008: Volcanism and the Little Ice Age. *PAGES News*, No. 2, PAGES International Project Office, Bern, Switzerland, 22–23.
- Donohoe, A., and D. Battisti, 2011: Atmospheric and surface contributions to planetary albedo. *J. Climate*, **24**, 4402–4418, <https://doi.org/10.1175/2011JCLI3946.1>.
- Ferreira, D., J. Marshall, and J.-M. Campin, 2010: Localization of deep water formation: Role of atmospheric moisture transport and geometrical constraints on ocean circulation. *J. Climate*, **23**, 1456–1476, <https://doi.org/10.1175/2009JCLI3197.1>.
- Free, M., and A. Robock, 1999: Global warming in the context of the Little Ice Age. *J. Geophys. Res.*, **104**, 19057–19070, <https://doi.org/10.1029/1999JD900233>.
- Frierson, D. M. W., I. M. Held, and P. Zurita-Gotor, 2006: A gray-radiation aquaplanet moist GCM. Part I: Static stability and eddy scale. *J. Atmos. Sci.*, **63**, 2548–2566, <https://doi.org/10.1175/JAS3753.1>.
- Gent, P., and J. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, [https://doi.org/10.1175/1520-0485\(1990\)020<0150:IMIOCM>2.0.CO;2](https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2).
- Geoffroy, O., D. Saint-Martin, D. J. L. Olivé, A. Voldoire, G. Bellon, and S. Tytéca, 2013: Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J. Climate*, **26**, 1841–1857, <https://doi.org/10.1175/JCLI-D-12-00195.1>.
- Gleckler, P. J., T. M. L. Wigley, B. D. Santer, J. M. Gregory, K. AchutaRao, and K. E. Taylor, 2006: Volcanoes and climate: Krakatoa’s signature persists in the ocean. *Nature*, **439**, 675, <https://doi.org/10.1038/439675a>.
- Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dyn.*, **16**, 501–515, <https://doi.org/10.1007/s003820000059>.
- , and P. M. Forster, 2008: Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.*, **113**, D23105, <https://doi.org/10.1029/2008JD010405>.
- , T. Andrews, P. Good, T. Mauritsen, and P. M. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dyn.*, **47**, 3979–3991, <https://doi.org/10.1007/s00382-016-3055-1>.
- Hansen, J., G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone, 1985: Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, **229**, 857–859, <https://doi.org/10.1126/science.229.4716.857>.
- Hegerl, G. C., T. J. Crowley, S. K. Baum, K. Kim, and W. T. Hyde, 2003: Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophys. Res. Lett.*, **30**, 1242, <https://doi.org/10.1029/2002GL016635>.
- , —, W. T. Hyde, and D. J. Frame, 2006: Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature*, **440**, 1029–1032, <https://doi.org/10.1038/nature04679>.
- Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate*, **23**, 2418–2427, <https://doi.org/10.1175/2009JCLI3466.1>.
- Kara, A. B., P. A. Rochford, and H. E. Hurlburt, 2000: An optimal definition for ocean mixed layer depth. *J. Geophys. Res.*, **105**, 16803–16821, <https://doi.org/10.1029/2000JC900072>.
- Klinger, B. A., and J. Marshall, 1995: Regimes and scaling laws for rotating deep convection in the ocean. *Dyn. Atmos. Oceans*, **21**, 227–256, [https://doi.org/10.1016/0377-0265\(94\)00393-B](https://doi.org/10.1016/0377-0265(94)00393-B).
- Kostov, Y., K. Armour, and J. Marshall, 2014: Impact of the Atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophys. Res. Lett.*, **41**, 2108–2116, <https://doi.org/10.1002/2013GL058998>.
- Lindzen, R. S., and C. Giannitsis, 1998: On the climatic implications of volcanic cooling. *J. Geophys. Res.*, **103**, 5929–5941, <https://doi.org/10.1029/98JD00125>.
- Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997a: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, **102**, 5753–5766, <https://doi.org/10.1029/96JC02775>.
- , C. Hill, L. Perelman, and A. Adcroft, 1997b: Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *J. Geophys. Res.*, **102**, 5733–5752, <https://doi.org/10.1029/96JC02776>.
- Merlis, T. M., I. M. Held, G. L. Stenchikov, F. Zeng, and L. W. Horowitz, 2014: Constraining transient climate sensitivity using coupled climate model simulations of volcanic eruptions. *J. Climate*, **27**, 7781–7795, <https://doi.org/10.1175/JCLI-D-14-00214.1>.
- Mignot, J., M. Khodri, C. Frankignoul, and J. Servonnat, 2011: Volcanic impact on the Atlantic Ocean over the last millennium. *Climate Past*, **7**, 1439–1455, <https://doi.org/10.5194/cp-7-1439-2011>.
- Miller, G. H., and Coauthors, 2012: Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophys. Res. Lett.*, **39**, L02708, <https://doi.org/10.1029/2011GL050168>.
- Neukom, R., and Coauthors, 2014: Inter-hemispheric temperature variability over the past millennium. *Nat. Climate Change*, **4**, 362–367, <https://doi.org/10.1038/nclimate2174>.
- Otterå, O. H., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for Atlantic multidecadal variability. *Nat. Geosci.*, **3**, 688–694, <https://doi.org/10.1038/ngeo955>.
- Redi, M., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158, [https://doi.org/10.1175/1520-0485\(1982\)012<1154:OIMBCR>2.0.CO;2](https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2).
- Robock, A., 2000: Volcanic eruptions and climate. *Rev. Geophys.*, **38**, 191–219, <https://doi.org/10.1029/1998RG000054>.
- Santer, B. D., and Coauthors, 2001: Accounting for the effects of volcanoes and ENSO in comparisons of modeled and observed temperature trends. *J. Geophys. Res.*, **106**, 28 033–28 059, <https://doi.org/10.1029/2000JD000189>.

- Schurer, A. P., G. C. Hegerl, and S. P. Obrochta, 2015: Determining the likelihood of pauses and surges in global warming. *Geophys. Res. Lett.*, **42**, 5974–5982, <https://doi.org/10.1002/2015GL064458>.
- Sigl, M., and Coauthors, 2015: Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, **523**, 543–549, <https://doi.org/10.1038/nature14565>.
- Soden, B. J., R. T. Wetherald, G. L. Stenchikov, and A. Robock, 2002: Global cooling after the eruption of Mount Pinatubo: A test of climate feedback by water vapor. *Science*, **296**, 727–730, <https://doi.org/10.1126/science.296.5568.727>.
- Stenchikov, G., T. L. Delworth, V. Ramaswamy, R. J. Stouffer, A. Wittenberg, and F. Zeng, 2009: Volcanic signals in oceans. *J. Geophys. Res.*, **114**, D16104, <https://doi.org/10.1029/2008JD011673>.
- Stouffer, R. J., 2004: Time scales of climate response. *J. Climate*, **17**, 209–217, [https://doi.org/10.1175/1520-0442\(2004\)017<0209:TSOCR>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0209:TSOCR>2.0.CO;2).
- Tsutsui, J., 2017: Quantification of temperature response to CO₂ forcing in atmosphere–ocean general circulation models. *Climatic Change*, **140**, 287–305, <https://doi.org/10.1007/s10584-016-1832-9>.
- Wigley, T. M. L., C. M. Ammann, B. D. Santer, and S. C. Raper, 2005: Effect of climate sensitivity on the response to volcanic forcing. *J. Geophys. Res.*, **110**, D09107, <https://doi.org/10.1029/2004JD005557>.
- Williams, R. G., J. C. Marshall, and M. A. Spall, 1995: Does Stommel’s mixed layer “demon” work? *J. Phys. Oceanogr.*, **25**, 3089–3102, [https://doi.org/10.1175/1520-0485\(1995\)025<3089:DSMLW>2.0.CO;2](https://doi.org/10.1175/1520-0485(1995)025<3089:DSMLW>2.0.CO;2).
- Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate*, **23**, 2333–2344, <https://doi.org/10.1175/2009JCLI3139.1>.
- Yokohata, T., S. Emori, T. Nozawa, Y. Tsushima, T. Ogura, and M. Kimoto, 2005: Climate response to volcanic forcing: Validation of climate sensitivity of a coupled atmosphere–ocean general circulation model. *Geophys. Res. Lett.*, **32**, L21710, <https://doi.org/10.1029/2005GL023542>.



AMS
American Meteorological Society

Supplemental Material

[© Copyright 2018 American Meteorological Society](#)

Permission to use figures, tables, and brief excerpts from this work in scientific and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is determined to be “fair use” under Section 107 of the U.S. Copyright Act or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108) does not require the AMS’s permission. Republication, systematic reproduction, posting in electronic form, such as on a website or in a searchable database, or other uses of this material, except as exempted by the above statement, requires written permission or a license from the AMS. All AMS journals and monograph publications are registered with the Copyright Clearance Center (<http://www.copyright.com>). Questions about permission to use materials for which AMS holds the copyright can also be directed to permissions@ametsoc.org. Additional details are provided in the AMS Copyright Policy statement, available on the AMS website (<http://www.ametsoc.org/CopyrightInformation>).

Supplementary Information

A. The 2-box model solution

We solve the 2-box model equations, Eq. (1) and (2) repeated below, for a delta function forcing $F(t) = V\delta(t)$ where V is the integrated amount of energy instantaneously extracted from the system.

$$\rho c_w h_1 \frac{dT_1}{dt} = -\lambda T_1 + q(T_2 - T_1) + V\delta(t). \quad (1)$$

$$\rho c_w h_2 \frac{dT_2}{dt} = q(T_1 - T_2). \quad (2)$$

We start by first normalizing the parameters λ , q and F as follows:

$$\begin{bmatrix} \lambda_0 \\ q_0 \\ V_0 \end{bmatrix} = \frac{1}{\rho c_w h_1} \begin{bmatrix} \lambda \\ q \\ V \end{bmatrix}. \quad (S1)$$

We also define the parameter r which represents the ratio between the heat capacities (and depths) of the top and bottom boxes

$$r = h_1/h_2. \quad (S2)$$

Eq. (1) and (2) can then be written in matrix form as follows:

$$\begin{bmatrix} \dot{T}_1(t) \\ \dot{T}_2(t) \end{bmatrix} = \begin{bmatrix} -q_0 - \lambda_0 & q_0 \\ r q_0 & -r q_0 \end{bmatrix} \begin{bmatrix} T_1(t) \\ T_2(t) \end{bmatrix} + \begin{bmatrix} V_0 \delta(t) \\ 0 \end{bmatrix} \quad (S3)$$

The homogenous ($V_0 = 0$) solution of Eq. (S3) is of the form:

$$T_1(t) = T_{1f} e^{-t/\tau_f} + T_{1s} e^{-t/\tau_s}, \quad (S4a)$$

$$T_2(t) = T_{2f} e^{-t/\tau_f} + T_{2s} e^{-t/\tau_s} \quad (S4b)$$

where $1/\tau_f$ and $1/\tau_s$ are the inverses of the fast and slow timescales respectively and represent the eigenvalues of the above matrix. The eigenvalues satisfy the following characteristic equation:

$$(1/\tau_{f,s})^2 + (q_0 r + \lambda_0 + q_0)(1/\tau_{f,s}) + r q_0 \lambda_0 = 0. \quad (\text{S5})$$

For convenience we define α_0 and γ_0 as:

$$\alpha_0 = q_0(1 + r) + \lambda_0, \quad (\text{S6})$$

$$\gamma_0 = r q_0 \lambda_0 \quad (\text{S7})$$

Solving Eq. (S5) for $1/\tau_{f,s}$ then gives:

$$\frac{1}{\tau_f} = \frac{\alpha_0 + \sqrt{\alpha_0^2 - 4\gamma_0}}{2}, \quad (\text{S8a}) \quad \frac{1}{\tau_s} = \frac{\alpha_0 - \sqrt{\alpha_0^2 - 4\gamma_0}}{2}. \quad (\text{S8b})$$

To find the constants T_f and T_s we impose the following initial conditions:

$$\dot{T}_1(0) = V_0 \delta(t), \quad (\text{S9})$$

$$T_2(0) = 0. \quad (\text{S10})$$

Eq. (S9) states that all the impulse forcing is delivered to the mixed layer box. Eq. (S10) states that the temperature of the deep ocean box is initially zero.

Integrating Eq. (S9) in time gives, using Eq. (S4a):

$$T_f + T_s = V_0, \quad (\text{S11})$$

and Eq. (S10) implies:

$$T_{f2} + T_{s2} = 0. \quad (\text{S12})$$

We now seek a relationship between T_f , T_s , T_{f2} and T_{s2} . Substituting Eq. (S4a) and (S4b) into Eq. (2), gives:

$$T_{f2} = \frac{r q_0}{r q_0 - 1/\tau_f} T_f = \alpha_1 T_f, \quad (\text{S13})$$

$$T_{s2} = \frac{r q_0}{r q_0 - 1/\tau_s} T_s = \alpha_2 T_s. \quad (\text{S14})$$

Now using Eq. (S13) and (S14) with boundary conditions Eq. (S11) and (S12) yields:

$$T_f = \frac{\alpha_2}{\alpha_2 - \alpha_1} V_0, \quad (S15a) \quad T_s = \frac{-\alpha_1}{\alpha_2 - \alpha_1} V_0, \quad (S15b)$$

$$T_{f2} = \frac{\alpha_1 \alpha_2}{\alpha_2 - \alpha_1} V_0, \quad (S16a) \quad T_{s2} = \frac{-\alpha_1 \alpha_2}{\alpha_2 - \alpha_1} V_0. \quad (S16b)$$

Limiting cases

We now explore approximate forms of Eq. (S8) and Eq. (S15) in the limiting case that r is small.

(a) The fast timescale τ_f

Writing out Eq. (S8a) gives:

$$\frac{2}{\tau_f} = (1+r)q_0 + \lambda_0 + \sqrt{q_0^2(1+r)^2 + 2q_0\lambda_0(1+r) + \lambda_0^2 - 4rq_0\lambda_0}. \quad (S17)$$

In the limit that $r \ll 1$, Eq. (S17) readily simplifies to:

$$\tau_f \approx \frac{1}{\lambda_0 + q_0}. \quad (S18)$$

Eq. (S18) is equivalent to Eq. (7a) in the main text. It is appropriate when $T_2 = 0$ and $V_0 = 0$ in Eq. (1), which then reduces to:

$$\frac{dT_1}{dt} \approx -\frac{\lambda + q}{\rho c_w h_1} T_1, \quad (S19)$$

supporting a decaying solution with the timescale given by Eq. (S18), with λ_0 and q_0 defined in Eq. (S1).

(b) The slow timescale τ_s

Writing out Eq. (S8b) gives:

$$\frac{2}{\tau_s} = (1+r)q_0 + \lambda_0 - \sqrt{q_0^2(1+r)^2 + 2q_0\lambda_0(1+r) + \lambda_0^2 - 4rq_0\lambda_0}, \quad (S20)$$

which in the limit $r \ll 1$ simplifies to:

$$\frac{1}{\tau_s} \approx \frac{rq_0\lambda_0}{q_0 + \lambda_0} \quad (\text{S21})$$

This approximate form of the long timescale can be arrived at by setting $r \ll 1$ in Eq. (S3), giving:

$$T_1 \approx \frac{q_0 T_2}{\lambda_0 + q_0}. \quad (\text{S22})$$

Using Eq. (S22) in Eq. (2) then gives:

$$\frac{dT_2}{dt} \approx -\frac{T_2}{\rho c_w h_2} \frac{\lambda q}{\lambda + q}, \quad (\text{S23})$$

which supports decaying solutions with a timescale given by Eq. (S21), with λ_0 and q_0 defined in Eq. (S1).

(c) The coefficients T_f and T_s

Simplified expressions for the coefficients T_f and T_s can be obtained by assuming $r \ll 1$ in Eq. (S15). Using the timescales given by Eq. (S18) and (S21) approximate forms for the parameters α_1 and α_2 become:

$$\alpha_1 \approx \frac{-rq_0}{q_0 + \lambda_0}, \quad (\text{S24a}) \quad \alpha_2 \approx \frac{q_0 + \lambda_0}{q_0} \quad (\text{S24b})$$

Substituting Eq. (S24) in to Eq. (S15) gives:

$$T_f \approx \frac{(\lambda + q)^2}{(\lambda + q)^2 + rq^2} V_0, \quad (\text{S25a}) \quad T_s \approx \frac{rq^2}{(\lambda + q)^2 + rq^2} V_0. \quad (\text{S25b})$$

Peak cooling

When the forcing is a perfect impulse, all the energy is delivered to the mixed layer. This gives:

$$T_c = \frac{V}{\rho c_w h_1} = V_0. \quad (4)$$

However, if the forcing lasts a small but finite time Δt , some of that cooling is also sequestered into the deeper ocean as well as being damped by atmospheric feedbacks. By assuming that Δt is small enough that throughout the duration of the forcing the deep layer temperature is small ($T_2 \ll T_1$), we can use Eq. (15), repeated below:

$$\rho c_w h_1 \frac{dT_1}{dt} \approx -(\lambda + q)T_1 + F(t). \quad (15)$$

Integrating Eq. (15) for the duration of the forcing gives:

$$\rho c_w h_1 T_c \approx -(\lambda + q) \int_0^{\Delta t} T_1 dt + V, \quad (S26)$$

where V is defined in Eq. (5) in the main text.

Approximating the integral by:

$$\int_0^{\Delta t} T_1 dt \approx \frac{\Delta t T_c}{2} \quad (S27)$$

yields the following estimate of the peak cooling due to a forcing of duration Δt :

$$T_c \approx \frac{V}{\rho c_w h_1 + \frac{\lambda \Delta t (1 + \mu)}{2}}. \quad (13)$$

B. Accumulated response from uniform & regular eruptions

Here we derive Eq. (18) and (19) in the main text. We consider a series of uniform eruptions spaced at a regular interval τ . Each eruption has a response of the form given by Eq. (3) and we assume the system to be perfectly linear. Our aim is to obtain an expression for the envelope of the response, as plotted in Figure S1. It is easiest to first consider the problem in a discrete framework and obtain the temperature of the mixed layer T_1 after each eruption. For simplicity, we start with the response consisting of a single timescale τ_m and a peak cooling T_c . We then extend our method for the two-timescale response.

With a single timescale response, the peak response after N eruptions can be written as follows:

$$T_N = T_c \left(1 + e^{-\frac{\tau}{t_m}} + e^{-\frac{2\tau}{t_m}} + \dots + e^{-\frac{N\tau}{t_m}} \right). \quad (\text{S28})$$

This is a geometric series which when evaluated yields:

$$T_N = T_c \frac{1 - e^{-(N+1)\tau/t_m}}{1 - e^{-\tau/t_m}}. \quad (\text{S29})$$

We can obtain a continuous expression by setting:

$$t = N\tau \quad (\text{S30})$$

and so Eq.(29) becomes

$$T_{en}(t) = T_c \frac{1 - e^{-(t+\tau)/t_m}}{1 - e^{-\tau/t_m}}. \quad (\text{S31})$$

In the two timescale case, we can think of Eq. (3) as being composed of two independent responses with timescales τ_f and τ_s , and peaks T_f and T_s , respectively. Assuming linearity, we sum their respective envelopes to give:

$$T_{en}(t) = T_f \frac{1 - e^{-(t+\tau)/\tau_f}}{1 - e^{-\tau/\tau_f}} + T_s \frac{1 - e^{-(t+\tau)/\tau_s}}{1 - e^{-\tau/\tau_s}}. \quad (19)$$

Then letting $t \rightarrow \infty$ in Eq. (19) gives:

$$T_{\infty} = \frac{T_f}{1 - e^{-\tau/\tau_f}} + \frac{T_s}{1 - e^{-\tau/\tau_s}}. \quad (20)$$

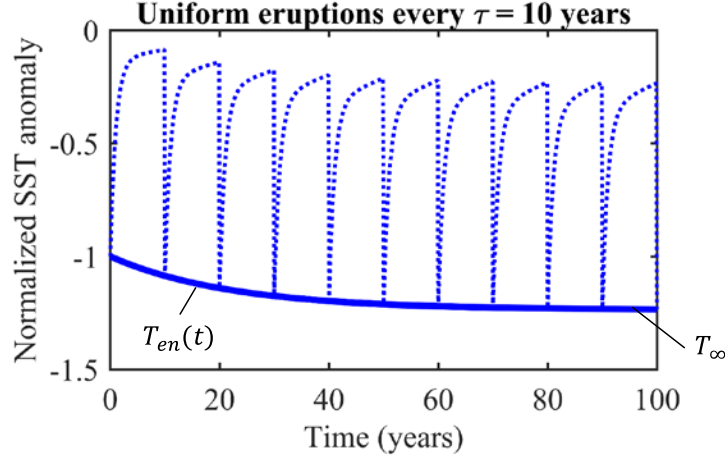


Figure S1: Response of the 2-box model to a series of uniform eruptions spaced regularly every 10 years (dashed line). The solid line represents the envelope of the response T_{en} derived analytically in Eq. (13). The envelope asymptotes to the value T_{∞} .

As shown in Figure S1, the response approaches a cyclostationary equilibrium after several iterations, with the response pattern repeating at interval τ . In this equilibrium state, the amount of energy provided by the impulse in a given cycle must equal the energy taken out by the climatic feedbacks over that cycle. This can be expressed as follows:

$$\lambda \int_x^{x+\tau} T_1(t) dt = V, \quad (S32)$$

where x is a time long enough for the cyclostationary equilibrium to have developed and V is the energy extracted by a volcanic eruption persisting for one year. This then gives the equilibrium cycle average temperature:

$$\bar{T}_{\infty} = \frac{V}{\lambda\tau}. \quad (21)$$