UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

MSc in Physics

Reservoir computer-based detection of AMOC tipping

When the Atlantic Ocean Circulation may transition under the influence of stochastic freshwater forcing

Alicja Barbara Kałucka
Advisor: Peter Ditlevsen and Henk Dijkstra

Submitted: May 22, 2023

Abstract

The undeniable role of the Atlantic Meridional Overturning Circulation (AMOC) in the global climate system is recognized. The potential consequences of stopping or reversing the AMOC are of great concern. The aim of this study is to assess whether a tipping point in the AMOC could be forecast using modern machine learning methods.

An idealized coupled ocean-atmosphere model was used to train a reservoir computer, into which a bifurcation channel was incorporated. This channel, which can identify complex system behaviour, allows critical transitions to be detected and system collapse to be forecast.

The majority of the predictions made by our reservoir computer were found to align with the outputs of our model, with most values falling within three standard deviations. These findings suggest that collapses in low-order climate models might potentially be forecast using this machine-learning approach. This could lead to an enhanced understanding of abrupt climate changes and the ability to predict them.

Contents

1 Introduction

Thermohaline Circulation

Circulation of any liquid is a quite common concept. Also one that's pretty intuitive for everyone. It can be seen when a person's circling teaspoon causes a circulation in a teacup when they want to make the sugar dissolve in it or a bigger scale circulation of water from clouds to land, to the ocean and then back to clouds - a hydrological cycle. As people's circulatory system keeps people alive, the Earth's system of oceanic circulation is making sure that the planet stays healthy. Oceanic circulation is a result of many different reasons. It can be caused by wind, tides, the Earth's rotating, shape of land and differences between temperature and salinity. Most importantly 9/10 of the currents of the world are driven by density differences between warm and cold water. Warm water is less dense than cold water. Warm water rises while cold one sinks and that creates the major driving force of Earth's deep ocean circulation.

The other part that is a driving force for Earth's currents is salinity. As a rule of thumb - in the whole ecosystem everything wants to come to equilibrium at some point. The same goes for salinity differences between different regions. It's intuitive to know that water in a sea near an influx of a river is less salty than water in the middle of the Atlantic Ocean. Salty water is denser than freshwater creating a density difference that drives ocean circulation.

Thermohaline circulation, also known as the ocean conveyor belt or the meridional overturning circulation, is a global-scale ocean current system that plays a critical role in regulating the Earth's climate. It is influenced by density gradients created by surface heat and freshwater fluxes. Part *thermo* refers to temperature and *-haline* to salinity.

The process begins in the North Atlantic, where cold, dense water sinks to the bottom and flows southward into the deep ocean. This deep water then circulates the globe, gradually warming and rising as it moves through the ocean basins. In the Southern Ocean, the warm water rises to the surface and flows back toward the North Atlantic, completing the circuit.

The scheme of it can be seen in Figure 1.

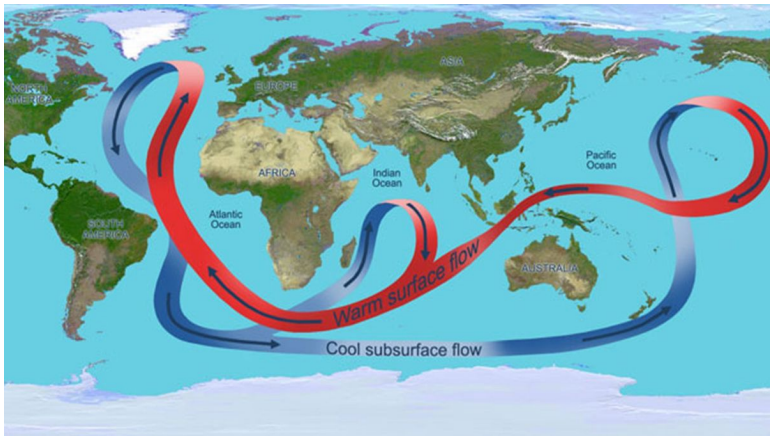

Figure 1: Schematic graph of Thermohaline Circulation.[1]

Thermohaline circulation is important for several reasons. It helps to distribute heat and regulate the Earth's climate, by transferring warm water from the tropics to higher latitudes and bringing cold water back toward the equator. It also plays a critical role in the transport of nutrients and dissolved gases throughout the ocean. It brings nutrient-rich water from the deep ocean to the surface, where it supports the growth of phytoplankton and other marine organisms that form the basis of the ocean's food web. It also helps to transport dissolved gases, such as carbon dioxide, between the ocean and the atmosphere, which is important for regulating the Earth's carbon cycle. This process helps to regulate global weather patterns, including the intensity and frequency of storms, droughts, and heat waves. It's thought to have an impact on how much energy Earth loses to outer space and receives from Sun.

As mentioned before - the thermohaline circulation is driven by differences in temperature and salinity in the ocean, which create density gradients that cause water to move in a circular pattern. Changes in these density gradients, such as through the melting of ice sheets or changes in ocean salinity, could potentially slow down or disrupt the circulation.

However, the thermohaline circulation is a highly complex system that is influenced by many factors, including winds, tides, and ocean currents. Changes in these factors could also impact the circulation, making it difficult to predict the future behaviour of the system with certainty. Hence in this

thesis, it was decided to take into consideration one of the components of it - AMOC.

AMOC

A very important part of the thermohaline circulation is called AMOC which stands for Atlantic Meridional Oceanic Circulation. It is a large-scale circulation pattern in the Atlantic Ocean. Its characteristic features are that there's the northward flow of warm, salty water and simultaneously there's the southward flow of cold, deeper water. Warm water flows northward until it reaches Greenland. There water is being cooled by wind and low temperature. The wind that blows on the surface of the water also increases evaporation resulting in lower temperature. Worth mentioning here is the fact that evaporation influences only water molecules leaving the remaining water more salty, hence even denser. Thus water sinks creating a deep water basin in North Atlantic Ocean. The high latitude cooling and low latitude warming move the whole deep water flow southward. The flow slowly moves towards the Pacific and gradually rises there.

The climate of northwest Europe, as well as the entire Northern Hemisphere, is deeply impacted by the conveyance of heat and salt by ocean currents from tropical areas to subpolar regions. The transfer of heat from the ocean to the atmosphere in the subpolar North Atlantic plays a significant role in maintaining the relatively mild climate in northwest Europe, resulting in temperatures up to 6°C higher than similar coastal climates adjacent to the Pacific (refer to Palter [29] for a comprehensive review). In the Atlantic Ocean, the northward movement of heat extends throughout the entire water column, leading to the remarkable phenomenon where heat is transported towards the equator in the South Atlantic Ocean.

The average strength of the AMOC at 26.5°N, as observed by the RAPID array since 2004, is approximately 17.0 Sv (1 Sv $\equiv 10^6$ m^3/s). The time series data reveals variability across all timescales, including a 15% weakening trend over the recorded period (Smeed et al.[30]). Numerous climate models exhibit variability in the AMOC on decadal and multidecadal timescales (e.g., Muir & Fedorov [31]), and the majority of models foresee a progressive reduction in the strength of the AMOC over the course of the 21st century, attributable to human-induced factors (e.g., Cheng et al. [32]). Dima and Lohmann [33], Rahmstorf et al. [34], and Caesar et al. [35] argue that a gradual slowdown of the AMOC has already commenced and is evident in proxy records dating back to the mid-twentieth century.

Deep convection in the subpolar North Atlantic Ocean, particularly in the Labrador and Nordic Seas, is believed to be a crucial process for sustaining the AMOC (e.g., Marshall & Schott [36]). During winter, the cooling of relatively saline surface waters initiates convective instabilities, resulting in the removal of heat from the water column and the formation of cold and

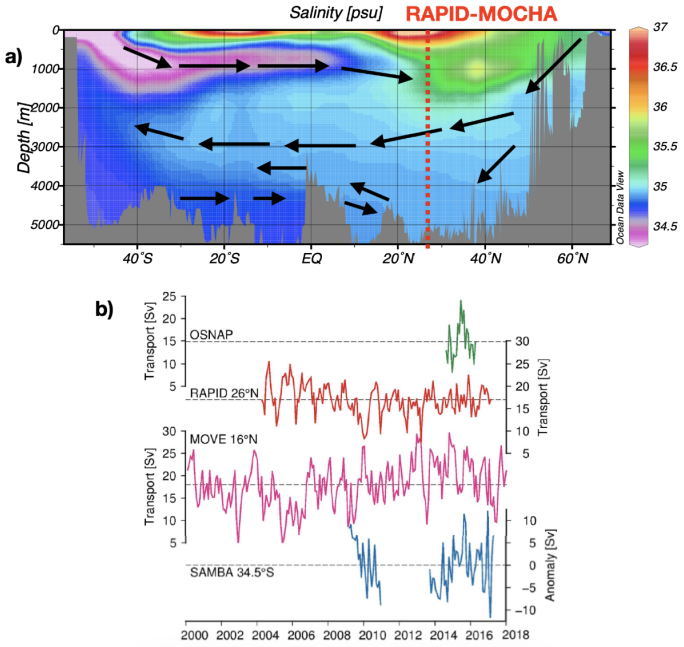

Figure 2: a) A cross-sectional representation of the salinity distribution in the western Atlantic Ocean along a meridional axis. The data collected by RAPID is represented by the red line, which corresponds to the 26°N latitude. [48] b) The graph displays the monthly MOC_z which is AMOC strength determined in depth coordinates for four observing arrays: OSNAP (green), RAPID 26°N (red), MOVE 16°N (magenta), and SAMBA 34.5°S (blue). In the case of SAMBA, the transports are depicted as anomalies. The black dashed line represents the corresponding means, with zero indicating the mean for SAMBA. [47]

saline North Atlantic Deep Water (NADW). It also plays a role in the distribution of nutrients and marine life in the Atlantic Ocean. Thus it affects the entire ecosystem. It is an important part of the world's climate being a result of atmospheric and thermohaline drivers.

The salinity surplus in the upper North Atlantic, which is notably higher than in other basins, is sustained through a combination of factors. This includes a net excess of evaporation over precipitation, river runoff, and ice melt (Emile-Geay et al. [37]; Schmitt [38]; Warren [39]), as well as oceanic transports such as the import of salty Indian Ocean water through Agulhas Leakage (De Ruijter et al. [40]; Gordon [41]). These processes play a role in preconditioning the Atlantic for deep convection (e.g., Marsh et al. [42]). However, an important contributor is the AMOC itself, as it transports high-salinity water from the subtropical to the subpolar North Atlantic, thereby maintaining elevated salinity conditions in the convection regions. This results in a robust feedback mechanism, known as the salt-advection feedback, which has contributed to the persistence of a strong AMOC throughout the Holocene.

The Atlantic Meridional Overturning Circulation (AMOC) has been an important component of Earth's climate system for millions of years. However, its stability and variability over time are still not fully understood. Scientists have used various methods to reconstruct past changes in the AMOC, including proxy data from sediment and ice cores, as well as climate models.

Research suggests that the AMOC has gone through periods of stability and instability in the past. During the last ice age, for example, the AMOC was weaker than it is today. This is thought to have been caused by the formation of large ice sheets in North America and Eurasia, which altered the circulation patterns of the North Atlantic.

In contrast, during the Holocene period, which began about 11,700 years ago and continues to the present day, the AMOC appears to have been relatively stable. However, there is evidence of short-term variations, such as the "Little Ice Age" that occurred between the 16th and 19th centuries, when temperatures in the North Atlantic region were significantly cooler than they are today.

Proxy data also suggest that the AMOC may have experienced abrupt changes in the past, such as during the Younger Dryas period around 12,000 years ago [24]. During this time, there was a sudden cooling in the North Atlantic region that lasted for about 1,000 years, which some scientists believe was caused by a disruption of the AMOC [25].

How the collapse may happen?

> "The basic idea is that the salt-advection feedback that currently maintains a strong AMOC can also work against it: a weakened AMOC would transport less salt into the subpolar North At-

> lantic, leading to reduced convection and even more weakening of the AMOC. A state would result in which no NADW is formed, and the deep Atlantic is either stagnant or experiences basinwide upwelling of abyssal waters formed around Antarctica." [4]

Multiple experiments on global climate models have been conducted to observe AMOC's potential slowdown or collapse and its impacts. Jackson et al., 2015 [7] found that freshwater flux applied in North Atlantic resulted in the cessation of convection over deep water formation regions and a big decrease in AMOC's strength. Moreover, in this experiment, AMOC hasn't recovered for the first 100 years after the freshwater input has been removed. Moreover, the experiment suggests that due to weaker deep water circulation and AMOC's reversal in the South Atlantic, there may be cooling of North Atlantic and the atmosphere over it. Temperature estimations are 2-5°C cooler in subtropical gyre and 5-10°C in subpolar gyre.

The reason those experiments are important to be conducted and analyzed is that paleo-proxies data shows such temperature variations in the past. Researchers analyzing ice core data drilled from Greenland ([7], [8]) found that in δ^{18}O stable isotopes and methane records are giving credence to AMOC's collapses being responsible for rapid climate changes during the Pleistocene era of ice ages.

Stommel's box model

As paleoclimatologists were looking for an explanation of rapid change swings deducted from ice core data from Greenland in the 1980s, they realized the relevance of Stommel's ocean model. He described it in 1961. His model developed a two-box model to show convection in a driven by salinity and temperature system. The ocean-box model represents a simple model of a single ocean basin, specifically the North Atlantic. The basin is divided into two boxes: one located at the equator and another at the North Pole. Within each box, the water is assumed to be perfectly mixed, resulting in constant temperature and salinity. However, there may be variations between the boxes. This configuration gives rise to a circulation between the boxes, which represents the Thermohaline Circulation (THC). In this circulation, water evaporates from the equatorial box and precipitates into the polar box.

Consequently, the salinity difference is maintained due to the disparity in heat flux from the sun. This leads to a circulation where the temperature and salinity drive the flow in opposite directions.

Stommel showed the existence of multiple equilibria in a system as well as the possibility of bistability on a single hemisphere. The model represents the ocean as two boxes, one representing the North Atlantic and the other representing the rest of the world ocean.

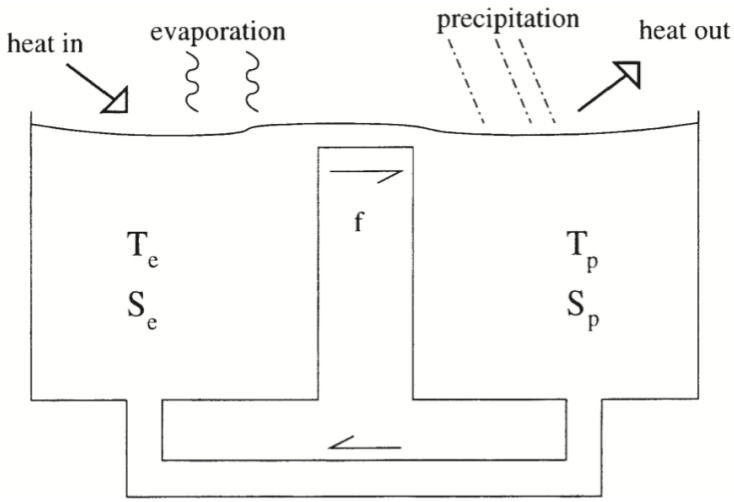

Figure 3: Scheme of Stommel two-box model.[16]

Fig. 3 shows what the model looks like. The volume of water is kept equal, but variations in density may exist between the boxes. By employing a linearized equation of state and incorporating assumptions on the damping, dynamical equations for the temperature difference $T = T_e - T_p$ and salinity difference $S = S_e - S_p$ can be derived

$$\dot{T} = k_a(T_a - T) - |f(T,S)|T - k_w T, \tag{1}$$

$$\dot{S} = \delta - |f(T,S)|S - k_w S, \tag{2}$$

$$f = \omega T - \xi S, \tag{3}$$

where k_a is the coefficient of heat exchange between ocean and atmosphere, k_w is the coefficient of internal diffusion and ω and ξ derive from the linearised equation of state. Thermohaline Circulation (THC) is represented as the flow f. The inclusion of the absolute value in equations 1 and 2 was implemented by Stommel, who argued that the water mixing should remain independent of the flow direction. A more straightforward derivation of the equations of motion for a simplified ocean model associated with the box model confirms that this assumption holds, albeit with the quadratic term arising instead of a piecewise linear form [20].

The model assumes that the North Atlantic box receives more freshwater input from precipitation and river runoff than it loses through evaporation, while the rest of the world ocean box loses more freshwater through evaporation than it gains through precipitation and river runoff. The freshwater imbalance in the North Atlantic box leads to a reduction in the water density, which inhibits the sinking of the water mass and therefore, reduces the strength of the thermohaline circulation.

To further investigate how the model reacts to changes in salinity and freshwater forcing, the hosing experiment has been conducted. Freshwater flux is being applied to the model resulting in the system going to the negative branch of Ψ^*. After that, the perturbation is removed and the system comes back to a bistable regime, but this time on the negative branch. This shows the possibility of two steady states for the same salinity forcing.

The Stommel two-box model predicts that a small freshwater perturbation in the North Atlantic can lead to a reduction of the thermohaline circulation, which can result in significant changes in the climate system, such as the cooling of the North Atlantic and adjacent land areas. The model also suggests that a threshold exists beyond which the thermohaline circulation may collapse, leading to a more permanent state of reduced circulation.

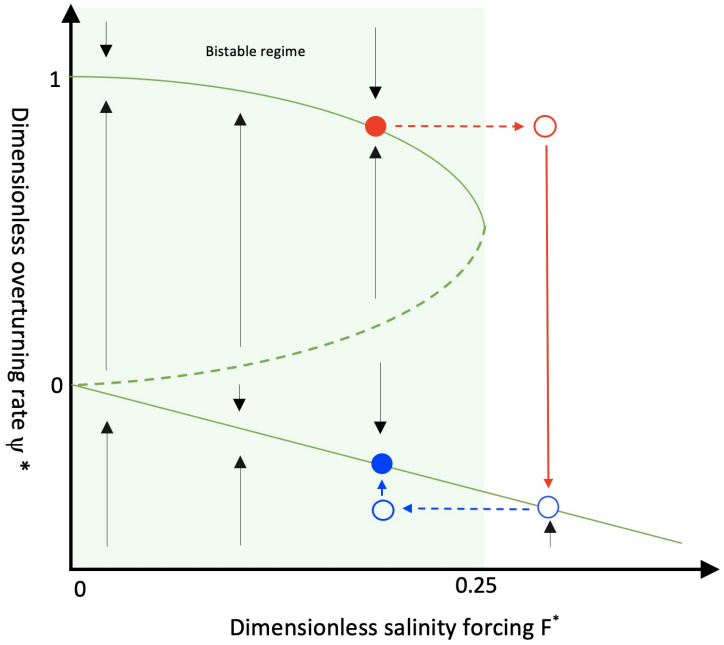

Figure 4: δ^{18}O data for 5 chronological records.[9]

Early Warning Signals

As the hypothesis of AMOC's collapse appeared, research on how can we possibly predict its potential collapse has started. The possibility to have early warning signals before possibly abrupt AMOC transitions and for those to be generated for accessible data made a lot of researchers very interested in them.

> "Numerical modelling experiments suggest that without a vigorous AMOC, surface air temperature in the North Atlantic region would cool by around 1–3°C, with enhanced local cooling of up to 8 °C in regions with large sea-ice changes. Substantial weakening of the AMOC would also cause a southward shift of the inter-tropical convergence zone, encouraging Sahelian drought, and dynamic changes in sea level of up to 80 cm along the coasts of North America and Europe. Theoretical arguments, numerical models of varying complexity and evidence from palaeoclimate proxy records support the existence of two stable AMOC states—'on' and 'off'. A reduction in density of the surface waters of the North Atlantic (through an increase in freshwater input or surface warming) can inhibit the formation of deep water and weaken the AMOC." [3]

The easiest to think of warning signals are changes in temperature and salinity as AMOC is driven by those.

- If the AMOC slows down or weakens, it can cause an increase in surface salinity in the North Atlantic. This is because the AMOC is responsible for transporting freshwater from the surface to the deep ocean. If this transport is reduced, more freshwater will remain at the surface and increase salinity.
- The AMOC plays a crucial role in regulating global heat transport. As the AMOC weakens, it can cause an increase in the heat content of the ocean in the North Atlantic. This can also lead to an increase in sea surface temperatures.

Other worth mentioning are reduced variability, changes in atmospheric circulation and increased carbon storage.

- The AMOC can become less variable before a collapse. This means that the flow of water through the AMOC becomes more consistent, with less fluctuation from year to year.
- Changes in atmospheric circulation patterns can also be an early warning signal for the potential collapse of the AMOC. As the AMOC weakens, it can cause changes in the atmospheric circulation patterns in the North Atlantic region, which can in turn impact weather patterns in other parts of the world.
- As the AMOC slows down or weakens, it can also impact the ocean's ability to absorb and store carbon dioxide. This can cause an increase in atmospheric CO2 concentrations, which can have implications for global climate change.

It is important to note that these early warning signals are very hard to estimate and that they rather provide an indication that the AMOC may be at risk of collapsing in the future.

> "The search for early warning signals of an impending AMOC collapse is an important but complicated endeavor. Even though methods might allow us to detect a change in the AMOC, the challenge is to distinguish the start of a collapse from longer-term variability or a gradual weakening. Models may help us to understand what natural variability looks like and help to detect signals of impending AMOC collapse [22]; however, there is a wide range of variability mechanisms, magnitudes, and timescales across models [21]. **At present no metrics have been found in models** to distinguish multidecadal variability from a systematic downward trend of the AMOC. There has also been no analysis to establish which properties would precede a collapse and at the same time do not precede a downward trend as, for instance, simulated by the CMIP5 multimodel ensemble [23] in response to global warming." [4]

Time series analysis described in [3] shows autocorrelation and variance are present in a fully coupled atmosphere-ocean general circulation model and they are proposed to be early warning signals.

Autocorrelation refers to the degree to which a time series is correlated with a lagged version of itself. In the context of early warning signals, increasing autocorrelation is considered an indicator of an approaching critical transition. As a system approaches a tipping point, the dynamics of the system become more sensitive to its own past states, leading to an increase in the persistence or memory of the system. This increased memory is reflected in the increased autocorrelation of the time series.

Increased variance, on the other hand, refers to an increase in the amplitude of fluctuations in the time series. As a system approaches a tipping point, it becomes increasingly unstable, leading to larger and more frequent fluctuations. This increased variability is often accompanied by a decrease in the damping of the system, which can be observed as an increase in the variance of the time series.

Both autocorrelation and increased variance are relatively easy to measure from time series data and be reliable indicators of approaching tipping points in a wide range of systems. However, it is important to note that these signals are not foolproof and may not always be present before a critical transition occurs. Therefore, they should be used in conjunction with other methods to increase the reliability of early warning systems.

In the article time series data from FAMOUS model [5] is being analysed. The early warning signals were searched in annual and decadal mean time

series at 33.75°S - 58.75°N (at every 2.5°). The hosing experiment has been conducted on the model where the freshwater flux is being applied in North Atlantic. That results in a lower density of the surface waters and weakens the AMOC. The forcing is gradually increased causing AMOC to turn to an 'off' state after about 800 years. The flux is then turned off allowing AMOC to recover. The results of this are shown in Figure 5. The AMOC transport at a specific location (26.25°N and 1,000 m depth) is plotted against the imposed freshwater input. The decadal mean AMOC is represented by the solid black line, which corresponds to a transient experiment where the freshwater input increases from 0 to 1 Sv over a period of 2,000 years. Similarly, the solid grey line represents the decadal mean AMOC for a transient experiment where the freshwater input decreases from 1 to -0.4 Sv at the same rate. The filled circles indicate the equilibrium AMOC transport achieved in a series of constant-forcing simulations. These simulations are initialized from the respective points of the transient (increasing forcing) simulation.

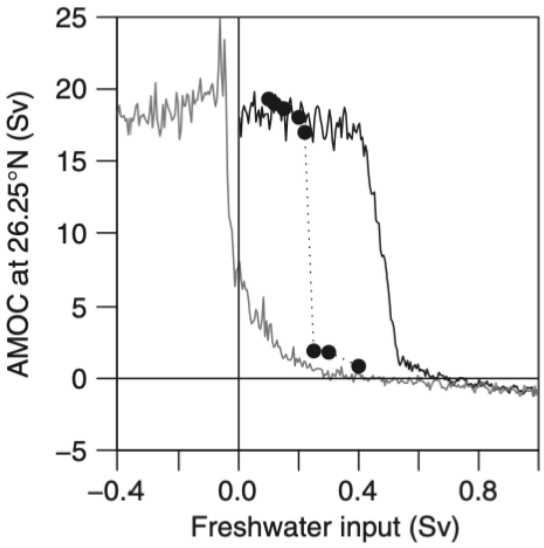

Figure 5: Hysteresis of AMOC.[3]

After that early warning signals were searched at a wide range of latitudes in the model Atlantic basin (33.75°S - 58.75°N) as there varying AMOC strength is shown. Increasing autocorrelation and variance are found at all latitudes. The outcome of this analysis is that the strongest upward trends in autocorrelation and variance are found towards the southern boundary of the Atlantic and just north of the equator respectively. Later to check the significance of those comparisons to the null model was conducted. According to the article the most reliable indicator was found from autocorrelation in the mid-high northern latitudes and at the southern boundary of

the Atlantic and from variance in the equatorial North Atlantic and in parts of the sub-polar gyre. Early warning signal from variance was found to be less reliable than increasing autocorrelation. The outcome of this analysis is shown in Figure 6.

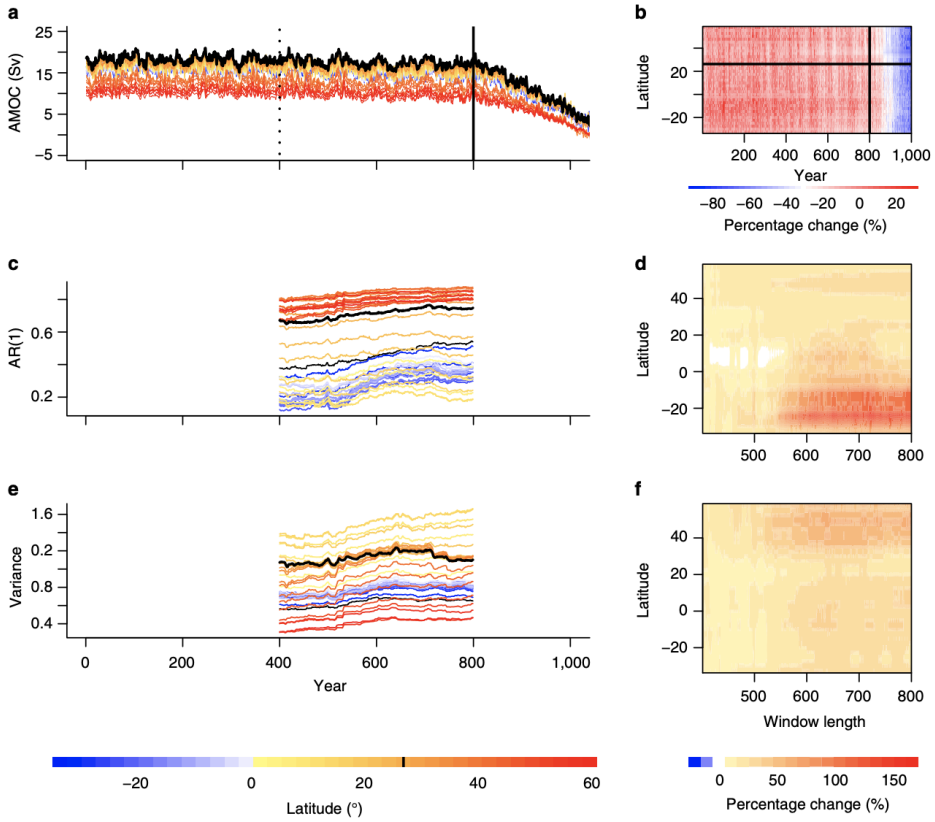

Figure 6: (a) The annual time series of AMOC (Sv) at each latitude are displayed, but they are truncated before the onset of collapse, which occurs at 800 years. A solid vertical line marks this cutoff point. (b) Additionally, the time series are presented as a contour plot in the time-latitude plane. To perform the analysis, the data is detrended, and the focus is on the residuals. The estimation of candidate early warning signals is based on a sliding window of 400 years, with the dotted vertical line indicating the end of the first window. (c,d) The AR(1) coefficient and (e,f) variance are examined as part of the analysis. The time series are color-coded according to their respective latitudes, with 26.25°N represented by the color black. Additionally, contour plots depicting the percentage change are included.

The disruption of the Atlantic Meridional Overturning Circulation (AMOC) due to significant influxes of freshwater, which result in changes to seawater density, has been noted. A subsequent decision was made to further investi-

gate any implications indicating an impending collapse of the AMOC. The strategy adopted involved predicting the freshwater flux value that would cause the AMOC to collapse.

This was accomplished by training a reservoir computer with an additional bifurcation parameter channel incorporated into the existing model, and subsequently predicting the freshwater flux value at which a transition is observed. The significance of accurately predicting this value lies in its potential to expedite the processes of global climate models.

Moreover, it should serve as a valuable indicator for future studies where past data, rather than models, may be utilized for training the reservoir. This underscores the importance of this study in contributing to an improved understanding of the potential collapse of the AMOC and, by extension, of the global climate system.

The question of which tipping point is the right one to look at and when any potential threshold may be crossed, so when the tipping may happen, remains unanswered for the time being. Research on early warning signals is still ongoing.

2 Methods

Reservoir computer framework

The reservoir computer used in this research was shown by Kong et al. [12]. The authors presented 3 different nonlinear and complex systems for which their reservoir computer predicted the transition point accurately. This study focuses on the development of a model-free machine learning-based method aimed at predicting the system's state, specifically whether it is currently in, or will transition into, a state preceding collapse.They came up with a powerful tool that incorporated bifurcation parameters as parameter input channel.

Why their research is so important?

> "Due to global warming and climate change, some natural systems may have already been in a transient state awaiting a catastrophic collapse to occur. A reliable determination at the present that the system has already passed the critical transition or a "tipping" point to a transient state would send a clear message to policy makers and the general public that actions must be taken immediately to avoid the otherwise inevitable catastrophic collapse." [12]

The way their reservoir computer is implemented is schematically presented in Figure 7. In the green area, the system exists normally with its chaotic attractor and periodic windows. The moment the chaotic attractor is destroyed through a transition is shown as the end of the green region

and that's when the bifurcation occurs. In the vicinity of a critical parameter value, the system exhibits transient chaos, ultimately culminating in a collapse state.

The goal of this reservoir is to

- assess whether the system will enter a chaotic regime leading to a collapse state.
- If yes, then to check when the collapse will happen, meaning how long will the system survive.

They show that when the system is trained with 3 different bifurcation parameter values, the reservoir predicts not only the collapse but also transient chaos after the collapse takes place. The machine can also predict how long the system will survive before collapse as the probability distribution of the transient lifetimes of trajectories generated by the machine aligns with the true distribution.

There are 3 layers the machine consists of:

- input layer consisting of two components:
  - $\mathbf{u(t)}$ D_{in}-dimensional input data, W_{in} is weighted, $D_r \times D_{in}$ input matrix that maps u(t) to D_r dimensional state vector r(t)
  - W_b is matrix $D_r \times D_b$ that specifies the connection weights between the input bifurcation parameter to r(t)
- hidden layer: $\mathcal{A}$ is adjacency matrix $D_r \times D_r$ that transforms hidden state r(t) to r(t+Δt), the dynamics of each node in the system are characterized by an internal state and governed by a hyperbolic tangent activation function, $\mathbf{r(t)}$ is a high-dimensional hidden state
- output layer: W_{out} is output matrix $D_{out} \times D_r$ transforming r(t) back to v(t), $\mathbf{v(t)}$ is D_{out}-dimensional time series data

D_{in} and D_{out} were set to be equal, D_b is set to 1 and D_r is the number of nodes n (one of later described hyperparamters).

The dynamical evolution of the reservoir is described by:

$$r(t + \Delta t) = (1 - \alpha)r(t) + \alpha tanh[\mathcal{A} \cdot r(t) + W_{in} \cdot u(t) + k_b W_b(b + b_0)], \quad (4)$$

$$v(t) = W_{out} \cdot r(t) \quad (5)$$

where α is the leakage parameter, which controls the decay of the reservoir state over time, b is the bifurcation parameter that belongs to the target system, k_b and b_0 are scaling and the bias of bifurcation parameter b into the reservoir network.

Preparation phase

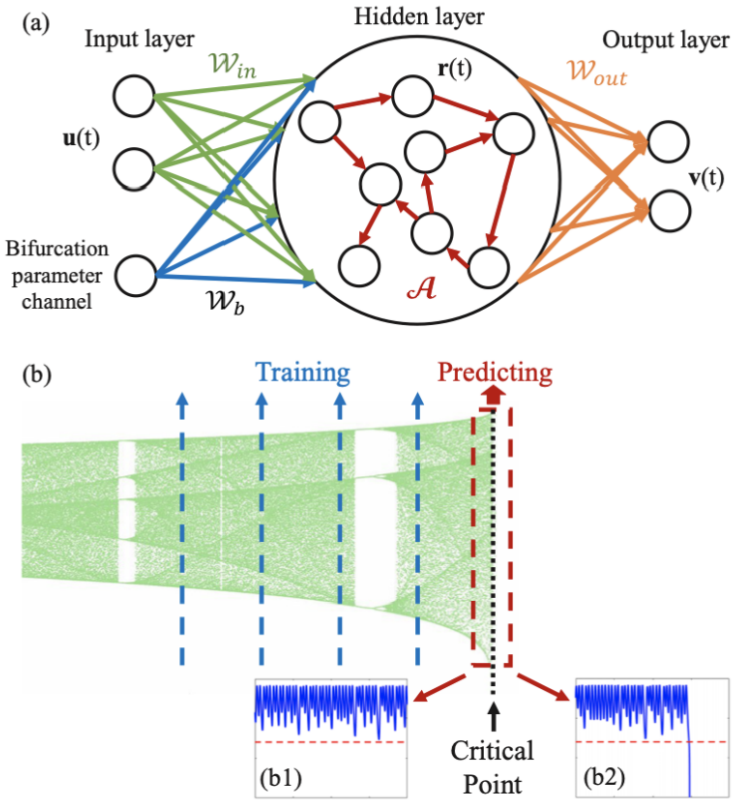

Figure 7: Reservoir Computing Scheme and Illustration of Predicting Critical Transition and Collapse Idea [12]

Elements of W_{in} and W_b are generated randomly before the training and prediction part takes place. Their distributions are determined by hyperparameters (n, d, ρ and k_{in}) of the reservoir where n is the number of neurons in the hidden layer, which controls the capacity of the reservoir, d and ρ are the average degree and spectral radius of adjacency matrix $\mathcal{A}$ and k_{in} is a scaling factor of the elements of the input matrices W_{in} and W_b. The nonzero elements of W_{in} and W_b are generated from uniform distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

in the interval [-k_{in}, k_{in}].

Transfer of input to hidden layer

After the preparation phase is over, so the data is generated, the machine starts the process of feeding-forwarding the data from the input to the hidden state.

The time series input, represented as the vector $u(t)$, is subsequently propagated forward to the hidden layer of the model, and its evolution over time is recorded. The equation 5 is applied to obtain $r(t+\Delta t)$ records. The initial state of the network is set to be equal to $r(t=0)=0$. The biggest advantage of this reservoir is that there's a possibility to train the target system at multiple bifurcation parameters. After the system is trained with one, the time and initial states are reset and the training part is repeated with different values of the bifurcation parameter. As a result, multiple recordings of $r(t)$ are obtained, regardless of the order of training.

In order to facilitate accurate predictions beyond the training parameter region, it is crucial for the reservoir machine to undergo thorough training to acquire a comprehensive understanding of the diverse dynamical behaviors displayed by the target system across all training bifurcation parameter values. The training objectives entail enabling the reservoir machine to learn the distinct attractors exhibited by the target system and establishing an association between each attractor and its corresponding value of the bifurcation parameter.

Throughout the research, it was observed that a reservoir machine, when appropriately trained and applied to representative target systems governed by nonlinear ordinary or partial differential equations, exhibited remarkable accuracy in predicting dynamics across all training parameter points. The relative errors consistently remained below 5%, a threshold maintained for approximately four or five Lyapunov times. Furthermore, the machine demonstrated reliable and accurate predictions of critical points and transient chaos. Hence, it is justifiable to establish this criterion of prediction accuracy for all training parameter values when estimating hyperparameter values based on the training data.

Regression

Upon completion of the training phase, the regression phase is initiated. This phase involves determining the elements of the output weight matrix W_{out} by performing regression between the input vector $u(t)$ and the network state vector $r(t)$. The recorded states $r(t)$ obtained for different values of the bifurcation parameter are temporally concatenated, forming a vector function denoted as $r_{all}(t)$. To ensure consistency, the concatenated vector $r_{all}(t)$ is substituted with $r'_{all}(t) = r_{all}(t)$.

Subsequently, linear regression is employed, incorporating L2 regularization, between the concatenated input vector $u_{all}(t)$ and the transformed state vector $r'_{all}(t)$. This regression is accomplished by minimizing the loss function, which is defined as:

$$\mathcal{L} = \sum_t ||\mathbf{u}_{all}(t) - W_{out} \cdot \mathbf{r}'_{all}(t)||^2 + \beta ||W_{out}||^2 \tag{6}$$

where β is the regularization coefficient. The regression can be accomplished by calculating:

$$W_{out} = \mathcal{U} \cdot \mathcal{R}'^{\mathcal{T}} (\mathcal{R}' \cdot \mathcal{R}'^{\mathcal{T}} + \beta I)^{-1} \tag{7}$$

where I is D_r dimensional matrix, $\mathcal{U}$ and $\mathcal{R}'$ are the matrix forms of $\mathbf{u}_{all}(t)$ and $\mathbf{r}'_{all}(t)$ respectively. Each column is a record of the values of the vector at a certain t and different rows representing different dimensions. $\mathcal{U}$ as well as $\mathcal{R}'$ consist of $n_b(T_{train} - T_{cut})$ columns, where n_b is the number of different bifurcation parameters values that are used in order to train the reservoir, T_{train} is the number of training time steps and T_{cut} is the number of removed time steps in order to make sure that the reservoir system has passed the transient phase.

Validation

TTo validate the training process, the set of training bifurcation parameters is employed to predict the dynamical evolution of the target system at the same parameter values. The original input data vector is substituted by the output vector $\mathbf{v}(t)$, resulting in a modification of Equation 5 as follows:

$$\mathbf{r}(t + \Delta t) = (1 - \alpha)\mathbf{r}(t) + \alpha tanh[\mathcal{A} \cdot \mathbf{r}(t) \tag{8}$$

$$+ \mathcal{W}_{in} \cdot \mathcal{W}_{out} \cdot \mathbf{r}'(t) + k_b \mathcal{W}_b (b + b_0)]. \tag{9}$$

In this equation, the reservoir system transforms into a self-evolving dynamical system augmented by additional parameter inputs. During the validation process, the state vector of the hidden nodes at the final training step, denoted as $\mathbf{r}_{\text{last}}$, is utilized as the initial condition for each bifurcation parameter value. The reservoir system is then allowed to evolve according to Equation 9 for a specified number of time steps. The average Root Mean Square Error (RMSE) is computed by comparing the predicted output vector with the true vector for the corresponding time steps.

During validation, the training bifurcation parameters are employed to calculate the RMSE for each parameter value, resulting in the generation of

three RMSE values. The validation error is determined by selecting the highest RMSE value among the three. In the case of spatiotemporal dynamical systems, the RMSE is computed at each time step, and the minimum time required for the RMSE to reach a predefined tolerance threshold is recorded. This minimal time serves as an indicator of the training performance and aids in reducing the computational burden.

Testing

Testing, in contrast to validation, involves evaluating the reservoir machine's performance on bifurcation parameter values that have not been encountered during training. This ensures that the reservoir machine can effectively generalize and accurately predict transient chaos and system collapse. During testing, the dynamical evolution of the reservoir machine continues to be governed by Equation 9, albeit with a different value assigned to the bifurcation parameter, denoted as b. To establish the initial condition for the reservoir machine during testing, a short segment of the real-time series from the target system (e.g., several oscillation cycles) is utilized. The warm-up phase of the reservoir machine involves input from an arbitrary value of the bifurcation parameter within the sustained chaos regime. Subsequently, the hidden states at the last time step, denoted as $\mathbf{r}_{last}$, serve as the initial condition for testing.

The optimization of hyperparameters becomes crucial to ensure accurate predictions for all training parameter values within approximately four or five Lyapunov times. Learning multiple distinct attractors simultaneously poses a more significant challenge for the reservoir machine compared to learning a single attractor. Therefore, Bayesian optimization is employed to optimize seven hyperparameters of the reservoir computing machine. These hyperparameters include the average degree d of the reservoir hidden network, the spectral radius ρ of the reservoir hidden network, the scaling factor k_in of the input matrix W_{in}, the regularization parameter β used during the training of the output matrix W_{out}, as well as parameters k_b, b_0, and α.

Furthermore, fluctuations in the validation and prediction performance can arise due to the stochastic nature of the input and hidden layers. To mitigate this variability, five different random realizations of the reservoir machines are simultaneously trained, and the one exhibiting the lowest validation error is selected and retained for further analysis.

The accurate prediction of transient chaos and system collapse is of paramount importance for a well-trained reservoir machine. In this study, "well trained" refers to the reservoir's ability, following training, to predict the dynamic evolution of the target system with a relative error of less than 5% for multiple Lyapunov times and across all selected bifurcation parameter values during the validation process. The influence of hyperparameters on prediction results is widely recognized [44] and holds particular significance in this study, given the reservoir machine's requirement to exhibit predictive capability across a range of bifurcation parameter values.

To enable the reservoir machine to make accurate predictions for bifurcation parameter values beyond those encountered during training, the optimization of hyperparameters becomes necessary. For this purpose, the Bayesian optimization method [25], implemented using the "skopt" package in Python [45], was employed. However, the optimization algorithm often yields multiple sets of hyperparameter values. To address this, multiple reservoirs were trained using these different hyperparameter values, and the average validation Root Mean Square Error (RMSE) was computed and provided as feedback to the Bayesian algorithm. The training and validation processes were repeated multiple times with different random reservoir realizations to reduce RMSE fluctuations.

Following several iterations of the Bayesian algorithm, hyperparameter values associated with the lowest validation RMSE across all iterations were selected, not necessarily from the last iteration. It is important to note that the optimization process employed training and validation data exclusively based on bifurcation parameter values prior to the transition point, within the sustained chaos regime. Thus, the system remained unaware of the impending crisis or potential system collapse throughout the optimization process.

One limitation of the Bayesian optimization algorithm is its susceptibility to generating suboptimal solutions when trapped in local minima of the cost function landscape, particularly in cases where the validation RMSE exhibits substantial fluctuations.

Low-order climate models

The climate model that is used in this research is mainly based on van Veen et al. [16] with some of Cessi's [17] results of research implemented. First, the introduction of van Veen's model is being done.

On a time scale of days or weeks, the atmospheric component of the Earth's climate system assumes dominance. Consequently, for short-range weather forecasts, the oceanic variables, such as the sea surface temperature, can often be regarded as fixed. However, over much longer time scales, the dynamics of the ocean can exert significant influence. When examining decadal climate variability or anthropogenic influences like the enhanced greenhouse effect, it becomes crucial to consider the role of the ocean. To address these aspects, state-of-the-art climate models, possessing millions of degrees of freedom, are commonly employed. The results of experiments conducted with such models are subjected to statistical analysis since they surpass the scope of conventional dynamical systems theory.

While significant insights into atmospheric models have been obtained through the examination of extremely low-dimensional truncations, our objective is to pursue a similar approach for coupled models. Specifically, we aim to explore the interplay between the short-time-scale variability inherent

in the atmospheric component, which exhibits chaos, and the long-time-scale behaviour of the oceanic component. In climatological terms, the focus lies on discerning whether the ocean is passive or active. The presence of activity in the ocean can give rise to notable feedback effects on atmospheric dynamics. Despite the inherent weakness in the ocean's forcing on the atmosphere, it can lead to decadal variability in atmospheric observables, exemplifying the potential impact of this interaction.

In van Veen's article, the atmosphere-ocean interaction in a low-dimensional coupled model applicable to midlatitudes is examined. It is demonstrated that the ocean's behaviour can be either passive or active, contingent upon small changes in the coupling parameters. These findings may help elucidate the contrasting conclusions drawn in studies utilizing more realistic and complex models.

The low-order model investigated in this study is based on a proposal by Roebber [14]. Roebber coupled the Lorenz-84 model [49], which serves as a metaphor for the general circulation of the atmosphere, with Stommel's box model for a single ocean basin [16]. Roebber employed numerical integrations and power spectra to characterize the dynamics of the coupled system, without exploring bifurcation analysis and other tools of dynamical systems theory.

In the fully coupled system, subject to periodic forcing, Roebber identified increased energy in low-frequency atmospheric modes when compared to one-way coupling, where there is no oceanic feedback to the atmosphere. In this paper, a detailed investigation of the coupled dynamics is conducted, going beyond Roebber's initial findings.

Lorenz-84 model

Similar to the Lorenz-63 model [18], the Lorenz-84 model can be associated with a Galerkin truncation of the Navier-Stokes equations. While the Lorenz-63 model characterizes convection, the Lorenz-84 model provides a simplified approximation of the general atmospheric circulation at mid-latitudes. This approximation is specifically applicable to an f-plane situated over the North Atlantic Ocean.

The variables of the Lorenz-84 model can be interpreted physically:

- x represents the intensity of the westerly circulation
- y and z correspond to the sine and cosine components of a large travelling wave.

The time derivatives are given by:

$$\dot{x} = -y^2 - z^2 - ax + aF, \tag{10}$$

$$\dot{y} = xy - bxz - y + G, \tag{11}$$

$$\dot{z} = bxy + xz - z, \tag{12}$$

where F is the forcing term due to the average north-south temperature contrast and G is the forcing term due to the average earth-sea temperature contrast. The value for a=$\frac{1}{4}$ and b=4.

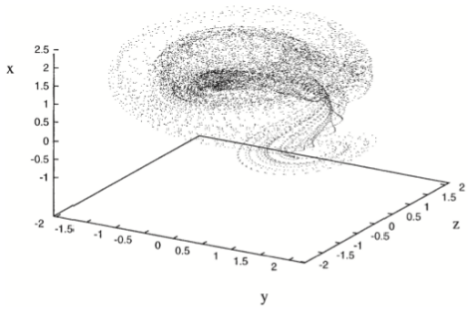

Fig. 1. Chaotic motion in the Lorenz-84 model for $(F, G) = (8, 1)$: about 10^4 points on the attractor.

Figure 8: Chaotic motion in the Lorenz-84 model for (F,G)=(8,1).[16]

The bifurcation diagram of this model exhibits a rich diversity of behaviours, including equilibrium points, periodic and quasi-periodic orbits, and chaotic motion. The qualitative behaviour can be understood by examining the energy transfer between the westerly circulation and the travelling wave. Solar heating drives the growth of energy content in the westerly circulation. However, once a certain threshold is surpassed, this circulation becomes unstable, leading to energy transfer to travelling waves, which are subsequently dissipated. The energy content of the westerly circulation rapidly decreases, and the cycle repeats itself in a periodic or irregular manner. Fig. 8 illustrates the orbit spiralling around the x-axis towards a critical value of x, followed by a descent towards the y-z plane.

At parameter values (F,G)=(6,1), the Lorenz-84 model exhibits two coexisting stable periodic solutions, which are referred to as summer conditions. When the parameter values are (F,G)=(8,1), the behaviour becomes chaotic, corresponding to winter conditions. Lorenz [19] argues that the north-south temperature contrast, F, is larger during winters, resulting in significant baroclinic wave activity, which manifests as chaotic motion in the Lorenz-84 model. Under summer conditions, with a smaller north-south temperature contrast, the periodic motion reflects less turbulent large-scale dynamics.

In van Veen's investigations, perpetual winter conditions will be focused on by fixing the forcing parameters to (F,G)=(8,1) in the coupled model described below. This choice ensures the emergence of complex dynamics when varying the coupling parameters. Equilibrium points and periodic solutions are observed when the coupling parameters are varied.

Coupled equations

Stommel's box model was introduced in eq. 1-3. When coupling the box model to the Lorenz-84 model, the estimates in Roebber [14] are used for the parameters in eq. 1-3. The volume of the deep ocean box, which is not included in van Veen's model, is determined by the interaction between the polar and equatorial boxes.

As mentioned before, the absolute value in eq. 1 and 2 should be quadratic [20]. If the quadratic term is assumed in the coupled model described below, significant changes are observed in the average values of T and S. However, the qualitative behaviour remains the same.

After introducing these simple models for atmospheric and oceanic circulation, along with their physical interpretation of variables, three mechanisms by which they interact can be identified:

1. The equilibrium between the atmospheric pole-equator temperature contrast and the strength of the zonal wind x is assumed to be maintained indefinitely, denoted as $T_a \propto x$. Additionally, the forcing of the atmosphere by the north-south temperature contrast in equation 10 is modified by the ocean temperature contrast, resulting in $F \to F_0 + F_1 T$. This represents the simplest form of geostrophic equilibrium, where a north-south temperature gradient drives an east-west atmospheric circulation.

2. The inhomogeneous forcing due to the land-sea temperature contrast described in equation 11 is expected to decrease as the temperature difference T increases. It is assumed that in the polar region, the seawater temperature exceeds the temperature over land, while in the equatorial region, it is lower. A larger temperature difference T corresponds to a reduced land-sea temperature contrast. This influence is represented as a fluctuation upon a fixed forcing: $G \to G_0 + G_1(T_{av} - T)$.

3. The water transport through the atmosphere is assumed to be linearly dependent on the energy content of the travelling wave: $\delta \to \delta_0 + \delta_1(y^2 + z^2)$

By combining equations 10 to 3 with the suggested coupling terms, the resulting set of equations is obtained.

$$\dot{x} = -y^2 - z^2 + ax + a(F_0 + F_1 T) \tag{13}$$
$$\dot{y} = xy - bxz - y + G_0 + G_1(T_{av} - T) \tag{14}$$
$$\dot{z} = bxy + xz - z \tag{15}$$
$$\dot{T} = k_a(\gamma x - T) - |f(T,S)|T - k_w T \tag{16}$$
$$\dot{S} = \delta_0 + \delta_1(y^2 + z^2) - |f(T,S)|S - k_w S, \tag{17}$$

with f as in eq. 3. The introduction of the coupling has led to the inclusion of new constants. These constants are denoted as T_{av}, representing the standard temperature constant between the polar and equatorial box; γ, the proportionality constant between the westerly wind strength and the

north-south temperature contrast; and δ_1, a measure of the rate of water transport through the atmosphere.

During the exploration of the model's dynamical behaviour, F_1 and G_1 are treated as free parameters. As previously justified, we focus on a scenario of small coupling to the atmosphere model. This corresponds to taking $(F_1, G_1) \in [0, 0.1] \times [0, 0.1]$. The parameter values are listed in Table 1.

In this scaling, one unit of time in the model corresponds to the typical damping time scale of the planetary waves, estimated to be between five to ten days.

a	0.25	δ_0	$7.8 \cdot 10^{-7}$
b	4	k_w	$1.8 \cdot 10^{-5}$
F_0	8	k_a	$1.8 \cdot 10^{-4}$
G_0	1	ξ	$1.1 \cdot 10^{-3}$
γ	30	ω	$1.3 \cdot 10^{-4}$
δ_1	$9.6 \cdot 10^{-8}$	T_{av}	30

Table 1: The constants of the coupled model [16].

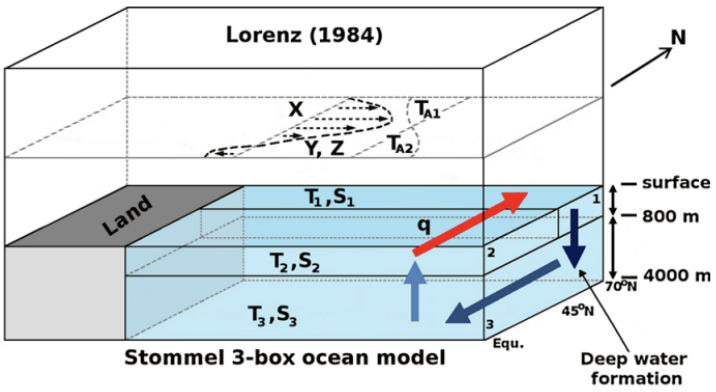

Figure 9: Schematic representation of the low-order climate model[13].

The atmospheric model comprises five variables, namely the tropospheric zonal flow strength (or meridional gradient of tropospheric temperature), the amplitudes of cosine and sine phases representing large-scale transient eddies superimposed on the mean flow, and the air temperatures at high and low latitudes. Interactions within the internal flow are manifested through the amplification of eddies from the zonal flow strength (referred to as baroclinic eddy growth), the displacement of eddies by the zonal flow, and mechanical damping.

On the other hand, the ocean model adopts a three-box representation of the North Atlantic basin, with two boxes representing the subtropical and

subpolar upper ocean, and the third box representing the deep ocean (see Fig. 1b). The temperature in the upper ocean is influenced by diffusion with the atmosphere, while both temperature and salinity within each box are determined through advection by the Meridional Overturning Circulation (MOC) and diffusion between the upper and deep ocean. The strength of the MOC is linearly linked to the meridional gradients of upper-ocean temperature and salinity.

Cessi's variant

The buoyancy-driven ocean circulation is represented using a modified Stommel two-box model, chosen as a minimal representation. In this model, when the temperature relaxation is fast, the salinity evolves while the temperature remains fixed at the prescribed ambient value. The box model itself lacks inherent variability, consisting of two linearly stable equilibria and one unstable equilibrium. To transition from one stable equilibrium to the other, a finite perturbation is required. The minimum amplitude and duration of the perturbation in time are calculated.

To account for the influence of changes in the global hydrological cycle caused by weather variations, a stochastic component of freshwater flux forcing is incorporated. This stochastic forcing introduces extrinsic time dependence. The evolution of the salinity gradient follows an equation analogous to the trajectory of a viscous particle within a double-welled potential, subject to Brownian agitation. If the amplitude of the stochastic driving surpasses a specific threshold, there is a finite probability of transitioning between stable equilibria. The threshold variance and the average residence time in each equilibrium are determined. For timescales approaching or exceeding the average residence time, the box model exhibits behaviour akin to a random telegraph process.

A reexamination of the box model proposed by Bryan and Hansen [26] is conducted, focusing on the regime where the stochastic forcing exhibits intermediate amplitude. Box models are regarded as coarse finite difference approximations of the continuum system, with each grid point represented by a box. They provide a straightforward framework for studying stochastically forced systems.

Fig. 10 illustrates the most basic Global Climate Model (GCM), featuring the model ocean with only two grid points. This model captures the vertically averaged circulation within a single hemisphere. Additional boxes can be included to simulate interhemispheric exchanges or to enhance resolution. However, the minimal model considered here exhibits multiple equilibria. Box 1 represents the low-latitude warm water, characterized by the temperature $T_1(t)$, while box 2 represents the high-latitude cold water, with temperature $T_2(t)$. The salinities are denoted as $S_1(t)$ and $S_2(t)$, respectively. The density is determined by the linearized equation of state,

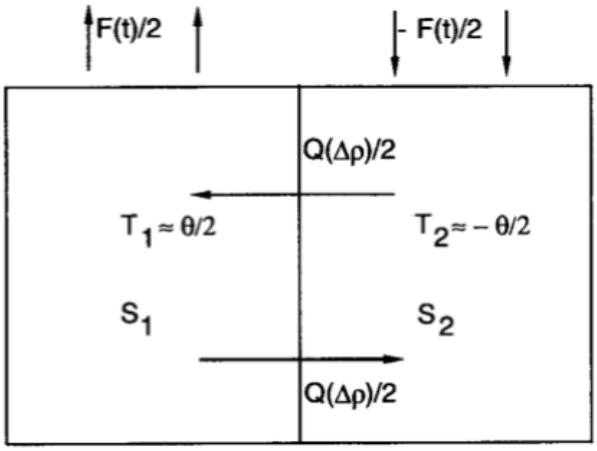

FIG. 1. The two-box model of Stommel (1961). The boxes represent two control volumes at different latitudes. Box 1 is the low-latitude box where the relaxation temperature is $\theta/2$, and box 2 is the high-latitude box where the relaxation temperature is $-\theta/2$.

Figure 10: The two-box model of Stommel[17].

relating it to the temperature and salinity.

$$\rho/\rho_0 = 1 + \alpha_S(S - S_0) - \alpha_T(T - T_0). \tag{18}$$

The index 0 denotes the reference state around which the linearization is done. The values of the parameters used are: ρ_0=1029 kg m^{-3}, T_0=5°C, S_0=35 psu, α_S=0.75 × 10^{-3} and α_T=0.17 × $10^{-3\circ}C^{-1}$. The conservation equations for the temperature and salinity are

$$\dot{T}_1 = -t_r^{-1}(T_1 - \frac{\theta}{2}) - \frac{1}{2}Q(\Delta\rho)(T_1 - T_2) \tag{19}$$

$$\dot{T}_2 = -t_r^{-1}(T_2 - \frac{\theta}{2}) - \frac{1}{2}Q(\Delta\rho)(T_2 - T_1) \tag{20}$$

$$\dot{S}_1 = \frac{F(t)}{2H}S_0 - \frac{1}{2}Q(\Delta\rho)(S_1 - S_2) \tag{21}$$

$$\dot{S}_2 = -\frac{F(t)}{2H}S_0 - \frac{1}{2}Q(\Delta\rho)(S_2 - S_1) \tag{22}$$

The modelling of mass exchange between the two boxes is accomplished using the "exchange function" denoted as Q. It is assumed that the exchange function relies solely on the density difference between the two boxes $\Delta\rho = \alpha_S(S_1 - S_2) - \alpha_T(T_1 - T_2)$. Additionally, it is assumed that Q is positive definite, ensuring that the transport of temperature, salinity, and density all occur in a downgradient manner.

The temperature variables are subject to relaxation conditions towards prescribed values θ with a certain time constant t_r, while the salinities are

forced using prescribed flux $F(t)$. Notably, there is an asymmetry in the linear forcing functions employed: relaxation forcing is used for T_1 and T_2 to capture atmospheric feedbacks, while the salinities are forced with prescribed fluxes representing imbalances between evaporation, precipitation, and runoff for each box. The depth of the model ocean is denoted as H.

By subtracting the eq 19-22, a coupled pair of equations is obtained, describing the differences in salinity and temperature between the boxes, $\Delta S = S_1 - S_2$ and $\Delta T = T_1 - T_2$,

$$\frac{d}{dt}\Delta T = -t_r^{-1}(\Delta T - \theta) - Q(\Delta\rho)\Delta T \tag{23}$$

$$\frac{d}{dt}\Delta S = \frac{F(t)}{H}S_0 - Q(\Delta\rho)\Delta S \tag{24}$$

The final form of Stommel's two-box model is achieved by combining the aforementioned system with the equation of state and a model for $Q(\Delta\rho)$.

The model used for Q is

$$Q = t_d^{-1} + V^{-1}q(\Delta\rho)^2 \tag{25}$$

because in a particular parameter range, it was discovered by Cessi and Young [27] that Boussinesq convection gives rise to a nonlinear exchange function with the mentioned form.

The system described by eq. 23 and 24 is most effectively addressed using nondimensional variables. The definitions given below are employed:

$$x = \frac{\Delta T}{\theta} \tag{26}$$

$$y = \frac{\alpha_S \Delta S}{\alpha_T \theta} \tag{27}$$

$$t = t_d t'. \tag{28}$$

The equations 23 and 24 become:

$$\dot{x} = -\alpha(x-1) - x[1 + \mu^2(x-y)^2] \tag{29}$$

$$\dot{y} = p(t) - y[1 + \mu^2(x-y)^2]. \tag{30}$$

and explicitly:

$$\dot{T} = -\alpha(T-1) - T[1 + \mu^2(T-S)^2] \tag{31}$$

$$\dot{S} = p(t) - S[1 + \mu^2(T-S)^2]. \tag{32}$$

The ratio of the diffusive timescale (t_d) to the temperature relaxation timescale (t_r), denoted as $\alpha = t_d/t_r$, provides an estimation of the order of magnitude for the nondimensional parameter α, as well as for μ^2 and p. A typical value for t_r is set at approximately 25 days, while the diffusion time t_d is calculated

as $t_d = L^2/(\pi k_H)$, where L represents the meridional scale and k_H is the coefficient of horizontal diffusion with a value of $1000 m^2 s^{-1}$. Taking $L = 8250$ km, the diffusion time amounts to 219 years, resulting in $\alpha \approx 3.6 \times 10^3$.

The parameter $\mu^2 = q t_d (\alpha_T \theta)^2 / V$ corresponds to the ratio of the diffusive timescale (t_d) to the advective timescale (t_a^{-1}), where $t_a^{-1} = q(\alpha_T \theta)^2 / V$. The control volume V is estimated as $V = L \delta_w H$, with δ_w representing the typical width of the western boundary current. This estimation is based on the hypothesis that the advective timescale is determined by the transport of the deep western boundary current rather than the interior flow. For the chosen North Atlantic context, a suitable value for the transport is $12 \times 10^6 m^3 s^{-1}$. With $H = 4500$ m and $\delta_w = 300$ km, the advective timescale is calculated as $t_d \approx 29$ years, corresponding to a boundary current velocity of approximately $0.9 \times 10^{-2} m s^{-1}$. These estimations yield $\mu^2 \approx 7.5$.

If the control volume is estimated using the entire width of the North Atlantic, the advective timescale is determined to be $t_a = 320$ years, which is on the same order of magnitude as the diffusive timescale. This result contradicts the notion that present-day transport is predominantly influenced by advective processes. The function

$$p = \frac{\alpha_S S_0 t_d}{\alpha_T \theta H} F(t) \tag{33}$$

is the freshwater flux. The nondimensional freshwater flux, denoted as $\tilde{p}$, is further divided into a time-averaged component and a stochastic component, represented as $p'(t)$, which exhibits temporal fluctuations. The evolution of this quantity poses challenges due to insufficient observations of precipitation and evaporation over the oceans. The magnitude of the fluctuating component varies depending on the timescale under consideration. Schmitt et al. [28] suggest that $p'/\tilde{p}$ exceeds 3 on seasonal timescales. However, in Cessi's study, the focus is on the scenario where $p'(t)$ is modelled as stochastic white noise. It is important to note that this choice is undoubtedly unrealistic.

The Cessi-Van Veen model

As indicated earlier, the climate model used in this research was mainly based on a combination of Van Veen's and Cessi's equations. There were slight changes implemented to better approximate the real conditions but at the same time keeping the model fairly low order. The final form of the

equations:

$$\dot{x} = -y^2 - z^2 + ax + aF_a \tag{34}$$

$$\dot{y} = xy - bxz - y + G_a \tag{35}$$

$$\dot{z} = bxy + xz - z \tag{36}$$

$$\dot{T} = -\alpha(T - 1 + \gamma x) - T[1 + \mu(T - S)^2] \tag{37}$$

$$\dot{S} = F_S + \delta(z^2 + y^2) - S[1 + \mu(T - S)^2]. \tag{38}$$

The primary differences from Van Veen's work are eq. 37 and 38, which stem from Cessi's eq. 31 and 32. In pursuit of simplicity, the temporal component of the freshwater flux in eq. 32 has been replaced with a fixed value.

3 Results

3.1 Tipping behaviour

Establishing the appropriate values for the parameters gamma γ and delta δ in the model presented a significant challenge. In the study conducted by Van Veen, a value of $7.8{\cdot}10^{-8}$ was assigned to δ, but there was no prior use of γ. The parameter γ signifies the vertical diffusivity of momentum within the model, indicating the speed of vertical momentum mixing in the ocean. A larger value for γ implies more rapid vertical mixing, which usually simplifies the mathematical solutions.

Meanwhile, the parameter δ encapsulates the essence of turbulent vertical mixing occurring within the model. This quantity is associated with the vertical diffusivity of two essential elements, heat and salt. It gives a measure of the mixing rate of these components, thereby indicating the speed of potential disruptions. Similar to γ, a larger value for δ suggests more vigorous vertical mixing of these scalars, essentially denoting a higher level of chaotic behavior.

These parameters are instrumental in defining the vertical transport of momentum, heat, and other oceanic properties in the model. Typically, their values are fine-tuned or derived based on empirical observations or theoretical considerations to accurately depict the oceanic mixing processes and their influence on the model's dynamics and thermodynamics. After careful consideration, it was opted to test parameter values of δ=0.01 and γ=0.1, as they were deemed most reflective of actual conditions. In the event that the reservoir faced difficulties in achieving adequate training (i.e., if the root mean square error exceeded a 10% threshold), a decision would be made to reduce the values of γ and δ. This adjustment is intended to diminish the noisiness of the model, thereby facilitating more accurate predictions. The final parameters used for subsequent investigations are outlined in Table 2.

a	0.25	δ	0.01
b	4	α	300
F_0	8	μ	6.25
G_0	1	γ	0.1

Table 2: Values of parameters used in the model.

Subsequently, the model was utilized to determine the freshwater flux value at which a transition occurs. The parameters from Table 2 were incorporated into the model. The model was executed several times, each time with varied freshwater flux inputs, and with each successive simulation, the flux value was increased.

The model's behavior was tracked, with a specific focus on identifying any drastic alterations. Salinity emerged as the most sensitive component to these changes, and thus, became the primary focus of this investigation. Significant shifts or instability in salinity values could potentially signify a bifurcation or a considerable weakening of the AMOC.

The freshwater flux value that coincided with the point of bifurcation or any abrupt alteration was postulated as the tipping point for the potential collapse of the AMOC within the context of this model. The graphs in Fig. 11 and 12 showed that the calculated transition happened at freshwater flux equal to 0.99 for the salinity component, which was the value that it had been hoped to obtain from the prediction phase.

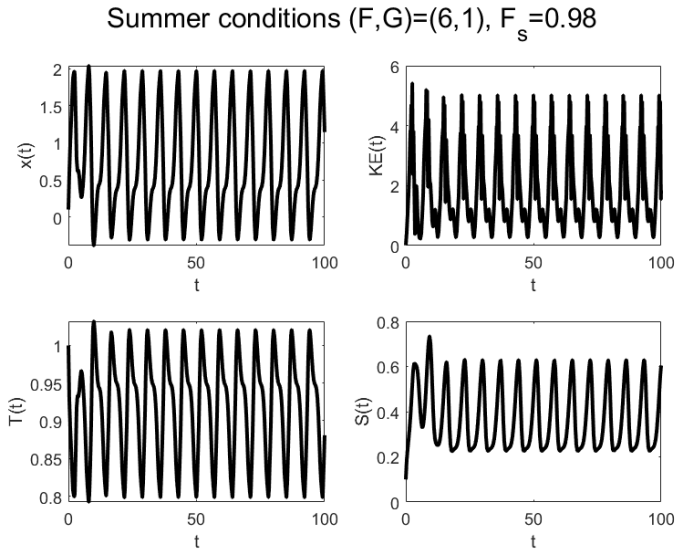

Figure 11: Model run for parameters listed in Table 2 and freshwater flux=0.98.

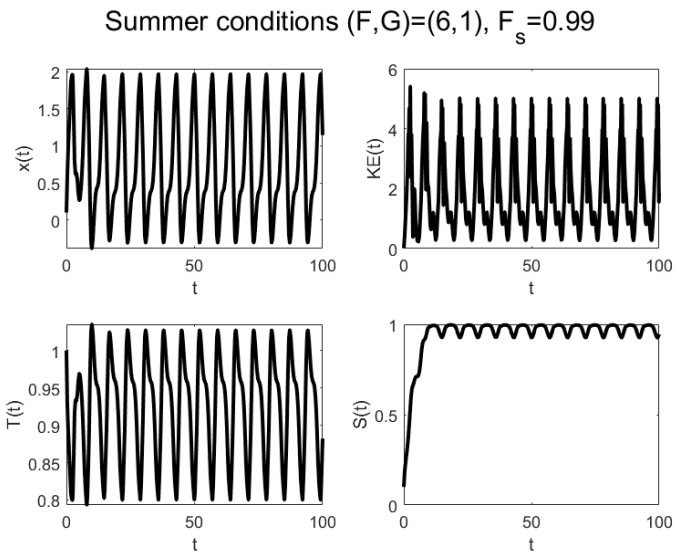

Figure 12: Model run for parameters listed in Table 2 and freshwater flux=0.99.

3.2 Training of the reservoir

In order to optimize the reservoir's training and subsequent prediction accuracy, adjustments were made to the hyperparameters. This optimization process involved the utilization of both grid search and random search methods. Notably, the optimization was conducted by considering a single set of bifurcation parameters [0.65, 0.7, 0.75], which proved to be representative of the overall performance. The results obtained from training the reservoir with these optimized hyperparameters were consistent across different sets of bifurcation parameters when additional input channels were introduced, resulting in satisfactory outcomes with an RMSE below the defined threshold of 10%.

The random search was performed using values specified in Table 3, while later the grid was executed with the parameter values outlined in Table 4. For the random search, a total of 50 iterations were conducted to explore the hyperparameter space thoroughly. The selected hyperparameter values determined through this optimization process were utilized for the summer conditions, as enumerated in Table 5.

To enhance the accuracy of the results, a decision was made to employ Bayesian optimization, as implemented in the original code [12]. The optimization process comprised 50 iterations, and each iteration was repeated three times to mitigate any potential fluctuations. Similar to the grid search approach, this optimization was conducted for a single set of bifurcation parameters. However, it yielded promising outcomes when applied to other

n	[1000, 2500, 5000, 7500, 10000]
k	[100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]
ρ	[0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25, 3.5]
W_{in}	[5, 6, 7, 8]
tp_W	[0.6, 0.7, 0.8, 0.9]
tp_{bias}	[-1.5, -1.0, -0.5, 0.0]
a	[0.2, 0.3, 0.4, 0.5]
β	[1e-8, 1e-7, 1e-6, 1e-5]

Table 3: Hyperparameters ranges in random search.

n	[1500, 2000, 2500, 3000, 3500, 4000]	tp_W	1.25
k	[800, 1000, 1200, 1400]	tp_{bias}	-0.5
ρ	[5, 6, 7]	a	0.9
W_{in}	10	β	5e-6

Table 4: Hyperparameters ranges in grid search.

n	3500	tp_W	1.25
k	800	tp_{bias}	-0.5
ρ	6	a	0.9
W_{in}	10	β	5e-6

Table 5: Hyperparameters values for the reservoir for summer conditions.

bifurcation parameter combinations as inputs to the reservoir.

Following the completion of the optimization process, the hyperparameters that yielded favorable results under summer conditions were determined and are documented in Table 6.

n	3000	tp_W	0
k	1217.32	tp_{bias}	0.61
ρ	2.8	a	0.85
W_{in}	2.04	β	4e-4

Table 6: Hyperparameters values after Bayesian optimization for the reservoir for summer conditions.

The time step used in all of the cases to prepare training data was set to Δt=1 and the training length for each value of the bifurcation parameter is $t_{train} = 3200$. The validation time length is $t_{validating} = 100$ which is approximately equivalent to 750 days.

The way the training and prediction took place is shown in Fig. 13. The reservoir machine is trained using time series data from a few distinct parameter values, all within the normal or safe regime where a chaotic attractor is still present. With the inclusion of an additional parameter input channel, the machine trained with data from different parameter values gains the ability to "sense" variations in the parameter and the associated changes in the time-series data. In this way, a well-trained machine is created, serving as a high-dimensional representation of the original dynamical system. To predict the occurrence of transient chaos and subsequent collapse with a specific parameter drift, the parameter value of interest is simply inputted into the parameter input channel.

The training data has been produced using parameters listed in Table 2 The reservoir has been trained with hyperparameters listed in Table 5. The results of the training phase are shown in Fig. 14 and RMSE values are enlisted in Table 5. Later, the optimised hyperparameters listed in Table

F_s values used for training	RMSE
[0.65, 0.7, 0.75]	0.003295
[0.7, 0.75, 0.8]	0.000617
[0.75, 0.8, 0.85]	0.006202
[0.8, 0.85, 0.9]	0.002763
[0.85, 0.9, 0.95]	0.004392
[0.9, 0.95, 1]	0.008396

Table 7: RMSE values for each bifurcation parameters set used in training. The reservoir has been trained with hyperparameters listed in Table 5.

6 were used in order to train the reservoir and the RMSE values obtained

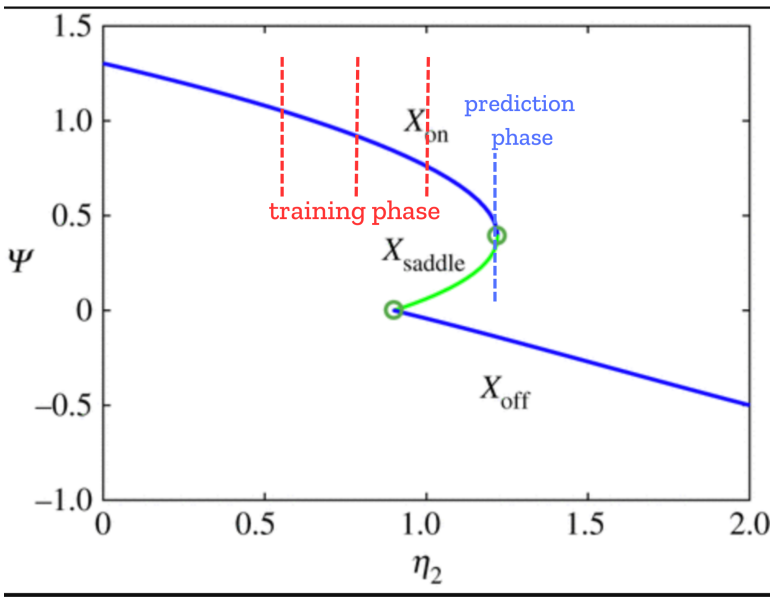

Figure 13: Schematic representation of training and predicting AMOC tipping [46].

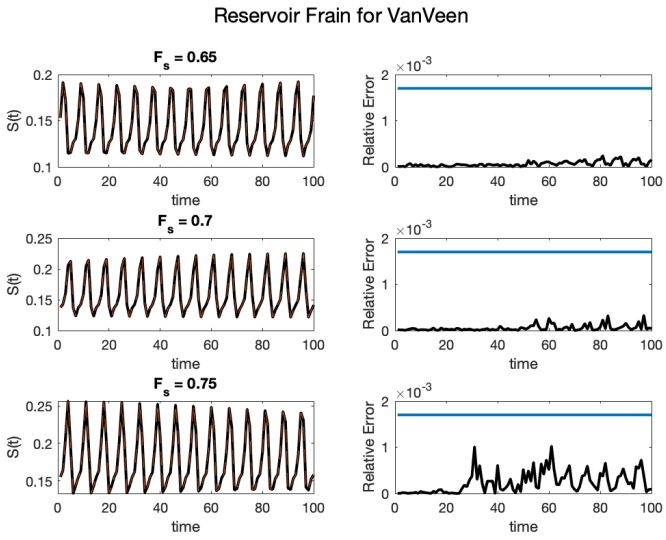

Figure 14: Training phase of the reservoir for bifurcation parameters 0.65, 0.7 and 0.75. Training was done with hyperparameters listed in Table 5.

that way are listed in Table 6.

F_s values used for training	RMSE
[0.65, 0.7, 0.75]	0.000360
[0.7, 0.75, 0.8]	0.001100
[0.75, 0.8, 0.85]	0.000120
[0.8, 0.85, 0.9]	0.000092
[0.85, 0.9, 0.95]	0.000157
[0.9, 0.95, 1]	0.000092

Table 8: RMSE values for each bifurcation parameters set used in training. The reservoir has been trained with hyperparameters listed in Table 6.

As can be observed, the RMSE values acquired from the reservoir trained with optimized hyperparameters are markedly (tenfold) smaller than those acquired from the reservoir trained with hyperparameters obtained from grid and random search. It must be kept in mind that a random factor exists in training the reservoir, which can cause fluctuations. Consequently, it is considered a sound approach to execute the same iteration multiple times, thereby mitigating randomness influence.

3.3 Prediction of tipping point

The major objective of this research was to assess the predictive accuracy of the model in anticipating system transitions by examining a range of parameters prior to the occurrence of these transitions. To carry out this phase of the research, the following steps were undertaken:

- The reservoir was trained using the bifurcation parameters of 0.65, 0.7 and 0.75 while other parameters were kept the same as listed in Table 2 and the hyperparameters listed in Table 5).
- Following the training process, the Root Mean Square Error (RMSE) was evaluated to ensure that it fell below the predefined threshold of 10%.
- A prediction phase was conducted for the bifurcation parameters within the range of 0.9 to 1.2, where the model predicts the transition at 0.99 for γ=0.1 and δ=0.01. This phase lasted for 150 time units, during which the final observed salinity value was recorded.
- A visualization was created in the form of a graph that displayed the ratio $\frac{(S-\bar{S})^2}{\bar{S}^2}$ along the y-axis, with S representing the predicted salinity value and $\bar{S}$ representing the true salinity value calculated by the model. The bifurcation parameter value that was hoped to be

predicted, represented by F_s^P, was plotted along the x-axis. The midpoint of the training bifurcation parameter set, which equated to 0.7, was used as a reference for which the true value $\bar{S}$ has been calculated. Upon examination of the graph, the first record where the difference in the ratio surpassed a value of 1 was identified. This signified the bifurcation parameter at which the predicted salinity value matched the true salinity value. The corresponding value of the bifurcation parameter, denoted as F_s^R, was noted for further reference.

- To minimize the influence of random variations in the results, five trials of this experiment were conducted, and the mean value and standard deviation were calculated for each dataset.

This procedure was iterated for various sets of bifurcation parameters, specifically [0.7, 0.75, 0.8], [0.75, 0.8, 0.85], [0.8, 0.85, 0.9], [0.85, 0.9, 0.95], and [0.9, 0.95, 1.0]. The purpose was to assess the quality of predictions in relation to the extent of reservoir training prior to the occurrence of the transition. To facilitate this evaluation, an error graph was constructed.

Subsequently, a comprehensive analysis was conducted on the specified hyperparameters outlined in Table 6, maintaining consistent model parameters. To augment the error graph, an additional data point was appended subsequent to training the reservoir with bifurcation parameters [0.95, 1.0, 1.05] during the transient chaos phase, with the intention of assessing the machine's ability to accurately predict the freshwater flux despite being trained with parameters in a state of transient chaos. Upon confirming that the RMSE satisfied the predetermined threshold of 10%, the subsequent phase involved conducting the prediction process. With the anticipation of the transition occurring at a freshwater flux value of 0.99 showed by the original model, the reservoir model was subjected to the prediction phase within the range of 0.9 to 1. This range was visually represented in Figures 15 and 16, showcasing the span within which the transition was expected to manifest. As can be discerned, the transition occurs within this range, with salinity values doubling once the transition has taken place.

A graph illustrating the error pattern in relation to the proximity from a transition point has been compiled. The y-axis of the graph represented the normalized difference between the predicted freshwater flux value for the transition (F_S^R) and the actual value from the model (F_S^a), expressed as $\frac{|F_S^R - F_S^a|}{F_a}$, where F_S^a denotes the actual freshwater flux value. On the x-axis, the normalized difference between the bifurcation parameter used during reservoir training (F_S^t) and the actual value from the model (F_S^a) was plotted, denoted as $\frac{|F_S^t - F_S^a|}{F_S^a}$. The value of F_S^t was set to be the middle of each bifurcation parameters set used in training phase.

The error graph was constructed to investigate the relationship between the quality of predictions and the extent of reservoir training preceding the

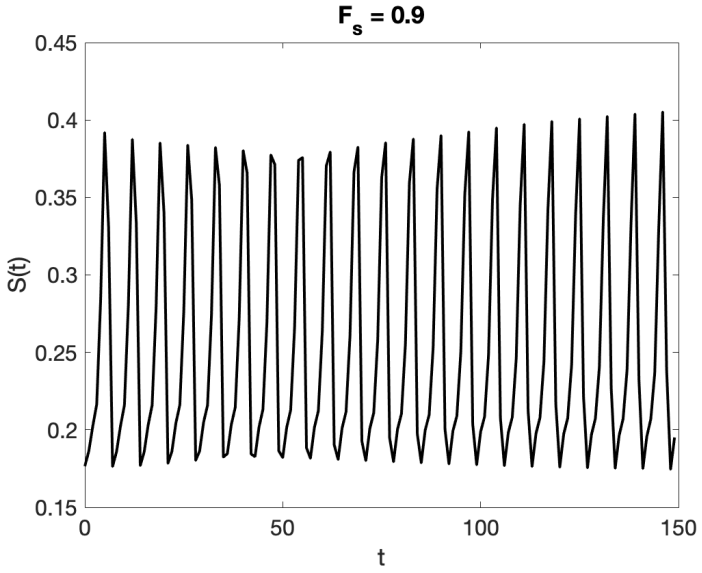

Figure 15: Prediction phase of the trained reservoir for bifurcation parameter 0.9. Training was done with hyperparameters enlisted in Table 5.

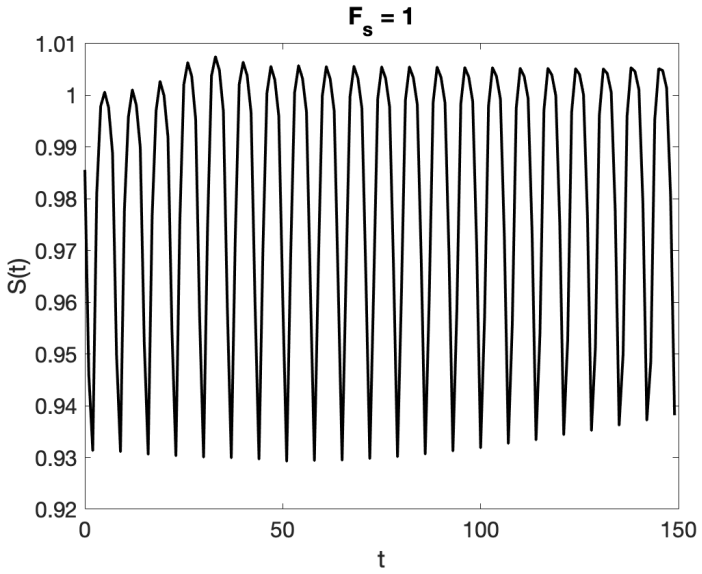

Figure 16: Prediction phase of the trained reservoir for bifurcation parameter 1. Training was done with hyperparameters enlisted in Table 5.

transition occurrence.

In order to futher validate the accuracy of each predicted value, a rigorous 3σ test was performed to determine whether the predicted value aligned with the corresponding true value. Adherence was mandatory for every predicted value to the prescribed rule

$$|F_S^R - F_S^a| < 3\sigma \tag{39}$$

where σ is the calculated standard deviation. The results are shown in Fig. 17.

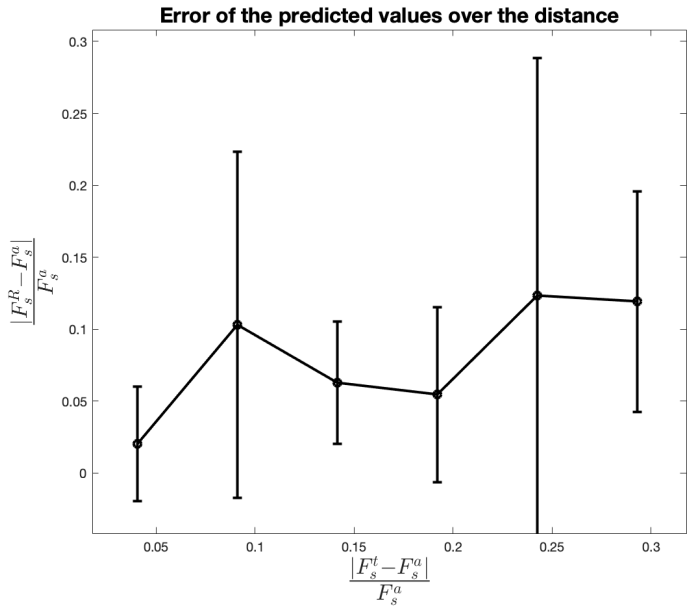

Figure 17: Error in prediction phase against the distance from the transition phase.

The values used to produce the graph are listed in Table 9.

Training bifurcation parameter	Predicted value	Error of the trained value
[0.65, 0.7, 0.75]	1.11	0.08
[0.7, 0.75, 0.8]	1.11	0.16
[0.75, 0.8, 0.85]	1.04	0.06
[0.8, 0.85, 0.9]	1.05	0.04
[0.85, 0.9, 1]	1.09	0.11
[0.9, 0.95, 1]	0.97	0.04

Table 9: Predicted freshwater flux values after 5 trials for each training bifurcation parameters set and their error.

For the optimized hyperparameters listed in Table 6 a graph shown in Fig 18 has been prepared in order to see how much better the reservoir did using optimised hyperparameters.

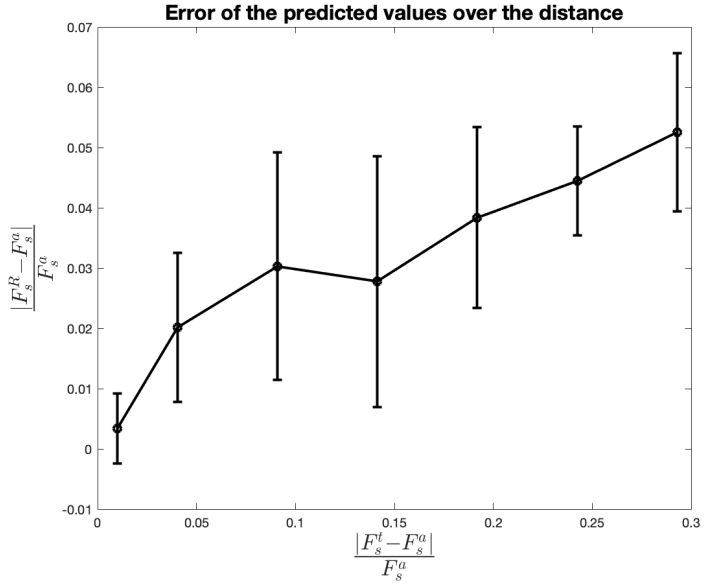

Figure 18: Error of the predicted values against the distance from the transition phase after optimisation of hyperparameters.

The values used to produce the graph are listed in Table 10.

Training bifurcation parameter	Predicted value	Error of the trained value
[0.65, 0.7, 0.75]	0.938	0.042
[0.7, 0.75, 0.8]	0.946	0.034
[0.75, 0.8, 0.85]	0.952	0.028
[0.8, 0.85, 0.9]	0.962	0.017
[0.85, 0.9, 1]	0.960	0.020
[0.9, 0.95, 1]	0.970	0.010
[0.95, 1, 1.05]	0.986	0.006

Table 10: Predicted freshwater flux values after 5 trials for each training bifurcation parameters set and their error after optimisation of hyperparameters.

Table 11 presents the results obtained from five trial runs for each training bifurcation parameter with hyperparameters from Table 5. The table includes the predicted bifurcation parameters, their corresponding errors, the disparity between the predicted values and the actual values derived from the model, three times the error, and a verification of whether the

predicted values fall within the range of three standard deviations from the actual values. The value described as training bifurcation parameter is the middle value of each of the set used in training phase.

t_p	F_S^a	F_S^R	$\sigma_{F_F^R}$	$\lvert F_S^a\text{-}F_S^R\rvert$	$3\sigma_{F_F^R}$	3σ test
0.7	0.99	1.108	0.075	0.118	0.227	Satisfies
0.75	0.99	1.112	0.163	0.122	0.490	Satisfies
0.8	0.99	1.044	0.060	0.054	0.180	Satisfies
0.85	0.99	1.052	0.042	0.062	0.126	Satisfies
0.9	0.99	1.092	0.119	0.102	0.357	Satisfies
0.95	0.99	0.970	0.039	0.020	0.118	Satisfies

Table 11: Predicted bifurcation parameters analysis and deviation verification.

Subsequently, an analysis was conducted using the same model, but with the hyperparameters set obtained from Table 6. The previously employed method, as outlined in Table 11, was utilized to assess the accuracy of the predicted values following the implementation of the optimized hyperparameters. The resulting outcomes are presented in Table 12.

t_p	F_S^a	F_S^R	$\sigma_{F_F^R}$	$\lvert F_S^a\text{-}F_S^R\rvert$	$3\sigma_{F_F^R}$	3σ test
0.7	0.99	0.938	0.013	0.052	0.039	Doesn't satisfy
0.75	0.99	0.946	0.009	0.044	0.027	Doesn't satisfy
0.8	0.99	0.952	0.014	0.038	0.044	Satisfies
0.85	0.99	0.963	0.021	0.028	0.061	Satisfies
0.9	0.99	0.960	0.019	0.030	0.056	Satisfies
0.95	0.99	0.970	0.012	0.020	0.037	Satisfies
1	0.99	0.986	0.006	0.003	0.017	Satisfies

Table 12: Predicted bifurcation parameters analysis and deviation verification after optimization of the hyperparameters.

4 Discussion

The conducted training phase, as depicted in Figure 14, exhibited favorable outcomes across all parameter sets. The attainment of low RMSE values instilled confidence in the adequacy of the reservoir's training. Subsequently, the prediction phase demonstrated the reservoir's ability to anticipate the occurrence of a transition, albeit with a slightly earlier forecast compared to the computed value from the model. This discrepancy may be attributed to the specific hyperparameter values employed during the training phase. In an effort to mitigate this possibility, a two-fold approach was adopted:

initially, a grid search and random search were performed to explore the hyperparameter space. Due to the time-intensive nature of the optimization process, it was not feasible to conduct multiple runs of the optimization. Consequently, a Bayesian optimization approach was subsequently employed to fine-tune the hyperparameters. It is noteworthy to mention that the optimization process was not executed for several hundred iterations as described in the original code [12]. This deviation was necessary due to the increased complexity of the model employed in this research, which entails a lengthier training phase.

As previously mentioned, it was anticipated that the error graphs in Fig 17 and 18 would exhibit a diminishing trend as the bifurcation parameter approached the transition point. This expectation stemmed from the understanding that predicting further into the future becomes increasingly challenging due to the inherent chaotic nature of the system. Encouragingly, both of the presented graphs align with this anticipated pattern, thus reinforcing the reliability and credibility of the obtained results.

Both the graph shown in Fig 17 and the accompanying table 11 showcase the reservoir's remarkable accuracy in predicting the freshwater flux value to be 0.99. Notably, all predicted values for each bifurcation parameter consistently fall within the 3 standard deviation range, thus providing compelling evidence of the efficacy and reliability of the employed methodology. These findings affirm that the method yields highly satisfactory outcomes.

An evident observation is that the predictions stemming from the bifurcation parameters furthest from the transition phase consistently fall within the 3 standard deviation range. Consequently, these predictions can be deemed unreliable and should be disregarded.

Upon comparing the outcomes presented in the two tables, a clear distinction arises from the application of Bayesian optimization to the hyperparameters. The resulting predicted values in the optimized reservoir exhibit significantly smaller discrepancies when compared to the values calculated from the model. This substantial improvement in accuracy highlights the efficacy of Bayesian optimization in refining the reservoir's performance.

It is worth noting that in the first case, where all predicted values fall within the three standard deviation range of the actual values, the relatively larger calculated standard deviations play a crucial role. This observation emphasizes the importance of appropriately tuned hyperparameters in achieving more reliable predictions.

The notable advantage of Bayesian optimization should be acknowledged, despite its potential time-consuming nature, especially when executed on computers with limited processing capabilities or a small number of cores. Nonetheless, the obtained results demonstrate the significant benefit of employing Bayesian optimization in enhancing prediction accuracy and warrant the investment of computational resources required for its execution.

For further research, one compelling avenue to pursue entails the exploration of a 'winter conditions' scenario, which has presented significant challenges to the computational resources at hand due to its chaotic nature. Such a system necessitates a meticulously optimised reservoir, with an emphasis on the careful selection of hyperparameters, to ensure effective training and subsequent accurate predictions.

This study has shown that, given the hyperparameters γ and δ that were utilized in this research, achieving RMSE values less than 10% of the trained values was unfeasible. This finding underscores the sensitivity of the system to the choice of hyperparameters and the subsequent requirement for their judicious selection and tuning. The current limitations faced in this context provide a strong impetus for further research into refining computational methods for handling complex, chaotic systems such as the 'winter conditions' scenario.

5 Conclusions

The objective of this research was to investigate the feasibility of predicting a low-order idealized ocean-atmosphere model using reservoir computing. The results presented in the previous sections demonstrate the successful attainment of this goal, affirming that reservoir computing can indeed yield reasonable predictions. Notably, reservoir computing exhibits significant potential in accurately predicting transition phases within chaotic systems.

Considering that Global Climate Models (GCMs) often require extensive computational resources and prolonged execution times, reservoir computing holds promise as a valuable tool for expediting GCM simulations or providing estimations regarding specific components of the climate system approaching transition phases. The selection of appropriate hyperparameters for the reservoir is crucial to ensure its optimal performance. Moreover, it is important to acknowledge that the computational time required by the reservoir increases as more complex systems are fed into it. In this context, parallel computing emerges as a useful approach to expedite the computation process.

Parallel computing entails the simultaneous execution of multiple computational tasks by dividing them into smaller subtasks that can be executed concurrently. By harnessing multiple processing units or cores, parallel computing enhances computational efficiency and reduces overall execution time. In this study, a maximum of two nodes were employed concurrently. However, it is reasonable to expect that utilizing a greater number of nodes would yield improved results within a reasonable timeframe.

Overall, the findings of this research underscore the potential of reservoir computing in predicting complex systems and highlight the importance of hyperparameter selection.

References

[1] https://worldwarzero.com/magazine/2022/06/the-collapse-of-a-major-atlantic-current-would-cause-worldwide-disasters/

[2] Crivellari, Stefano. (2018). Effects of abrupt changes in the Atlantic meridional overturning circulation over the Amazon Basin: an isotopic and elemental approach. 10.13140/RG.2.2.32595.20000.

[3] Boulton, Chris & Allison, Lesley & Lenton, Timothy. (2014). Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model. Nature communications. 5. 5752. 10.1038/ncomms6752.

[4] Weijer, W., Cheng, W., Drijfhout, S. S., Federov, A.V., Hu, A., Jackson, L. C., et al. (2019). 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

[5] Smith, R. S., Gregory, J. M., and Osprey, A.: A description of the FAMOUS (version XDBUA) climate model and control run, Geosci. Model Dev., 1, 53–68, https://doi.org/10.5194/gmd-1-53-2008, 2008.

[6] Jackson, L.C., Kahana, R., Graham, T. et al. Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM. Clim Dyn 45, 3299–3316 (2015). https://doi.org/10.1007/s00382-015-2540-2

[7] Abrupt Climate Changes: How Freshening of the Northern Atlantic Affects the Thermohaline and Wind-Driven Oceanic Circulations Marcelo Barreiro, Alexey Fedorov, Ronald Pacanowski, S. George Philander Annual Review of Earth and Planetary Sciences 2008 36:1, 33-58

[8] Blunier T, Brook EJ. Timing of millennial-scale climate change in Antarctica and Greenland during the last glacial period. Science. 2001 Jan 5;291(5501):109-12. doi: 10.1126/science.291.5501.109. PMID: 11141558.

[9] Dansgaard, W., Johnsen, S., Clausen, H. et al. Evidence for general instability of past climate from a 250-kyr ice-core record. Nature 364, 218–220 (1993). https://doi.org/10.1038/364218a0

[10] https://en.wikipedia.org/wiki/Multiple_equilibria_in_the_Atlantic_meridional_overturning_circulation

[11] Michel SLL, Swingedouw D, Ortega P, Gastineau G, Mignot J, McCarthy G, Khodri M. Early warning signal for a tipping point suggested by a millennial Atlantic Multidecadal Variability reconstruction. Nat Commun. 2022 Sep 2;13(1):5176. doi: 10.1038/s41467-022-32704-3. PMID: 36056010; PMCID: PMC9440003.

[12] Kong, Ling-Wei & Fan, Hua-Wei & Grebogi, Celso & Lai, Ying-Cheng. (2020). Machine learning prediction of critical transition and system collapse.

[13] Tardif, R., Hakim, G.J. & Snyder, C. Coupled atmosphere–ocean data assimilation experiments with a low-order climate model. Clim Dyn 43, 1631–1643 (2014). https://doi.org/10.1007/s00382-013-1989-0

[14] Paul J. Roebber (1995) Climate variability in a low-order coupled atmosphere-ocean model, Tellus A: Dynamic Meteorology and Oceanography, 47:4, 473-494, DOI: 10.3402/tellusa.v47i4.11534

[15] Stommel, Henry (1961) Thermohaline Convection with Two Stable Regimes of Flow. Tellus, 13 (2). 224-230 doi:10.3402/tellusa.v13i2.9491

[16] Lennaert Van Veen, Theo Opsteegh & Ferdinand Verhulst (2001) Active and passive ocean regimes in a low-order climate model, Tellus A: Dynamic Meteorology and Oceanography, 53:5, 599-615, DOI: 10.3402/tellusa.v53i5.12229

[17] CESSI, Paola. A simple box model of stochastically forced thermohaline flow. Journal of Physical Oceanography, 1994, 24.9: 1911-1920.

[18] Lorenz, E.N. (1963) Deterministic Nonperiodic Flow. Journal of Atmospheric Sciences, 20, 130-141. http://dx.doi.org/10.1175/1520-0469(1963)020¡0130:dnf¿2.0.co;2

[19] E. N. Lorenz (1990) Can chaos and intransitivity lead to interannual variability?, Tellus A: Dynamic Meteorology and Oceanography, 42:3, 378-389, DOI: 10.3402/tellusa.v42i3.11884

[20] Maas, Leo R. M.. "A simple model for the three-dimensional, thermally and wind-driven ocean circulation." Tellus A 46 (1994): 671-680.

[21] Buckley, Martha W., and John Marshall. "Observations, Inferences, and Mechanisms of the Atlantic Meridional Overturning Circulation: A Review." Reviews of Geophysics 54, no. 1 (January 26, 2016): 5–63.

[22] [Roberts, Jackson, and McNeall(2014)]2014GeoRL..41.3204RRoberts, C.D., Jackson, L., and McNeall, D.: 2014, Geophys. Res. Lett. 41, 3204. doi:10.1002/2014GL059473.

[23] Reintges, A., Martin, T., Latif, M. et al. Uncertainty in twenty-first century projections of the Atlantic Meridional Overturning Circulation in CMIP3 and CMIP5 models. Clim Dyn 49, 1495–1511 (2017). https://doi.org/10.1007/s00382-016-3180-x

[24] Lehman, S., Keigwin, L. Deep circulation revisited. Nature 358, 197–198 (1992). https://doi.org/10.1038/358197b0

[25] Fairbanks, R. A 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. Nature 342, 637–642 (1989). https://doi.org/10.1038/342637a0

[26] Bryan, K., and F.C. Hansen, 1993: A toy model of North Atlantic climate variability on a decade to century time-scale, The Natural Variability of the Climate System on 10-100 Year Time Scales, U.S. Natl. Acad. of Sci.

[27] Cessi, P., & Young, W. (1992). Multiple equilibria in two-dimensional thermohaline circulation. Journal of Fluid Mechanics, 241, 291-309. doi:10.1017/S0022112092002040

[28] Schmitt, Raymond W., Philip Bogden and Clive E. Dorman. "Evaporation Minus Precipitation and Density Fluxes for the North Atlantic." Journal of Physical Oceanography 19 (1989): 1208-1221.

[29] The Role of the Gulf Stream in European Climate Jaime B. Palter Annual Review of Marine Science 2015 7:1, 113-137

[30] Smeed, David & Josey, Simon & Beaulieu, Claudie & Johns, W. & Moat, Ben & Frajka-Williams, Eleanor & Rayner, D. & Meinen, Christopher & Baringer, M. & Bryden, H. & McCarthy, G.. (2018). The North Atlantic Ocean Is in a State of Reduced Overturning. Geophysical Research Letters. 45. 10.1002/2017GL076350.

[31] Muir, L.C., Fedorov, A.V. How the AMOC affects ocean temperatures on decadal to centennial timescales: the North Atlantic versus an interhemispheric seesaw. Clim Dyn 45, 151–160 (2015). https://doi.org/10.1007/s00382-014-2443-7

[32] Cheng, Wei & Chiang, John & Zhang, Dongxiao. (2013). Atlantic Meridional Overturning Circulation (AMOC) in CMIP5 models: RCP and historical simulations. Journal of Climate. 26. 7187-7197. 10.1175/JCLI-D-12-00496.1.

[33] Dima, M., & Lohmann, G. (2010). Evidence for two distinct modes of large-scale ocean circulation changes over the last century. Journal of Climate, 23(1), 5–16

[34] Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., & Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. Nature Climatic Change, 5(5), 475.

[35] Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., & Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. Nature, 556(7700), 191.

[36] Marshall, J., & Schott, F. (1999). Open-ocean convection: Observations, theory, and models. Reviews of Geophysics, 37(1), 1–64

[37] Emile-Geay, J., Cane, M. A., Naik, N., Seager, R., Clement, A. C., & van Geen, A. (2003). Warren revisited: Atmospheric freshwater fluxes and why is no deep water formed in the North Pacific. Journal of Geophysical Research, 108(C6), 9–1. https://doi.org/10.1029/2001JC001058

[38] Schmitt, R. W. (2008). Salinity and the global water cycle. Oceanography, 21(1), 12–19.

[39] Warren, B. A. (1983). Why is no deep water formed in the North Pacific?. Journal of Marine Research, 41(2), 327–347.

[40] Weijer, W., De Ruijter, W. P. M., Dijkstra, H. A., & van Leeuwen, P. J. (1999). Impact of interbasin exchange on the Atlantic overturning circulation. Journal of Physical Oceanography, 29, 2266–2284

[41] Gordon, A. L. (1986). Interocean exchange of thermocline water. Journal of Geophysical Research, 91(C4), 5037–5046.

[42] Marsh, R., Hazeleger, W., Yool, A., & Rohling, E. J. (2007). Stability of the thermohaline circulation under millennial CO2 forcing and two alternative controls on Atlantic salinity. Geophysical Research Letters, 34, L03605. https://doi.org/10.1029/2006GL027815

[43] An, S.-I., Kim, H.-J., and Kim, S.-K., "Rate-Dependent Hysteresis of the Atlantic Meridional Overturning Circulation System and Its Asymmetric Loop", ¡i¿Geophysical Research Letters¡/i¿, vol. 48, no. 1, 2021. doi:10.1029/2020GL090132.

[44] N. D. Haynes, M. C. Soriano, D. P. Rosin, I. Fischer, and D. J. Gauthier, Reservoir computing with a single time-delay autonomous Boolean node, Phys.Rev.E91, 020801(R) (2015).

[45] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, 1st ed. (Addison-Wesley Longman Publish- ing Co., Inc., USA, 1989).

[46] Alkhayuon H, Ashwin P, Jackson LC, Quinn C, Wood RA. Basin bifurcations, oscillatory instability and rate-induced thresholds for Atlantic meridional overturning circulation in a global oceanic box model. Proc Math Phys Eng Sci. 2019 May;475(2225):20190051. doi: 10.1098/rspa.2019.0051. Epub 2019 May 15. PMID: 31236059; PMCID: PMC6545045.

[47] Frajka-Williams, Eleanor & Ansorge, Isabelle & Baehr, Johanna & Bryden, Harry & Chidichimo, Maria & Cunningham, Stuart & Danabasoglu,

Gokhan & Dong, Shenfu & Donohue, Kathleen & Elipot, Shane & Heimbach, Patrick & Holliday, N. & Hummels, Rebecca & Jackson, Laura & Karstensen, Johannes & Lankhorst, Matthias & Bras, Isabela & Lozier, M. & Mcdonagh, Elaine & Wilson, Chris. (2019). Atlantic Meridional Overturning Circulation: Observed Transport and Variability. Frontiers in Marine Science. 6. 10.3389/fmars.2019.00260.

[48] Lozier, M.. (2012). Overturning in the North Atlantic. Annual review of marine science. 4. 291-315. 10.1146/annurev-marine-120710-100740.

[49] Lorenz, Edward Norton. "Irregularity: a fundamental property of the atmosphere*." Tellus A 36 (1984): 98-110.