https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Determining Optimal Noise Induced Tipping Transition Paths Using Deep Reinforcement Learning

Francesco Guardamagna[1] and Henk A. Dijkstra[1]

[1]Institute for Marine and Atmospheric research Utrecht, Department of Physics, Utrecht University, the Netherlands

Correspondence: Henk A. Dijkstra (h.a.dijkstra@uu.nl)

Abstract. Noise-induced transitions between coexisting climate model states are rare in the weak-noise regime. Their most probable pathways can be characterized by instantons, which minimize a large-deviation action. Traditional methodologies for instanton computation present important limitations when applied to high-dimensional dynamical systems as they suffer from the curse of dimensionality. We here propose a Deep Reinforcement Learning approach to compute instantons which uses short model simulations to estimate segment costs and does not require derivatives of the model drift. To demonstrate its performance, we apply this method to determine optimal transition paths between strong and collapsed Atlantic Meridional Overturning Circulation states in two conceptual models. We compare these instanton results with those from the geometric Minimum Action Method and find good quantitative agreement. The results demonstrate feasibility of instanton computation by Deep Reinforcement Learning methods for low-dimensional conceptual models and motivate further work on sampling efficiency, noise structure, and validation before application to high-dimensional climate models.

1 Introduction

According to dynamical systems theory (Ashwin et al., 2012), transitions can occur through at least three distinct mechanisms: bifurcation-induced tipping (B-tipping), rate-induced tipping (R-tipping), and noise-induced tipping (N-tipping). In the first two cases, tipping occurs as a consequence of changes in external forcing on timescales comparable to (B-tipping) or faster than (R-tipping) the system's intrinsic relaxation time. In contrast, N-tipping occurs in multistable systems when stochastic fluctuations induce transitions between coexisting stable states. Because N-tipping arises solely from noise and may occur even under constant external forcing, it is generally more difficult to identify and predict than B- and R-tipping. This is particularly true in the low-noise regime, where noise-induced transitions become rare events. Large deviation theory characterizes their most probable pathway through minimization of the Freidlin–Wentzell action (Freidlin and Wentzell, 1998). The resulting instanton identifies both the trajectory most likely followed by the system during a transition and the optimal realization of the stochastic forcing responsible for triggering it. Instantons therefore provide a useful tool for characterizing noise-induced transitions and understanding the mechanisms through which stochastic fluctuations drive N-tipping.

Computing an instanton replaces the sampling of many rare transitions by an optimization problem over trajectories. For additive Gaussian noise, the action measures the accumulated, covariance-weighted departure of a path from the deterministic drift. Minimizing this cost identifies the most probable route subject to a specified endpoint or event constraints. The noise

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

covariance is therefore part of the problem: changing which variables are forced, or their relative forcing amplitudes, can change the optimal path. The Minimum Action Method (MAM) discretizes a trajectory over a prescribed time interval and minimizes its discretized Freidlin–Wentzell action (Weinan et al., 2004). For transitions between equilibria, however, the optimal duration can become infinite, and a uniform time grid may allocate many points near equilibria while poorly resolving the transition itself. The geometric minimum action method (gMAM) addresses this difficulty by eliminating the time parametrization and optimizing a curve in state space, with redistribution of discretization points along its arclength (Heymann and Vanden-Eijnden, 2008). This makes it well suited to computing the transition geometry without prescribing a transition duration.

Related approaches exploit different structures of the variational problem. For gradient dynamics with isotropic noise, string methods evolve a discretized curve toward a minimum-energy path using a 'force' perpendicular to the curve and repeated re-parametrization (Weinan et al., 2002). Such energy-based constructions do not directly characterize general non-gradient systems, for which the action formulation is needed. Alternatively, the stationarity conditions of the action can be written as coupled equations for the state and a conjugate variable representing the optimal forcing. Numerical solutions of these instanton equations, including iterative forward–backward integration, have been developed for stochastic fluid dynamical models (Grafke et al., 2015). All these methods differ in their computational requirements, but suffer from the curse of dimensionality. In addition, the standard gMAM relaxation and instanton-equation solvers all involve derivatives of the drift. This limits their applicability to complex and high-dimensional systems.

In this paper, we use the Deep Deterministic Policy Gradient (DDPG) approach, a Deep Reinforcement Learning method, to compute instantons associated with noise-induced tipping in climate models. This method was originally developed to identify optimal transition paths between metastable states in molecular systems (Lin et al., 2025). It does not suffer from the curse of dimensionality and does not need information on the derivatives. We apply the DDPG framework here to compute instantons of two conceptual ocean circulation models and benchmark the resulting transition paths against those computed with gMAM. Such conceptual models provide controlled settings in which numerical approaches to these rare transitions can be compared. Our objective is to test approximation accuracy and identify practical limitations; scalability to high-dimensional climate models is not established by these experiments.

2 Models and minimum-action formulation

2.1 Conceptual models

The Atlantic Meridional Overturning Circulation (AMOC) has long been recognized as a climate tipping element. Its potential multistability (Weijer et al., 2019; Jackson and Wood, 2018; van Westen et al., 2024) and the climatic consequences of a transition has motivated studies of a noise-induced collapse (Castellana et al., 2019; Cini et al., 2024; Oh et al., 2025). The cornerstone AMOC model is the Stommel two-box model (Stommel, 1961), representing the overturning flow in the North Atlantic between subtropical and polar ocean reservoirs through non-dimensional meridional temperature and salinity contrasts

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

T and S. Following Lohmann et al. (2025), the dimensionless equations of a stochastic extension of this model are

$$dT_t = (\eta_1 - T - |T - S|T)\, dt + \sqrt{\epsilon}\, dW_t^T, \tag{1a}$$

$$dS_t = (\eta_2 - \eta_3 S - |T - S|S)\, dt + \sqrt{\epsilon}\, dW_t^S, \tag{1b}$$

where $q = T - S$ measures the overturning strength. Here η_1 represents the temperature forcing, η_2 the salinity forcing, η_3 is a ratio of surface adjustment time scales of temperature and salinity. ϵ represent the noise strength and the Wiener processes W_t^T and W_t^S are independent.

While still a conceptual model, the Wood model is a box model of the global ocean. It represents the North Atlantic, Atlantic thermocline, Southern Ocean, Indo-Pacific thermocline, and deep ocean by five interacting boxes (Wood et al., 2019). We use the deterministic formulation as presented in Soons et al. (2024), where the state vector is four dimensional (consisting of salinities in the different boxes). However, we replace their degenerate, box-dependent stochastic forcing as in Soons et al. (2024) by independent additive noise of equal amplitude in all boxes. This defines a methodological benchmark rather than a physically calibrated stochastic AMOC model, and its equations appear in appendix A. Consequently, comparison with the published instanton of Soons et al. (2024) is qualitative; the quantitative benchmark here uses gMAM for the same isotropic-noise system as DDPG.

2.2 Instantons and geometric action

Both the Stommel and Wood model can be formulated as a general stochastic dynamical system described by

$$dX_t = f(X_t)\, dt + \sqrt{\epsilon}\sigma\, dW_t, \tag{2}$$

where $X_t \in \mathbb{R}^d$ is the state vector, $f : \mathbb{R}^d \to \mathbb{R}^d$ defines the deterministic drift, $\sigma = \mathrm{I}_d$ is the noise (identity) matrix, and W_t is a d-dimensional Wiener process.

Consider a trajectory ϕ such that $\phi(0) = x$ and $\phi(\tau) = y$, where τ indicates a chosen end time. This transition may be impossible in the absence of noise ($\epsilon = 0$) but becomes possible for $\epsilon > 0$. In the limit $\epsilon \to 0$, the instanton is the trajectory ϕ^* that minimizes the Freidlin–Wentzell action functional S_t subject to the boundary conditions, i.e.,

$$\phi^*(t) = \underset{\phi(0)=x,\ \phi(\tau)=y}{\operatorname{argmin}} S_\tau[\phi]. \tag{3}$$

with

$$S_\tau[\phi] = \frac{1}{2}\int_0^\tau \Big\langle \dot\phi(t) - f(\phi(t)), a^{-1}\big(\dot\phi(t) - f(\phi(t))\big)\Big\rangle\, dt, \tag{4}$$

and $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product, the dot indicates differentiation to time, and a^{-1} is the inverse of the covariance matrix of the noise.

To eliminate the explicit dependence on time, gMAM (Heymann and Vanden-Eijnden, 2008) reformulates the minimization problem geometrically in state space. The transition trajectory is represented as a curve $\phi(s)$, parametrized by the normalized

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

arclength parameter $s \in [0,1]$. The Freidlin–Wentzell action functional can then be rewritten as

$$S_g[\phi] = \int_0^1 \left[\sqrt{\langle \phi'(s), a^{-1}\phi'(s) \rangle \langle f(\phi(s)), a^{-1} f(\phi(s)) \rangle} - \langle \phi'(s), a^{-1} f(\phi(s)) \rangle \right] ds, \tag{5}$$

where $\phi'(s) = d\phi/ds$. The instanton corresponds to the curve minimizing this geometric action functional.

The minimization problem is solved numerically using the geometric minimum action method (gMAM), with slightly different numerical implementations for the Stommel and Wood models. For the Stommel model, the Euler–Lagrange equations associated with S_g are derived, and an artificial optimization time (τ) is introduced. The resulting relaxation equations are numerically integrated in artificial time to progressively evolve the transition path toward the minimum-action trajectory. For the Wood model, instead, the geometric action is discretized using midpoint quadrature, and the minimization is performed directly through a gradient-based optimization procedure.

In both cases, the path is discretized into a finite number of points, $(\phi_0, \phi_1, \ldots, \phi_N)$, with fixed boundary conditions $(\phi_0 = x)$ and $(\phi_N = y)$. Starting from an initial guess, typically a straight-line interpolation connecting the two stable states, the discretized path is iteratively updated to minimize S_g. During the optimization, the path is periodically re-parameterized in terms of normalized arc length to maintain an approximately uniform distribution of discretization points along the trajectory. Both approaches require evaluating the Jacobian matrix of the deterministic drift. The resulting converged trajectory corresponds to the gMAM approximation of the instanton.

3 Reinforcement-learning approach

Consider again the stochastic dynamical system (2) and assume that, in the absence of noise ($\epsilon = 0$), the system possesses two stable states, denoted by A and B. In the low-noise limit $\epsilon \to 0$, transitions from A to B concentrate around the instanton, denoted by $\phi^\star$, which minimizes the Freidlin–Wentzell action functional (4).

3.1 Reformulation of the minimization problem

To reformulate this minimization problem in a discrete setting, we introduce, for a small parameter $\gamma > 0$, a path space C_γ consisting of continuous piecewise linear functions. Any $\phi \in C_\gamma$ connecting the stable states A and B is represented by a finite sequence of states $(z_0, \ldots, z_N)$ with corresponding time points $0 = t_0 < t_1 < \cdots < t_N = T$, satisfying

$$z_0 = A, \qquad z_N = B, \tag{6a}$$

$$|z_{i+1} - z_i| = \gamma, \qquad 0 \le i \le N-2, \tag{6b}$$

$$|z_i - z_{i-1}| \le \gamma, \qquad i = N. \tag{6c}$$

On each interval $[t_i, t_{i+1}]$, the function $\phi(t)$ is defined as the straight line connecting z_i and z_{i+1}, i.e.,

$$\phi(t) = z_i + \frac{t - t_i}{h_i}(z_{i+1} - z_i), \qquad t_i \le t \le t_{i+1}, \tag{7}$$

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

where $h_i = t_{i+1} - t_i$.

The minimization problem can then be reformulated in the discrete space C_γ as

$$\phi^\star = \underset{\phi \in C_\gamma}{\arg\min} \sum_{i=0}^{N-1} L(z_i, z_{i+1}; h_i) \; ; \; L(z_i, z_{i+1}; h_i) = \int_0^{h_i} \left| \frac{z_{i+1} - z_i}{h_i} - f\big(\phi^i(t)\big) \right|^2 dt, \tag{8}$$

where $\phi^i(t) = \phi(t + t_i), \qquad t \in [0, h_i]$. Approximating the integral using a midpoint numerical quadrature and minimizing with respect to h_i, we obtain

$$\min_{h_i} L(z_i, z_{i+1}; h_i) \approx \min_{h_i} h_i \left| \frac{z_{i+1} - z_i}{h_i} - f(z_{i+\frac{1}{2}}) \right|^2 \tag{9}$$

where

$$z_{i+\frac{1}{2}} = \frac{z_i + z_{i+1}}{2}, \qquad \Delta z_i = z_{i+1} - z_i.$$

Minimizing (9) with respect to h_i yields the cost

$$R(z_i, z_{i+1}) = 2|z_{i+1} - z_i|\,|f(z_{i+\frac{1}{2}})| - 2\left\langle z_{i+1} - z_i, f(z_{i+\frac{1}{2}}) \right\rangle. \tag{10}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The minimization problem can be reformulated as

$$\phi^\star = \underset{(z_0, \ldots, z_N)}{\arg\min} \sum_{i=0}^{N-1} R(z_i, z_{i+1}). \tag{11}$$

In Lin et al. (2025), the drift term $f\left(z_{i+\frac{1}{2}}\right)$ is replaced by an effective force $F_\epsilon\left(z_{i+\frac{1}{2}}\right)$ in order to handle rough potential landscapes. This quantity is approximated by integrating (2) forward in time over a small interval $[0, h]$, starting from $x_0 = z_{i+\frac{1}{2}}$ and averaging over M independent noisy trajectories in the small-noise regime:

$$F_\epsilon\left(z_{i+\frac{1}{2}}\right) \approx \frac{1}{M} \sum_{j=1}^{M} \frac{x_j^e - z_{i+\frac{1}{2}}}{h}, \tag{12}$$

where x_j^e denotes the state at time h for the j-th realization. Replacing $f\left(z_{i+\frac{1}{2}}\right)$ with $F_\epsilon\left(z_{i+\frac{1}{2}}\right)$ leads to the modified cost

$$R_\epsilon(z_i, z_{i+1}) = 2|z_{i+1} - z_i| \left| F_\epsilon\left(z_{i+\frac{1}{2}}\right) \right| - 2\left\langle z_{i+1} - z_i, F_\epsilon\left(z_{i+\frac{1}{2}}\right) \right\rangle. \tag{13}$$

The corresponding minimization problem becomes

$$\phi^\star = \underset{(z_0, \ldots, z_N)}{\arg\min} \sum_{i=0}^{N-1} R_\epsilon(z_i, z_{i+1}). \tag{14}$$

For a detailed discussion on the construction of the effective force $F_\epsilon\left(z_{i+\frac{1}{2}}\right)$, and the motivation for replacing the drift term $f\left(z_{i+\frac{1}{2}}\right)$, we refer the reader to Lin et al. (2025).

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

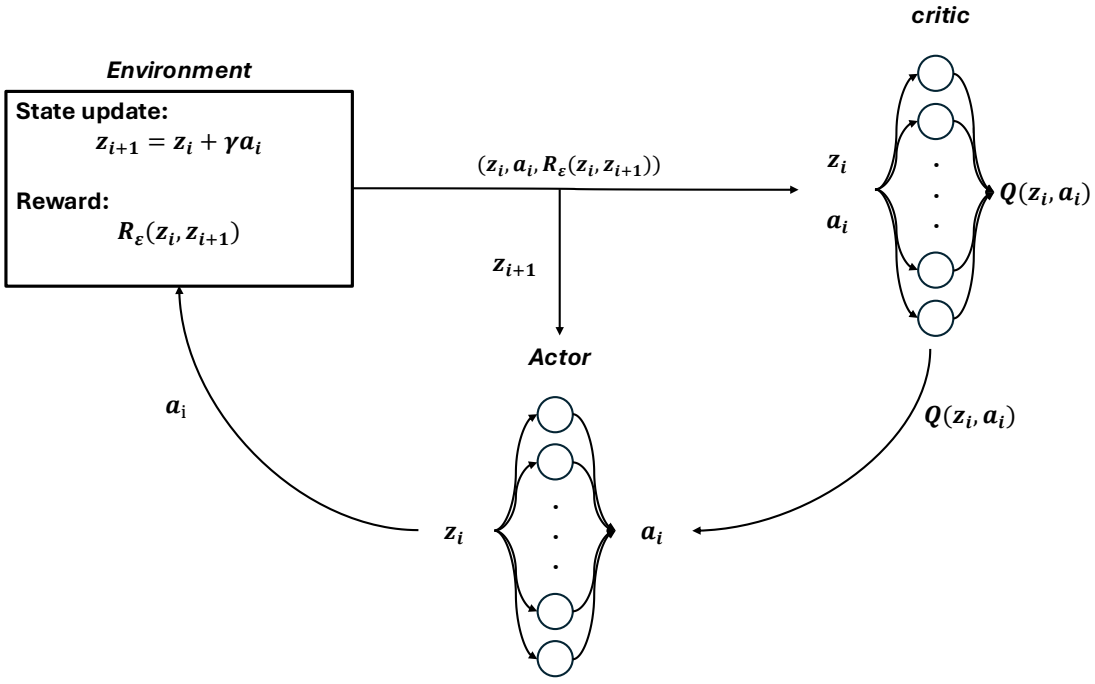

Figure 1. Simplified schematic representation of the DDPG framework. For clarity, the replay buffer and target networks are omitted.

3.2 DDPG solution of the minimization problem

To solve the minimization problem (14), a DDPG-based framework is adopted. The framework consists of two neural networks, commonly referred to as the *actor* and the *critic*, which are trained jointly (Fig. 1). The actor network receives as input the current system state $z_i \in \mathbb{R}^d$ and outputs an action a_i, belonging to the continuous action space $\mathcal{A} = \{x \in \mathbb{R}^d : |x| = 1\}$. The selected action a_i is then used to update the system state as $z_{i+1} = z_i + \gamma a_i$, where $\gamma > 0$ is a hyperparameter selected a priori. After the state update, the actor receives a cost $r_i = R_\epsilon(z_i, z_{i+1})$, computed according to (13).

The actor's behavior can be described by a policy π. The return associated with a state z_i is defined as the cumulative future cost $R_i = r_i + \cdots + r_N$, where r_N denotes the cost at the final step. The objective is to learn an optimal policy π^* minimizing the return for every state $z \in \mathbb{R}^d$. The state–action function $Q(z, a)$ associated with the optimal policy π^* gives the expected cumulative return obtained after taking action a_i in state z_i and subsequently following the optimal policy π^*. In deep reinforcement learning, the function $Q(z, a)$ is usually approximated using the recursive Bellman equation

$$Q(z_i, a_i) = r_i + \min_{b \in \mathcal{A}} Q(z_{i+1}, b). \tag{15}$$

where z_{i+1} is the next state obtained after applying action a_i at state z_i.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

The critic network

$$\tilde{Q}_\theta(z,a) : \mathbb{R}^d \times \mathcal{A} \to [0,1], \tag{16}$$

where θ denotes the set of trainable network parameters, receives as input the current state z_i together with the action a_i predicted by the actor, and outputs a scalar value in the interval $[0,1]$ through a sigmoid activation function. The action-value function is then parameterized as

$$Q_\theta(z,a) = \begin{cases} \lambda \tilde{Q}_\theta(z,a), & z \notin \Omega_\delta, \\ R_\epsilon(z,B), & z \in \Omega_\delta, \end{cases} \tag{17}$$

where Ω_δ denotes a neighborhood of the target state B, defined as

$$\Omega_\delta = \left\{ x \in \mathbb{R}^d : |x - B| < \delta \right\} \tag{18}$$

and $\lambda > 0$ is a hyperparameter defining the output range of the critic function as $[0,\lambda]$.

The actor function is defined so as to produce output actions mapped to unit vectors in $\mathbb{R}^d$:

$$\mu_\theta(z) = \Theta[\cos(\tilde{\mu}_\theta(z))], \tag{19}$$

where $\tilde{\mu}_\theta(z) \in \mathbb{R}^d$ denotes the raw output of the actor network, while $\Theta[v] = v/|v|$ is the normalization function for an input vector $v \in \mathbb{R}^d$. The cosine function is applied component-wise to constrain the entries of the raw actor output to the interval $[-1,1]$ before normalization.

Both the actor and critic networks are implemented as multilayer perceptrons (MLPs) with two hidden layers, each containing n_{nodes} neurons. The DDPG training procedure is conducted in episodes. At the beginning of each episode, the system state is initialized as $z_0 = A$. The state is subsequently updated according to the actor network $\mu_\theta(z)$, until either the actor enters the neighborhood Ω_δ of the target state B defined in Eq. (18) or a maximum number of steps N_s is reached. To encourage exploration of the state space and reduce the risk of converging to local minima, exploration noise is introduced during training according to

$$\tilde{a}_i = \begin{cases} \cos(\tilde{\mu}_\theta(z_i)), & \text{with probability } p_1, \\ f(z_i), & \text{with probability } p_2, \\ \cos(\tilde{\mu}_\theta(z_i) + \xi_i), & \text{with probability } 1 - p_1 - p_2, \end{cases} \tag{20}$$

where $\xi_i \in \mathbb{R}^d$ is sampled from a Gaussian distribution with mean 0 and standard deviation $\pi/4$, $p_1 > 0$, $p_2 > 0$, and $1 - p_1 - p_2 > 0$. The final action is then obtained through normalization as

$$a_i = \Theta[\tilde{a}_i]. \tag{21}$$

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

During training, N_p episodes are executed in parallel. At each state update, every parallel episode generates a transition of the form (z_i, a_i, r_i, z_{i+1}), which is stored in a replay buffer $\mathcal{R}$ of fixed capacity N_r. Once the buffer is full, the oldest transitions are discarded as new ones are added. The replay buffer enables off-policy training by allowing the actor and critic networks to reuse transitions collected during previous stages of training. After all parallel episodes complete one environment step, a mini-batch $\mathcal{B} = \{(z_i, a_i, r_i, z_{i+1})\}_{i=1}^{b_s}$, of size b_s is sampled from $\mathcal{R}$ and used to perform one update of both the actor and critic networks. An epoch is completed once all parallel episodes satisfy the terminal condition. The overall training procedure terminates after n_{ep} epochs.

To improve training stability, target actor and critic networks, denoted respectively by $\mu'_\theta(z)$ and $Q'_\theta(z, a)$, are introduced. The target networks are initialized as copies of the actor and critic networks and are subsequently updated after each training epoch by copying the current parameters of the actor and critic networks. The target networks are used to update the critic network through the loss function

$$L_Q(\theta) = \frac{1}{b_s} \sum_{(z_i, a_i, r_i, z_{i+1}) \in \mathcal{B}} |Q_\theta(z_i, a_i) - y_i|^2, \tag{22}$$

where $y_i = r_i + Q'_\theta\left(z_{i+1}, \mu'_\theta(z_{i+1})\right)$. The actor network is updated using the gradient of the average return J_μ over the batch $\mathcal{B}$, defined as

$$\nabla_\theta J_\mu = \frac{1}{b_s} \sum_{z_i \in \mathcal{B}} \nabla_a Q_\theta(z_i, a)\big|_{a=\mu_\theta(z_i)} \cdot \nabla_\theta \mu_\theta(z_i). \tag{23}$$

In both cases, the network parameters are optimized using the Adam optimizer with learning rate lr (Kingma and Ba, 2014). The complete set of hyperparameters of the DDPG-based framework therefore includes the state update step γ, the terminal threshold δ, the critic scaling parameter λ, the exploration probabilities p_1 and p_2, the learning rate lr, the integration horizon h used in the estimation of the effective force F_ϵ, the number M of integrated trajectories adopted to estimate F_ϵ, the batch size b_s, the replay buffer size N_r, the maximum number of steps per episode N_s, the number of parallel episodes N_p, the number of training epochs n_{ep}, and the number of neurons n_{nodes} in the hidden layers of the actor and critic networks.

4 Results

4.1 Stommel model

For the Stommel model, we fix $\eta_2 = 1$, $\eta_3 = 0.3$ and following Lohmann et al. (2025), the parameter η_1 is used as the control parameter. In the deterministic limit ($\epsilon = 0$), the model exhibits bistability over a range of η_1 (Fig. 2), characterized by the coexistence of two stable equilibria corresponding to a strong AMOC ON state ($q > 0$) and a collapsed or reversed AMOC OFF state ($q \leq 0$). Decreasing η_1 from 3.5 causes the stable AMOC ON state, corresponding to a strong overturning circulation ($q > 0$), to disappear through a saddle-node bifurcation occurring approximately at $\eta_1 = 2.53$. Above $\eta_1 = 3.35$, the OFF equilibrium disappears, leaving only the stable AMOC ON state. The model therefore exhibits bistability for $2.53 < \eta_1 < 3.35$, where stable ON and OFF equilibria coexist. For the instanton computation, we consider $\eta_1 = 2.57$, 2.8, and 3.1, all within the bistable regime.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

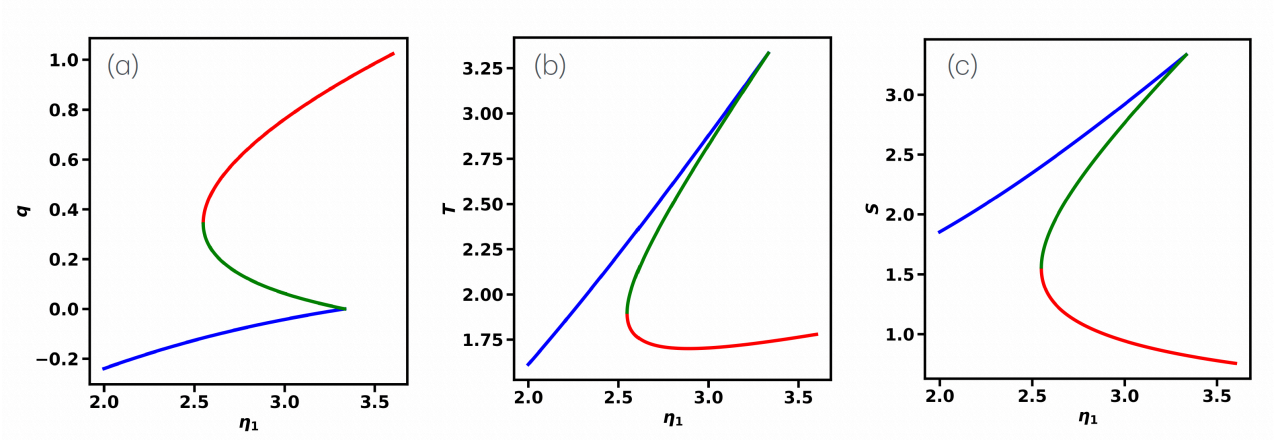

Figure 2. Bifurcation diagrams of the Stommel model for $\eta_2 = 1.0$ and $\eta_3 = 0.3$. Panels (a), (b), and (c) show the equilibrium values of overturning strength $q = T - S$, the temperature T and salinity S, respectively, as functions of the atmospheric temperature forcing parameter η_1. The red branch corresponds to the stable AMOC ON state ($q > 0$), the blue branch to the stable AMOC OFF state ($q < 0$), and the green branch to the unstable equilibrium.

The state vector ($z \in \mathbb{R}^2$), used as input to both the Actor and Critic networks within the DDPG framework, is defined by the two dynamical variables (S) and (T). In the stochastic integration employed to estimate the effective force (F_ϵ), we adopt a noise amplitude $\sqrt{\epsilon} = 10^{-4}$. The hyperparameters of the DDPG framework were selected through manual tuning, and the complete set of adopted values is reported in Appendix C. Since the DDPG algorithm relies on extensive exploration of the state space before converging to an optimal policy, random exploration noise is added during training (see section 3.2). While essential for efficient exploration, this noise can introduce instabilities and cause the solution to temporarily deviate from the optimum. To mitigate this issue, we retain the best solution encountered during training, defined as the one that minimizes the loss function in (11), rather than using the policy obtained at the final training epoch. We empirically found that a maximum of $n_{ep} = 1000$ training epochs was sufficient to obtain the reported best paths in the considered cases; however, this does not guarantee global optimality.

Figure 3 shows the estimated instanton trajectories between the AMOC ON and OFF states obtained with the DDPG-based framework, together with the corresponding instantons computed using the gMAM. The instanton trajectories computed by the DDPG framework closely match the gMAM solutions for $\eta_1 = 2.57$ and $\eta_1 = 3.1$. For $\eta_1 = 2.8$, the DDPG framework still provides a good approximation of the instanton trajectory over most of the transition path, but discrepancies emerge in the vicinity of the OFF state. In particular the algorithm has difficulty converging exactly to the instanton trajectory and exhibits oscillatory behavior around the OFF state for several iterations before terminating. Nevertheless, the predicted trajectory remains close to the gMAM instanton throughout the transition. The larger discrepancy found for $\eta_1 = 2.8$ is somewhat unexpected. Indeed, one would anticipate the most pronounced difficulties for $\eta_1 = 3.1$, since for this parameter value the AMOC ON state is more stable and transitions toward the OFF state are therefore less likely and more difficult to realize.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

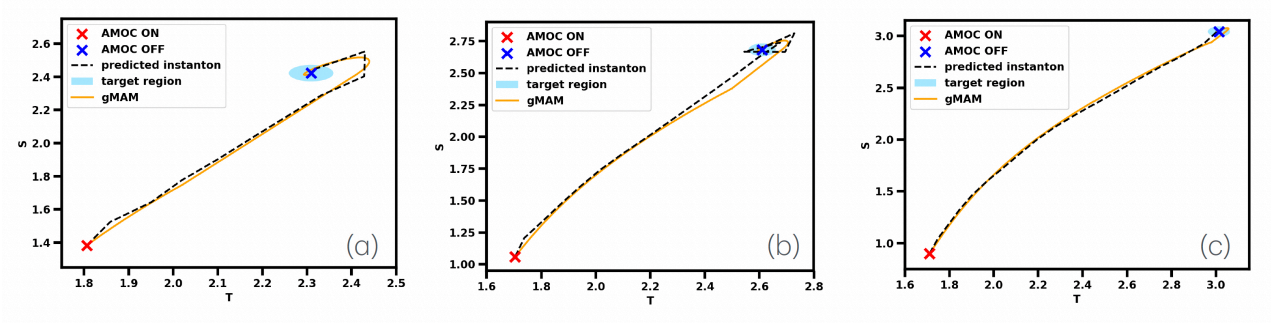

Figure 3. Instanton trajectories between the AMOC ON and OFF states for the Stommel model in the (T, S) phase space, for (a) $\eta_1 = 2.57$, (b) $\eta_1 = 2.8$, and (c) $\eta_1 = 3.1$. In each panel, the red and blue dots denote the stable AMOC ON and OFF states, respectively. The orange curve corresponds to the instanton computed using the gMAM method, while the black dashed curve represents the instanton determined by the DDPG-based framework. The blue area indicates the distance threshold δ adopted as the terminal condition for terminating a training episode (see Sect. 3).

To quantify the agreement between the instantons obtained with the DDPG framework and those computed using gMAM, we evaluate the relative path error

$$e_\phi = \frac{\|\phi_\theta - \phi\|}{\|\phi\|}, \tag{24}$$

where ϕ_θ denotes the DDPG-determined instanton and ϕ the corresponding gMAM solution. Since the two trajectories may be discretized using a different number of points, both paths are first reparametrized using the normalized arclength coordinate $\alpha \in [0, 1]$ and subsequently interpolated onto a common uniform grid. The norm appearing in (24) is defined as

$$\|\phi\| = \left(\frac{1}{n_e} \sum_{i=1}^{n_e} |\phi_i|^2 \right)^{1/2}, \tag{25}$$

where ϕ_i denotes the value of the trajectory at the i-th interpolation point of the common arclength grid, and n_e is the total number of interpolation points, here chosen as $n_e = 500$. Consequently, e_ϕ provides a global measure of the relative discrepancy between the DDPG and gMAM instantons. It is independent of the original point counts after interpolation, but dependent on coordinate scaling and the arclength convention.

The resulting relative errors are $e_\phi = 9.21 \times 10^{-3}$ for $\eta_1 = 2.57$, $e_\phi = 5.19 \times 10^{-2}$ for $\eta_1 = 2.8$, and $e_\phi = 2.20 \times 10^{-2}$ for $\eta_1 = 3.1$. These values confirm the visual comparison of the trajectories: the agreement between the two methodologies is excellent for $\eta_1 = 2.57$ and $\eta_1 = 3.1$, while the largest discrepancy is found for $\eta_1 = 2.8$. However, even in this case, the relative error remains below 6%, indicating that the DDPG framework is able to capture the transition path determined by gMAM. The DDPG-based framework hence provides a good approximation of the ON $\rightarrow$ OFF instanton trajectories for the three η_1 values considered.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.2 Wood model

For the Wood model (Section 2.1 and Appendix A), the parameter values adopted for $(V_i, F_i, K_i, T_S, T_0, \mu, \lambda, \eta, \gamma)$ as well as the total salinity C are shown in appendix B. For this parameter configuration, the deterministic system ($\epsilon = 0$) exhibits the coexistence of two stable equilibria, corresponding to the AMOC ON and OFF states. The corresponding bifurcation diagrams can be found in Fig. 3 of Alkhayuon et al. (2019) and is not shown here, because we only consider one parameter configuration. As for the Stommel model, the instanton trajectories which are used as benchmark solutions for evaluating the DDPG-based approximation are computed using the gMAM.

The state vector $z \in \mathbb{R}^4$ is defined by the four state variables $(\phi_1, \phi_2, \phi_3, \phi_4)$, which fully characterize the state of the Wood model. The same DDPG framework formulation adopted for the Stommel model is also used for the Wood model, with one minor modification. Before being provided as input to the Actor and Critic networks, the system states z are normalized according to

$$z_{\text{norm}} = \frac{z - \dfrac{z_{\text{ON}} + z_{\text{OFF}}}{2}}{|z_{\text{OFF}} - z_{\text{ON}}|}, \tag{26}$$

where z_{ON} and z_{OFF} denote the stable AMOC ON and OFF equilibrium states, respectively. This normalization centers the state space around the midpoint between the two equilibria and rescales it by their distance, improving the numerical stability of the training procedure and facilitating learning for the Actor and Critic networks. The DDPG-framework hyperparameters have been optimized manually and the full list of parameters has been reported in appendix D. A noise amplitude $\sqrt{\epsilon} = 1 \times 10^{-5}$ was adopted as representative of the small-noise regime for the stochastic model integration used in the approximation of the effective force F_ϵ.

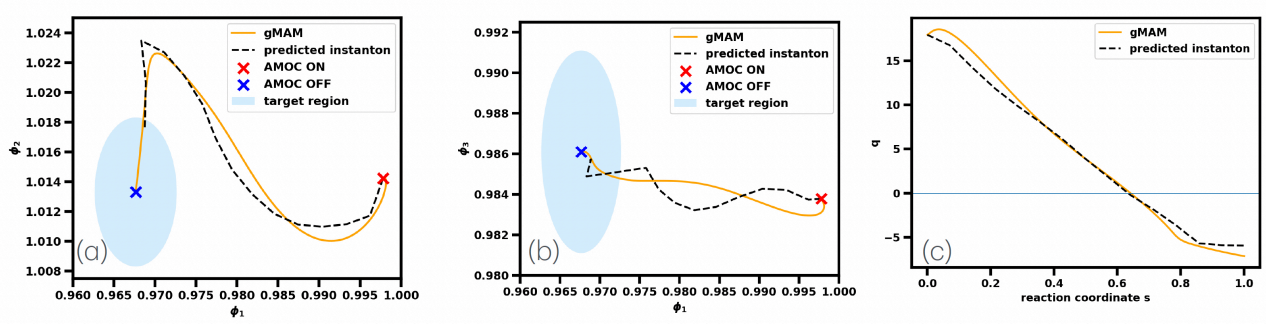

Figure 4. Instanton trajectories between the AMOC ON and OFF states for the Wood model projected onto the (a) (ϕ_1, ϕ_2) and (b) (ϕ_1, ϕ_3) phase spaces. In both panels, the red and blue dots denote the stable AMOC ON and OFF equilibria, respectively. The orange curve corresponds to the instanton computed using the gMAM, while the black dashed curve represents the instanton predicted by the DDPG-based framework. The blue area indicates the distance threshold δ adopted as the terminal condition for terminating a training episode (see Sect. 3). Panel (c) shows the corresponding AMOC strength q as a function of the reaction coordinate s, where $s \in [0, 1]$ parameterizes the instanton trajectory along its arc length in the gMAM formulation.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Figure 4 shows the instanton trajectories computed by the DDPG framework for the Wood model, together with the benchmark instantons computed using the gMAM methodology. The corresponding AMOC strength q as a function of the arclength s is also shown, where $s \in [0, 1]$ parameterizes the instanton trajectory in the gMAM, with $s = 0$ corresponding to the AMOC ON state and $s = 1$ to the AMOC OFF state (see section 2.2). Again, the instanton trajectories obtained with the gMAM methodology are qualitatively consistent with those reported in Soons et al. (2024). The DDPG-based framework provides a very good approximation of the instanton trajectory in the (ϕ_1, ϕ_2) phase space. The approximation in the (ϕ_1, ϕ_3) phase space is less accurate, but it still closely follows the benchmark instanton trajectory. This behavior is also reflected in the corresponding AMOC strength q, which is reproduced with good accuracy by the DDPG-based approximation (Figure 4c). The quantitative evaluation is consistent with the qualitative comparison shown in Fig. 4. The relative error, as measured by the metric e_ϕ, is equal to 1.19×10^{-3}, confirming the very high level of agreement between the DDPG-derived trajectory and the benchmark instanton.

5 Summary and Discussion

When a system is in a bistable regime, transitions between its two stable states under constant forcing can be induced by internal stochastic fluctuations, usually referred to as N-tipping. In the low-noise limit, these transitions become rare events. According to large deviation theory, such transitions are expected to occur preferentially along a particular trajectory, known as the instanton, which represents the most probable path connecting the two stable states. The instanton characterizes also the optimal stochastic forcing required to drive the transition.

Motivated by limitations of traditional instanton computation methods (Weinan et al., 2004; Heymann and Vanden-Eijnden, 2008; Weinan et al., 2002), we here applied an instanton computation methodology based on a Deep Reinforcement Learning method, the DDPG-framework, originally developed for molecular systems (Lin et al., 2025). We demonstrated its applicability to computing instantons associated with noise-induced tipping events, using models of the Atlantic Meridional Overturning Circulation (AMOC) as case studies. We showed that the DDPG framework can successfully approximate instanton trajectories in two conceptual AMOC models of increasing complexity: the Stommel and Wood box models. The resulting trajectories are in close agreement with those obtained using the gMAM, as measured by a relative path metric. This metric is sensitive to coordinate scaling and, for the Wood model, to the large background salinity values. It does not establish comparable accuracy in ac tion barriers, transition probabilities, or individual projected components. Endpoint oscillations in the Stommel model also show that neighborhood termination and local policy behavior affect the recovered path. Multiple-seed variability, action errors, and discretization convergence would provide further validation, but are out of the scope of this paper.

These results demonstrate that the DDPG framework can provide a reliable approximation of the instanton trajectory without requiring any prior knowledge of the derivatives of the underlying dynamical system. This advantage, together with the framework's strong modularity, opens the possibility of extending the methodology to higher-dimensional climate systems. A next step is to implement the methodology for the degenerate noise case, such as in Soons et al. (2024), which is more realistic for AMOC models, as the noise acts through the ocean-atmosphere interface. The DDPG framework can also be improved.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

For example, the Actor and Critic components could be replaced by neural network architectures better suited to analyze high-dimensional spatial climate data, such as convolutional neural networks, graph neural networks, or transformer-based architectures.

One of the remaining limitations is the large number of training transitions typically required by Deep Reinforcement Learning algorithms (Moerland et al., 2023), which currently restricts their direct applicability to General Circulation Models (GCMs) due to the associated computational cost. This limitation is also likely to extend also to Earth System Models of Intermediate Complexity (EMICs) such as CLIMBER (Willeit et al., 2022) and PlaSim (Holden et al., 2016), where the large number of model evaluations required during training would still result in substantial computational expenses. Even in the low-dimensional conceptual models considered in this study, the DDPG-framework may require several days of computation on a standard CPU to converge to the optimal solution, substantially longer than the few hours typically needed by established instanton computation methods such as the gMAM. Consequently, the primary advantage of the DDPG-framework does not lie in its computational efficiency, but rather in its ability to identify instanton trajectories without requiring explicit knowledge of the model's derivatives.

Nevertheless, ongoing developments in next-generation climate and ocean modeling frameworks, natively implemented in modern programming languages such as Python and Julia (Häfner et al., 2018; Silvestri et al., 2023), together with recent advances in AI-based global ocean emulators (Dheeshjith et al., 2025; El Aouni et al., 2025; Wu et al., 2026), offer promising opportunities for overcoming these limitations. These emulators can operate orders of magnitude faster than traditional Fortran-based GCMs, potentially making Deep Reinforcement Learning instanton methodologies feasible for climate models of substantially greater complexity than those considered here.

Appendix A: Wood model equations

Following Soons et al. (2024), we define the nondimensional version of the Wood model by introducing the nondimensional salinities ϕ_i and nondimensional volumes V_i as

$$S_N = S_0\phi_1, \qquad V_N = V_0V_1, \tag{A1a}$$

$$S_T = S_0\phi_2, \qquad V_T = V_0V_2, \tag{A1b}$$

$$S_S = S_0\phi_3, \qquad V_S = V_0V_3, \tag{A1c}$$

$$S_{IP} = S_0\phi_4, \qquad V_{IP} = V_0V_4, \tag{A1d}$$

$$S_B = S_0\phi_5, \qquad V_B = V_0V_5. \tag{A1e}$$

Time is nondimensionalized according to $t = t_d\tau$. The deterministic drift is defined by the following equations; isotropic noise is added in the four independent coordinates as in Eq. (**??**). The box-flow fraction γ below is distinct from the DDPG

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

step length

$$\frac{d\phi_1}{d\tau} = \frac{t_d}{V_0 V_1}\Big(q\big[\theta(q)(\phi_2-\phi_1)-\theta(-q)(\phi_5-\phi_1)\big] + K_N(\phi_2-\phi_1)-F_N\Big), \tag{A2a}$$

$$\frac{d\phi_2}{d\tau} = \frac{t_d}{V_0 V_2}\Big(q\big[\theta(q)\big(\gamma\phi_3+(1-\gamma)\phi_4-\phi_2\big)-\theta(-q)(\phi_1-\phi_2)\big] + K_S(\phi_3-\phi_2)+K_N(\phi_1-\phi_2)-F_T\Big), \tag{A2b}$$

$$\frac{d\phi_3}{d\tau} = \frac{t_d}{V_0 V_3}\Big(\gamma q\big[\theta(q)(\phi_5-\phi_3)-\theta(-q)(\phi_2-\phi_3)\big] + K_{IP}(\phi_4-\phi_3)+K_S(\phi_2-\phi_3)+\eta(\phi_5-\phi_3)-F_S\Big), \tag{A2c}$$

$$\frac{d\phi_4}{d\tau} = \frac{t_d}{V_0 V_4}\Big((1-\gamma)q\big[\theta(q)(\phi_5-\phi_4)-\theta(-q)(\phi_2-\phi_4)\big] + K_{IP}(\phi_3-\phi_4)-F_{IP}\Big). \tag{A2d}$$

The dimensional AMOC strength q is defined as

$$q = \kappa\left(\alpha(T_S - T_0) + \beta S_0(\phi_1-\phi_3)\right). \tag{A3}$$

The nondimensional salinity of the B box is obtained from conservation of total salinity as

$$\phi_5 = \frac{C}{V_0 S_0 V_5} - \frac{V_1}{V_5}\phi_1 - \frac{V_2}{V_5}\phi_2 - \frac{V_3}{V_5}\phi_3 - \frac{V_4}{V_5}\phi_4. \tag{A4}$$

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix B: Wood Model Parameter Configuration

Parameter	Value
V_0 ($10^6\,\mathrm{m}^3$)	10^{10}
V_1	3.261
V_2	7.777
V_3	8.897
V_4	22.02
V_5	86.49
t_d (s)	3.1536×10^9
S_0 (psu)	35
C ($10^6\,\mathrm{m}^3\,\mathrm{psu}$)	$4.446304026 \times 10^{13}$
α ($\mathrm{kg\,m^{-3}\,K^{-1}}$)	0.12
β ($\mathrm{kg\,m^{-3}}$)	0.79
γ	0.39
η (Sv)	74.492
K_N (Sv)	5.456
K_S (Sv)	5.447
K_{IP} (Sv)	96.817
F_N (Sv)	0.384
F_T (Sv)	-0.723
F_S (Sv)	1.078
F_{IP} (Sv)	-0.739
T_S (K)	4.773
T_0 (K)	2.65
μ ($10^{-6}\,\mathrm{m^{-3}\,s\,K}$)	0.055
λ ($10^6\,\mathrm{m^6\,kg^{-1}\,s^{-1}}$)	34.197

Table B1. Wood model parameter values adopted for the Wood model experiments.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix C: DDPG Hyperparameters for the Stommel Model Experiments

Hyperparameter	$\eta_1 = 2.57$	$\eta_1 = 2.8$	$\eta_1 = 3.1$
γ	0.15	0.15	0.15
δ	0.05	0.05	0.05
λ	1000	1000	1000
p_1	1/3	1/3	1/3
p_2	1/3	1/3	1/3
lr	0.001	0.001	0.001
h	0.0005	0.0005	0.0005
M	50	50	50
bs	5000	5000	5000
N_r	5×10^5	5×10^5	5×10^5
N_s	1000	1000	1000
N_p	500	500	500
n_{ep}	1000	1000	1000
n_{nodes}	50	50	50

Table C1. Hyperparameter values used in the DDPG-based framework (see Sect. 3) for the Stommel model experiments at the three considered η_1 values.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix D: DDPG Hyperparameters for the Wood Model Experiments

Hyperparameter	Value
γ	0.003
δ	0.005
λ	1000
p_1	1/3
p_2	1/3
lr	0.001
h	0.00015
M	100
bs	5000
N_r	5×10^5
N_s	1000
N_p	500
n_{ep}	1000
n_{nodes}	200

Table D1. Hyperparameter values used in the DDPG-based framework (see Sect. 3) for the Wood model experiments.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

Code and data availability. The code required to reproduce the results presented in this study is publicly available on Zenodo (Guardamagna, 2026).

Author contributions. FG developed the algorithms, ran the simulations and prepared the figures. Both authors discussed the results, contributed to writing, and reviewed the final manuscript

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The work of Francesco Guardamagna was supported by the Netherlands Organization for Scientific Research (NWO) under Grant OCENW.M20.277. The work of Henk Dijkstra was also supported by the European Research Council through the ERC-AdG project TAOC (PI: Dijkstra, project 101055096). AI tools were used in coding and to polish the text of the manuscript.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Alkhayuon, H., Ashwin, P., Jackson, L. C., Quinn, C., and Wood, R. A.: Basin bifurcations, oscillatory instability and rate-induced thresholds for Atlantic meridional overturning circulation in a global oceanic box model, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475, 20190051, https://doi.org/10.1098/rspa.2019.0051, publisher: Royal Society, 2019.

Ashwin, P., Wieczorek, S., Vitolo, R., and Cox, P.: Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370, 1166–1184, 2012.

Castellana, D., Baars, S., Wubs, F. W., and Dijkstra, H. A.: Transition probabilities of noise-induced transitions of the Atlantic Ocean circulation, Scientific Reports, 9, 20284, 2019.

Cini, M., Zappa, G., Ragone, F., and Corti, S.: Simulating AMOC tipping driven by internal climate variability with a rare event algorithm, npj Climate and Atmospheric Science, 7, 31, 2024.

Dheeshjith, S., Subel, A., Adcroft, A., Busecke, J., Fernandez-Granda, C., Gupta, S., and Zanna, L.: Samudra: An AI Global Ocean Emulator for Climate, Geophysical Research Letters, 52, e2024GL114318, https://doi.org/https://doi.org/10.1029/2024GL114318, e2024GL114318 2024GL114318, 2025.

El Aouni, A., Gaudel, Q., Regnier, C., Van Gennip, S., Le Galloudec, O., Drevillon, M., Drillet, Y., and Lellouche, J.-M.: GLONET: Mercator's End-to-End Neural Global Ocean Forecasting System, Journal of Geophysical Research: Machine Learning and Computation, 2, e2025JH000686, https://doi.org/https://doi.org/10.1029/2025JH000686, e2025JH000686 2025JH000686, 2025.

Freidlin, M. I. and Wentzell, A. D.: Random perturbations, Springer, 1998.

Grafke, T., Grauer, R., and Schäfer, T.: The instanton method and its numerical implementation in fluid mechanics, Journal of Physics A: Mathematical and Theoretical, 48, 333001, https://doi.org/10.1088/1751-8113/48/33/333001, 2015.

Guardamagna, F.: Determining Optimal Noise Induced Tipping Transition Paths Using Deep Reinforcement Learning, https://doi.org/10.5281/zenodo.24096036, 2026.

Häfner, D., Jacobsen, R. L., Eden, C., Kristensen, M. R., Jochum, M., Nuterman, R., and Vinter, B.: Veros v0. 1–A fast and versatile ocean simulator in pure Python, Geoscientific Model Development, 11, 3299–3312, 2018.

Heymann, M. and Vanden-Eijnden, E.: The geometric minimum action method: A least action principle on the space of curves, Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 61, 1052–1117, 2008.

Holden, P. B., Edwards, N. R., Fraedrich, K., Kirk, E., Lunkeit, F., and Zhu, X.: PLASIM–GENIE v1. 0: a new intermediate complexity AOGCM, Geoscientific Model Development, 9, 3347–3361, 2016.

Jackson, L. and Wood, R.: Hysteresis and resilience of the AMOC in an eddy-permitting GCM, Geophysical Research Letters, 45, 8547–8556, 2018.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv:1412.6980 [cs], http://arxiv.org/abs/1412.6980, arXiv: 1412.6980, 2014.

Lin, B., Zhong, Y., and Ren, W.: Computing transition pathways for the study of rare events using deep reinforcement learning, Journal of Computational Physics, 527, 113812, 2025.

Lohmann, J., Hansen, A. B., Lovo, A., Chapman, R., Bouchet, F., and Lucarini, V.: The role of edge states for early warning of tipping points, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 481, 2025.

https://doi.org/10.5194/egusphere-2026-7043
Preprint. Discussion started: 27 November 2026
© Author(s) 2026. CC BY 4.0 License.



Moerland, T., Broekens, J., Plaat, A., and Jonker, C.: Model-based reinforcement learning: A survey, Foundations and Trends in Machine Learning, 16, 1–118, 2023.

Oh, J.-H., Kug, J.-S., Shin, Y., Geng, X., Wang, S., Jin, F.-F., An, S.-I., Xie, S.-P., and Liu, W.: Noise-induced tipping of Atlantic Meridional Overturning Circulation under climate mitigation scenarios, Nature communications, 2025.

Silvestri, S., Wagner, G., Hill, C., Ardakani, M. R., Blaschke, J., Campin, J.-M., Churavy, V., Constantinou, N., Edelman, A., Marshall, J., et al.: Oceananigans. jl: A model that achieves breakthrough resolution, memory and energy efficiency in global ocean simulations, arXiv preprint arXiv, 2309, 2023.

Soons, J., Grafke, T., and Dijkstra, H. A.: Optimal transition paths for AMOC collapse and recovery in a stochastic box model, Journal of Physical Oceanography, 54, 2537–2552, 2024.

Stommel, H.: Thermohaline Convection with Two Stable Regimes of Flow, Tellus A, 13, https://doi.org/10.3402/tellusa.v13i2.9491, 1961.

van Westen, R. M., Kliphuis, M., and Dijkstra, H. A.: Physics-based early warning signal shows that AMOC is on tipping course, Science Advances, 10, eadk1189, https://doi.org/10.1126/sciadv.adk1189, 2024.

Weijer, W., Cheng, W., Drijfhout, S. S., Fedorov, A. V., Hu, A., Jackson, L. C., Liu, W., McDonagh, E. L., Mecking, J. V., and Zhang, J.: Stability of the Atlantic Meridional Overturning Circulation: A Review and Synthesis, Journal of Geophysical Research: Oceans, 124, 5336–5375, https://doi.org/10.1029/2019JC015083, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JC015083, 2019.

Weinan, E., Ren, W., and Vanden-Eijnden, E.: String method for the study of rare events, Physical Review B, 66, 052 301, 2002.

Weinan, E., Ren, W., and Vanden-Eijnden, E.: Minimum action method for the study of rare events, Communications on pure and applied mathematics, 57, 637–656, 2004.

Willeit, M., Ganopolski, A., Robinson, A., and Edwards, N. R.: The Earth system model CLIMBER-X v1.0 – Part 1: Climate model description and validation, Geoscientific Model Development, 15, 5905–5948, https://doi.org/10.5194/gmd-15-5905-2022, publisher: Copernicus GmbH, 2022.

Wood, R. A., Rodríguez, J. M., Smith, R. S., Jackson, L. C., and Hawkins, E.: Observable, low-order dynamical controls on thresholds of the Atlantic meridional overturning circulation, Climate Dynamics, 53, 6815–6834, 2019.

Wu, S., Chen, Y., Pu, G., Sun, X., Chen, Y., Qi, J., Kong, M., Yang, K., Xu, L., Wang, W., et al.: AxiomOcean: Forecasting the Three-Dimensional Structure of the Upper Ocean, arXiv preprint arXiv:2605.10455, 2026.