

Triggering global climate transitions through volcanic eruptions

Article

Accepted Version

Gupta, M., Marshall, J. and Ferreira, D. ORCID: https://orcid.org/0000-0003-3243-9774 (2019) Triggering global climate transitions through volcanic eruptions. Journal of Climate, 32 (12). pp. 3727-3742. ISSN 1520-0442 doi: 10.1175/JCLI-D-18-0883.1 Available at https://centaur.reading.ac.uk/83783/

It is advisable to refer to the publisher's version if you intend to cite from the work. See <u>Guidance on citing</u>.

To link to this article DOI: http://dx.doi.org/10.1175/JCLI-D-18-0883.1

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the <u>End User Agreement</u>.

www.reading.ac.uk/centaur

CentAUR



Central Archive at the University of Reading

Reading's research outputs online

1	Triggering global climate transitions through volcanic eruptions
2	Mukund Gupta *, John Marshall [†] , David Ferreira [‡]
3	Massachusetts Institute of Technology, Cambridge, MA/USA

⁴ *Corresponding author address: Mukund Gupta, Department of Earth, Atmosphere and Planetary

- ⁵ Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA/USA
- 6 E-mail: guptam@mit.edu
- ⁷ [†]Department of Earth, Atmosphere and Planetary Sciences, Massachusetts Institute of Technology,
- ⁸ 77 Massachusetts Avenue, Cambridge, MA/USA
- ³ [‡]Department of Meteorology, University of Reading, Reading, UK

ABSTRACT

A coupled climate model with idealized representations of atmosphere, 10 ocean, sea ice and land is used to investigate transitions between global cli-11 mate equilibria. The model supports the presence of climates with limited ice 12 cover (Warm), a continuum of climates in which sea ice extends down into the 13 mid-latitudes and the tropics (Cold), together with a completely ice-covered 14 earth (Snowball). Transitions between these states are triggered through vol-15 canic eruptions, where the radiative effect of stratospheric sulfur emissions is 16 idealized as a 1-year impulse reduction in incoming solar radiation. Snow-17 ball transitions starting from the Cold state are more favorable than from the 18 Warm state, because less energy must be extracted from the system. However, 19 even when starting from a Cold climate, Toba-like volcanic events (cooling 20 of order -100 Wm^{-2}) must be sustained continuously for several decades to 2 glaciate the entire planet. When the deep ocean is involved, the volcanic re-22 sponse is characterized by relaxation timescales spanning hundreds to thou-23 sands of years. If the interval between successive eruptions is significantly 24 shorter (years to decades) than the ocean's characteristic timescales, the cu-25 mulative cooling can build over time and initiate a state transition. The model 26 exhibits a hysteresis loop in which the Cold equilibrium branch is not directly 27 accessible from the Snowball state, but when starting from a Cold climate, the 28 system can transition back to an ice-free state without hysteresis. 29

30 1. Introduction

Paleoclimate reconstructions have shown that Earth has experienced a wide array of climates 31 over the last billion years. At its most extreme, the climate has ranged between hothouse worlds 32 with little or no ice at the poles and deep ocean temperatures upwards of 20° C (e.g. Tripati and 33 Elderfield (2005)), to Snowball periods in which ice covered most of the Earth's surface (e.g. 34 Hoffman et al. (1998)). The ability of the climate to exist in a number of dynamical states was 35 discussed by Stommel (1961) with regards to the bi-stability of the oceanic meridional overturning 36 circulation (MOC). This work suggested that the MOC could exhibit two stable states, namely a 37 vigorous circulation mode driven by temperature and a weak mode controlled by salinity. Another 38 source of instability in the climate is the ice-albedo feedback, which was first investigated by 39 Budyko (1969) and Sellers (1969) in simple energy balance models (EBMs). These studies found 40 that the Earth could exist in a stable state with either a small amount of ice at the poles or with 41 complete ice cover. Rose and Marshall (2009) modified the EBM to demonstrate that an additional 42 state with a large ice cap can be stabilized by the meridional convergence of heat transported by 43 the ocean subtropical cells. Indeed several General Circulation Models (GCMs) have shown that 44 climates with ice extending down to the mid-latitudes or the tropics are realizable when ocean heat 45 transport (OHT) can arrest the advance of sea ice at those latitudes (e.g. Poulsen and Jacob (2004), 46 Langen (2004), Ferreira et al. (2011) and Rose et al. (2013)) or when the meridional gradient in 47 ice albedo can limit the strength of the positive ice-albedo feedback (Abbot et al. (2011)). 48

Transitions between these multiple equilibria require global-scale forcings that push the climate away from its original equilibrium to the next stable state. Many of the long-lasting climatic shifts that occurred in the last billion years were governed by the balance between weathering of silicate rocks and the input of carbon dioxide by volcanic eruptions in the atmosphere, which typically spans $10^5 - 10^6$ year timescales (Walker et al. (1981)). However, large perturbations in the atmosphere's radiation budget acting over much shorter (~ decadal) timescales could have also caused severe episodes of climate change. Perhaps one of the largest forcings ever experienced by the Earth's climate was the Chicxulub asteroid impact 66 millions years ago, during which the amount of dust and sulfate aerosols in the atmosphere could have blocked more than half of the solar radiation reaching the Earth's surface for several years, subsequently leading to extreme cooling and the beginning of a mass extinction (Brugger et al. (2017), Kaiho and Oshima (2017)).

The volcanic emission of sulfur particles into the stratosphere is also known to cool the sur-60 face of the Earth for several years after the eruption (Robock (2000)). Over the last few thousand 61 years, these volcanic forcings were too weak and short-lived to have caused a drastic shift in cli-62 matic state. However, there have been recorded instances of volcanic activity much larger than 63 that of the Holocene, which may have triggered a significant glaciation (e.g. Rampino and Self 64 (1992), Prueher and Rea (2001)). The Toba super-eruption around 73.5 ka before present had a 65 sulfur loading that was two orders of magnitude larger than the 1991 Mt Pinatubo eruption, and 66 could have reduced incoming solar radiation by a third for several years after the event (Timm-67 reck et al. (2010)). While GCM simulations with these types of forcings did not yield widespread 68 glaciation (Jones et al. (2005)), some authors have argued that a sequence of super-eruptions sep-69 arated by a short time interval could eject enough sulfur particles into the stratosphere to trigger 70 a significant shift in Earth's history, such as a mass extinction (Baresel et al. (2017)) or a Snow-71 ball Earth (Stern et al. (2008), Macdonald and Wordsworth (2017)). In particular, Macdonald and 72 Wordsworth (2017) argue that starting from the cool Neoproterozoic background climate, a suc-73 cession of super-eruptions occurring over several decades could have initiated the Snowball event 74 of the Sturtian (717 - 616 Ma). While this estimate was obtained using a single-column climate 75

⁷⁶ model with a heat capacity corresponding to the mixed layer depth of the ocean (~ 50 m), Voigt ⁷⁷ and Marotzke (2010) and Voigt et al. (2011) argue that the entire ocean must cool to the freezing ⁷⁸ point to initiate a Snowball climate. In the present study, we explore the extent to which the ocean ⁷⁹ constrains the forcing magnitude and timescale required for global climate transitions using a cou-⁸⁰ pled Atmosphere-Ocean-Ice GCM (which we refer to here as the 'climate model'), and explore ⁸¹ whether volcanic eruptions could successfully initiate a drastic climatic state change, such as a ⁸² Snowball transition.

The presence of multiple equilibria in the global climate can also affect the transition behavior 83 of the system. While Voigt and Marotzke (2010) did not find intermediate stable states between 84 their reference 'warm' climate and the Snowball, the following work of Voigt et al. (2011), Voigt 85 and Abbot (2012), Yang et al. (2012), and others did find such states. The GCM used in this study 86 allows the existence of a continuum of climates with sea ice extending from the mid-latitudes to 87 the tropics (45° to 25° latitude), owing to the stabilizing nature of meridional heat transport by 88 subtropical oceanic cells (e.g. Rose and Marshall (2009), Ferreira et al. (2011)). We initiate tran-89 sitions starting from these different states and evaluate how the proximity to a transition threshold 90 affects the timescale required to tip the climate to a new state. Moreover, a number of studies have 91 suggested that a Waterbelt (or Slushball) climate with ice extending down to the tropics could 92 better explain the survival of oceanic life during the Sturtian (717 - 616 Ma) and Marinoan (650 -93 635 Ma) glaciations events, as opposed to a complete Snowball (e.g. Hoffman et al. (2017)). The 94 climate model used in this study shows that a Waterbelt-like climate may occur for a significantly 95 smaller threshold forcing than that required for complete glaciation, but unlike the Snowball, this 96 state exhibits no hysteresis when the forcing is removed. 97

⁹⁸ Our paper is set out as follows: Section 2 introduces the coupled climate model; Section 3 ⁹⁹ presents its equilibrium states. Section 4 focuses on the transition behavior, first from step and ¹⁰⁰ then from impulse volcanic forcings. Section 5 concludes.

2. The coupled climate model and modeling framework

Numerical experiments are carried out with the coupled atmosphere, ocean, sea ice and land 102 model based on the MITgcm (Marshall et al. (1997a), Marshall et al. (1997b)). The particular 103 setup of the model used here comprises a 'boomerang' configuration, which consists of two 45° 104 wide flat continents separated by a 90° angle such that the ocean is split into a narrow Atlantic-105 like basin and a wide Pacific-like basin connected to an unblocked 'Southern Ocean' in the South 106 (see Ferreira et al. (2018)). The ocean is flat-bottomed with a uniform depth of 3000 m. The 107 atmospheric scheme is based on the simplified parameterizations of primitive-equation dynamics 108 (SPEEDY) from Molteni (2003) at five level vertical resolution. It includes a four-band radiation 109 scheme, a parameterization of moist convection, diagnostic clouds and a boundary layer scheme. 110 In the reference states, the solar constant (divided by 4) is 341.5 Wm^{-2} . Despite the idealized 111 nature of the model, it captures many of the essential large-scale features of the climate such as an 112 overturning circulation dominated by the small basin, atmospheric storm tracks and hydrological 113 cycle, gyres and a circumpolar current in the oceans, and a seasonal sea ice cycle. Additional 114 modelling details are provided in the Supplementary Information (SI). 115

Perturbation experiments are conducted by reducing the TOA insolation. The forcing F is diagnosed as the net radiative imbalance at the TOA, averaged over the first year of the simulation. For step forcing experiments, the insolation is changed abruptly at t = 0 and subsequently held constant in time. For volcanic forcing experiments, we crudely emulate the effects of stratospheric aerosol injection by uniformly reducing the insolation for one year and subsequently turning the
 forcing off (as in Gupta and Marshall (2018)). We also conduct simulations comprising successive
 eruptions separated by a specified time interval.

3. Stable climate states of a coupled climate model

124 a. Equilibrium solutions

Figure 1 shows the three equilibrium branches found in our model: Warm, Cold and Snowball. 125 Note that we reserve the term 'state' for individual points on a given equilibrium branch. The 126 Warm reference state has some sea ice in the Southern hemisphere (SH) but none in the Northern 127 hemisphere (NH). It has a globally-averaged ocean potential temperature $\theta = 8.3$ °C and a global 128 mean sea surface temperature (SST) of 19.6 °C. The Cold equilibrium consists of a continuum of 129 states with sea ice latitude ranging from 45° to 25° and θ between 0.5 and -1.7 °C. The mildest 130 states on this Cold branch have an ice cap extending down to the mid-latitudes, which is remi-131 niscent of the glacial climate of the Pleistocene (Ferreira et al. (2018)). The most extreme states 132 on the Cold branch have sea ice reaching down to the tropics, which is analogous to Waterbelt or 133 Slushball states found by other studies (e.g. Rose (2015) and Abbot et al. (2011)). Since these 134 Waterbelt-like climates are plausible analogs for the state of the Earth during the Neoproterozoic 135 glaciations, it is instructive to treat them somewhat separately than the rest of the Cold branch, 136 particularly in the context of state transitions. In what follows, we thus choose to refer to Cold 137 states with $\theta < -1.5$ °C as Waterbelts, and others simply as Cold. Our discussion does not depend 138 strongly on this threshold, as long as it is chosen close enough to the freezing point. Finally, the 139 Snowball is fully ice covered with $\theta = -1.8$ °C (the mean freezing point of the ocean). Simulations 140

that reach this state are stopped when the whole ocean reaches the freezing temperature, to avoid runaway ice growth in the absence of geothermal heat flux.

Figure 1 also presents the various equilibrium states in terms of their energy levels E, approximated by the sum of ocean heat content and latent energy stored in the sea ice, all measured relative to the Warm state.

$$E = \rho c H_{tot}(\theta - \theta_w) + A^{-1} L_f(M - M_w) \approx \rho c H_{tot}(\theta - \theta_w), \tag{1}$$

where ρ is the sea water density, c is the thermal heat capacity of water, H_{tot} is the globally area-146 averaged ocean depth, θ is the ocean mean potential temperature, A is the Earth's surface area, 147 L_f is the latent heat capacity of ice, M is the total mass of ice and the subscript 'w' indicates the 148 Warm reference state. The energy stored in the ice only represents at most 10% of E. Therefore, θ 149 and E can be used almost equivalently to describe the energy content of the system. E is an energy 150 per unit area, which we express in $[Wm^{-2}yr]$ (= $3.1 \times 10^7 Jm^{-2}$) to aid the comparison with the 151 energy extracted by volcanic eruptions. For reference, the 1991 Mount Pinatubo eruption imposed 152 a globally-averaged forcing of approximately -4 Wm^{-2} at the TOA for approximately one year, 153 thus equivalent to an energy extraction of $-4 Wm^{-2}yr$. 154

Figure 1 shows that the Warm and Cold reference states have a large energetic gap (~ 2500 $Wm^{-2}yr$), which is a significant barrier to overcome for any transition between them. The energy gap between the Cold reference state and Waterbelt climates is smaller (~ 700 $Wm^{-2}yr$) but significant enough to distinguish them energetically, owing to the existence of a tropical thermocline in the former, but not in the latter. On the other hand, the Waterbelt and Snowball states are not energetically distinguishable, because most of the ocean is at freezing temperature in both the
 states.

Figure 2 shows cross-sections of ocean potential temperature and overturning streamfunction, 162 along with the sea ice edge latitude and ocean heat transport (OHT) for various stable states of 163 our model. The shallow subtropical cells (STCs) are driven by the trade winds, which act upon 164 the temperature gradient of the thermocline to transport heat to the mid-latitudes. This heat flux 165 allows the stabilization of the sea ice edge and facilitates the realization of multiple equilibria in 166 the model (see Rose and Marshall (2009), Ferreira et al. (2011) and Rose (2015)). In the Warm 167 state, the meridional heat transport by the oceans reaches high latitudes in the NH, but only mid-168 latitudes in the SH. The temperature of the deep ocean (below 1000 m) is everywhere above 6 169 $^{\circ}$ C, including at the poles where fresh surface waters allow a temperature inversion with a surface 170 colder than the abyss. In the Cold state, all the waters below 1200 m and polewards of 50° latitude 171 in both hemispheres ($\sim 50\%$ of the total ocean volume) are at the freezing point. This results from 172 temperature-induced convection having to bring the entire water column to the freezing point 173 before ice can grow at the surface (see Rose et al. (2013)). The thermocline, however, remains 174 well stratified, supporting poleward heat transport by the subtropical cells, which arrests the sea 175 ice in the mid-latitudes. The maximum strength of poleward heat transport in both the Warm and 176 Cold states is order 2 PW, peaking around 20° latitude. In the Waterbelt state, the thermocline 177 stratification is eroded and the sea ice extends further equatorward, until only a shallow (~ 300 178 m) surface layer of the tropical ocean remains above freezing and the ice edge reaches around 25° 179 latitude in both hemispheres. The strength of poleward heat transport falls to 1.2 PW and the peak 180 moves equatorward to 10° latitude. Finally, in the Snowball, OHT is essentially reduced to zero 181

and the entire ocean is at the freezing point. In the following section, we explore the timescales
 associated with changes between these states.

¹⁸⁴ b. Triggering transitions between equilibrium states

In this section, we conduct step (constant) forcing simulations in our climate model with the 185 aim of initiating transitions between the various climate states presented in Section 3(a). Figure 186 3 shows these results on a bifurcation diagram, where each scatter represents the NH sea ice 187 latitude in the equilibrated state of a given simulation with forcing F. This forcing magnitude is 188 expressed with respect to the reference solar constant $S_0 = 341.5 Wm^{-2}$. Simulations start from 189 either the Warm (red circles), Cold (blue squares), Waterbelt (green triangle) or Snowball (black 190 diamond) states. The solid lines describe the range over which these states exist, with the same 191 color convention. The individual states labelled 'W' (Warm), 'C' (Cold) and 'S' (Snowball) are 192 chosen as reference climates because they exist for F = 0. 193

194 (i) Warm start

Starting from 'W', a small step forcing of -1.5 Wm^{-2} induces a slow transition to the Cold 195 equilibrium over several thousand years. This implies that the Warm reference state is at the edge 196 of a critical threshold and is thus particularly sensitive to changes in its radiative budget. Larger 197 forcings of -7 Wm^{-2} and -11 Wm^{-2} both lead to a Cold equilibrium and a forcing of -36 Wm^{-2} 198 produces a Snowball. We also force the Warm reference state with a positive forcing of $+3 Wm^{-2}$, 199 which leads to a Very Warm ('VW') climate that has no ice anywhere on the globe, a global 200 mean ocean potential temperature of 14 °C and a global mean SST of 24 °C. We find no stable 201 states between 90° and 45° NH sea ice extent. The absence of a stable climate with a small ice 202 cap, known as the small ice instability (SICI), has been discussed in the context of simple EBMs 203

(e.g. Held and Suarez (1975) and North (1985)) and more sophisticated models (e.g. Huang
and Bowman (1992), Matteucci (1993), Lee and North (1995), Morales Maqueda et al. (1998)
and Langen (2004)). While these studies concluded that the SICI tends to disappear with better
representation of continental geometry, unforced variability, ice sheets and ice dynamics, further
work would be required to investigate the lack of a small ice cap climate in the model used here.

209 (ii) Cold start

Starting from 'C', we find that step forcings of -2, -3, -5 and -9 Wm^{-2} lead to progressively 210 colder climates without abrupt state transitions. For $F_{c/s} = -20 Wm^{-2}$, the climate initially transi-211 tions to a Cold climate that remains stable for about 150 years, but then collapses into a Snowball. 212 Between $F_{wb} = -5 Wm^{-2}$ and $F_{c/s}$, we find Cold climates with $\theta < -1.5$ °C and tropical sea ice, 213 which we refer to as Waterbelts. Starting from such a Waterbelt climate ('WB') and relaxing the 214 forcing to 0, we find that the system returns to 'C'. Finally, we conduct warming experiments 215 starting from 'C' with F = +1.5 and +8 Wm^{-2} , which lead back to a Warm climate. We conclude 216 that there is no identifiable hysteresis loop between the Cold and Warm branches. 217

218 (iii) Snowball start

Starting from a Snowball, we relax F to 0 and find that the climate remains in a Snowball state 219 ('S'). Thus, the Cold branch is not directly accessible from a Snowball state. Starting from 'S', 220 we attempt to exit the Snowball state by imposing a forcing of +7 Wm^{-2} , but find that the surface 221 remains frozen everywhere. We expect that a larger positive forcing could overcome the strong ice-222 albedo feedback and bring the system back to the Warm branch, but the magnitude of this forcing 223 is not quantified in this study. Note that if this threshold large enough however, the climate may 224 not stabilize into a Warm state, and instead transition directly to a moist or runaway greenhouse 225 (Yang et al. (2017)). 226

The plot on the right of Figure 3 summarizes these results in a simplified schematic (not to scale). 227 The arrows show two hypothetical paths along the bifurcation diagram. Starting from Warm (red 228 arrows) and slowly reducing the solar constant, the system would reach the Warm/Cold transition 229 threshold $F_{w/c} \sim -1.5 Wm^{-2}$ and move down to the Cold branch. The system would then smoothly 230 transition to a Waterbelt-like climate around $F_{wb} \sim -5 Wm^{-2}$, before reaching the Cold/Snowball 231 threshold $F_{c/s} \sim -20 \ Wm^{-2}$. From a Snowball state, a large increase in the solar constant would 232 be required to switch back to the Warm state and complete the loop. When starting from the Cold 233 branch (blue arrows) and increasing the solar constant, the system can move to the Warm branch 234 without any hysteresis. In the Appendix, we further interpret this bifurcation diagram in terms 235 of a 1D EBM developed by Rose (2015), which shows characteristics that are reminiscent of our 236 climate model. 237

4. Transition timescales between equilibrium states

a. Transitions in response to step forcings

Figure 4 shows state transition times in the climate model for the step simulations starting from 240 the Warm (red) and the Cold (blue) reference states as a function of the forcing magnitude. These 241 are the same experiments as used in Figure 3. State transitions to both a Waterbelt climate and 242 a Snowball are considered. As expected, the transition time shortens as the forcing magnitude 243 increases, and there is a minimum threshold forcing below which no transition occurs. Transitions 244 from the Warm state take longer than from Cold (for the same forcing) because more energy must 245 be extracted in the process. Voigt and Marotzke (2010) (subsequently VM2010) argued that as 246 the planet approaches a Snowball, deep convection and the transport of cold waters into the abyss 247

eventually lead to a well-mixed ocean. This prompted them to adopt the following 1-box model
 (here expressed in terms of anomalies) to estimate state transition times:

$$\rho c H \frac{dT(t)}{dt} = F(t) - \lambda T(t), \qquad (2)$$

where F is the TOA forcing, T is the temperature anomaly relative to the starting state, λ is the climate feedback parameter and H is the ocean depth relevant to the transition. The model does not take account of the change in albedo as the transition occurs since all parameters are constant in time. The solution for the step response (constant F) of Eq. (2) is:

$$T(t) = T_{eq}(1 - e^{-t/\tau})$$
(3)

where the system timescale τ is given by:

$$\tau = \frac{\rho c H}{\lambda} \tag{4}$$

²⁵⁵ and the equilibrium climate sensitivity is

$$T_{eq} = \frac{F}{\lambda},\tag{5}$$

²⁵⁶ where here F is negative, corresponding to extraction of energy.

As we now explore, state transition times in our climate model are well captured by the 1box model, when considering the difference in potential temperature between the initial and final states $\Delta\theta$. The time taken by the 1-box model to achieve a temperature difference $\Delta\theta$ through a step forcing F is obtained from Eq. (3) as follows:

$$t_f = \tau \times ln \Big(\frac{T_{eq}}{T_{eq} - \Delta \theta} \Big), \tag{6}$$

From Eq. (5), the forcing required to obtain a temperature response $\Delta \theta$ at equilibrium is:

$$F_{min} = \lambda \Delta \theta, \tag{7}$$

It follows that forcings weaker than F_{min} cannot achieve this temperature change, even when sustained indefinitely. Therefore, F must be larger than or equal to F_{min} to achieve a temperature change $\Delta\theta$.

One can also estimate the efficiency of energy transfer required to achieve $\Delta \theta$, as a function of forcing magnitude. The total energy E_{tot} supplied by the forcing during the transition is:

$$\Delta E_{tot} = F \times t_f,$$

where we ignore the energy stored in the atmosphere, land and sea ice. From Eq. (1), the net energy ΔE_{net} extracted from the system for a temperature change $\Delta \theta$ is:

$$\Delta E_{net} = \rho c H \Delta \theta$$
.

The efficiency of energy extraction ε is then:

$$\varepsilon = \frac{\Delta E_{net}}{\Delta E_{tot}} = \frac{\Delta \theta}{T_{eq}} ln \Big(\frac{T_{eq}}{T_{eq} - \Delta \theta} \Big)^{-1}.$$

As the forcing amplitude increases (F $\rightarrow -\infty$), the efficiency ε tends to 1, but drops to 0 as F tends toward F_{min} . For large forcings, climatic feedbacks have limited time to act, and hence ²⁷² the energy extraction occurs efficiently. Conversely, as the forcing magnitude decreases, the cli-²⁷³ matic feedbacks have enough time to counter the effect of the forcing, and as F drops below F_{min} , ²⁷⁴ the feedbacks become too strong to allow a temperature change $\Delta\theta$ altogether. In the following ²⁷⁵ paragraphs, we investigate to what extent these simple ideas apply to the climate model transitions.

For simulations starting from Warm, the temperature anomalies spread into every part of the 276 ocean, including the abyss. Thus, the relevant depth H of the 1-box model is $H_{tot} = 2416$ m (the 277 total area-averaged depth of the ocean). The climate sensitivity parameter in the Warm reference 278 state is found using the method outlined in Gregory et al. (2004) (details in SI), which gives $\lambda_w =$ 279 $0.72 Wm^{-2}K^{-1}$. This value is associated with the SST response of the model, but the assumption 280 of a well-mixed ocean allows us to extend it to the ocean-mean temperature. For a transition to 281 a Snowball, the initial temperature is $\theta_i = 8.3^\circ$ C and the final temperature is $\theta_f = -1.8^\circ$ C, the 282 freezing point of water. Inserting these values into Eq. (6) gives the red curve in Figure 4, which 283 is a reasonable approximation to the coupled model results. 284

Figure 4 also shows (in orange squares) the Snowball transition times obtained by VM2010, when starting from a modern-day climate in a comprehensive GCM. To facilitate comparison with our simulations starting from Warm, we scale the VM2010 timescales by a factor f_{vm} , to account for the different heat capacities of the two models:

$$f_{vm} = \frac{H_{tot}(\theta_w - \theta_f)}{H_{vm}(\theta_{vm} - \theta_f)},$$

where $H_{vm} = 2603$ m is a realistic globally-averaged depth for the modern ocean (including land surfaces) and $\theta_{vm} = 4.4$ °C, as reported in VM2010. This gives a scaling parameter $f_{vm} =$ 1.5. Figure 4 shows that the transition timescales obtained by VM2010 are larger than the ones

corresponding to the Warm start simulations and this difference is accentuated by the factor f_{vm} . 292 Using the 1-box model with the parameters provided by VM2010 (orange line), one finds the 293 effective climate feedback parameter in their model to be $\lambda_{vm} = 3.3 Wm^{-2}K^{-1}$ (see SI). This is 294 significantly larger than λ_w and results in longer transition times. Note that λ_{vm} is larger than the 295 climate feedback parameters typically reported by state-of-the-art GCMs (~ 1 $Wm^{-2}K^{-1}$) such 296 as the one used in VM2010. However, a direct comparison may not be possible because λ_{vm} is 297 obtained from a 1-box model fit to the Snowball transition times associated with θ , rather than 298 from the more conventional 'Gregory' method associated with surface temperatures. Therefore, 299 λ_{vm} may encapsulate different climate processes than typical feedback parameters. In particular, 300 the representation of non-linear atmospheric feedbacks (e.g. clouds, water vapor, etc) and the 301 restructuring of the ocean may differ. 302

In the simulations starting from Cold, the temperature anomaly spreads into the part of the ocean 303 that is not already at the freezing temperature, namely the tropical thermocline. The relevant ocean 304 depth is the top 1000 m (approximately), which has an average potential temperature of 3.1 °C. 305 The climate feedback parameter corresponding to the Cold reference state is again obtained by the 306 'Gregory' method and found to be somewhat higher than for Warm: $\lambda_c = 0.95 W m^{-2} K^{-1}$ (see 307 details in SI). This corresponds to a timescale of 159 years (see Section 1). Using these values 308 along with $\theta_i = 3.1 \circ C$ and $\theta_f = -1.8 \circ C$ in Eq. (6) gives the solid green line in Figure 4. Overall, 309 the box model can capture the climate model's transition times starting from Cold, but tends to 310 underestimate them as F approaches the threshold forcing. 311

313 1) SINGLE ERUPTIONS

We now investigate the behavior of the climate model following large idealized volcanic erup-314 tions. The largest eruption simulated from the Warm state has a magnitude of F = -105 Wm^{-2} , 315 which is on the same order as estimates of the Toba mega-eruption (e.g. Jones et al. (2005)). In 316 the year following the simulated eruption, the global-mean SST drops by 5.4 °C, the global sea ice 317 fraction increases by 2% and both quantities recover their original values within 75 years. We also 318 conduct a simulation of comparable magnitude (F = -89 Wm^{-2}) starting from Cold, which leads 319 to a global-mean SST drop of 3.1 °C, a shift in the ice edge latitude from 44.4° to 42.2° (a 4% 320 increase in sea ice fraction) and a recovery within 400 years. Despite the large magnitude of the 321 forcings considered here, the net energy extracted from the system ($\sim 100 Wm^{-2}yr$) is not large 322 enough to cause a shift in climate equilibrium (see Figure 1). The longer recovery time for SST 323 and sea ice fraction in the Cold start eruption is likely due to a stronger ice albedo feedback, since 324 the Cold state has a larger sea ice fraction (33%) than the Warm state (11%). However, we now 325 show that unlike the recovery times of surface quantities, the timescales associated with the global 326 mean ocean potential temperature θ are larger in the Warm state than the Cold one. 327

Figure 5 shows timeseries of θ for all the single eruption simulations conducted in this study. The experiments are run until θ relaxes back to within the 2σ levels of unforced variability. One can estimate the peak cooling θ_{peak} at the end of the first year by equating the energy extracted by the forcing to the change in heat content of the ocean, and neglecting the effects of the feedbacks as follows:

$$\theta_{peak} \approx \frac{F\Delta t}{\rho c H_{tot}},$$
(8)

where Δt is the forcing duration, which is assumed small compared to the timescale associated with climatic feedbacks. Eq. (8) is a good approximation for all the simulations shown in Figure 5, irrespective of how deep the anomaly penetrates into the ocean.

³³⁶ For Cold starts, we consider two eruptions with F = -89 and $-29 Wm^{-2}$ (blue and purple curves ³³⁷ respectively). Because the deep ocean is already at the freezing temperature in the starting state, ³³⁸ the temperature anomaly remains within the tropical thermocline. When normalized with respect ³³⁹ to θ_{peak} (Figure 5b), both responses follow the same decay, which can be fit by a decaying expo-³⁴⁰ nential with an e-folding time of $\tau_c = 159$ years (blue dotted line). According to the 1-box model, ³⁴¹ the ocean depth associated with this timescale is (see Eq. (4)):

$$H_c = \frac{\lambda_c \tau_c}{\rho c},\tag{9}$$

with $\lambda_c = 0.95 W m^{-2} K^{-1}$ corresponding to the climate feedback for the Cold state. This gives $H_c = 1149m$ (roughly 47 % of the total ocean depth).

For Warm start simulations, Figure 5(b) shows that forcing pulses of magnitude F = -35 and -105 Wm^{-2} lead to a relaxation back to equilibrium over several thousand years (orange and red curve, respectively). The normalized plot shows that the two curves follow a similar path to recovery, but the one corresponding to $F = -35 Wm^{-2}$ reaches background noise levels within 2000 years, versus at least 5000 years for $F = -105 Wm^{-2}$. We therefore choose to fit to the $F = -105 Wm^{-2}$ experiment and find that the following two-exponential model gives a good approximation of the impulse response (red dotted line):

$$T_i^{(w)}(t) = T_1 e^{-t/\tau_{w1}} + T_2 e^{-t/\tau_{w2}},$$
(10)

with $T_1 = -0.10^{\circ}$ C, $\tau_{w1} = 500$ years, $T_2 = -0.13^{\circ}$ Cand $\tau_{w2} = 5911$ years.

The millennial-scale decay τ_{w2} cannot be captured by the 1-box model, since the longest timescale that can arise from Eq. (4) is 443 years, when using the full ocean depth H = 2416 m and the inferred climate sensitivity parameter $\lambda_w = 0.72 Wm^{-2}K^{-1}$. To better understand the origin of τ_{w2} , we conduct an additional three numerical experiments with: (i) a smaller magnitude F, (ii) a positive F, and (iii) a large negative F starting from the Very Warm state (see Figure 3):

(i) When F is reduced to $-11 Wm^{-2}$ with a Warm start (yellow curve), the magnitude of the perturbation is not large enough for the θ anomaly to remain outside the range of natural variability for more than a couple of decades. The ocean temperature anomalies remain in the top 1000 m and are rapidly dissipated by climatic processes near the surface. This suggests that there is a minimum amount of forcing required (somewhere between -11 and -35 Wm^{-2}) to induce the millennial timescale of decay in θ .

(ii) A positive forcing experiment was conducted with $F = +31 Wm^{-2}$ (black curve), which likely does not correspond to a particular geophysical mechanism, but is useful for assessing the asymmetry in our model's response between warming and cooling. Unlike in the corresponding negative forcing case (-35 Wm^{-2}), the decay of θ does not have a millennial timescale as it reaches background noise levels within less than 1000 years.

(iii) To assess the sensitivity to the starting state, a mega-eruption simulation with F = -108 Wm^{-2} was conducted from the Very Warm state (pink curve). In this experiment, θ responded with an e-folding time of 437 years, which corresponds to the decay timescale of the 1-box model when considering the full depth of the ocean (pink dotted line).

In the following paragraph, we briefly explore aspects of ocean dynamics that can lead to the 372 millennial decay timescale τ_{w2} . Figure 6 displays vertical profiles of ocean potential temperature 373 for a subset of the single eruptions simulations conducted for this study. Panel (a) shows the F 374 = -105 Wm^{-2} simulation starting from Warm. After 100 years, the temperature profile exhibits 375 two peaks in cooling; the first around 500 m depth and the second at the bottom of the ocean. The 376 shallower peak is likely a result of anomalously cold water parcels being driven down into the ther-377 mocline by Ekman pumping in the mid-latitudes. This mechanism can shield the cold temperature 378 anomaly away from atmospheric damping processes acting at the surface for decades to centuries 379 (e.g. Stouffer (2004), Stenchikov et al. (2009), Gupta and Marshall (2018)). The peak cooling in 380 the abyss is set by enhanced deep convection around the poles, particularly at the sea ice margin 381 in the SH (see Figure S2). After 500 years, the peak cooling in the thermocline is damped away 382 by a combination of atmospheric feedbacks and mixing into the deeper ocean. At that time, the 383 temperature anomaly profile is almost linear with depth and the mixed layer temperature anomaly 384 reaches unforced variability levels (-0.1 $^{\circ}$ C). The subsequent evolution of the temperature profile 385 occurs over 1000-year timescales, as the deep ocean temperature anomaly slowly diffuses back up 386 toward the surface. This upward diffusion process is illustrated using a simple 1D ocean diffusive 387 model with constant diffusivity $\kappa = 3 \times 10^{-5} m^2 s^{-1}$ (the uniform background value used in our 388 coupled model) and boundary conditions ensuring no flux at the bottom of the ocean and a tem-389 perature anomaly fixed to 0 at the top. The latter implicitly assumes that climatic feedbacks at the 390 surface act on much shorter timescales than the various ocean processes bringing the temperature 391 anomaly back up. We initialize the 1D model with a linear temperature profile, roughly corre-392

³⁹³ sponding to the temperature anomaly profile found in our climate model at t = 500 years. The 1D ³⁹⁴ model temperature evolution (in the dotted lines) shows a roughly similar relaxation behavior to ³⁹⁵ the coupled model and has a timescale that scales with H^2/κ , which is O(1000 years).

³⁹⁶ Panel (b) shows θ anomaly profiles for the F = -35 Wm^{-2} impulse simulation starting from ³⁹⁷ Warm. At year 100, there is again a peak cooling around 500 m depth and a small peak in the abyss. ³⁹⁸ Figure S2 shows signatures of Ekman pumping transporting cold waters into the thermocline and ³⁹⁹ enhanced deep convection in the SH high latitudes, although not as large as in (a). Subsequently, ⁴⁰⁰ diffusive processes flatten the temperature profile and the deep temperature anomalies reach back ⁴⁰¹ up to the surface over several thousands of years.

Panel (c) shows temperature anomaly profiles for $F = +31 Wm^{-2}$ starting from Warm. After 100 years, mid-latitude Ekman pumping has driven warm temperature anomalies down into the thermocline. However, a smaller fraction of the anomaly than in (b) penetrates into the deep ocean, due to the higher buoyancy of these anomalously warm waters. This results in an overall ocean relaxation timescale of several hundred years only.

Panel (d) shows temperature anomaly profiles for $F = -108 Wm^{-2}$ starting from Very Warm. After 100 years, there is a large peak cooling at the surface and a small one in the abyss. Some of the temperature anomaly has penetrated into the thermocline, but the surface cooling has not been damped as significantly as in (a). To explain this behavior, one should note that the Very Warm state does not contain ice anywhere on the globe. The impulse cooling promotes sea ice formation in the SH, but the resulting deep convection at the ice margin is not strong enough to carry a large volume of cold waters into the abyss (see Figure S2). In the mid-latitudes and the tropics, a fraction of the temperature anomaly does penetrate into the deeper ocean, but it is a smaller share of the total cooling than in (a) and (b).

In summary, the millennial decay timescale τ_{w2} is a result of cold temperature anomalies reaching the deepest layers of the ocean and slowly diffusing upward over several thousand years. These anomalies can reach the seafloor by enhanced deep convection in the SH high latitudes over the first few decades after the eruption. Notably, the millennial timescale does not arise for a positive impulse or when starting from the Very Warm state.

421 2) SUCCESSION OF ERUPTIONS

We now investigate how the long oceanic timescales can favor a large build up of the cooling response and potentially initiate state transitions when forced by repeated volcanic eruptions. We first establish whether the transition behavior for volcanic eruptions of magnitude F and duration Δt (assumed 1 year) repeated every τ_i years is equivalent to that of a step simulation with average forcing F_{av} :

$$F_{av} = \frac{F\Delta t}{\tau_i}.$$
(11)

Figure 7 shows the globally-averaged θ and sea ice fraction evolution for two simulations in which the model is forced by an idealized eruption every $\tau_i = 10$ years, with F = -105 Wm^{-2} starting from Warm and F = -89 Wm^{-2} starting from Cold. For comparison, we also consider step simulations with $F_{av} = -11 Wm^{-2}$ starting from Warm and $F_{av} = -9 Wm^{-2}$ starting from Cold. The results show that the θ and sea ice fraction timeseries obtained from the repeated eruption simulations closely follow the corresponding step simulations, for both Warm and Cold starts. All simulations eventually transition to a Waterbelt state with approximately 60% ice fraction and the sea ice edge stabilizes around 27° latitude in both hemispheres. The step response is a good approximation for repeated eruptions because the interval between them is much smaller than the ocean relaxation timescales, and because each incremental cooling from a given eruption is much smaller than the overall $\Delta\theta$.

Figure 7 also shows that the 1-box model of Eq. (2) faithfully represents the evolution of θ in 438 all the cases considered. There are non-linearities associated with the transition, but since those 439 only significantly affect θ near the freezing point ($\theta_f = -1.8$ °C), the 1-box model provides a 440 good estimate of the state transition times. For the Cold start simulations, both single eruption 441 and step simulations were well approximated by the 1-box model with timescale $\tau_c = 159$ years 442 and a climate feedback parameter $\lambda_c = 0.95 \ Wm^{-2}K^{-1}$ (see Sections and a and 1). For Warm 443 start simulations, however, there is an apparent discrepancy between the two-timescale model 444 used in Figure 5 and the single timescale model used in Figure 7. This is resolved by noting 445 that here, a state transition occurs after 550 years, which is too short to activate the millennial 446 timescale. Therefore, the 1-box model is a good approximation for transitions occurring within 447 several centuries, but otherwise the 2-timescale model is more appropriate. 448

In Figure 8, we explore the state transition times for repeated uniform eruptions over a range of impulse forcings F (lasting 1 year) and eruption intervals τ_i . One can estimate the transition threshold by assuming a step response with constant forcing F_{av} because the system timescales are much longer than the interval between eruptions. This implies that scenarios with the same F_{av} have the same transition time, which results in the straight contours in Figure 8. The dotted lines show the following transition threshold boundaries: Warm/Cold in black ($F_{w/c} = -1.5 Wm^{-2}$), Cold/Waterbelt in red ($F_{wb} = -5 Wm^{-2}$) and Cold/Snowball in white ($F_{c/s} = -20 Wm^{-2}$).

For the Cold start case, we assume a 1-box model with timescale $\tau_c = 160$ years and can thus es-456 timate transition times to either a Waterbelt or Snowball climate using Eq. (6) and the parameters 457 calculated in Section 4(a), with $F = F_{av}$. As noted that Section, the box model tends to underes-458 timate the GCM transition timescales for Cold start simulations when F is of low magnitude, so 459 the results of Figure 8(a) should be regarded as a lower bound. When the repeated eruptions start 460 from Warm, the transition times can be estimated similarly, but in this case there are two timescales 461 $(\tau_{w1} \text{ and } \tau_{w2})$ involved in the response. When the forcing is large enough for a transition to occur 462 within several hundred years, Figure 7 showed that the 1-box model was a suitable approximation. 463 However, a more general estimate of the step response can be obtained by integrating Eq. (10) and 464 scaling it appropriately to obtain the two-timescale step response from Warm: 465

$$T_s^{(w)}(t) = \frac{F}{F_{ref}} \left[T_1 \tau_{w1} (1 - e^{-t/\tau_{w1}}) + T_2 \tau_{w2} (1 - e^{-t/\tau_{w2}}) \right],$$
(12)

where $F_{ref} = -105 Wm^{-2}$. The time taken to transition from an initial temperature $\theta_i = 8.3^{\circ}$ C to the final temperature θ_f is calculated numerically using Eq. (12) and shown in Figure 8(b). For transitions to either a Waterbelt or Snowball climate, the final temperature is $\theta_f = -1.8^{\circ}$ C (the freezing point), whereas for transitions to the Cold state it is $\theta_f = 0.5^{\circ}$ C. Note that we only consider transition times smaller than 10⁴ years, because over longer timescales, greenhouse gas emissions from volcanic eruptions may start to overcome the cooling (e.g. Walker et al. (1981)).

Figure 8 shows that transition to a Snowball requires a sequence of almost continuous Toba-like events ($\sim -100 Wm^{-2}$) for at least several decades, even when starting from a Cold climate. This differs from the results of Macdonald and Wordsworth (2017), who suggested that such a transition could occur within 3 years for volcanic eruptions with a peak forcing as weak as $-10 Wm^{-2}$. This difference arises because the heat capacity of their model consists of the mixed layer depth of the

ocean (50 m), whereas in our climate model, the entire tropical thermocline (~ 1000 m) plays a 477 role when starting from Cold. Transitions to a Waterbelt climate may occur for a weaker F_{av} , but 478 over several hundred years for a Cold start and around one or two thousand years for a Warm start. 479 In this climate model, such Waterbelt-like states do not exist for F = 0 (see Figure 3), so the system 480 would relax back to Cold when the eruptions stopped. Nevertheless, the hysteresis associated with 481 Waterbelt states may differ between models (e.g. Rose (2015)), and therefore the possibility of 482 a long-term transition to such a state is still plausible. Finally, volcano-induced transitions from 483 Warm to Cold may occur for small magnitude forcings ($\sim -1.5 Wm^{-2}$), due to the proximity of the 484 Warm reference state to a transition threshold. However, such transitions take several thousands 485 of years, during which the build-up of greenhouse gases from intense volcanic activity could start 486 mitigating the cooling effect. 487

488 5. Conclusions

⁴⁸⁹ Our study has explored transitions between the global equilibria of a complex coupled climate ⁴⁹⁰ model with representations of atmosphere, ocean, sea ice and land. We study whether cooling ⁴⁹¹ associated with volcanic eruptions - acting alone (single pulse) or in concert (repeating pulse or ⁴⁹² step) - could trigger transitions between equilibrium states. Although the models employed are ⁴⁹³ rather idealized, we offer the following general conclusions that hopefully do not depend on those ⁴⁹⁴ idealizations:

⁴⁹⁵ 1. A series of Toba-like events (F \sim -100 Wm^{-2}) must occur almost continuously for a min-⁴⁹⁶ imum of several decades to initiate a transition to a Snowball climate by volcanic cooling ⁴⁹⁷ acting alone. This is a significantly longer forcing period than reported by Macdonald and ⁴⁹⁸ Wordsworth (2017), because over the large part of the global ocean, the entire water column ⁴⁹⁹ (rather than just the mixed layer) must reach the freezing temperature of sea water before

500	ice can form at the surface. This result is consistent with the work of Voigt and Marotzke
501	(2010) and Voigt et al. (2011). The relevant system timescales are associated with the full
502	ocean depth when starting from a Warm climate and the tropical thermocline when starting
503	from a Cold one. Assuming that volcanic forcings are no greater in magnitude than Toba-like
504	events, the interval between eruptions must be on the order of years to decades (much smaller
505	than the oceanic timescales) to successfully initiate any state transition between them. Con-
506	sequently, the transition behavior can be predicted from the time-average forcing F_{av} without
507	knowledge of the detailed history of volcanic forcing.

2. The presence of multiple equilibria in the climate system could in theory enable a multistage transition to a Snowball. Indeed, if the climate were already in a stable Cold state with sea ice extending down into the mid-latitudes, then volcanic cooling could more readily initiate a Snowball transition, since only the tropical thermocline would remain to be brought to the freezing temperature. Nevertheless, decades of intense volcanic cooling would still be required, even in this most favorable case.

⁵¹⁴ 3. The proximity of the equilibrium state to a transition threshold controls the amount of forcing ⁵¹⁵ required to initiate a transition. In our model, the Snowball transition threshold is $F_{c/s} = -20$ ⁵¹⁶ Wm^{-2} (a 5.8% reduction in incoming solar radiation with respect to the modern constant). ⁵¹⁷ This is comparable to the values found in other studies (e.g. Voigt and Marotzke (2010), ⁵¹⁸ Voigt et al. (2011), Yang et al. (2012)). The Warm reference state, which resides close to a ⁵¹⁹ critical threshold, may evolve into a Cold state for a forcing as small as -1 Wm^{-2} , although ⁵²⁰ this forcing must be sustained for several thousands of years.

4. The most extreme climates on the Cold branch of our bifurcation diagram are analogous to
 Waterbelt (or Slushball) climates, in which sea ice extends down to the tropics. The energy

required to access such states is similar to that of a Snowball, since both have a mean ocean temperature close to freezing. The threshold forcing for the Waterbelt state ($F_{wb} = -5 Wm^{-2}$) is significantly lower than for a Snowball ($F_{c/s} = -20 Wm^{-2}$), but if the forcing is relaxed after a Waterbelt transition, the system smoothly returns to a climate with no tropical sea ice. This differs from the results of Rose (2015), who find two distinct branches for the Cold and Waterbelt states, and a significant amount of hysteresis associated with the latter.

5. A simple 1-box model can capture transition times for both Warm and Cold start simulations
 relatively well. However, close to the Warm/Cold transition threshold, the climate model
 exhibits a millennial decay timescale that cannot be explained by the assumption of a well mixed ocean. This is caused by enhanced deep convection in the SH, which produces cold
 temperature anomalies in the abyssal ocean that diffuse back up to the surface over thousands
 of years. In this case, a 2-box model that includes this millennial decay timescale provides
 better estimates of transition times.

Acknowledgments. The authors acknowledge support from the MIT-GISS collaborative agreement (NASA) and are grateful for informative conversations with Brian E. J. Rose.

538

APPENDIX

539

Interpretation in terms of a Budyko-Sellers energy balance model

A simple Budyko-Sellers energy balance model (EBM) modified to include the effect of OHT (as in Rose and Marshall (2009)) can be used to illuminate the multiple equilibria found in the climate model. We employ the following model, developed by Rose (2015), which describes the zonal-mean energy balance of the planet in terms of a characteristic surface temperature T:

$$Q_0(1+s_2P_2(x))a(x,T) - (A+BT(x)) + K_a \frac{d}{dx} \left((1-x^2)\frac{dT(x)}{dx}\right) + F_o(x) = 0,$$
(A1)

where $x = sin(\phi)$ and ϕ is latitude, T(x) is the zonal-mean equilibrium surface temperature [° C], K_a is a coefficient of large-scale heat diffusion for the atmosphere $[Wm^{-2}K^{-1}]$, B is the longwave radiative damping $[Wm^{-2}K^{-1}]$, Q_0 is the solar constant (divided by 4), $F_o(x)$ is OHT convergence and a(x,T) is a non-linear co-albedo function that depends strongly on T. The term $Q_0(1+s_2P_2(x))$ gives a reasonable spatial distribution of the incoming radiation, where $P_2(x)$ is the second Legendre polynomial and $s_2 = -0.48$. The OHT is held constant in time and parameterized as follows:

$$H_o(x) = \Psi sin(\phi) cos(\phi)^{2N} = \Psi x (1 - x^2)^N,$$
 (A2)

where N is a positive integer and Ψ is a constant (in units of PW) setting the amplitude of OHT. The parameter N controls the shape of the OHT profile, with higher values shifting the peak of OHT and its convergence closer to the equator and thus helping the stabilization of the ice edge at lower latitudes. The ocean heat convergence is then:

$$F_o(x) = -\frac{\Psi}{2\pi^2 R} (1 - x^2)^{N-1} (1 - x^2(2N+1)).$$
(A3)

The co-albedo switches abruptly from ice-free value (a_0) to the ice-covered (a_i) , depending on the local temperature as follows:

$$a(x,T) = \begin{cases} a_0 \colon \mathbf{T}(\mathbf{x}) < T_{xi} \\ a_i \colon \mathbf{T}(\mathbf{x}) > T_{xi}. \end{cases}$$

⁵⁵⁷ A key feature is that the threshold T_{xi} is not tied to a fixed temperature, but depends on the local ⁵⁵⁸ energy balance at the ice-water interface as follows:

$$T_{xi} = T_f - \frac{h_{min}}{k_i} F_o(x_i). \tag{A4}$$

Eq. (A4) is derived by equating the vertical heat flux through the ice to the ocean heat convergence at the ice edge. The parameter h_{min} is the minimum sea ice thickness to have a substantial effect on the albedo and k_i is the thermal conductivity of sea ice. Therefore, T_{xi} represents the maximum surface temperature possible to maintain an ice thickness larger than h_{min} for a given ocean heat convergence F_o , and is used as the crude threshold between ice-free and ice-covered conditions in the EBM. Incorporating this piece of ice physics helps stabilize climates with intermediate ice covers.

Figure A1 shows a range of bifurcation diagrams obtained by solving the EBM with typical 566 parameter values, as in Rose (2015). The analytical solution is obtained using the method detailed 567 in North (1975) and the numerical solution is obtained with a Crank-Nicholson scheme. The 568 numerical estimates only capture the stable (positive slope) states whereas the analytical solution 569 gives all equilibria. Solutions are plotted for increasing values of $\delta = \frac{K_a}{R}$, which measures the 570 efficiency of atmospheric meridional heat transport relative to radiative damping. Figure A1 shows 571 that as δ increases, the slopes become steeper and thus more prone to instability, consistent with 572 the analytical results of Held and Suarez (1975). However, as argued by Rose and Marshall (2009), 573 the presence of OHT can help stabilize the ice edge as δ increases. Figure A1 also shows the OHT 574 profile for $\Psi = 4$ PW and N = 4, which has a peak value of 0.91 PW that occurs at 22° latitude, not 575 untypical of our own climate. We find that δ values between 0.4 and 0.5 give a bifurcation diagram 576 with somewhat similar characteristics to the one obtained from the climate model, including steep 577 slopes and a continuous set of stable equilibria at intermediate latitudes. The $\delta = 0.5$ curve also 578 gives a bifurcation point close to $Q_0 = 341.5 Wm^{-2}$, which is a feature seen in our climate model. 579

However, the stable states of this curve are not actually accessible from the warm state, as cooling
 from the Warm branch directly leads to a Snowball.

582 **References**

Abbot, D. S., A. Voigt, and D. Koll, 2011: The Jormungand global climate state and implications for Neoproterozoic glaciations. *Journal of Geophysical Research*, **116 (D18)**, doi: 10.1029/2011JD015927, URL http://doi.wiley.com/10.1029/2011JD015927.

⁵⁸⁶ Baresel, B., H. Bucher, B. Bagherpour, M. Brosse, K. Guodun, and U. Schaltegger, 2017:
 ⁵⁸⁷ Timing of global regression and microbial bloom linked with the Permian-Triassic bound ⁵⁸⁸ ary mass extinction: implications for driving mechanisms. *Scientific Reports*, **7**, 43 630, doi:
 ⁵⁸⁹ 10.1038/srep43630, URL http://www.nature.com/articles/srep43630.

Brugger, J., G. Feulner, and S. Petri, 2017: Baby, it's cold outside: Climate model simulations of the effects of the asteroid impact at the end of the cretaceous. *Geophysical Research Letters*, 44 (1), 419–427, doi:10.1002/2016GL072241, URL http://doi.wiley.com/10.
 1002/2016GL072241.

⁵⁹⁴ Budyko, M. I., 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*,
⁵⁹⁵ **21** (5), 611–619.

Ferreira, D., J. Marshall, T. Ito, and D. McGee, 2018: Linking Glacial-Interglacial States to Mul tiple Equilibria of Climate. *Geophysical Research Letters*, doi:10.1029/2018GL077019, URL
 http://doi.wiley.com/10.1029/2018GL077019.

Ferreira, D., J. Marshall, and B. Rose, 2011: Climate Determinism Revisited: Multiple Equilibria in a Complex Climate Model. *Journal of Climate*, **24** (**4**), 992–1012, doi:10.1175/

⁶⁰¹ 2010JCLI3580.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3580.1.

602	Gregory, J. M., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate
603	sensitivity. Geophysical Research Letters, 31 (3), doi:10.1029/2003GL018747, URL http://doi.
604	wiley.com/10.1029/2003GL018747.

- ⁶⁰⁵ Gupta, M., and J. Marshall, 2018: The Climate Response to Multiple Volcanic Eruptions Mediated
- ⁶⁰⁶ by Ocean Heat Uptake: Damping Processes and Accumulation Potential. *Journal of Climate*,
- ⁶⁰⁷ **31** (21), 8669–8687, URL http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0703.1.

Held, I., and M. Suarez, 1975: Simple albedo feedback models of the icecaps. *Tellus*, **26**, 613–629.

⁶⁰⁹ Hoffman, P., A. Kaufman, G. Halverson, and D. Schrag, 1998: A neoproterozoic snowball earth.

Science, **281**, pp. 1342–1346, doi:10.1126/science.281.5381.1342.

- Hoffman, P. F., and Coauthors, 2017: Snowball earth climate dynamics and cryogenian geology geobiology. *Science Advances*, 3 (11), e1600 983, doi:10.1126/sciadv.1600983, URL http:
 //advances.sciencemag.org/lookup/doi/10.1126/sciadv.1600983.
- Huang, J., and K. P. Bowman, 1992: The small ice cap instability in seasonal energy balance
 models. *Climate Dynamics*, 7 (4), 205–215, doi:10.1007/BF00206862, URL http://link.springer.
 com/10.1007/BF00206862.
- Jones, G. S., J. M. Gregory, P. A. Stott, S. F. B. Tett, and R. B. Thorpe, 2005: An AOGCM simula-
- tion of the climate response to a volcanic super-eruption. *Climate Dynamics*, **25** (**7-8**), 725–738, doi:10.1007/s00382-005-0066-8, URL http://link.springer.com/10.1007/s00382-005-0066-8.
- Kaiho, K., and N. Oshima, 2017: Site of asteroid impact changed the history of life on Earth: the
 low probability of mass extinction. *Scientific Reports*, 7 (1), doi:10.1038/s41598-017-14199-x,
- URL http://www.nature.com/articles/s41598-017-14199-x.

Langen, P. L., 2004: Multiple equilibria and asymmetric climates in the CCM3 coupled to an oceanic mixed layer with thermodynamic sea ice. *Geophysical Research Letters*, **31** (**4**), doi: 10.1029/2003GL019039, URL http://doi.wiley.com/10.1029/2003GL019039.

Lee, W.-H., and G. North, 1995: Small ice cap instability in the presence of fluctuations. *Climate Dynamics*, (11), 242–246, doi:10.1007/BF00215010.

Macdonald, F. A., and R. Wordsworth, 2017: Initiation of Snowball Earth with volcanic sulfur
 aerosol emissions. *Geophysical Research Letters*, doi:10.1002/2016GL072335, URL http://doi.
 wiley.com/10.1002/2016GL072335.

Marshall, J., A. Adcroft, C. Hill, Perelman, and Heisey, 1997a: A finite-volume, incompressible
 Navier Stokes model for studies of the ocean on parallel computers. *Journal of geophysical research: Oceans*, doi:doi.org/10.1029/96JC02775.

Marshall, J., C. Hill, Perelman, and A. Adcroft, 1997b: Hydrostatic, quasi-hydrostatic, and
 non-hydrostatic ocean modelling. *Journal of Geophysical Research: Oceans*, doi:10.1029/
 96JC02776.

Matteucci, G., 1993: Multiple equilibria in a zonal energy balance climate model: The thin ice
 cap instability. *Journal of Geophysical Research: Atmospheres*, **98** (D10), 18515–18526, doi:
 10.1029/93JD01740, URL http://doi.wiley.com/10.1029/93JD01740.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametriza tions. I: model climatology and variability in multi-decadal experiments. *Climate Dynam- ics*, **20** (2), 175–191, doi:10.1007/s00382-002-0268-2, URL http://link.springer.com/10.1007/
 s00382-002-0268-2.

Morales Maqueda, M. A., A. J. Willmott, J. L. Bamber, and M. S. Darby, 1998: An investigation
of the small ice cap instability in the Southern Hemisphere with a coupled atmosphere-sea iceocean-terrestrial ice model. *Climate Dynamics*, 14 (5), 329–352, doi:10.1007/s003820050227,
URL http://link.springer.com/10.1007/s003820050227.

North, G., 1975: Theory of energy-balance models. *Journal of the atmospheric sciences*, 32 (11),
 2033–2043.

⁶⁵⁰ North, G., 1985: The small ice cap instability in diffusive models. *Journal of the Amospheric* ⁶⁵¹ *Sciences*, **41 (23)**, 3390–3395, doi:10.1175/1520-0469(1984)041(3390:TSICII)2.0.CO;2.

Poulsen, C. J., and R. L. Jacob, 2004: Factors that inhibit snowball Earth simulation. *Paleoceanog- raphy*, **19** (**4**), doi:10.1029/2004PA001056, URL http://doi.wiley.com/10.1029/2004PA001056.

⁶⁵⁴ Prueher, L., and D. Rea, 2001: Volcanic triggering of late pliocene glaciation: evidence from the

flux of volcanic glass and ice-rafted debris to the north pacific ocean. *Palaeogeography*, *Palaeo-*

climatology, Palaeoecology, 173 (3-4), 215–230, doi:https://doi.org/10.1016/S0031-0182(01)
 00323-6.

Rampino, M., and S. Self, 1992: Volcanic winter and accelerated glaciation floowing the toba
 super-eruption. *Nature*, **359**, 50–52.

Robock, A., 2000: Volcanic eruptions and climate. *Reviews of Geophysics*, 38, doi:10.1029/
 1998RG000054.

Rose, B. E. J., 2015: Stable Waterbelt climates controlled by tropical ocean heat transport: A
 nonlinear coupled climate mechanism of relevance to Snowball Earth. *Journal of Geophys- ical Research: Atmospheres*, **120** (4), 1404–1423, doi:10.1002/2014JD022659, URL http:
 //doi.wiley.com/10.1002/2014JD022659.

⁶⁶⁶ Rose, B. E. J., D. Ferreira, and J. Marshall, 2013: The Role of Oceans and Sea Ice in Abrupt Transitions between Multiple Climate States. *Journal of Climate*, **26** (**9**), 2862–2879, doi:10.1175/
 ⁶⁶⁷ JCLI-D-12-00175.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00175.1.

Rose, B. E. J., and J. Marshall, 2009: Ocean Heat Transport, Sea Ice, and Multiple Climate States:
 Insights from Energy Balance Models. *Journal of the Atmospheric Sciences*, 66 (9), 2828–2843,
 doi:10.1175/2009JAS3039.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009JAS3039.
 1.

Sellers, W. D., 1969: A global climatic model based on the energy balance of the earth-atmosphere
 system. *Journal of applied meteorology*, doi:10.1175/1520-0450(1969)008(0392:AGCMBO)2.
 0.CO;2.

Stenchikov, G., T. L. Delworth, V. Ramaswamy, R. J. Stouffer, A. Wittenberg, and F. Zeng,
2009: Volcanic signals in oceans. *Journal of Geophysical Research*, **114** (**D16**), doi:10.1029/
2008JD011673, URL http://doi.wiley.com/10.1029/2008JD011673.

Stern, R. J., D. Avigad, N. Miller, and M. Beyth, 2008: From volcanic winter to Snowball Earth:
 An alternative explanation for Neoproterozoic biosphere stress. *Links Between Geological Processes, Microbial Activities&Evolution of Life*, Springer, 313–337.

Stommel, H., 1961: Thermohaline convection with two stable regimes of flow. *Tellus*, XI11, 224–
 230, URL https://doi.org/10.1111/j.2153-3490.1961.tb00079.x.

⁶⁸⁴ Stouffer, R. J., 2004: Time Scales of Climate Response. Journal of Climate, 17 (1), 209–

⁶⁸⁵ 217, doi:10.1175/1520-0442(2004)017(0209:TSOCR)2.0.CO;2, URL http://journals.ametsoc.

org/doi/abs/10.1175/1520-0442%282004%29017%3C0209%3ATSOCR%3E2.0.CO%3B2.

687	Timmreck, C., HF. Graf, S. J. Lorenz, U. Niemeier, D. Zanchettin, D. Matei, J. H. Jungclaus,
688	and T. J. Crowley, 2010: Aerosol size confines climate response to volcanic super-eruptions.
689	Geophysical Research Letters, 37 (24), doi:10.1029/2010GL045464, URL http://doi.wiley.com/
690	10.1029/2010GL045464.

- ⁶⁹¹ Tripati, A., and H. Elderfield, 2005: Deep-sea temperature and circulation changes at the ⁶⁹² paleocene-eocene thermal maximum. *Science*, **308** (**5730**), 1894–1898, doi:10.1126/science. ⁶⁹³ 1111724, URL http://www.sciencemag.org/cgi/doi/10.1126/science.1111724.
- Voigt, A., and D. S. Abbot, 2012: Sea-ice dynamics strongly promote Snowball Earth initiation
 and destabilize tropical sea-ice margins. *Climate of the Past*, 8 (6), 2079–2092, doi:10.5194/
 cp-8-2079-2012, URL https://www.clim-past.net/8/2079/2012/.
- ⁶⁹⁷ Voigt, A., D. S. Abbot, R. T. Pierrehumbert, and J. Marotzke, 2011: Initiation of a Marinoan
 ⁶⁹⁸ Snowball Earth in a state-of-the-art atmosphere-ocean general circulation model. *Climate of the* ⁶⁹⁹ *Past*, 7 (1), 249–263, doi:10.5194/cp-7-249-2011, URL http://www.clim-past.net/7/249/2011/.
- Voigt, A., and J. Marotzke, 2010: The transition from the present-day climate to a modern
 Snowball Earth. *Climate Dynamics*, 35 (5), 887–905, doi:10.1007/s00382-009-0633-5, URL
 http://link.springer.com/10.1007/s00382-009-0633-5.
- Walker, J. C., P. B. Hays, and J. F. Kasting, 1981: A negative feedback mechanism for the long term stabilization of Earth's surface temperature. *Journal of Geophysical Research: Oceans*,
 86 (C10), 9776–9782.
- Yang, J., F. Ding, R. M. Ramirez, W. R. Peltier, Y. Hu, and Y. Liu, 2017: Abrupt climate transition
- ⁷⁰⁷ of icy worlds from snowball to moist or runaway greenhouse. *Nature Geoscience*, **10** (**8**), 556–
- ⁷⁰⁸ 560, doi:10.1038/ngeo2994, URL http://www.nature.com/doifinder/10.1038/ngeo2994.

Yang, J., W. R. Peltier, and Y. Hu, 2012: The Initiation of Modern Soft Snowball and Hard
Snowball Climates in CCSM3. Part II: Climate Dynamic Feedbacks. *Journal of Climate*, **25 (8)**, 2737–2754, doi:10.1175/JCLI-D-11-00190.1, URL http://journals.ametsoc.org/doi/abs/
10.1175/JCLI-D-11-00190.1.

LIST OF FIGURES

714 715 716 717 718 719 720 721	Fig. 1.	A graphical representation of the 3 equilibria of our climate model: Warm, Cold and Snow- ball. On the x-axis, we plot the latitude of NH sea ice extent (bottom axis) and ice fraction (top axis). On the y-axis, we plot the ocean-mean potential temperature, θ , (left axis) and energy level (right axis). The thermal energy stored in the ocean dominates over all other sources. The unstable (dotted) lines are linear interpolations between the stable states found in the model. In these ranges of sea ice latitudes and θ , no stable solution can be found. The globes show the SST and sea ice extent of the Warm reference, Cold reference, Waterbelt and Snowball states (from right to left).	. 39
722 723 724 725 726 727 728	Fig. 2.	Zonally-averaged ocean potential temperature (color shading) for the Warm reference, Cold reference, Waterbelt and Snowball states expressed on a stretched depth scale to highlight the thermal structure of the upper ocean. The annual-mean sea ice cover in each hemisphere is denoted by the thick black lines. The zonally-averaged ocean overturning streamfunction (in Sv) is shown with clockwise circulation represented by solid lines and anticlockwise circulation by dotted lines, with a contour interval of 10 Sv. The zonally-averaged meridional ocean heat transport (in PW) is shown above each panel. Note the different scales of OHT.	. 40
729 730 731 732 733 734 735 736 737 738 739 740 741	Fig. 3.	(left) Bifurcation diagram of the climate model expressed in terms of the NH sea ice latitude (left axis) versus the forcing F (bottom axis). F is calculated relative to the reference solar constant $S_0 = 341.5 Wm^{-2}$. The right axis indicates the approximate global mean SST and the top axis shows the solar constant. The thick lines show the range of stable Warm (red), Cold (blue) and Snowball (black) states. The black dotted lines indicate unstable NH sea ice latitudes. Each red circle, blue square, green triangle and black diamond shows the equilibrated state of a simulation started from the Warm (W), Cold (C), Waterbelt (WB) and Snowball (S) states, respectively. (right) Schematic plot (not to scale) representing the stable (solid) and unstable (dotted) branches of the bifurcation diagram. The red arrows outline a possible path along the diagram starting from a Warm state and the blue arrows shows a different path starting from a Cold state. $F_{w/c}$ and $F_{c/s}$ are transition thresholds between the Warm/Cold and the Cold/Snowball equilibria respectively. F_{wb} (= -5 Wm^{-2}) is the forcing below which the Cold states become Waterbelt-like.	. 41
742 743 744 745 746 747 748 749	Fig. 4.	Transition times to reach either the Waterbelt (circles) or Snowball (squares) climate for step simulations starting from the Warm reference state (red) and Cold reference state (blue). The Snowball transition times obtained by VM2010 are also plotted (in orange squares) and scaled by a factor $f_{vm} = 1.5$ for ease of comparison with simulations starting from the Warm reference state. Solid lines are corresponding estimates from the 1-box model with the following parameter values: for a Warm start, $\lambda_w = 0.72 Wm^{-2}$ and H = 2416 m (red line); for a Cold start, $\lambda_c = 0.95 Wm^{-2}$ and H = 1149 m (blue line); for VM2010, $\lambda_{vm} = 3.3 Wm^{-2}$, H = 2603 m and $f_{vm} = 1.5$ (orange line).	. 42
750 751 752 753 754 755 756 757	Fig. 5.	Time evolution of the whole-ocean mean potential temperature anomaly following 1-year long eruptions in our coupled model for a range of forcing magnitudes and starting states. (left) Absolute temperature anomaly response, where the shading represents the 2σ range of unforced variability in the control run. (right) Normalized responses with respect to the peak cooling, which in each case occurs at the end of the first year, the time at which the forcing is turned off. Solid lines are simulation results from the climate model, whereas dotted lines are exponential fits to the following simulations: $F = -105 Wm^{-2}$ run from Warm (red), $F = -89 Wm^{-2}$ from Cold (blue) and $F = -108 Wm^{-2}$ from Very Warm (pink).	. 43
758 759	Fig. 6.	Global-mean ocean potential temperature anomaly, plotted as a function of depth, follow- ing impulse cooling forcings lasting 1 year, with a 50-year time averaging window. The	

760 761 762 763 764 765		shading represents the 2σ level of unforced variability estimated from the relevant control simulation. The panels present results from the following simulations: (a) F = -105 Wm ⁻² starting from Warm, (b) F = -35 Wm ⁻² starting from Warm, (c) F = +31 Wm ⁻² starting from Warm, and (d) F = -108 Wm ⁻² starting from Very Warm. The dotted lines in panel (a) show temperature profiles obtained from a 1D diffusive model with a diffusivity $\kappa = 3 \times 10^{-5}$ m ² s ⁻¹ .	44
766 767 768 769 770 771	Fig. 7.	(left) The response of global-mean ocean temperature, θ , to repeated 1-year eruptions every $\tau_i = 10$ years beginning from a Warm start with $F = -105 Wm^{-2}$ (red) and a Cold start with $F = -89 Wm^{-2}$ (dark blue). Step simulation responses are also shown for a Warm start with $F = -11 Wm^{-2}$ (orange) and for a Cold start with $F = -9 Wm^{-2}$ (light blue). Box model step responses for a Warm start with $F = -11 Wm^{-2}$ (black dotted) and for a Cold start with $F = -9 Wm^{-2}$ (green dotted). (right) Sea ice fraction response in the climate model simulations.	45
772 773 774 775 776 777 778 779	Fig. 8.	Estimates of state transition times for forcings consisting of repeated volcanic eruptions lasting 1 year, every τ_i years. The transition times are estimated using the step response under an average forcing F_{av} , for a range of forcing magnitudes F and eruption intervals τ_i . The points below the white dotted line are transitions to a Snowball. Points between the white and red dotted lines are transitions to the Waterbelt climate. Points between the red and black dotted lines are transitions to Cold. The straight contours represent constant values of F_{av} . Panel (a) and (b) show transition times when starting from the Cold and Warm reference climates, respectively. Note the different scales for the two panels.	46
780 781 782 783 784	Fig. A1.	Bifurcation diagrams obtained from the EBM for a range of δ values. All other parameters are fixed to: A = 210 Wm ⁻² , B = 1.5 Wm ⁻² K ⁻¹ , $a_0 = 0.7$, $a_i = 0.4$, $s_2 = -0.48$, $h_{min} = 0.67$ m, $k_i = 2 Wm^{-2}K^{-1}$, $T_f = 0$ °C, $\Psi = 4$ PW and N = 4. Solid curves are obtained from the analytical model, whereas the dotted lines are results from the numerical solution of the EBM, which only picks out the stable branches of the diagram.	47



FIG. 1. A graphical representation of the 3 equilibria of our climate model: Warm, Cold and Snowball. On the x-axis, we plot the latitude of NH sea ice extent (bottom axis) and ice fraction (top axis). On the y-axis, we plot the ocean-mean potential temperature, θ , (left axis) and energy level (right axis). The thermal energy stored in the ocean dominates over all other sources. The unstable (dotted) lines are linear interpolations between the stable states found in the model. In these ranges of sea ice latitudes and θ , no stable solution can be found. The globes show the SST and sea ice extent of the Warm reference, Cold reference, Waterbelt and Snowball states (from right to left).



FIG. 2. Zonally-averaged ocean potential temperature (color shading) for the Warm reference, Cold reference, Waterbelt and Snowball states expressed on a stretched depth scale to highlight the thermal structure of the upper ocean. The annual-mean sea ice cover in each hemisphere is denoted by the thick black lines. The zonally-averaged ocean overturning streamfunction (in Sv) is shown with clockwise circulation represented by solid lines and anticlockwise circulation by dotted lines, with a contour interval of 10 Sv. The zonally-averaged meridional ocean heat transport (in PW) is shown above each panel. Note the different scales of OHT.



FIG. 3. (left) Bifurcation diagram of the climate model expressed in terms of the NH sea ice latitude (left 798 axis) versus the forcing F (bottom axis). F is calculated relative to the reference solar constant $S_0 = 341.5$ 799 Wm^{-2} . The right axis indicates the approximate global mean SST and the top axis shows the solar constant. The 800 thick lines show the range of stable Warm (red), Cold (blue) and Snowball (black) states. The black dotted lines 801 indicate unstable NH sea ice latitudes. Each red circle, blue square, green triangle and black diamond shows the 802 equilibrated state of a simulation started from the Warm (W), Cold (C), Waterbelt (WB) and Snowball (S) states, 803 respectively. (right) Schematic plot (not to scale) representing the stable (solid) and unstable (dotted) branches 804 of the bifurcation diagram. The red arrows outline a possible path along the diagram starting from a Warm state 805 and the blue arrows shows a different path starting from a Cold state. $F_{w/c}$ and $F_{c/s}$ are transition thresholds 806 between the Warm/Cold and the Cold/Snowball equilibria respectively. F_{wb} (= -5 Wm^{-2}) is the forcing below 807 which the Cold states become Waterbelt-like. 808



FIG. 4. Transition times to reach either the Waterbelt (circles) or Snowball (squares) climate for step simulations starting from the Warm reference state (red) and Cold reference state (blue). The Snowball transition times obtained by VM2010 are also plotted (in orange squares) and scaled by a factor $f_{vm} = 1.5$ for ease of comparison with simulations starting from the Warm reference state. Solid lines are corresponding estimates from the 1-box model with the following parameter values: for a Warm start, $\lambda_w = 0.72 Wm^{-2}$ and H = 2416 m (red line); for a Cold start, $\lambda_c = 0.95 Wm^{-2}$ and H = 1149 m (blue line); for VM2010, $\lambda_{vm} = 3.3 Wm^{-2}$, H = 2603 m and $f_{vm} =$ 1.5 (orange line).



FIG. 5. Time evolution of the whole-ocean mean potential temperature anomaly following 1-year long eruptions in our coupled model for a range of forcing magnitudes and starting states. (left) Absolute temperature anomaly response, where the shading represents the 2σ range of unforced variability in the control run. (right) Normalized responses with respect to the peak cooling, which in each case occurs at the end of the first year, the time at which the forcing is turned off. Solid lines are simulation results from the climate model, whereas dotted lines are exponential fits to the following simulations: $F = -105 Wm^{-2}$ run from Warm (red), $F = -89 Wm^{-2}$ from Cold (blue) and $F = -108 Wm^{-2}$ from Very Warm (pink).



FIG. 6. Global-mean ocean potential temperature anomaly, plotted as a function of depth, following impulse cooling forcings lasting 1 year, with a 50-year time averaging window. The shading represents the 2σ level of unforced variability estimated from the relevant control simulation. The panels present results from the following simulations: (a) F = -105 Wm^{-2} starting from Warm, (b) F = -35 Wm^{-2} starting from Warm, (c) F = +31 Wm^{-2} starting from Warm, and (d) F = -108 Wm^{-2} starting from Very Warm. The dotted lines in panel (a) show temperature profiles obtained from a 1D diffusive model with a diffusivity $\kappa = 3 \times 10^{-5} m^2 s^{-1}$.



Repeated eruptions versus step forcing

⁸²⁹ FIG. 7. (left) The response of global-mean ocean temperature, θ , to repeated 1-year eruptions every $\tau_i = 10$ ⁸³⁰ years beginning from a Warm start with F = -105 Wm^{-2} (red) and a Cold start with F = -89 Wm^{-2} (dark blue). ⁸³¹ Step simulation responses are also shown for a Warm start with F = -11 Wm^{-2} (orange) and for a Cold start with ⁸³² F = -9 Wm^{-2} (light blue). Box model step responses for a Warm start with F = -11 Wm^{-2} (black dotted) and for ⁸³³ a Cold start with F = -9 Wm^{-2} (green dotted). (right) Sea ice fraction response in the climate model simulations.



FIG. 8. Estimates of state transition times for forcings consisting of repeated volcanic eruptions lasting 1 year, every τ_i years. The transition times are estimated using the step response under an average forcing F_{av} , for a range of forcing magnitudes F and eruption intervals τ_i . The points below the white dotted line are transitions to a Snowball. Points between the white and red dotted lines are transitions to the Waterbelt climate. Points between the red and black dotted lines are transitions to Cold. The straight contours represent constant values of F_{av} . Panel (a) and (b) show transition times when starting from the Cold and Warm reference climates, respectively. Note the different scales for the two panels.



Fig. A1. Bifurcation diagrams obtained from the EBM for a range of δ values. All other parameters are fixed to: A = 210 Wm⁻², B = 1.5 Wm⁻²K⁻¹, $a_0 = 0.7$, $a_i = 0.4$, $s_2 = -0.48$, $h_{min} = 0.67$ m, $k_i = 2 Wm^{-2}K^{-1}$, $T_f = 0$ °C, $\Psi = 4$ PW and N = 4. Solid curves are obtained from the analytical model, whereas the dotted lines are results from the numerical solution of the EBM, which only picks out the stable branches of the diagram.