POLITECNICO DI TORINO
Repository ISTITUZIONALE

Limits to predictability of the asymptotic state of the Atlantic Meridional Overturning Circulation in a conceptual climate model

*Original*
Limits to predictability of the asymptotic state of the Atlantic Meridional Overturning Circulation in a conceptual climate model / Mehling, O., Börner, R., Lucarini, V.. - In: PHYSICA D-NONLINEAR PHENOMENA. - ISSN 0167-2789. - STAMPA. - 459:(2024). [10.1016/j.physd.2023.134043]

*Availability:*
This version is available at: 11583/2984756 since: 2024-01-02T11:03:48Z

*Publisher:*
Elsevier

*Published*
DOI:10.1016/j.physd.2023.134043

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

27 November 2026

Physica D 459 (2024) 134043



Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd





# Limits to predictability of the asymptotic state of the Atlantic Meridional Overturning Circulation in a conceptual climate model

Oliver Mehling [a,*], Reyk Börner [b,c], Valerio Lucarini [b,c,d]

[a] DIATI – Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[b] Department of Mathematics and Statistics, University of Reading, Reading, RG6 6AX, United Kingdom
[c] Centre for the Mathematics of Planet Earth, University of Reading, Reading, RG6 6AX, United Kingdom
[d] School of Mathematical and Computational Sciences, University of Leicester, Leicester, LE1 7RH, United Kingdom

## ARTICLE INFO

Dataset link: https://github.com/omehling/amoc-edgetrack, https://doi.org/10.5281/zenodo.10370900

*Keywords:*
Chaotic saddle
Melancholia state
Atlantic Meridional Overturning Circulation
Transient chaos
Tipping point
Conceptual climate model

## ABSTRACT

Anticipating critical transitions in the Earth system is of great societal relevance, yet there may be intrinsic limitations to their predictability. For instance, the asymptotic state of a dynamical system possessing multiple chaotic attractors depends sensitively on the initial condition in the proximity of a fractal basin boundary. Here, we approach the problem of final-state sensitivity of the Atlantic Meridional Overturning Circulation (AMOC) using a conceptual climate model, composed of a slow bistable ocean coupled to a fast chaotic atmosphere. First, we explore the occurrence of long chaotic transients in the monostable regime, which can mask a loss of stability near bifurcations. In the bistable regime, we explicitly construct the chaotic saddle using the edge tracking technique. We quantify the final-state sensitivity through the maximum Lyapunov exponent and the lifetime of the saddle and find that the system exhibits a fractal basin boundary with almost full phase space dimension, implying vanishing predictability of the second kind near the basin boundary. Our results demonstrate the usefulness of studying non-attracting chaotic sets in the context of predicting climatic tipping points, and provide guidance for the interpretation of critical transitions in higher-dimensional climate models.

## 1. Introduction

Like many systems in nature, several elements of the Earth system are thought to be multistable: for a given forcing, they may possess multiple competing attractors that can be reached from different initial conditions [1]. Multistability has been demonstrated in rather complex physical models of the Greenland [2] and West Antarctic [3] ice sheets, the Amazon rainforest [4], the Atlantic Meridional Overturning Circulation (AMOC) [5], and even Earth as a whole [6], supported by paleoclimatic evidence for abrupt climate changes in the past [7,8]. The proposed multistability of the AMOC has been intensively studied for several decades [9–11] and has recently attracted renewed attention due to a suggested loss of stability over the past century seen in observation-based indicators [12]. The paradigm of a multistable AMOC dates back to the seminal model introduced by Stommel [13], who showed that there can be two competing ("on" and "off") states for a given freshwater forcing due to the positive salt-advection feedback. This concept has been invoked to explain qualitative changes of the ocean circulation in past climates [14,15] as well as the response of climate models to changes in the hydrological cycle [16–18].

Collectively, the multistable climate subsystems mentioned above are subsumed under the term tipping elements [19], as anthropogenic global warming may trigger a possibly irreversible transition (tipping) to a qualitatively different state on timescales relevant for policy [20]. In turn, the concept of a "safe operating space" has been framed as keeping a control parameter (such as greenhouse gas forcing) below a critical value to avoid the transition into an undesired state of the Earth system [21]. However, in addition to the classical picture of crossing a bifurcation ("bifurcation tipping"), internal variability and the rate of the forcing can also lead to noise-induced and rate-induced tipping, respectively [22], which means that in reality the safe operating space is likely more complex than what can be described with a single critical value [23–25].

By incorporating the dominant physical processes into low-dimensional conceptual models, dynamical systems theory has proven extremely useful in advancing our understanding of the mechanisms of tipping points in the Earth system (see [8,26–29] for recent reviews on this topic). For example, the derivation of statistical early warning indicators [30–32], which can be applied to observational (e.g., [12,33]) or modeled (e.g., [34]) climatic time series, has increased the

* Corresponding author.
*E-mail address:* oliver.mehling@polito.it (O. Mehling).

https://doi.org/10.1016/j.physd.2023.134043
Received 30 August 2023; Received in revised form 18 December 2023; Accepted 21 December 2023
Available online 23 December 2023
0167-2789/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

prospect of anticipating that a bifurcation is approached. However, in a multistable chaotic system there are fundamental limitations to predictability of the final state even in the autonomous case. The question of whether the attractor reached from a given initial condition can be accurately predicted has been coined "predictability of the second kind" by Lorenz [35]. This is in contrast to predictability "of the first kind", which refers to the ability to predict the future state of a system at a given horizon, given the knowledge of the initial conditions with finite precision. The presence of a limited predictability of the first kind is a key characteristic of chaotic systems and is usually investigated through Lyapunov analysis [36,37].

Let us consider the simpler case of bistable systems. Uncertainty in our ability to predict the final state emerges from the (unavoidably) finite observation time. Beyond a bifurcation point, one may encounter so-called "ghost attractors" — states that are not asymptotically stable but feature transient chaos with finite yet possibly very long lifetimes, which depend sensitively on the initial conditions [38]. In practice, this means that the system appears to reside in a well-defined steady state until a sudden transition to the actual attractor occurs. This phenomenon is common in many areas of physics; for example, turbulence in pipe flows is often regarded as a very long chaotic transient rather than a genuine attractor (e.g., [39]). In this case, given a finite time of observation or simulation (such as in state-of-the-art climate models, which are typically integrated for decades to centuries), one might misjudge the stability properties of the system. This is particularly relevant when a control parameter is close to the critical value separating monostable and bistable behavior.

In the parametric region where bistability is observed, a separate instance of difficulty emerges in predicting the asymptotic state from the choice of the initial condition. In a chaotic, bistable $D$-dimensional system, limited final state predictability is directly linked to the presence of a fractal boundary between the two basins of attraction with fractal dimension $D-1 \leq D_b < D$ [40]: given that an initial condition $\mathbf{u}_0$ can only be determined to a precision $\varepsilon$, the fraction $f$ of (a bounded region of) phase space in which the outcome is uncertain (i.e., different attractors can be reached from within $\mathbf{u}_0 \pm \varepsilon$) scales like

$$f \propto \varepsilon^{\alpha}, \tag{1}$$

where $\alpha = D - D_b$ is the uncertainty exponent [40]. In practical terms, $\alpha \ll 1$ means that decreasing the uncertainty $\varepsilon$ only yields a very small improvement in final state predictability as given by $f$, and the phase space region around the boundary is essentially a "grey zone" in which the final state is almost unpredictable. The case $\alpha \ll 1$ is believed to be relatively common [41]. It has been conjectured that the fractal basin boundary dimension is linked to the Lyapunov spectrum and the lifetime of the chaotic saddle [42,43], with $\alpha \ll 1$ being associated with a chaotic instability on the saddle that is fast compared to its lifetime [44]. While their importance for the transient and asymptotic behavior of chaotic dynamical systems has long been recognized [38,45], both fractal basin boundaries and chaotic saddles remain understudied in the context of climatic tipping points. To our knowledge, the only study that explicitly determined the basin boundary dimension in the context of two competing climatic attractors has been by Lucarini & Bódai [46]. They used a climate model of intermediate complexity to investigate the "Snowball Earth" transition [47,48], in which the ice-albedo feedback drives almost every initial condition either to a fully glaciated or an ice-free climate, and found that the basin boundary has almost full dimension, with $\alpha \approx 0.02$. Regarding the AMOC, Lohmann & Ditlevsen [25] found that in the context of rate-induced tipping (occurring before the bifurcation point is crossed [22]), the final AMOC state in a simplified ocean general circulation model (GCM) depends sensitively on the initial condition. They qualitatively linked this behavior to the presence of a fractal basin boundary but did not assess its properties explicitly.

In this study, we quantitatively explore the limits of predictability of the final AMOC state in a conceptual atmosphere–ocean model inspired by Gottwald [49] (Section 2). The model mimics a key feature of more complex GCMs: it exhibits chaotic dynamics driven by a fast atmospheric component, while the oceanic component acts as a slow integrator. The oceanic component introduces the bistability in the system as a result of the coexistence of the AMOC "on" and "off" states. We investigate aspects related to both time and phase space that limit predictability of the second kind. First, we focus on the lifetime of chaotic transients which may effectively prevent the system from tipping on finite timescales, even when overshooting the bifurcation (Section 3). Then, building on the methods of [46], we use the edge tracking algorithm [50] to construct the chaotic saddle (also called *Melancholia state* in [46]) between two competing AMOC states, whose lifetime and Lyapunov exponents are directly linked to the dimension of the fractal basin boundary [42,43] (Section 4). Finally, we discuss the implications of both phenomena on the predictability of AMOC tipping and on the concept of a safe operating space (Section 5), and summarize our conclusions in Section 6.

## 2. The conceptual atmosphere–ocean model

First, we introduce the conceptual climate model that describes succinctly the main features of atmosphere–ocean flow that we will focus on in this study — a fast, chaotic atmosphere coupled to a slow, bistable ocean component. To this end, we use a coupled configuration of two seminal conceptual models for mid-latitude atmospheric flow and of the AMOC, the Lorenz-84 model [51] (L84 hereafter) and the Stommel model [13], respectively.

### 2.1. Model components

The L84 model for the atmospheric mid-latitude circulation is given by a set of three ordinary differential equations [51]:

$$\begin{aligned} \dot{x} &= -(y^2 + z^2) - a\,(x - F) \\ \dot{y} &= xy - bxz - (y - G) \\ \dot{z} &= bxy + xz - z, \end{aligned} \tag{2}$$

where $x$ is the strength of the westerly winds (a conceptual representation of the jet stream) given by thermal wind balance, and $y$ and $z$ are the amplitudes of the cosine and sine phases of a superimposed traveling wave. The parameters $F$ and $G$ are the background values for the meridional and zonal temperature gradients, respectively, to which $x$ and $y$ would relax in an uncoupled setting; $a$ and $b$ control the internal timescales of the model and will be kept at their original values $a = 0.25$ and $b = 4$ [51]. Despite its very simple representation of atmospheric flow, the model has a strong physical basis, as it can be derived by Galerkin truncation of a two-layer quasi-geostrophic model [52]. It also exhibits a complicated bifurcation structure on its own [53,54], but since different states of atmospheric flow are not the focus of our study, we remain in the regime $G = 1$ where the L84 model only has a single chaotic attractor depicted in Fig. 1.

The Stommel model [13] is the canonical conceptual model for the bistable AMOC and describes the thermohaline flow in the North Atlantic as a function of the meridional temperature gradient $T$ and the salinity gradient $S$ between two boxes representing the subtropical and subpolar North Atlantic. For several decades, Stommel's model has served as a key model to interpret qualitative changes in the AMOC [10,11,16]. Here, we use the version derived by Cessi [55] and follow the notation of Gottwald [49]:

$$\begin{aligned} \dot{T} &= -\frac{1}{\varepsilon_a}(T - \theta) - T - \mu|\Psi|T \\ \dot{S} &= \sigma - S - \mu|\Psi|S, \end{aligned} \tag{3}$$

where $\varepsilon_a = t_r/t_d$ is the ratio between the restoring timescale for temperature $t_r$ and the diffusive timescale of the ocean $t_d$, $\mu = t_{ad}/t_d$ is the

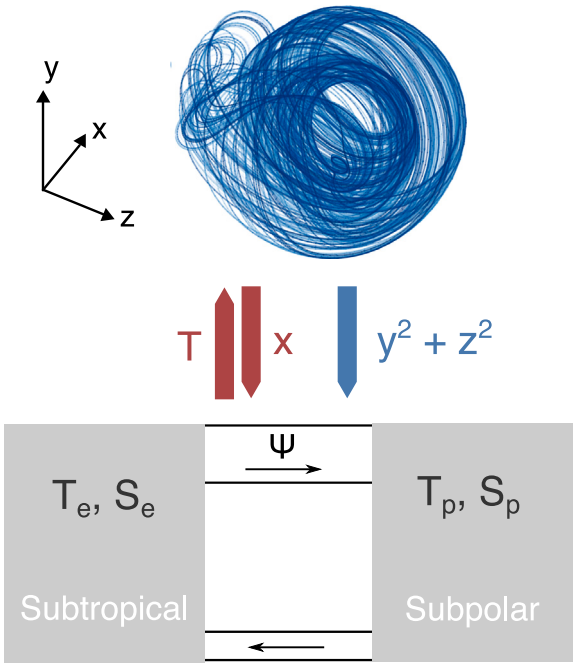



**Fig. 1.** Sketch of the coupled L84-Stommel model. In the Stommel model, $T_e$ and $S_e$ are temperature and salinity in the subtropical North Atlantic box, and $T_p$ and $S_p$ are temperature and salinity in the subpolar North Atlantic box, with $T \equiv T_e - T_p$ and $S \equiv S_e - S_p$. The atmosphere–ocean coupling terms via heat fluxes and freshwater fluxes are sketched in red and blue, respectively.

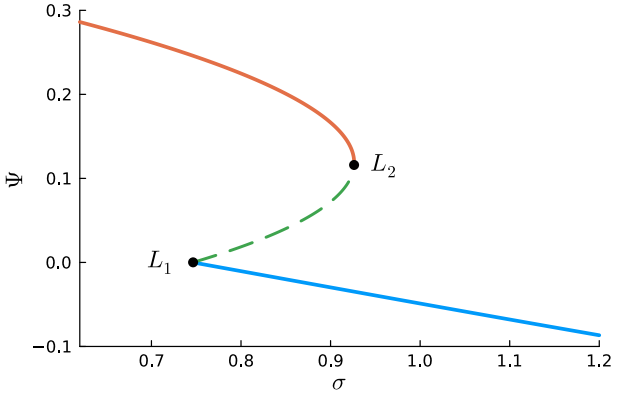



**Fig. 2.** Bifurcation diagram of the uncoupled Stommel model (Eq. (3)): fixed points of the AMOC strength $\Psi$ as a function of the control parameter $\sigma$. Solid lines indicate stable fixed points and the dashed line indicates the unstable fixed point. The saddle–node bifurcation points are labeled as $L_1$ and $L_2$.

ratio between the advective timescale $t_{ad}$ and the diffusive timescale, $\theta$ is the surface temperature gradient to which $T$ is restored, and $\sigma$ is the surface freshwater flux. The advective terms are proportional to the absolute value of the AMOC strength $\Psi = T - S$, where $\Psi > 0$ represents a thermally driven circulation like in the present-day climate, and $\Psi < 0$ represents a salinity-driven AMOC where the circulation would be reversed. Here, as in many previous studies, we consider the freshwater flux $\sigma$ as the main control parameter that determines the AMOC regime. The corresponding bifurcation diagram, which reveals a bistable region bounded by two saddle–node bifurcations, is shown in Fig. 2.

### 2.2. Coupled model equations

As in several previous studies [49,56,57] of conceptual atmosphere–ocean coupling, we couple the Stommel model to the L84 atmosphere through the temperature gradients and the freshwater flux (Fig. 1). Here, we follow the recent formulation of coupling terms by Gottwald [49] (without the sea ice component), who introduced an explicit timescale separation $\varepsilon_f$ between the atmospheric and oceanic components, similar to earlier studies [56,57]. The equations of the five-dimensional coupled model used in our study are given by:

$$\begin{aligned}
\varepsilon_f \dot{x} &= -\Delta - a\,(x - F_0 - F_1 T)\\
\varepsilon_f \dot{y} &= xy - bxz - (y - G_0)\\
\varepsilon_f \dot{z} &= bxy + xz - z\\
\dot{T} &= -\frac{1}{\varepsilon_a}(T - T_{\text{surf}}) - T - \mu|S - T|T\\
\dot{S} &= S_{\text{surf}} - S - \mu|S - T|S
\end{aligned} \tag{4}$$

with

$$\begin{aligned}
\Delta &= y^2 + z^2\\
T_{\text{surf}} &= \theta_0 + \theta_1 \frac{x - \bar{x}}{\sqrt{\varepsilon_f}}\\
S_{\text{surf}} &= \sigma_0 + \sigma_1 \frac{\Delta - \bar{\Delta}}{\sqrt{\varepsilon_f}}
\end{aligned}$$

Here, in a similar spirit as in [56,57], we use a two-way coupling: the fast atmospheric component drives the ocean via the restoring temperature and salinity/freshwater flux, while the slow ocean component couples to the atmosphere via the meridional temperature gradient $F$. For simplicity, we keep the zonal temperature gradient $G = G_0$ constant, noting that the chaotic properties of the L84 model as characterized by its maximum Lyapunov exponent already follow an intricate non-monotonic structure when only $F$ is varied [58].

The main coupling constant is $F_1$, which we set to $F_1 = 0.1$ following [57], who argued that the atmosphere should be weakly coupled to the ocean. Furthermore, the timescale separation parameter between atmosphere and ocean is chosen as $\varepsilon_f = 3 \cdot 10^{-4}$, which corresponds to defining a timescale of about 10 days for the atmosphere and of 100 years for the ocean. Note that this oceanic timescale defines the scale of the dimensionless time $t$, such that all times given here can be read in a "unit" of centuries. Our oceanic timescale is slightly shorter than the one given by Cessi [55], but we mostly aim at a realistic order of magnitude, as a smaller $\varepsilon_f$ needs to be traded off against a higher computational cost for the edge tracking. Except for $F_1$ and $\varepsilon_f$, we use the default parameters of [49] (see Appendix A for a full list of parameter definitions and values). The equations are solved numerically using a fifth-order Runge–Kutta scheme with a fixed timestep of $7.5 \cdot 10^{-6}$, which corresponds to a timestep of 1/40 of the atmospheric timescale as used in [56].

We remark that it is possible to perform a rigorous asymptotic analysis of Eq. (4) that, in the limit $\varepsilon_f \to 0$, leads to writing the homogenized dynamics for the slow variables $(T, S)$ in terms of a stochastic differential equation. The presence of a two-way coupling between the fast and the slow variables requires a more general approach [59] than the more common case of skew-symmetric systems, where the dynamics of the fast variables has an impact on the slow variables but not the other way around [60,61]. See Appendix B for a sketch of the asymptotic multiscale analysis of the system given in Eq. (4) and a derivation of the corresponding homogenized evolution equations for the slow variables.

## 3. Chaotic transients

We first turn our attention to the limits of predictability caused by the sensitivity of the lifetime of chaotic transients to the initial condition. As an example, we select a set of initial conditions by

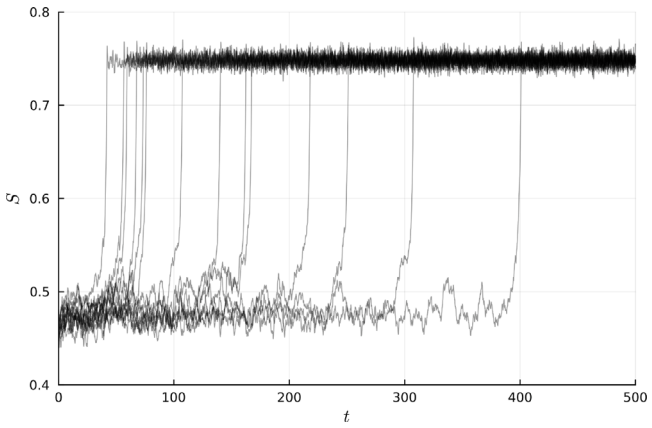



**Fig. 3.** An example of long transients for $\sigma_0 = 0.932$. Initial conditions were sampled from the "AMOC on" attractor for $\sigma_0 = 0.926$ and were integrated forward after the parameter change. The lifetime of the transients spans one order of magnitude and appears unpredictable, but eventually all trajectories convergence to the only remaining attractor.

sampling the "AMOC on" state for $\sigma_0 = 0.926$ every $\Delta t = 7.5$. As we will show explicitly below (cf. Fig. 8), at this value of $\sigma_0$ the "AMOC on" state is globally stable. Then, we integrate trajectories from these initial conditions but with the freshwater parameter set to $\sigma_0 = 0.932$ (Fig. 3), i.e. slightly larger than for the bifurcation point $L_2$ of the uncoupled Stommel model, $\sigma(L_2) = 0.9263$. We can assume that this value is also close to the bifurcation point of the coupled model because the coupling is weak (see Fig. 8).

The initial conditions in Fig. 3 remain in the vicinity of a "ghost attractor" reminiscent of the "AMOC on" state for a long time (up to 400 time units $\approx$ 40 000 years) before converging to the only remaining genuine attractor, the "AMOC off" state. In this section, we systematically explore the lifetime of such chaotic transients and show that they can be observed over a wide range of $\sigma_0$.

For an ensemble of initial conditions, the lifetime of chaotic transients, denoted as $\tau$, is expected to be exponentially distributed [62] with a mean lifetime $\langle\tau\rangle$, if $\tau$ is sufficiently large [63]. Indeed, this is the case for a 100-member initial condition ensemble from which the trajectories in Fig. 3 were drawn, with $\langle\tau\rangle \approx 200$ time units (20 000 years). Grebogi et al. [63] then demonstrated that typically, the mean lifetime follows a power law as one approaches the critical value of the control parameter $\sigma_{0,c}$, such that we expect:

$$\langle\tau\rangle \propto |\sigma_0 - \sigma_{0,c}|^{-\gamma} \tag{5}$$

where $\gamma$ is the critical exponent.

Here, we characterize this region of transient chaos by evaluating the lifetimes of trajectories near the "ghost attractor" using an ensemble of initial conditions for different values of $\sigma_0$. The initial conditions are sampled over $t = 200$ from the "off" or "on" state near the bifurcation points of the underlying Stommel model, and the results are not sensitive to the exact location of the sampling. Here and in the following, all sampling intervals are chosen to be at least several times larger than the Lyapunov timescale of the fast system ($\approx 0.002$ time units). Because $\langle\tau\rangle \to \infty$ as $\sigma_0 \to \sigma_{0,c}$, we need to choose a threshold time $T_{\max}$ after which we deem a trajectory "stable", even though it may still be a very long chaotic transient. This means that $\sigma_{0,c}$ would be biased if determined by direct simulation, as it would depend on the choice of $T_{\max}$. Therefore, we look to find a range of values for $\sigma_0$ with $\langle\tau\rangle$ sufficiently large for the values to be exponentially distributed, but smaller than $T_{\max}$. Using Eq. (5), we then determine $\sigma_{0,c}$ and $\gamma$ within this region via a weighted least-squares fit, taking into account that the parameter $\langle\tau\rangle$ is distributed as $2n\langle\hat{\tau}\rangle/\chi^2(2n)$ [64], where $\langle\hat{\tau}\rangle$ is equal to the sample mean of the lifetimes, $n$ is the number of samples, and $\chi^2(2n)$ is the chi-square distribution with $2n$ degrees of freedom.

Fig. 4 shows the lifetime of the transients tracking the "ghost attractor" as a function of the distance to the critical value $|\sigma_0 - \sigma_{0,c}|$. Here, we only evaluate the range of $\sigma_0$ in which the lifetimes of an ensemble of $n = 100$ initial conditions approximately follow an exponential distribution, and in which $\tau < T_{\max} = 10^4$ for all trajectories. Using the fitted critical value, there is a clear power-law relation (a straight line in the log–log plot) between the lifetime and $|\sigma_0 - \sigma_{0,c}|$ near both boundary crisis points. We obtain $\sigma_{0,c1} = 0.816 \pm 0.008$ and $\sigma_{0,c2} = 0.9273 \pm 0.0009$ (cf. Fig. 9b), where index "1" denotes the transition from the monostable "AMOC on" to the bistable regime, and index "2" denotes the transition from the bistable to the monostable "AMOC off" regime. These values are quite robust to the choice of the scaling law. For instance, fitting the non-power law scaling proposed by [65], we obtain $\sigma_{0,c2} = 0.9251 \pm 0.0013$, which agrees well with the estimate from the fit to Eq. (5).

The critical exponents are $\gamma_1 = 14 \pm 3$ at the first and $\gamma_2 = 7.2 \pm 1.5$ at the second bifurcation. This implies that long transients can be observed over a wider range of parameter values at the first compared to the second bifurcation [45]. For instance, the parameter range $|\sigma_0 - \sigma_{0,c}|$ for which $\langle\tau\rangle > 100$ is 0.04 at the first bifurcation and 0.005 at the second bifurcation. Nevertheless, both critical exponents are larger than those of many typical dynamical systems [45]. Using the scaling law from Eq. (5) and the quantile function of the exponential distribution, we can now calculate how far the "safe operating space" of the AMOC is extended beyond the genuinely bistable regime through the existence of long transients. In other words, up to which value of $\sigma_0$ will the desirable AMOC state (in our case, the "on" state) continue to act like a stable attractor, even though the bifurcation point has been crossed? The exponential distribution underlying the transient lifetime means that this safe operating space is inherently linked to a timescale of interest and an acceptable probability of survival on the desired AMOC state (or, equivalently, a tolerable probability of tipping to the undesired state). Fig. 5 shows this probability of survival on the "on" state as a function of $|\sigma_0 - \sigma_{0,c}|$ and the chosen timescale. For example, overshooting to $\sigma_0 \approx 0.934$ would yield a 5% probability of obtaining a transient lifetime of less than 100 years (red marker in Fig. 5). On the other hand, if the probability of a "transient collapse" should be limited to 0.1% on the same timescale, the safe operating space is limited to no more than $\sigma_0 \approx 0.931$. In this way, if the control parameter is changed sufficiently rapidly to bring the system back into the bistable regime, a transition might successfully be avoided despite "overshooting" the critical value (cf. [66]).

## 4. Chaotic saddle and fractal basin boundary

We now focus on the limits to predictability in the bistable regime. The long-lived transients make it difficult to estimate its boundaries precisely, such that we rely on the two extrapolated critical values from Section 3 to bound this regime ($\sigma_0 \in [0.82, 0.927]$ in the following), within which the "AMOC on" and "AMOC off" states are expected to coexist as genuine attractors. As outlined above, in the bistable regime predictability of the asymptotic state is limited by the dimension $D_b$ of the fractal basin boundary. $D_b$ is in turn fundamentally linked to the properties of the chaotic saddle between the two attractors [42–44]. However, the chaotic saddle is typically a set with Lebesgue measure zero that has one unstable and one stable manifold (the basin boundary) [45], and therefore cannot be obtained by direct simulation. Therefore, we start with constructing a *pseudo*-trajectory of the saddle (Section 4.1) which we use to determine its lifetime (Section 4.2) and Lyapunov exponents (Section 4.3) before assessing the resulting basin boundary dimension (Section 4.4).

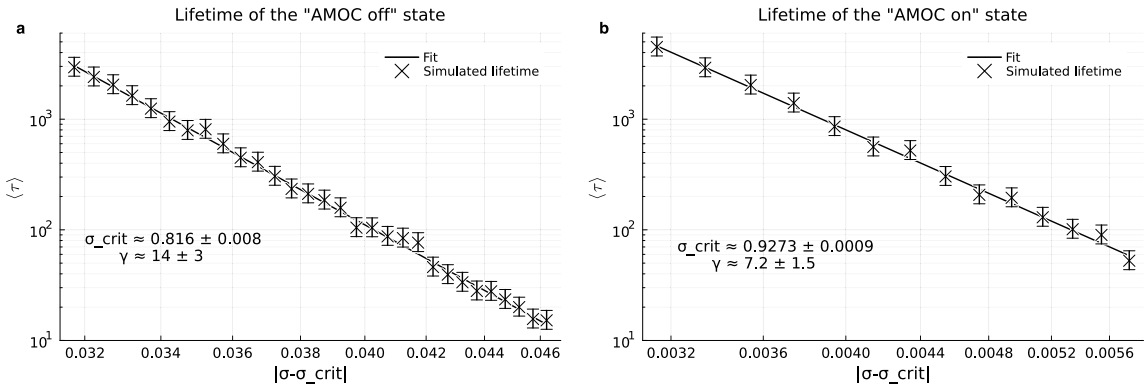



**Fig. 4.** Numerically estimated mean lifetime $\langle\tau\rangle$ of chaotic transients, corresponding 95% confidence intervals and power-law fit for different values of $\sigma_0$ for (a) the transient "AMOC off" state and (b) the transient "AMOC on" state outside the regime of bistability. Note that all axes are logarithmic and that the subplots use different $x$-axes.

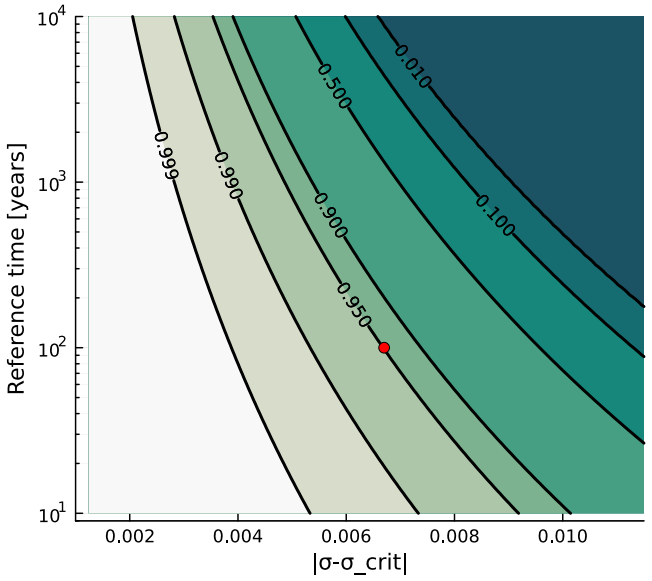



**Fig. 5.** Probability of survival in the region of chaotic transients beyond the critical value $\sigma_{0,c2} \approx 0.927$ as a function of the distance to the critical value $|\sigma_0 - \sigma_{0,c}|$ and the timescale of interest (reference time) during which tipping should be prevented. The red marker indicates $\sigma_0 \approx 0.934$ for which the transient tipping probability is limited to 5% on a timescale of 100 years.

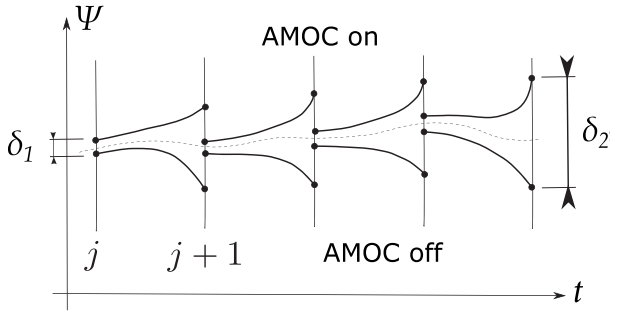



**Fig. 6.** Schematic depiction of the edge tracking algorithm. Note that in this study, the distances $\delta_{1,2}$ are calculated over the entire phase space instead of only $\Psi$.
*Source:* Adapted from Lucarini and Bódai [46], released under a Creative Commons license (CC-BY 3.0, https://creativecommons.org/licenses/by/3.0/).

### 4.1. Constructing the chaotic saddle: edge tracking algorithm

While there are different algorithms to construct a reliable approximation of a chaotic saddle [38], we follow [46] and use the edge tracking algorithm originally proposed in [67] and later re-formulated by [50]. This allows us to obtain an arbitrarily long pseudo-trajectory that tracks the chaotic saddle of the conceptual atmosphere–ocean model. The edge tracking algorithm has previously been applied successfully in both high- and low-dimensional chaotic systems (e.g., [46,50,68–70]). The algorithm is depicted schematically in Fig. 6 and described in detail in Sects. 4.2 and 5.1 of [46], such that we only summarize it briefly here.

We start from two initial conditions $\{\mathbf{u}_1^0, \mathbf{u}_2^0\}$ inside the two different basins of attraction. First, by applying the standard bisection method, we obtain a pair of states $\{\mathbf{u}_1^*, \mathbf{u}_2^*\}$ with $|\mathbf{u}_1^* - \mathbf{u}_2^*| < \delta_1$, where $\mathbf{u}_1^*$ and $\mathbf{u}_2^*$ still belong to the two different basins of attraction. By construction, these two states are both within a distance $\delta_1$ from the basin boundary. As a consequence, when the system is evolved from $\mathbf{u}_i(t_0) \equiv \mathbf{u}_i^*$, they are expected to diverge along the unstable direction of the basin boundary while tracking with the flow along its stable direction. Once $|\mathbf{u}_1(t) - \mathbf{u}_2(t)| \geq \delta_2$ for some $t > t_0$, the bisection is repeated such that the distance between the two "shadowing trajectories" $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ is again less than $\delta_1$, and the new states are evolved further. This procedure can be repeated any number of times to obtain a pseudo-trajectory $\mathbf{u}_S(t) = (\mathbf{u}_1(t) + \mathbf{u}_2(t))/2$ that follows the basin boundary and eventually converges to track the chaotic saddle itself.

In the following, we set $\delta_2 = 0.004$ and $\delta_1 = 0.0025$, such that only one bisection step is needed per iteration. However, as in more complex models, it is not clear a priori how the norm $|\mathbf{u}_1(t) - \mathbf{u}_2(t)|$ should be defined, as the state vector $\mathbf{u}$ comprises different model components with possibly different magnitudes for mean and variability. Here, we define the norm by rescaling the atmospheric variables with a scaling factor $\xi < 1$ (we choose $\xi = 1/5000$). It can be viewed as a hyperparameter of our method and does not alter the dynamics of the dynamical system. This enables us to calculate the divergence of trajectories over the full phase space (instead of just the $T - S$ subspace), but ensures that the divergence is not dominated by the fast chaotic motion. We run the edge tracking algorithm for 2000 iterations, generating around 50 time units ($\approx$ 5000 years) of pseudo-trajectory.

Fig. 7 shows an example result of the edge tracking algorithm for $\sigma_0 = 0.9$. While the attractors and the saddle separate clearly in both $T$ and $S$ as expected from the phase space structure of the underlying Stommel model, the atmospheric motion differs little between the two attractors. However, close to the saddle the atmospheric energy $\frac{1}{2}(X^2 + Y^2 + Z^2)$ has longer tails, which might be due to the repulsion along the unstable direction, and therefore indicates that the weak coupling does exert a certain influence on the atmosphere.

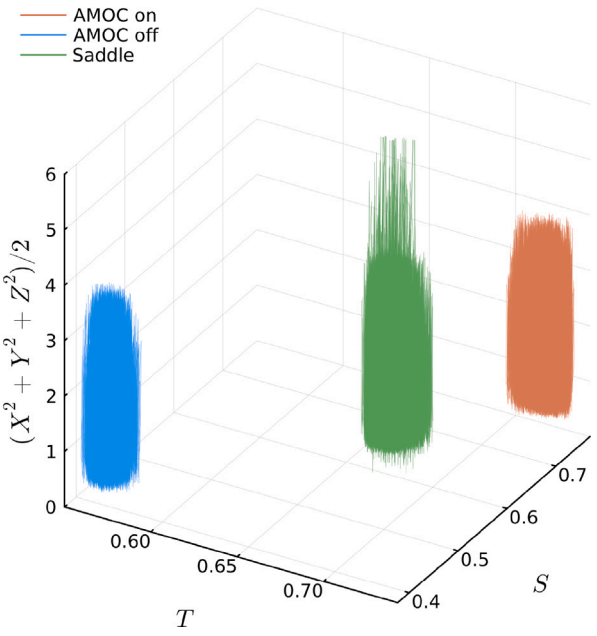



**Fig. 7.** Three-dimensional projection of a pseudo-trajectory of the saddle (green) and trajectories tracking the attractors (blue, orange) for $\sigma_0 = 0.9$. Axis show the slow variables $T$ and $S$ and the atmospheric energy $E = (X^2 + Y^2 + Z^2)/2$.

Applying the edge tracking algorithm for different values of $\sigma_0$, we can compute the unstable branch of the AMOC bifurcation diagram for the coupled atmosphere–ocean model (Fig. 8). Over most of the parameter range, our bifurcation diagram follows that of the underlying uncoupled Stommel model very closely, as expected from weak atmosphere–ocean coupling combined with a large timescale separation. Outside the bistable parameter range, we can also perform edge tracking in a region of parameter space where very long-lived chaotic transients exist; in this case, the algorithm detects an edge state lying in between the long-lived chaotic solution and the actual unique asymptotic solution (unfilled circles in Fig. 8). Note that in this regime, the approach is very similar to the one taken in turbulence [50], where one finds the unstable solution between the laminar fixed point and the long-lived turbulent state. An interesting feature of the bifurcation diagram in Fig. 8 is that the stable and unstable branches do not meet at the critical values $\sigma_{0,ci}$. This might put into question the extrapolated critical values from Section 3, but we have verified the transient nature of some of the unfilled points on the "AMOC on" branch of Fig. 8 via direct simulation (e.g., Fig. 3). In these cases, the transient lifetime is longer than the length of the edge tracking trajectory ($t \approx 50$). Therefore, the discrepancy seems to arise from the particular multiscale nature of our system, which merits further investigation in future studies.

### 4.2. Lifetime of the saddle

We now focus on two properties of the chaotic saddle, its lifetime and its spectrum of Lyapunov exponents (LEs) [45], for different values of the freshwater forcing $\sigma_0$ in the bistable regime. Both can be obtained by sampling a large number of initial conditions from the edge tracking trajectory obtained in Section 4.1. Since a state initialized on the trajectory $\mathbf{u}_S$ is very close to but not precisely on the actual saddle, it will remain in a phase space region $\mathcal{B}$ containing the saddle for some time before converging to one of the attractors.

Defining $\mathcal{B}$ as the bounding box of $\mathbf{u}_S$ after the initial spin-up, the number of remaining trajectories within $\mathcal{B}$ is expected to decay exponentially [45]:

$$N_S(t) = N_0\, e^{-t/\langle\tau\rangle},$$

**Table 1**
Lyapunov exponents for the two attractors (labeled as in Fig. 7) and the chaotic saddle for $\sigma_0 = 0.9$. Uncertainty estimates are 95% confidence intervals generated from bootstrap resampling of 100 trajectories each from the 2000-member ensemble.

| | "AMOC on" | "AMOC off" | Saddle |
|---|---|---|---|
| $\lambda_1$ | 535 ± 2 | 531 ± 2 | 538 ± 2 |
| $\lambda_2$ | −0.04 ± 0.02 | −0.015 ± 0.017 | 0.13 ± 0.07 |
| $\lambda_3$ | −0.67 ± 0.02 | −2.65 ± 0.03 | −0.56 ± 0.14 |
| $\lambda_4$ | −8.11 ± 0.13 | −2.91 ± 0.04 | −6.0 ± 0.3 |
| $\lambda_5$ | −1230 ± 2 | −1155 ± 2 | −1182 ± 4 |

where $N_0$ is the number of sampled initial conditions ($N_0 = 600$ in our numerical simulations), which are all within $\mathcal{B}$ by construction, and $\langle\tau\rangle$ is the mean lifetime. In practice, this exponential scaling is not expected to hold for small and large $t$ due to the non-uniform initial distribution in $\mathcal{B}$ and the finite size of the ensemble, respectively. However, we can obtain $\langle\tau\rangle$ via a least-squares fit of the slope of $\log N_S(t)/N_0$ against $t$ for intermediate values of $t$.

Fig. 9a shows $N_S(t)/N_0$ for several values of the freshwater forcing $\sigma_0$ in the bistable regime of the Stommel model. As expected, the number of remaining trajectories in $\mathcal{B}$ shows a clear exponential decay. The lifetime is of $\mathcal{O}(1)$ (about one century in physical units) but depends strongly on $\sigma_0$: approaching the saddle–node bifurcation of the underlying Stommel model, trajectories spend more time in the vicinity of the saddle. While the number of sampled $\sigma_0$ is too small to conclusively establish a functional relation $\langle\tau\rangle(\sigma_0)$, the lifetime increases clearly nonlinearly as the critical freshwater forcing parameter $\sigma_{0,c2}$ is approached (Fig. 9b).

### 4.3. Lyapunov spectrum of the saddle

Having explored the global instability of the coupled model characterized by the lifetime of the saddle, we now turn towards the local instability due to its chaotic dynamics. To this end, we compute the full spectrum of Lyapunov exponents of the saddle using the standard procedure of successive Gram–Schmidt orthonormalization [71] as implemented in [72]. The Lyapunov spectrum of the saddle is approximated by choosing 2000 initial conditions on $\mathbf{u}_S$ with the longest lifetimes in $\mathcal{B}$ and then averaging over the Lyapunov spectra of a subset of these individual trajectories while they remain in $\mathcal{B}$. For consistency, we tested the same method for the attractors, where averaging over 2000 finite-time LEs calculated over a similar duration ($t = 7.5$) yielded a comparable result to letting the algorithm converge along one long trajectory.

The Lyapunov spectra $\{\lambda_i\}, i = 1, \ldots, 5$ for the saddle and each of the two attractors are given in Table 1 for $\sigma_0 = 0.9$. For the attractors, we expect one positive LE characterizing the chaotic instability on the attracting set, a zero LE associated with the motion along the attractor, and otherwise negative LEs encapsulating the convergence of initial conditions to the attractor. Indeed, we find a large, positive maximum Lyapunov exponent (MLE) $\lambda_1$, while $\lambda_2$ is very close to zero and all remaining LEs are negative. Similarly, the saddle possesses one positive LE ($\lambda_1$), one LE close to zero ($\lambda_2$) and three negative LEs. Note that obtaining a good numerical estimate for the vanishing LE of a saddle is more difficult than for an attractor. This is because the trajectories considered for the evaluation of the LEs have a finite lifetime, as they unavoidably end up veering towards one of the competing attractors, realizing a positive stretching along the direction of the flow. We have verified (Fig. S1 in the Supplementary Information [SI]) that the estimate of $\lambda_2$ converges towards zero as we consider longer-lived trajectories for our estimates, so that the stretching along the flow is minimized (being zero if we could generate a trajectory living exactly on the saddle).

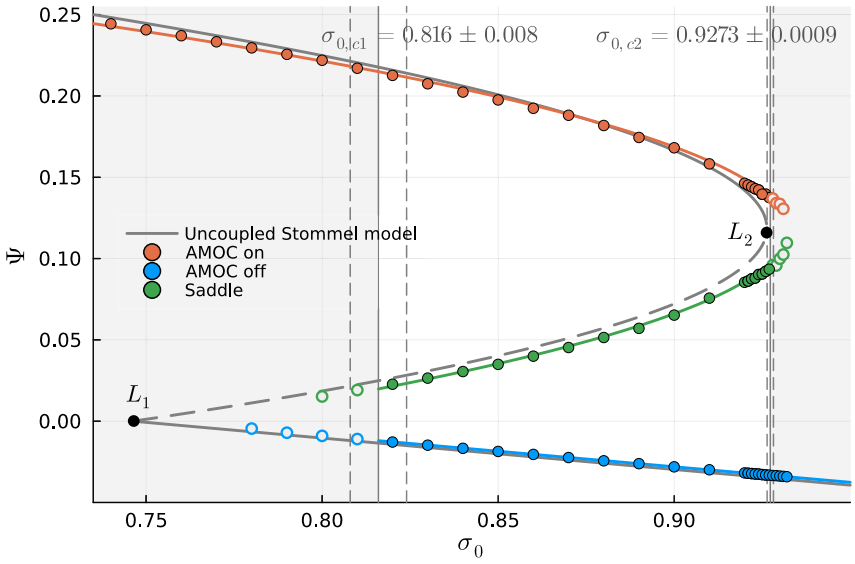



**Fig. 8.** Bifurcation diagram for AMOC strength $\Psi$ of the coupled model. Each dot represents one (pseudo-)trajectory averaged over about 50 time units after a spin-up period. Points in the transient regime (gray background) as determined in Section 3 are marked with unfilled circles. Vertical solid and dashed lines represent best estimates ± one standard deviation of the critical values $\sigma_{0,ci}$. The colored, solid lines represent square-root fits for the upper and the unstable branches and a linear fit for the lower branch. The bifurcation diagram of the uncoupled Stommel model (Eq. (3)) is shown in gray for comparison. Saddle–node bifurcation points for the Stommel model are labeled $L_1$ and $L_2$.

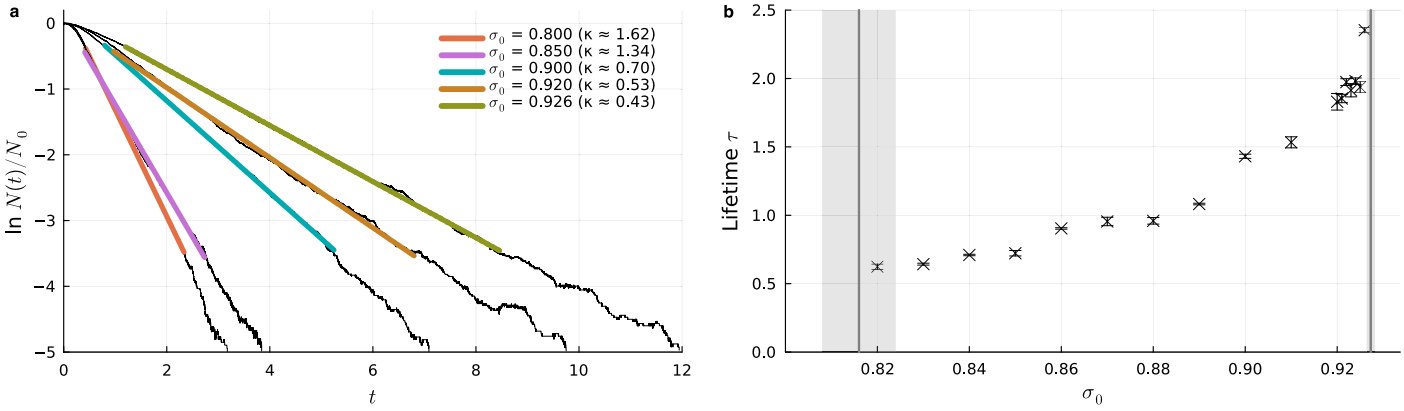



**Fig. 9.** Lifetimes of the chaotic saddle. (a) Fraction of trajectories $N(t)/N_0$ within the saddle bounding box $\mathcal{B}$ as a function of time for different values of $\sigma_0$. Colored lines indicate linear fits in lin–log space, whose slope is $-\kappa$, in the range $N/N_0 \in [0.03, 0.7]$. (b) Lifetime $\tau = 1/\kappa$ obtained from the fits in panel a for all simulated values of $\sigma_0$. Error bars reflect the uncertainty in $\tau$ with respect to reasonable changes in the fitting range in panel a. The vertical lines and shaded areas indicate estimates of the critical values $\sigma_{0,c}$ plus and minus one standard deviation, corresponding to the vertical solid and dashed lines in Fig. 8.

The MLE of our model can be attributed to the chaotic motion induced by the atmosphere via a simple scaling test. First, we note that the magnitude of the maximum and minimum LEs ($|\lambda_1|$ and $|\lambda_5|$) is large compared to the other LEs. Furthermore, the values of $\lambda_1$ and $\lambda_5$, respectively, are similar for all three invariant sets. This is consistent with the assumption that the atmospheric and oceanic components of our model are weakly coupled, and we may associate $\lambda_1$ and $\lambda_5$ with the maximal and minimal LE of the uncoupled L84 model, rescaled with the atmospheric timescale $1/\varepsilon_f$. This can be confirmed by computing the Lyapunov spectrum of the attractors for different values of $\varepsilon_f$ (Fig. S2 in the SI). We find that $\lambda_1$ and $\lambda_5$ scale in very good agreement with $\lambda_{i,\mathrm{L84}}/\varepsilon_f$, where $\lambda_{i,\mathrm{L84}}$ is the corresponding LE of the uncoupled L84 model from Eq. (2), while the second to fourth LEs are mostly independent of the timescale separation.

The results shown here exemplarily for $\sigma_0 = 0.9$ are qualitatively representative for all values of $\sigma_0$ which we have tested within the bistable regime, and there is no clear dependence of $\lambda_1$ or $\lambda_2$ on the freshwater flux. The MLE $\lambda_1$ is very similar across different $\sigma_0$ (mean: 536 ± 4) and the second LE is very small ($\lambda_2 < 0.3 \ll \lambda_1$) for all $\sigma_0$.

### 4.4. Dimension of the fractal basin boundary

Given the spectrum of Lyapunov exponents and the lifetime of the saddle, we can use the dimension formula proposed in [42,43] to relate these two invariants to the information dimension of the stable manifold of the saddle, i.e., the basin boundary. However, provided that the escape rate from the saddle $\kappa = 1/\tau$ is much smaller than the MLE $\lambda_1$, Bódai and Lucarini [44] showed that the formula for the dimension of the basin boundary $D_b$ can be simplified to (see their Eq. 25)

$$D_b = D - \frac{\kappa}{\lambda_1}\,. \tag{6}$$

This relation can be viewed as a generalization of the formula proposed earlier by Hsu et al. [73] to the case of rough basin boundaries [44]. Note that, strictly speaking, Eq. (6) gives the information dimension and not the box-counting dimension of the basin boundary which will be needed later for assessing the question of predictability, but the two are expected to yield very similar values here [42].

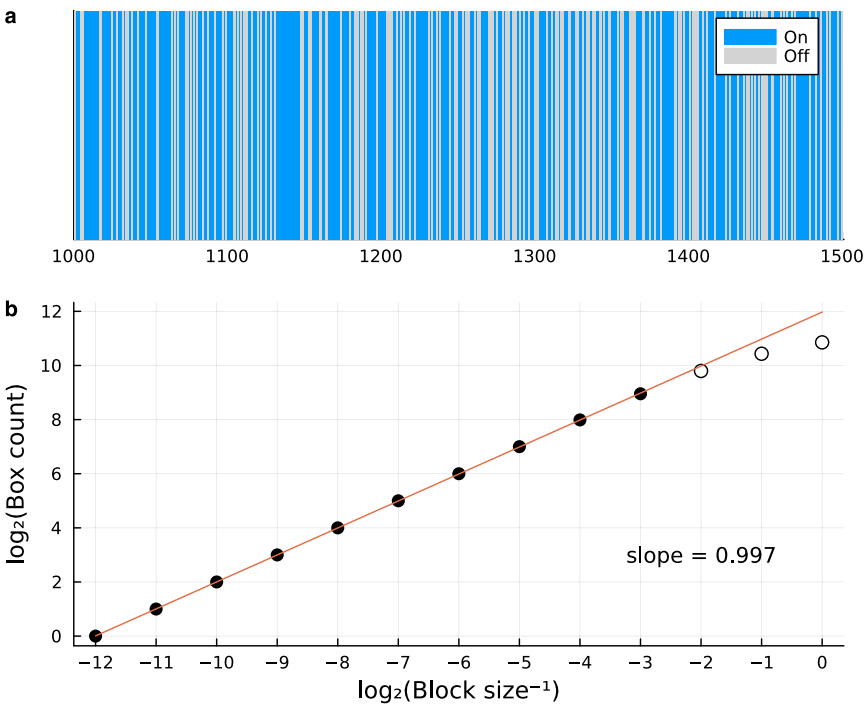



**Fig. 10.** Fractal basin boundary along a sample cross-boundary section for $\sigma_0 = 0.9$: (a) Outcome (final AMOC state, color-coded) of the forward integration of 512 evenly spaced initial conditions along a one-dimensional phase space section intersecting with the fractal basin boundary (see Fig. S4 in the SI for a visualization of all $2^{13}$ outcomes of the full section), (b) Computation of the box-counting dimension from $2^{13}$ initial conditions from the same boundary section. Here, the "pruning" parameter is $N_p = 3$, i.e., the three rightmost points (unfilled markers) are not taken into account for the linear fit.

Since the Lyapunov spectra and lifetimes obtained above clearly satisfy $\kappa < \lambda_1$, we can apply Eq. (6) to obtain a theoretical value for the basin boundary dimension. The resulting dimension $D_b$ is between 4.997 and 4.999 for different values of $\sigma_0$, with slightly increasing values as the critical values $\sigma_{0,c2}$ is approached (Fig. S3 in the SI). This means that the basin boundary dimension is extremely close to (within 0.1%) but strictly smaller than the phase space dimension $D = 5$. Following Eq. (6), this can be attributed to the escape rate being small compared to the MLE.

To verify these results numerically, we also compute the dimension of the basin boundary directly using the standard box-counting algorithm. We follow the methodology of [46] and [44] in sampling evenly spaced initial conditions along a line that intersects with the basin boundary, and determine the attractor to which each initial condition eventually converges. In Fig. 10a, we show a sample plot of the outcomes of 512 initial conditions, where the endpoints of the sampling interval are within $\delta_1 = 10^{-3}$ of the basin boundary (using the definition of the norm from Section 4.1), but away from the saddle. Repeating this procedure in different regions of phase space and then for each value of $\sigma_0$ considered previously, we find that the standard box-counting dimension ($0 < D_{\text{box}} \leq 1$) calculated from $2^{13}$ initial conditions, as shown exemplarily in Fig. 10b, is very close to one (between 0.988 and 0.999) for all $\sigma_0$. Both methods therefore confirm, regardless of whether the information or box-counting dimension is used, that a large scale separation between $\kappa$ and $\lambda_1$ yields a basin boundary with almost full phase space dimension.

The basin boundary dimension $D_b$ and the uncertainty exponent $\alpha$ that determines the *final state sensitivity* as given by Eq. (1) are linked by the simple relation $\alpha = D - D_b$ [40]. With the values of $D_b$ predicted from the saddle properties (Fig. S3 in the SI), $\alpha$ varies between 0.003 and about 0.001 close to the critical value $\sigma_{0,c2}$. From the viewpoint of Eq. (1), this means that an improvement in accuracy of the initial condition has a negligible effect on the uncertainty of the final climate state: if the accuracy of the initial condition is doubled, the phase space volume in which the outcome is uncertain decreases by about $1 - \left(\frac{1}{2}\right)^{\alpha} \approx 0.1$–$0.2\%$. Even if slightly larger values of $\alpha$ obtained from the box-counting algorithm are used, which are often on the order of $10^{-2}$, predictability of the final state only increases by 1%–2% for a doubling in the accuracy of the initial condition. Note that, formally, this vanishing predictability refers to the full five-dimensional system, but the two competing attractors separate mostly along the oceanic variables $(T, S)$ (cf. Fig. 7). Hence, we can interpret our results as vanishing predictability of the asymptotic AMOC state. In summary, the final AMOC state is essentially unpredictable in an extended region of phase space close to the basin boundary, to first order irrespective of the value of $\sigma_0$.

## 5. Discussion

Using a conceptual climate model that comprises a chaotic atmosphere and a bistable AMOC, we have investigated two fundamental limits to predictability of the asymptotic AMOC state: a fractal basin boundary in the bistable AMOC regime and chaotic transients in the monostable regime. In the bistable regime, we approximated the chaotic saddle (or *Melancholia state*) between the two competing AMOC attractors using the edge tracking algorithm and found a large timescale separation between the fast chaotic motion on the saddle and the slow escape rate from the saddle. This timescale separation implies a fractal basin boundary with close to full phase space dimension, which we verified by explicitly computing the box-counting dimension of the boundary. In the monostable regime, chaotic transients with exponentially distributed lifetimes arise close to the bifurcation points over a relatively wide range of the freshwater parameter $\sigma_0$.

This complex behavior complicates the assessment of a system's resilience in two distinct ways. In dynamical systems theory [74], one common definition of resilience relies on measuring the minimal "kick"

perturbation that causes the system to transition into a competing (undesired) attractor [75]. A fractal basin boundary, as quantified by Eq. (1), thus implies a complete loss of resilience in an extended region in phase space, where an arbitrarily small perturbation may cause a critical transition. In parameter space, resilience is often defined via the distance to critical points delimiting a "safe operating space" [21,74]. The presence of long chaotic transients on "ghost attractors" renders it essentially impossible to determine the exact position of such bifurcation points on finite timescales by observing or simulating only a single time series. Hence, defining a safe operating space requires a probabilistic definition due to the exponential distribution of transient lifetimes, and predictability depends crucially on the timescale of interest.

The two phenomena described here have recently been framed as *fractality-induced tipping* and *transient-reduced tipping*, respectively, by Kaszás et al. [76]. To our knowledge, this is the first study in which both phenomena have been systematically explored for a deterministic but chaotic (conceptual) climate model. In contrast to [76], we did not focus on individual *tipping probabilities*, but rather on a global characterization of phase space through the uncertainty exponent $\alpha$ (in the bistable regime) and the transient lifetime $\langle\tau\rangle$ (in the monostable regime). In the bistable regime, this can be achieved with a much smaller computational cost than the local phase space sampling approach taken by [76]. In fact, calculating $\alpha$ via the saddle properties and Eq. (6) also appears to yield more accurate results than repeatedly sampling the box-counting dimension of the basin boundary. While both global and local approaches would ideally complement each other, the construction of the saddle and computation of its properties might be computationally feasible in GCMs [46], while a sufficiently large ensemble might not.

To our knowledge, this is the first time that the edge tracking algorithm has been applied to find the chaotic saddle between two competing AMOC states. While it is not guaranteed that our results from a conceptual climate model will generalize to (much) higher-dimensional atmosphere–ocean models, we note several similarities with the results of Lucarini & Bódai [46], who used an intermediate-complexity climate model with $\mathcal{O}(10^4)$ degrees of freedom to study a different tipping point. From a technical standpoint, the edge tracking algorithm worked robustly, independently of the exact nature of the positive feedback and regardless of whether internal variability of the target variable was larger (as in our case) or smaller than $\delta_1$. This makes us confident that the edge tracking algorithm is a suitable method even for models with many more degrees of freedom in which bistability of the AMOC has been identified (e.g., [5,25,77,78]), even to those featuring large internal variability [78]. Indeed, we have recently implemented the edge tracking algorithm successfully in a coupled atmosphere–ocean GCM of intermediate complexity to investigate a Melancholia state separating the two competing AMOC states [79]. While other methods such as numerical continuation have previously been applied to find unstable AMOC equilibria [80], the edge tracking method only requires (sufficiently long) forward integration, which makes it easy to apply and very suitable for tipping elements in explicit models like typical GCMs.

A perhaps surprising similarity to [46] is that the basin boundary dimension is similarly close to full dimension despite the very different models and feedbacks. We believe that this is indeed the case for a rather wide class of models that feature (explicit or emergent) timescale separation, with the local instability governed by the fast component (atmosphere/weather) and the global instability governed by the slow component (ocean/climate). Thus, it seems plausible that sensitive and seemingly random dependence on the initial condition of the final AMOC state [25,81] or of the transient lifetime of a weak AMOC state [82] may be linked to the presence of a high-dimensional chaotic saddle and a fractal basin boundary.

So far, we have only investigated the autonomous case, which is approached when a parameter (here, the freshwater flux $\sigma_0$) is varied very slowly compared to the internal timescales of the system. This assumption is, however, invalid for the current anthropogenic warming and associated changes in the water cycle. Therefore, rate-dependent AMOC tipping could be observed, but the inertia of the system can also give rise to "safe overshoots" [66] beyond the effect of transient phenomena discussed here. It remains an intriguing open question how the barriers to final state predictability discussed here would be altered in a non-autonomous setting and how they would depend on the forcing rate.

## 6. Conclusions

Our results show that predictability of the asymptotic AMOC state is limited by a fractal basin boundary with almost full phase space dimension as well as long chaotic transients, which both arise from the chaotic multiscale nature of the coupled atmosphere–ocean system. We derived these findings from a conceptual climate model that consisted of four main ingredients for this behavior – bistability, chaotic motion, timescale separation and weak coupling –, such that our conclusions should generalize to more complex models with similar properties and even other components of the Earth system. This has practical implications for the boundaries of a safe operating space, which become intrinsically fuzzy both in a spatial and temporal sense. Starting from a state in the "gray zone" of phase space near a fractal basin boundary or near the "ghost attractor" in the parameter region of chaotic transients, the final outcome depends sensitively on the initial condition, which may give rise to non-monotonic and seemingly counter-intuitive outcomes of an initial condition ensemble — interpretable, however, via dynamical systems theory.

## CRediT authorship contribution statement

**Oliver Mehling:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Reyk Börner:** Conceptualization, Methodology, Software, Writing – review & editing. **Valerio Lucarini:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Code and data availability

Source code is available at https://github.com/omehling/amoc-edgetrack and archived at https://doi.org/10.5281/zenodo.10370900 [83]. Code and data to reproduce the figures can be obtained on request from the corresponding author.

## Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956170 (CriticalEarth) and under grant agreement No 820970 (TiPES). VL has received support from the University of Reading's RETF Open Science project CROPS. The authors would like to thank Tamás Bódai, Georg Gottwald, Calvin Nesbitt and Raphael Römer for discussions on different parts of this work. The authors also wish to thank two anonymous reviewers for their constructive comments which have improved the quality of the paper.

## Appendix A. Parameters of the coupled L84-Stommel model

| Parameter | Value | Physical meaning |
|---|---|---|
| $a$ | 0.25 | – |
| $b$ | 4 | – |
| $F_0$ | 8 | Atmospheric meridional temperature gradient |
| $G_0$ | 1 | Atmospheric zonal temperature gradient |
| $\varepsilon_a$ | 0.34 | Ratio between temperature restoring timescale and diffusive timescale of the ocean |
| $\mu$ | 7.5 | Ratio between advective and diffusive ocean timescales |
| $\theta_0$ | 1 | Sea surface temperature gradient |
| $\sigma_0$ | Is varied | Sea surface freshwater flux |
| $\theta_1$ | 0.0195 | Sea surface temperature gradient perturbation |
| $\sigma_1$ | 0.00934 | Sea surface freshwater flux perturbation |
| $\bar{x}$ | 1.0147 | Time-mean of $x$ of the uncoupled L84 |
| $\bar{\Delta}$ | 1.7463 | Time-mean of $\Delta$ of the uncoupled L84 |
| $F_1$ | 0.1 | Atmospheric meridional temperature gradient perturbation |
| $\varepsilon_f$ | $3 \cdot 10^{-4}$ | Ratio between atmospheric and oceanic timescale |

## Appendix B. Multiscale model reduction

As discussed in Section 2, $\varepsilon_f$ is a small parameter that controls the timescale separation between the dynamics of the atmospheric and oceanic component of the system. It is possible to cast Eq. (4) in the following form,

$$\begin{aligned}\dot{\mathbf{x}} &= \frac{1}{\epsilon^2} g(\mathbf{x},\mathbf{X}) = \frac{1}{\epsilon^2} g_1(\mathbf{x}) + \frac{1}{\epsilon^2} g_2(\mathbf{X}) \\ \dot{\mathbf{X}} &= a(\mathbf{x},\mathbf{X}) + \frac{1}{\epsilon} b(\mathbf{x},\mathbf{X}) = a(\mathbf{X}) + \frac{1}{\epsilon} b(\mathbf{x}),\end{aligned} \tag{B.1}$$

with $\epsilon^2 = \varepsilon_f$, where $\mathbf{x} = (x, y, z)$ and $\mathbf{X} = (T, S)$ are vectors comprising the fast and slow variables, respectively. Because of the non-vanishing term $g_2$, the system is not skew-symmetric. Hence, one cannot straightforwardly apply the homogenization theory presented in [60,61] to derive the effective stochastic differential equation (SDE) in Itô convention,

$$\mathrm{d}\mathbf{X} = \tilde{a}(\mathbf{X})\mathrm{d}t + \sigma(\mathbf{X})\mathrm{d}\mathbf{W}, \tag{B.2}$$

describing the properties of the slow variables $\mathbf{X}$ in the $\epsilon \to 0$ limit. Here $\tilde{a}(\mathbf{X}) \in \mathbb{R}^2$ is the drift term, $\sigma(\mathbf{X}) \in \mathbb{R}^{2\times p}$ is the matrix describing the noise law, and $\mathrm{d}\mathbf{W} \in \mathbb{R}^2$ is a vector whose components are the increments of 2 independent Wiener processes.

Instead, one needs to resort to a more general theory [59] that is able to deal with the conceptually more challenging case of two-way coupled systems for rather general functions $g(\mathbf{x},\mathbf{X})$, $a(\mathbf{x},\mathbf{X})$, and $b(\mathbf{x},\mathbf{X})$ in Eq. (B.1) above. We derive the following explicit expressions for the drift term and the noise law in Eq. (B.2):

$$\tilde{a} = a(\mathbf{X}), \qquad \sigma(\mathbf{X})\sigma(\mathbf{X})^T = \int_0^\infty dt \int d\mu_{\mathbf{X}}(\mathbf{x}) \left(b(\mathbf{x}) \otimes b(\mathbf{x}(t)) + b(\mathbf{x}(t)) \otimes b(\mathbf{x})\right), \tag{B.3}$$

where $\otimes$ indicates the tensor product, $\mu_{\mathbf{X}}(\mathbf{x})$ is the invariant measure of the system $\dot{\mathbf{x}} = g_1(\mathbf{x}) + g_2(\mathbf{X})$, where $\mathbf{X}$ is taken as a fixed parameter, and $\mathbf{x}(t)$ is the evolution at time $t$ of the initial condition $\mathbf{x}(t=0) = \mathbf{x}$. Hence, the correlation matrix of the noise is built starting from the properties of correlation of the fast system where the slow variables $\mathbf{X}$ are *frozen in*. While the general homogenization formulas presented in [59] are more complex, the rather simple result given in Eq. (B.3) comes from the fact that in our case the dependence on the fast and slow variables factors out in the functions $g$, $a$, and $b$. Indeed, one obtains the same result one would derive by naively using the theory developed for skew-symmetric systems, under the heuristic assumption that the slow variables can be treated as external parameters for the stochastic process associated with the fast variables.

We also wish to point out the fundamental difference between the original system given in Eq. (4) and its homogenized version for the slow variables given here in Eq. (B.2). In the parametric range where multistability is found for (4), it is by definition impossible for a trajectory initialized in the basin of attraction of one state to visit the competing state. By contrast, the stochastic differential equation given in Eq. (B.2) allows for (typically rare, when far from the bifurcation points) noise-induced transitions between the competing states of the slow variables (see discussion in [84–86]).

## Appendix C. Supplementary data

Supplementary material related to this article can be found online at https://ars.els-cdn.com/content/image/1-s2.0-S0167278923003974-mmc1.pdf.

## References


[1] U. Feudel, A.N. Pisarchik, K. Showalter, Multistability and tipping: From mathematics and physics to climate and brain—Minireview and preface to the focus issue, Chaos 28 (2018) 033501, http://dx.doi.org/10.1063/1.5027718.

[2] A. Robinson, R. Calov, A. Ganopolski, Multistability and critical thresholds of the Greenland ice sheet, Nature Clim. Change 2 (2012) 429–432, http://dx.doi.org/10.1038/nclimate1449.

[3] J. Garbe, T. Albrecht, A. Levermann, J.F. Donges, R. Winkelmann, The hysteresis of the Antarctic Ice Sheet, Nature 585 (2020) 538–544, http://dx.doi.org/10.1038/s41586-020-2727-5.

[4] M.D. Oyama, C.A. Nobre, A new climate-vegetation equilibrium state for Tropical South America, Geophys. Res. Lett. 30 (2003) 2199, http://dx.doi.org/10.1029/2003GL018600.

[5] E. Hawkins, R.S. Smith, L.C. Allison, J.M. Gregory, T.J. Woollings, H. Pohlmann, B. de Cuevas, Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport, Geophys. Res. Lett. 38 (2011) L10605, http://dx.doi.org/10.1029/2011GL047208.

[6] G. Margazoglou, T. Grafke, A. Laio, V. Lucarini, Dynamical Landscape and Multistability of a Climate Model, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 477 (2021) 20210019, http://dx.doi.org/10.1098/rspa.2021.0019.

[7] V. Brovkin, E. Brook, J.W. Williams, S. Bathiany, T.M. Lenton, M. Barton, R.M. DeConto, J.F. Donges, A. Ganopolski, J. McManus, S. Praetorius, A. de Vernal, A. Abe-Ouchi, H. Cheng, M. Claussen, M. Crucifix, G. Gallopín, V. Iglesias, D.S. Kaufman, T. Kleinen, F. Lambert, S. van der Leeuw, H. Liddy, M.-F. Loutre, D. McGee, K. Rehfeld, R. Rhodes, A.W.R. Seddon, M.H. Trauth, L. Vanderveken, Z. Yu, Past abrupt changes, tipping points and cascading impacts in the Earth system, Nat. Geosci. 14 (2021) 550–558, http://dx.doi.org/10.1038/s41561-021-00790-5.

[8] N. Boers, M. Ghil, T.F. Stocker, Theoretical and paleoclimatic evidence for abrupt transitions in the Earth system, Environ. Res. Lett. 17 (2022) 093006, http://dx.doi.org/10.1088/1748-9326/ac8944.

[9] H.A. Dijkstra, M. Ghil, Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach, Rev. Geophys. 43 (2005) http://dx.doi.org/10.1029/2002RG000122.

[10] T. Kuhlbrodt, A. Griesel, M. Montoya, A. Levermann, M. Hofmann, S. Rahmstorf, On the driving processes of the Atlantic meridional overturning circulation, Rev. Geophys. 45 (2007) RG2001, http://dx.doi.org/10.1029/2004RG000166.

[11] W. Weijer, W. Cheng, S.S. Drijfhout, A.V. Fedorov, A. Hu, L.C. Jackson, W. Liu, E.L. McDonagh, J.V. Mecking, J. Zhang, Stability of the Atlantic Meridional Overturning Circulation: A Review and Synthesis, J. Geophys. Res. Oceans 124 (2019) 5336–5375, http://dx.doi.org/10.1029/2019JC015083.

[12] N. Boers, Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation, Nat. Clim. Chang. 11 (2021) 680–688, http://dx.doi.org/10.1038/s41558-021-01097-4.

[13] H. Stommel, Thermohaline Convection with Two Stable Regimes of Flow, Tellus 13 (1961) 224–230, http://dx.doi.org/10.3402/tellusa.v13i2.9491.

[14] W.S. Broecker, D.M. Peteet, D. Rind, Does the ocean–atmosphere system have more than one stable mode of operation? Nature 315 (1985) 21–26, http://dx.doi.org/10.1038/315021a0.

[15] S. Rahmstorf, Ocean circulation and climate during the past 120,000 years, Nature 419 (2002) 207–214, http://dx.doi.org/10.1038/nature01090.

[16] S. Rahmstorf, Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle, Nature 378 (1995) 145–149, http://dx.doi.org/10.1038/378145a0.

[17] T.F. Stocker, A. Schmittner, Influence of CO2 emission rates on the stability of the thermohaline circulation, Nature 388 (1997) 862–865, http://dx.doi.org/10.1038/42224.

[18] S. Rahmstorf, M. Crucifix, A. Ganopolski, H. Goosse, I. Kamenkovich, R. Knutti, G. Lohmann, R. Marsh, L.A. Mysak, Z. Wang, A.J. Weaver, Thermohaline circulation hysteresis: A model intercomparison, Geophys. Res. Lett. 32 (2005) L23605, http://dx.doi.org/10.1029/2005GL023655.

[19] T.M. Lenton, H. Held, E. Kriegler, J.W. Hall, W. Lucht, S. Rahmstorf, H.J. Schellnhuber, Tipping elements in the Earth's climate system, Proc. Natl. Acad. Sci. 105 (2008) 1786–1793, http://dx.doi.org/10.1073/pnas.0705414105.

[20] D.I. Armstrong McKay, A. Staal, J.F. Abrams, R. Winkelmann, B. Sakschewski, S. Loriani, I. Fetzer, S.E. Cornell, J. Rockström, T.M. Lenton, Exceeding 1.5°C global warming could trigger multiple climate tipping points, Science 377 (2022) eabn7950, http://dx.doi.org/10.1126/science.abn7950.

[21] J. Rockström, W. Steffen, K. Noone, Å. Persson, F.S. Chapin, E.F. Lambin, T.M. Lenton, M. Scheffer, C. Folke, H.J. Schellnhuber, B. Nykvist, C.A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P.K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R.W. Corell, V.J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, J.A. Foley, A safe operating space for humanity, Nature 461 (2009) 472–475, http://dx.doi.org/10.1038/461472a.

[22] P. Ashwin, S. Wieczorek, R. Vitolo, P. Cox, Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system, Philos. Trans. R. Soc. A 370 (2012) 1166–1184, http://dx.doi.org/10.1098/rsta.2011.0306.

[23] V. Lucarini, S. Calmanti, V. Artale, Experimental mathematics: Dependence of the stability properties of a two-dimensional model of the Atlantic ocean circulation on the boundary conditions, Russ. J. Math. Phys. 14 (2007) 224–231, http://dx.doi.org/10.1134/S1061920807020124.

[24] H. Alkhayuon, P. Ashwin, L.C. Jackson, C. Quinn, R.A. Wood, Basin bifurcations, oscillatory instability and rate-induced thresholds for Atlantic meridional overturning circulation in a global oceanic box model, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 475 (2019) 20190051, http://dx.doi.org/10.1098/rspa.2019.0051.

[25] J. Lohmann, P.D. Ditlevsen, Risk of tipping the overturning circulation due to increasing rates of ice melt, Proc. Natl. Acad. Sci. 118 (2021) e2017989118, http://dx.doi.org/10.1073/pnas.2017989118.

[26] C. Kuehn, A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics, Physica D 240 (2011) 1020–1035, http://dx.doi.org/10.1016/j.physd.2011.02.012.

[27] M. Ghil, V. Lucarini, The physics of climate variability and climate change, Rev. Modern Phys. 92 (2020) 035002, http://dx.doi.org/10.1103/RevModPhys.92.035002.

[28] P.D.L. Ritchie, H. Alkhayuon, P.M. Cox, S. Wieczorek, Rate-induced tipping in natural and human systems, Earth Syst. Dyn. 14 (2023) 669–683, http://dx.doi.org/10.5194/esd-14-669-2023.

[29] H.A. Dijkstra, The role of conceptual models in climate research, Physica D 457 (2024) 133984, http://dx.doi.org/10.1016/j.physd.2023.133984.

[30] V. Dakos, S.R. Carpenter, W.A. Brock, A.M. Ellison, V. Guttal, A.R. Ives, S. Kéfi, V. Livina, D.A. Seekell, E.H. van Nes, M. Scheffer, Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data, PLoS One 7 (2012) e41010, http://dx.doi.org/10.1371/journal.pone.0041010.

[31] T.M. Lenton, Early warning of climate tipping points, Nature Clim. Change 1 (2011) 201–209, http://dx.doi.org/10.1038/nclimate1143.

[32] M. Santos Gutiérrez, V. Lucarini, On some aspects of the response to stochastic and deterministic forcings, J. Phys. A 55 (2022) 425002, http://dx.doi.org/10.1088/1751-8121/ac90fd.

[33] P. Ditlevsen, S. Ditlevsen, Warning of a forthcoming collapse of the Atlantic meridional overturning circulation, Nature Commun. 14 (2023) 4254, http://dx.doi.org/10.1038/s41467-023-39810-w.

[34] C.A. Boulton, L.C. Allison, T.M. Lenton, Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model, Nature Commun. 5 (2014) 5752, http://dx.doi.org/10.1038/ncomms6752.

[35] E.N. Lorenz, Climatic predictability, in: The Physical Basis of Climate and Climate Modelling, World Meteorological Organization, 1975, pp. 132–136.

[36] E. Kalnay, Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, Cambridge, 2003.

[37] A. Pikovsky, A. Politi, Lyapunov Exponents: A Tool to Explore Complex Dynamics, Cambridge University Press, Cambridge, 2016.

[38] Y.-C. Lai, T. Tél, Transient Chaos, Springer, New York, 2011.

[39] T. Tél, Y.-C. Lai, Chaotic transients in spatially extended systems, Phys. Rep. 460 (2008) 245–275, http://dx.doi.org/10.1016/j.physrep.2008.01.001.

[40] S.W. McDonald, C. Grebogi, E. Ott, J.A. Yorke, Fractal basin boundaries, Physica D 17 (1985) 125–153, http://dx.doi.org/10.1016/0167-2789(85)90001-6.

[41] C. Grebogi, E. Ott, J.A. Yorke, Chaos, Strange Attractors, and Fractal Basin Boundaries in Nonlinear Dynamics, Science 238 (1987) 632–638, http://dx.doi.org/10.1126/science.238.4827.632.

[42] B.R. Hunt, E. Ott, J.A. Yorke, Fractal dimensions of chaotic saddles of dynamical systems, Phys. Rev. E 54 (1996) 4819–4823, http://dx.doi.org/10.1103/PhysRevE.54.4819.

[43] D. Sweet, E. Ott, Fractal dimension of higher-dimensional chaotic repellors, Physica D 139 (2000) 1–27, http://dx.doi.org/10.1016/S0167-2789(99)00222-5.

[44] T. Bódai, V. Lucarini, Rough basin boundaries in high dimension: Can we classify them experimentally? Chaos 30 (2020) 103105, http://dx.doi.org/10.1063/5.0002577.

[45] E. Ott, Chaos in Dynamical Systems, second ed., Cambridge University Press, Cambridge, 2002.

[46] V. Lucarini, T. Bódai, Edge states in the climate system: Exploring global instabilities and critical transitions, Nonlinearity 30 (2017) R32, http://dx.doi.org/10.1088/1361-6544/aa6b11.

[47] M.I. Budyko, The effect of solar radiation variations on the climate of the Earth, Tellus 21 (1969) 611–619, http://dx.doi.org/10.3402/tellusa.v21i5.10109.

[48] W.D. Sellers, A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System, J. Appl. Meteorol. Climatol. 8 (1969) 392–400, http://dx.doi.org/10.1175/1520-0450(1969)008<0392:AGCMBO>2.0.CO;2.

[49] G.A. Gottwald, A model for Dansgaard–Oeschger events and millennial-scale abrupt climate change without external forcing, Clim. Dyn. 56 (2021) 227–243, http://dx.doi.org/10.1007/s00382-020-05476-z.

[50] J.D. Skufca, J.A. Yorke, B. Eckhardt, Edge of Chaos in a Parallel Shear Flow, Phys. Rev. Lett. 96 (2006) 174101, http://dx.doi.org/10.1103/PhysRevLett.96.174101.

[51] E.N. Lorenz, Irregularity: A fundamental property of the atmosphere, Tellus A 36 (1984) 98–110, http://dx.doi.org/10.3402/tellusa.v36i2.11473.

[52] L. van Veen, Baroclinic Flow and the Lorenz-84 Model, Int. J. Bifurcation Chaos 13 (2003) 2117–2139, http://dx.doi.org/10.1142/S0218127403007904.

[53] H. Broer, C. Simó, R. Vitolo, Bifurcations and strange attractors in the Lorenz-84 climate model with seasonal forcing, Nonlinearity 15 (2002) 1205–1267, http://dx.doi.org/10.1088/0951-7715/15/4/312.

[54] H. Wang, Y. Yu, G. Wen, Dynamical Analysis of the Lorenz-84 Atmospheric Circulation Model, J. Appl. Math. 2014 (2014) e296279, http://dx.doi.org/10.1155/2014/296279.

[55] P. Cessi, A Simple Box Model of Stochastically Forced Thermohaline Flow, J. Phys. Oceanogr. 24 (1994) 1911–1920, http://dx.doi.org/10.1175/1520-0485(1994)024<1911:ASBMOS>2.0.CO;2.

[56] P.J. Roebber, Climate variability in a low-order coupled atmosphere-ocean model, Tellus A 47 (1995) 473–494, http://dx.doi.org/10.3402/tellusa.v47i4.11534.

[57] L. Van Veen, T. Opsteegh, F. Verhulst, Active and passive ocean regimes in a low-order climate model, Tellus A 53 (2001) 616–628, http://dx.doi.org/10.3402/tellusa.v53i5.12229.

[58] J.G. Freire, C. Bonatto, C.C. DaCamara, J.A.C. Gallas, Multistability, phase diagrams, and intransitivity in the Lorenz-84 low-order atmospheric circulation model, Chaos 18 (2008) 033121, http://dx.doi.org/10.1063/1.2953589.

[59] M. Engel, M.A. Gkogkas, C. Kuehn, Homogenization of coupled fast-slow systems via intermediate stochastic regularization, J. Stat. Phys. 183 (2021) 25, http://dx.doi.org/10.1007/s10955-021-02765-7.

[60] D. Kelly, I. Melbourne, Deterministic homogenization for fast–slow systems with chaotic noise, J. Funct. Anal. 272 (2017) 4063–4102, http://dx.doi.org/10.1016/j.jfa.2017.01.015.

[61] G.A. Gottwald, I. Melbourne, Homogenization for deterministic maps and multiplicative noise, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 469 (2013) 20130201, http://dx.doi.org/10.1098/rspa.2013.0201.

[62] J.A. Yorke, E.D. Yorke, Metastable chaos: The transition to sustained chaotic behavior in the Lorenz model, J. Stat. Phys. 21 (1979) 263–277, http://dx.doi.org/10.1007/BF01011469.

[63] C. Grebogi, E. Ott, J.A. Yorke, Critical Exponent of Chaotic Transients in Nonlinear Dynamical Systems, Phys. Rev. Lett. 57 (1986) 1284–1287, http://dx.doi.org/10.1103/PhysRevLett.57.1284.

[64] B. Epstein, M. Sobel, Life Testing, J. Amer. Statist. Assoc. 48 (1953) 486–502, http://dx.doi.org/10.1080/01621459.1953.10483488.

[65] C. Grebogi, E. Ott, J.A. Yorke, Super persistent chaotic transients, Ergodic Theory Dynam. 5 (1985) 341–372, http://dx.doi.org/10.1017/S014338570000300X.

[66] P.D.L. Ritchie, J.J. Clarke, P.M. Cox, C. Huntingford, Overshooting tipping point thresholds in a changing climate, Nature 592 (2021) 517–523, http://dx.doi.org/10.1038/s41586-021-03263-2.

[67] P.M. Battelino, C. Grebogi, E. Ott, J.A. Yorke, E.D. Yorke, Multiple coexisting attractors, Basin boundaries and basic sets, Physica D 32 (1988) 296–305, http://dx.doi.org/10.1016/0167-2789(88)90057-7.

[68] P.A. Cassak, J.F. Drake, M.A. Shay, B. Eckhardt, Onset of Fast Magnetic Reconnection, Phys. Rev. Lett. 98 (2007) 215001, http://dx.doi.org/10.1103/PhysRevLett.98.215001.

[69] J. Vollmer, T.M. Schneider, B. Eckhardt, Basin boundary, edge of chaos and edge state in a two-dimensional model, New J. Phys. 11 (2009) 013040, http://dx.doi.org/10.1088/1367-2630/11/1/013040.

[70] M. Gelbrecht, V. Lucarini, N. Boers, J. Kurths, Analysis of a bistable climate toy model with physics-based machine learning methods, Eur. Phys. J. Spec. Top. 230 (2021) 3121–3131, http://dx.doi.org/10.1140/epjs/s11734-021-00175-0.

[71] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn, Lyapunov Characteristic Exponents for smooth dynamical systems and for Hamiltonian systems; A method for computing all of them. Part 2: Numerical application, Meccanica 15 (1980) 21–30, http://dx.doi.org/10.1007/BF02128237.

[72] G. Datseris, DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics, J. Open Source Softw. 3 (2018) 598, http://dx.doi.org/10.21105/joss.00598.

[73] G.-H. Hsu, E. Ott, C. Grebogi, Strange saddles and the dimensions of their invariant manifolds, Phys. Lett. A 127 (1988) 199–204, http://dx.doi.org/10.1016/0375-9601(88)90102-8.

[74] H. Krakovská, C. Kuehn, I.P. Longo, Resilience of dynamical systems, Eur. J. Appl. Math. (2023) http://dx.doi.org/10.1017/S0956792523000141.

[75] C.S. Holling, Resilience and Stability of Ecological Systems, Annu. Rev. Ecol. Syst. 4 (1973) 1–23, http://dx.doi.org/10.1146/annurev.es.04.110173.000245.

[76] B. Kaszás, U. Feudel, T. Tél, Tipping phenomena in typical dynamical systems subjected to parameter drift, Sci. Rep. 9 (2019) 8654, http://dx.doi.org/10.1038/s41598-019-44863-3.

[77] M. Willeit, A. Ganopolski, A. Robinson, N.R. Edwards, The Earth system model CLIMBER-X v1.0–Part 1: Climate model description and validation, Geosci. Model Dev. 15 (2022) 5905–5948, http://dx.doi.org/10.5194/gmd-15-5905-2022.

[78] O. Mehling, K. Bellomo, M. Angeloni, C. Pasquero, J. von Hardenberg, High-latitude precipitation as a driver of multicentennial variability of the AMOC in a climate model of intermediate complexity, Clim. Dyn. 61 (2023) 1519–1534, http://dx.doi.org/10.1007/s00382-022-06640-3.

[79] R. Börner, O. Mehling, J. von Hardenberg, V. Lucarini, Melancholia State of the Atlantic Meridional Overturning Circulation in an Intermediate-Complexity Climate Model, 2024, CC14C-1217, Ocean Sciences Meeting 2024, 18–23 February 2024, https://agu.confex.com/agu/OSM24/meetingapp.cgi/Paper/1488861.

[80] H.A. Dijkstra, W. Weijer, Stability of the Global Ocean Circulation: Basic Bifurcation Diagrams, J. Phys. Oceanogr. 35 (2005) 933–948, http://dx.doi.org/10.1175/JPO2726.1.

[81] M. Cini, G. Zappa, F. Ragone, S. Corti, Simulating AMOC tipping driven by internal climate variability with a rare event algorithm, npj Clim. Atmos. Sci. (2023) in press, Preprint https://doi.org/10.21203/rs.3.rs-3215995/v1.

[82] A. Romanou, D. Rind, J. Jonas, R. Miller, M. Kelley, G. Russell, C. Orbe, L. Nazarenko, R. Latto, G.A. Schmidt, Stochastic Bifurcation of the North Atlantic Circulation under a Midrange Future Climate Scenario with the NASA-GISS ModelE, J. Clim. 36 (2023) 6141–6161, http://dx.doi.org/10.1175/JCLI-D-22-0536.1.

[83] O. Mehling, R. Börner, Code for the coupled Stommel–L84 model, version 1.0, Zenodo [Software] (2023), https://doi.org/10.5281/zenodo.10370900.

[84] V. Lucarini, T. Bódai, Transitions across Melancholia States in a Climate Model: Reconciling the Deterministic and Stochastic Points of View, Phys. Rev. Lett. 122 (2019) 158701, http://dx.doi.org/10.1103/PhysRevLett.122.158701.

[85] V. Lucarini, T. Bódai, Global stability properties of the climate: Melancholia states, invariant measures, and phase transitions, Nonlinearity 33 (2020) R59, http://dx.doi.org/10.1088/1361-6544/ab86cc.

[86] V.M. Gálfi, V. Lucarini, F. Ragone, J. Wouters, Applications of large deviation theory in geophysical fluid dynamics and climate science, Riv. Nuovo Cimento 44 (2021) 291–363, http://dx.doi.org/10.1007/s40766-021-00020-z.