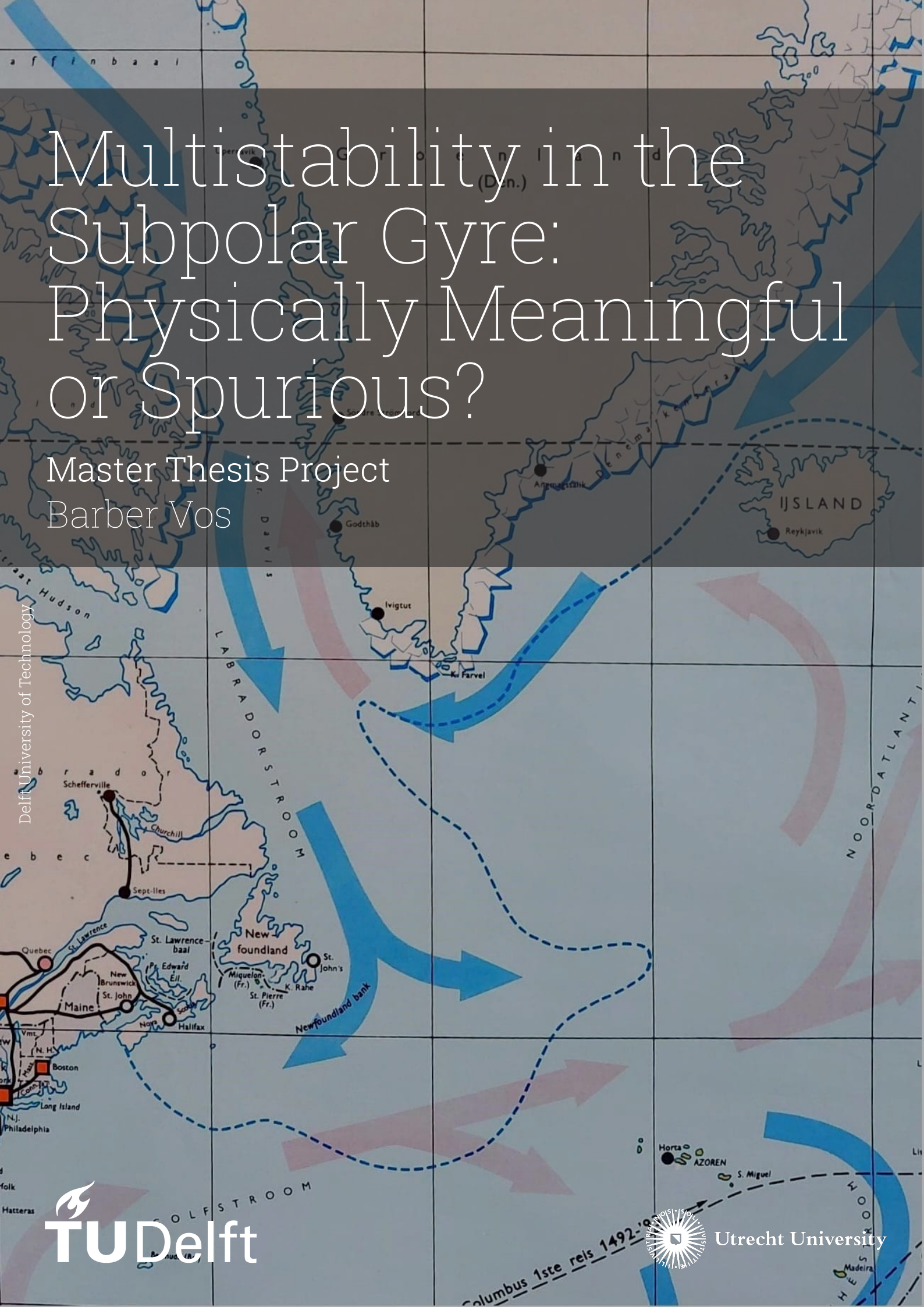
Multistability in the Subpolar Gyre: Physically Meaningful or Spurious?

Master Thesis Project

Barber Vos

Delft University of Technology



TU Delft

Utrecht University

Applied Mathematics
Computational Science and Engineering

Multistability in the Subpolar Gyre: Physically Meaningful or Spurious?

Author
Barber Vos
(5048826)

Thesis Committee

Responsible TU Delft Supervisor:	Prof. Dr. H.M. Schuttelaars
IMAU Supervisor:	Prof. Dr. A.S. Von der Heydt
IMAU Supervisor:	Prof. Dr. Ir. H.A. Dijkstra
TU Delft Committee Member:	Prof. Dr. C.A. Katsman

September 2025

Preface

This thesis is the final project of my Master's programme in Applied Mathematics at TU Delft, carried out in collaboration with the Institute for Marine and Atmospheric Research Utrecht (IMAU). Over the past eight months, I have had the opportunity to explore the fascinating and multidisciplinary topic of multistability in the Subpolar Gyre.

Working on this project has taught me not only about the complexities of climate and ocean models, but also about the process of conducting independent research. There were moments of confusion and frustration, but also of clarity and excitement when ideas finally came together. I am grateful to my supervisors, Henk Schuttelaars, Anna von der Heydt and Henk Dijkstra, for their guidance, encouragement, and insightful feedback throughout this journey. I would also like to thank other members of IMAU and TU Delft for many stimulating discussions, and experts such as Caroline Katsman and Bernd Krauskopf for their valuable input.

Finally, I want to thank my friends for the many long and much-needed coffee breaks, and my family for their continuous support. This thesis marks the conclusion of my time as a student, a period full of growth, learning, and great friendships.

Barber Vos
Delft, September 2025

Abstract

The Atlantic Meridional Overturning Circulation (AMOC) is a key component of the climate system, regulating heat and freshwater transport across the globe. Its stability is of particular concern because an AMOC weakening or collapse could trigger abrupt and potentially irreversible climate shifts. A critical region for this stability is the North Atlantic Subpolar Gyre (SPG), where deep convection hepls drive the AMOC. Yet, the literature presents contrasting explanations for multistability of convection in the SPG: the continuous horizontal box model of Bastiaansen suggests physically meaningful multistability, whereas the one-dimensional column model of Den Toom points to spurious multistability.

This thesis investigates whether the multistability observed in complex ocean models reflects genuine physical processes or arises as an artifact of modeling techniques. To bridge the contrasting perspectives by Bastiaansen and Den Toom, an overarching two-dimensional model is developed that contains elements of both approaches, allowing reproduction of each limiting case and systematic exploration of their combined effects.

Analysis shows that multistability in Bastiaansen's model depends critically on parameter choices, disappearing when physically realistic parameters are used. In Den Toom's model, multistability vanishes with increased vertical resolution, confirming that it is a numerical artifact. Since climate models lack sufficient vertical resolution, this artifact is relevant in practice. The overarching model demonstrates that a combination of both mechanics can introduce additional multistability on top of multistability similar to Den Toom, but these features vanish with higher horizontal resolution.

Overall, the results indicate that the multistability observed in these idealized ocean models is spurious, arising from unrealistic parameters or insufficient numerical resolution rather than reflecting true ocean dynamics. These findings highlight the importance of physically justified parameterizations and sufficient model resolution in interpreting bifurcation structures and potential tipping points in climate models.

Contents

1 Introduction

The Earth's climate system is regulated by a delicate balance of interacting components, an important one being the Atlantic Meridional Overturning Circulation (AMOC). This massive ocean current acts like a global conveyor belt, transporting warm water from the tropics to the North Atlantic and returning cold, dense water to the deep ocean. It plays a vital role in regulating global temperatures and weather patterns [1]. One well-known effect is its contribution to Europe's relatively mild climate compared to other regions at similar latitudes [2].

However, recent research suggests that the AMOC may be more vulnerable to disruption than previously assumed [3]. A particularly important region is the Subpolar Gyre (SPG) in the North Atlantic, see Figure 1. Here, deep-ocean convection helps drive the AMOC: surface waters cool, increase in density and sink [4, 5]. If this convection process is disrupted, it could destabilize the entire circulation system [3].

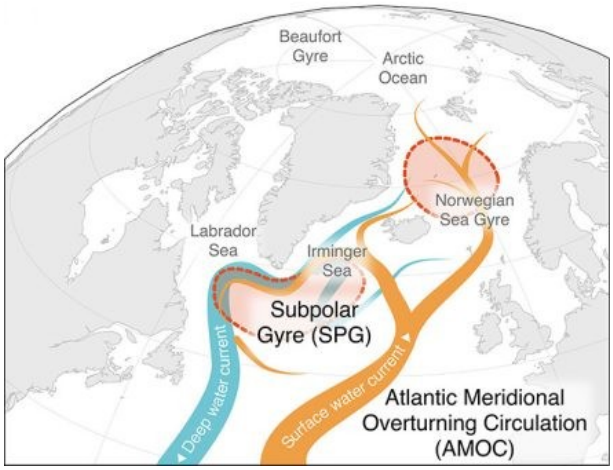


Figure 1: Map of the SPG and its connection to the AMOC [3]. The SPG is marked with red, containing the Labrador Sea and Irminger Sea. The AMOC is represented by the (warm) surface water current in orange and the (cold) deep water current in blue.

One way this disruption might occur is through freshwater input, for example from melting ice sheets. This reduces surface water density, making sinking less likely. This process, known as stratification, can block vertical mixing and potentially shut down deep convection. This mechanism can lead to multistability: the existence of multiple stable states under the same conditions. Such states can feature full, partial, or no deep convection [6].

Multistability is often visualized using bifurcation diagrams, which are mathematical tools that show how equilibrium states of a system change when a control parameter, such as freshwater input, is varied. Simulations in complex climate models reveal a variety of such stable states in the AMOC [7, 8], an illustrative example is given in Box 1. Transitions between these states could occur rapidly, potentially within a few decades, with significant implications for the climate system [9].

Box 1: Multistability in Complex Climate Models

Simulations by Lohmann [7] using Veros show rich multiplicity of the state of the AMOC over freshwater forcing. Veros is a versatile ocean simulator that supports realistic, high-resolution, global ocean simulations [10]. In the bifurcation diagram, see Figure 2, each dot represents a stable state, with colors marking different branches of stability.

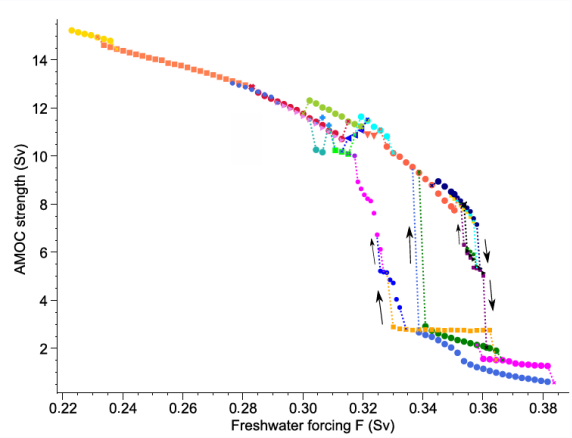


Figure 2: Bifurcation diagram of the state of the AMOC with the freshwater forcing as control parameter, showing multistability. The vertical axis shows the mean AMOC strength during the last 1000 years of simulation. Each dot represents a stable state, with colors marking different branches of stability [7].

The main hysteresis, the loop in Figure 2, displays the classical result of convection or no convection [11]. More smaller jumps between different stable branches are visible, which are labeled as "snaking" patterns. These coincide with changes in convection patterns [7], as illustrated in Figure 3. Here, maps of the modeled area are shown depicting the mixed layer depth (MLD) for different equilibrium simulations in Figure 2. The MLD is the depth until which the monthly average ocean temperature is within 0.5 K of the sea surface temperature. It gives an indication of the stratification.

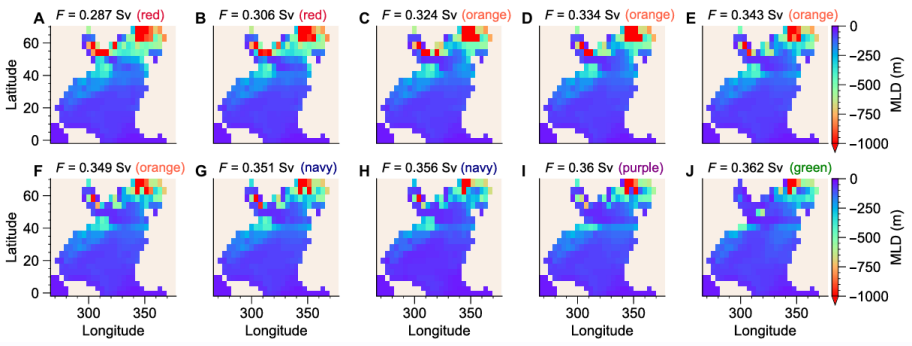


Figure 3: Spatial patterns of convection for different stable states depicting the 100-year average winter MLD [7]. The winter MLD is obtained by averaging over the first 4 months of the year. the color coding in the panel titles indicates the branches of attractors in Figure 2.

To explore this multistability, often idealized models are used. These are valuable tools for studying isolated mechanisms that govern stability and transitions in ocean circulation. Two idealized models that provide insight into the origins of multistability of the SPG are those developed and analyzed by Bastiaansen et al [12] and Den Toom et al [13].

The model of Bastiaansen et al [12] builds on the classical two-box convection model [14], which captures vertical mixing between surface and deep waters. By extending this framework to include horizontal spatial diffusion and spatially varying atmospheric forcing, Bastiaansen shows that multistability can arise from physical processes such as a spatially varying freshwater input from meltwater [12].

Den Toom et al [13] studies a one-dimensional vertical model with convection. This model also exhibits multistability. However, in this case, the behavior depends on the number of vertical layers used, indicating that the bifurcations might be numerical artifacts rather than physical mechanisms of convection [13].

The difference in origin of the multistability in the simplified models raises the following question:

Is the multistability observed in complex ocean models a true reflection of physical processes, or an artifact of modeling techniques?

Understanding this distinction is essential, not only for the credibility of numerical simulations but also for predicting tipping points in climate systems.

To tackle this question, the two models of Bastiaansen and Den Toom are analyzed in isolation, their model components are compared, and their characteristics that lead to multistability are combined into an overarching model. In Chapter 2, the existing models are introduced in more detail. It also outlines the key differences between the models of Bastiaansen and Den Toom, with a focus on identifying how they can be unified into one overarching model, introduced in Chapter 3. Here, the implementation of this model and the used analysis method are also explained. Next, in Chapter 4, the overarching model is first validated by checking the known limit cases studied by Bastiaansen and Den Toom. Then, these limit cases are analyzed focusing on the following subquestions:

1. How does multistability in Bastiaansen's model depend on its parameters? In particular, does the structure persist within a realistic parameter range?

2. How does the multistability in Den Toom's model depend on the number of vertical layers? Specifically, does the structure converge as the vertical resolution increases?

In Section 4.4.3, the two-dimensional overarching model is set up with parameters chosen to capture key SPG features and analyzed to investigate when multistability arises from physical mechanisms versus numerical artifacts, answering the following subquestion:

3. How does the bifurcation structure of the overarching model change under freshwater forcing?

Finally, the results are discussed in Chapter 5, and key conclusions are presented in Chapter 6.

2 Review of Convection Models

In this chapter, the existing models relevant to this thesis are reviewed. First of all the classical two-box convection model by Welander [14] is explained. This provides an understanding of how convective processes are modeled. Subsequently, Bastiaansen's extension of this model [12] is introduced and the physically meaningful multistability in the resulting model is discussed.

These box models simplify the system by focusing on integrated quantities and the relationships between them, thereby omitting the explicit representation of dynamics. A model that does capture these dynamics is the model developed by Weijer and Dijkstra [15], which is introduced in Section 2.3. This model incorporates dynamical behavior by solving the full momentum balance. A reduced version of the Weijer and Dijkstra model, proposed by Den Toom [13], is introduced next. The spurious multistability observed in this model is discussed.

Finally, the key differences between the models of Bastiaansen and Den Toom are identified, laying the foundation for their unification in the overarching model developed later in this thesis.

2.1 Modeling Convection

This section presents the classical two-box convection model developed by Welander [14], which serves to illustrate the basic principles involved in convection modeling.

Convection arises when denser water overlies less dense water, creating static instability. Then there is vertical mixing until static stability is restored. The density of water (ρ_*) depends on the temperature (T_*) and salinity (S_*) of the water. Their relation is described by the linearized equation of state:

$$\rho_* = \hat{\rho}(1 - \alpha_T(T_* - \hat{T}) + \alpha_S(S_* - \hat{S})), \tag{1}$$

where starred variables represent dimensional quantities. The parameters $\hat{\rho}$, $\hat{T}$ and $\hat{S}$ are reference quantities. The thermal expansion and haline contraction coefficients are denoted by α_T and α_S respectively, and convert the differences in temperature or salinity to differences in density. This equation shows that colder and more saline water has a higher density than fresher, less saline water.

Both temperature and salinity effects are relevant in the SPG. The cold atmosphere cools the surface water, increasing its density and potentially triggering convection. At the same time, input of precipitation and meltwater reduces surface salinity, thereby decreasing density and potentially suppressing convection.

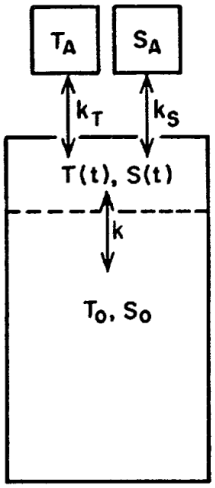


Figure 4: Schematic of Welander's two-box model for convection [14]. The surface box ($T(t)$, $S(t)$) exchanges heat and freshwater with the atmosphere (T_A, S_A), while convective exchange occurs between the surface and deep boxes (T_0, S_0) depending on density differences.

The box model of Welander is used to study this effect of temperature and salinity on the surface layer and its impact on convection. This model is illustrated in Figure 4 and is described by the following system of equations [14]:

$$\frac{\partial T}{\partial t} = k_T(T_A - T) - \kappa(\Delta\rho)(T - T_0), \tag{2a}$$

$$\frac{\partial S}{\partial t} = k_S(S_A - S) - \kappa(\Delta\rho)(S - S_0). \tag{2b}$$

Here, the temperatures and salinities are non-dimensionalized in line with the equation of state:

$$T = \alpha_T(T_* - \hat{T}), \quad S = \alpha_S(S_* - \hat{S}). \tag{3}$$

The model consists of a surface layer with a temperature $T(t)$ and salinity $S(t)$, which dynamically depend on the prescribed static atmosphere (T_A, S_A) and deep ocean layer (T_0, S_0). The temperatures and salinities are used to obtain the densities ρ and ρ_0 with Equation (1). The parameters k_T and k_S describe the rate of exchange of the temperature and salinity with the atmosphere box. The function $\kappa(\Delta\rho)$ represents the rate of exchange between the surface layer and the deep ocean due to convection.

The function $\kappa(\Delta\rho)$ should be large when there is static instability ($\rho > \rho_0$) and zero when there is static stability ($\rho < \rho_0$). Therefore it is modeled as an approximated step function depending on the density difference between the density of the surface box, ρ, and of the deep ocean box ρ_0: $\Delta\rho = \rho - \rho_0$, and reads:

$$\kappa(\Delta\rho) = \frac{\bar{\kappa}}{2}[1 + \tanh((\Delta\rho - \Delta\rho_{\text{ref}})/\bar{\epsilon})]. \tag{4}$$

Here $\bar{\kappa}$ is the maximum exchange rate during convection. For $\Delta\rho \ll \Delta\rho_{\text{ref}}$ the system is stably stratified and there is no convection. For $\Delta\rho \gg \Delta\rho_{\text{ref}}$ the exchange rate is $\kappa(\Delta\rho) = \bar{\kappa}$. Lastly, by reducing $\bar{\epsilon}$, Equation (4) approximates the step function, meaning that the region in $\Delta\rho$-space in which the system adjusts from no to maximum convection gets smaller. The effect of $\Delta\rho_{\text{ref}}$ and $\bar{\epsilon}$ on the step function is visualized in Figure 5.

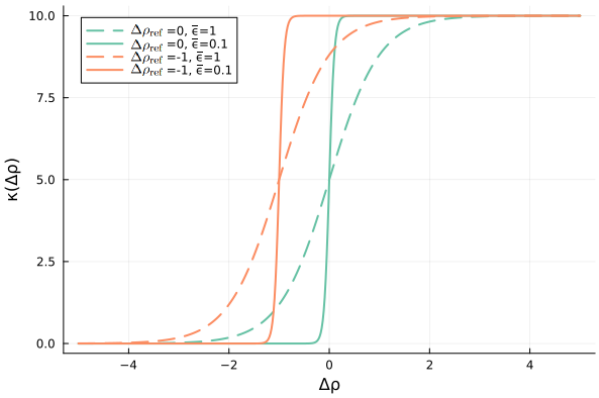


Figure 5: Rate of exchange $\kappa(\Delta\rho)$ expressing convection, varying convection threshold parameter $\Delta\rho_{\text{ref}}$ and steepness parameter $\bar{\epsilon}$.

2.2 Bastiaansen: Physically Meaningful Multistability

Bastiaansen [12] extends the box model by Welander spatially with diffusion in the horizontal direction described by x. The dimensionless equations describing this model are:

$$\frac{\partial T}{\partial t} = k_T(T_A(x) - T) - \kappa(\Delta\rho)(T - T_0) + D\frac{\partial^2 T}{\partial x^2}, \tag{5a}$$

$$\frac{\partial S}{\partial t} = k_S(S_A(x) - S) - \kappa(\Delta\rho)(S - S_0) + D\frac{\partial^2 S}{\partial x^2}, \tag{5b}$$

where D is the dimensionless diffusion coefficient defined by $D = \frac{\tau D_*}{L^2}$, with τ the characteristic timescale and L the horizontal length scale of the system. Bastiaansen sets $D = 0.01$. The exact scalings are not stated by Bastiaansen, but they should correspond to the following. The timescale is around $\tau = 75$ days, corresponding to the timescale of atmospheric temperature forcing [13]. The horizontal diffusion coefficient D_* of ocean systems is typically between 100 and 1000 m^2s^{-1}. This corresponds to an ocean basin of about 100-1000 km in size, aligning with the approximate width of the Labrador Sea, which is a key region for convection in the SPG, represented in the nondimensional domain $x \in [-1, 1]$. The prescribed atmospheric temperature, $T_A(x)$, and salinity, $S_A(x)$, are now a function of x. The other parameters, k_T, k_S and $\kappa(\Delta\rho)$ are the same as those in the Welander model. The parameter values used by Bastiaansen are given in Table 1.

$\bar{\kappa} = 10$
$\Delta\rho_{\text{ref}} = -1$
$k_T = 1$
$D = 0.01$
$\rho_0(x) = 0$
$x \in [-1, 1]$

Table 1: Nondimensional parameter values used by Bastiaansen [12]. While reproducing Figure 6, it was found that $\bar{\kappa} = 10$ and $\Delta\rho_{\text{ref}} = -1$ were actually used instead of the stated $\bar{\kappa} = 100$ and $\Delta\rho_{\text{ref}} = -0.5$.

Because the convection-driven exchange rate $\kappa(\Delta\rho)$ depends on $\Delta\rho$, these equations are rewritten to $\Delta\rho = \rho - \rho_0$. Since the variables have been non-dimensionalised using Equation (3), it follows from the equation of state that

$$\rho = \hat{\rho}^{-1}(\rho_* - \hat{\rho}) = S - T. \tag{6}$$

However, since both T and S evolve independently, an additional variable is introduced to capture their combined behavior. This variable is called the spiciness:

$$\mu = S + T, \tag{7}$$

with corresponding $\Delta\mu = \mu - \mu_0$, where μ_0 is the spiciness value of the deep ocean box.

To simplify the model, Bastiaansen assumes that $k_T = k_S$, meaning that the exchange rate of temperature and salinity with the atmosphere is the same. This is a strong and physically unrealistic assumption. In reality heat exchange and freshwater fluxes are distinct physical mechanisms and typically act on different timescales [16]. However, mathematically this assumption decouples the dynamics of $\Delta\rho$ from $\Delta\mu$. With the assumption $k_T = k_S$, and subtracting Equation (5a) from Equation (5b), it becomes clear that the equation for $\Delta\rho$ does not depend on $\Delta\mu$. This allows the system to be reduced to a single dimensionless equation:

$$\frac{\partial\Delta\rho}{\partial t} = D\frac{\partial^2\Delta\rho}{\partial x^2} + k_T(\Delta\rho_A(x) - \Delta\rho) - \kappa(\Delta\rho)\Delta\rho + D\frac{\partial^2\rho_0(x)}{\partial x^2}, \tag{8}$$

where $\Delta\rho_A(x) = \rho_A(x) - \rho_0$ and the atmospheric density is prescribed as $\rho_A(x) = S_A(x) - T_A(x)$.

At the boundaries in the horizontal direction, no flux (Neumann) boundary conditions are imposed, meaning there is no horizontal exchange of heat or salt across the edges of the domain. This assumption is common in idealized models and is reasonable for studying internal dynamics [17].

To further simplify the model, Bastiaansen sets the deep ocean density to $\rho_0(x) = 0$, implying a spatially uniform deep ocean. While in reality the temperature and salinity of the deep ocean do vary slightly in space and time, this simplification is justified due to the model's emphasis on surface processes over relatively short spatial and temporal scales.

The atmospheric density forcing is modeled as spatially varying by [12]:

$$\Delta\rho_A(x) = 2 + f[1 + \cos(\frac{\pi x}{2})], \tag{9}$$

where f is a bifurcation parameter, representing freshwater fluxes. Negative values of f simulate the addition of freshwater to the system, introducing a surface density minimum (freshwater maximum) at the center of the domain ($x = 0$). The horizontal direction corresponds to a zonal (east-west) transect. Such spatial freshwater forcing is physically plausible and observed in the subpolar North Atlantic, where freshwater accumulates due to precipitation and melting from sea ice and the Greenland Ice Sheet. Section 3.2 discusses this in more detail, supported by observational data.

In Figure 6, the steady state solutions are shown as a function of the parameter f, revealing how changes in freshwater forcing affect the surface layer density. In this bifurcation diagram the steady state solutions are represented by the mean of ρ over the

horizontal domain: $\langle\rho\rangle = \frac{\sum_{i=1}^{n_x} \rho_i}{n_x}$, where n_x is the number of grid cells in the horizontal direction. Solid lines indicate stable solutions, while dashed lines indicate unstable solutions. It can be seen that there is a possibility of coexisting states: there are multiple stable states for certain values of f. These states can be with full (red), partial (light blue) or absent (blue) deep convection. This multistability arises from the physical processes embedded in the equation, and is thus labeled physically meaningful.

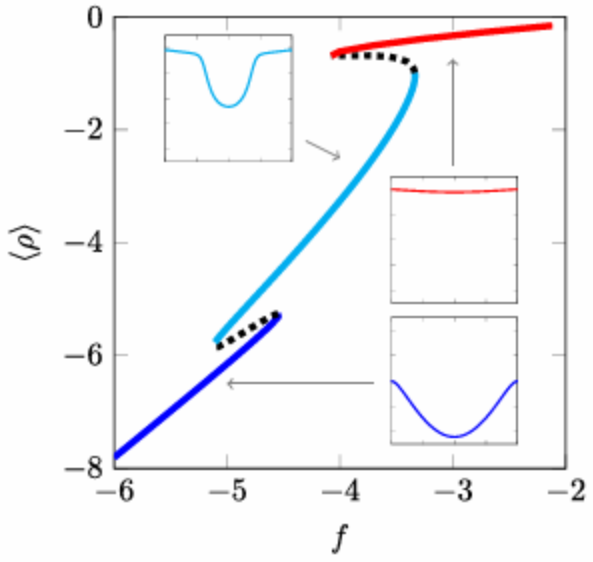


Figure 6: Bifurcation diagram from Bastiaansen [12]. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f. Solid lines indicate stable solutions, while dashed lines indicate unstable solutions. The solutions can be with full (red), partial (light blue) or absent (blue) deep convection. Plots of the equilibrium solutions are included in the boxes, where ρ is plotted over x.

2.3 Dynamical Extension of Convection Models

To move beyond the limitations of box models and capture the full dynamical behavior of ocean circulation, Weijer and Dijkstra [15] developed a two-dimensional (latitude-depth) model that explicitly conserves momentum, temperature, and salinity. This model serves as a more physically realistic foundation for studying density-driven circulation and forms the basis for Den Toom's reduced version discussed in the next section.

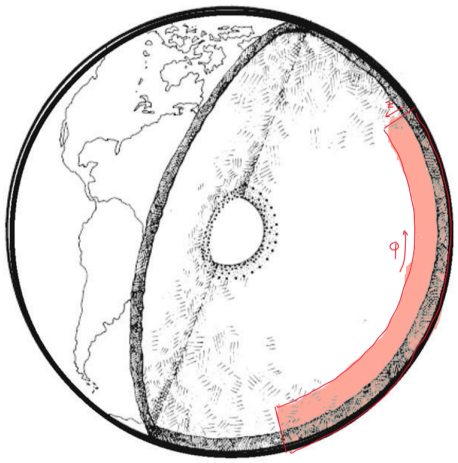


Figure 7: Modeled area by Weijer and Dijkstra [15] marked in red. The vertical axis represents depth on a domain of $z \in [-D, 0]$, while the horizontal axis represents latitude on a domain of $\phi \in [-\phi_N, \phi_N]$.

The model considers a purely buoyancy-driven flow on a non-rotating Earth in order to eliminate longitudinal dependence. The domain, illustrated in Figure 7, spans a vertical ocean slice with latitude and depth as its spatial dimensions: $([-\phi_N, \phi_N] \times [-\bar{D}, 0])$. At the surface, temperature and salinity are subjected to mixed boundary conditions: surface temperatures are relaxed towards a prescribed profile with a restoring timescale of $\tau = 75$ days, representing air-sea heat exchange, while salinity is forced by a net freshwater flux. Thus, unlike in Bastiaansen's model, ocean–atmosphere interactions influence temperature and salinity differently.

Mixing due to eddies is parametrized by anisotropic diffusion which means different diffusion rates apply in different directions. The horizontal and vertical eddy diffusivity are denoted by K_H and K_V, respectively. The key model parameters used in this study are summarized in Table 2.

$\bar{D} = 4.0 \times 10^3$ m	$\hat{\rho} = 1.0 \times 10^3$ kg m^{-3}
$\phi_N = 60°$	$\hat{T} = 15.0°$C
$K_H = 1.0 \times 10^3$ m^2 s^{-1}	$\hat{S} = 35.0$ psu
$K_V = 1.0 \times 10^{-4}$ m^2 s^{-1}	$\alpha_T = 1.0 \times 10^{-4}$ K^{-1}
$\tau = 75.0$ days	$\alpha_S = 7.6 \times 10^{-4}$ psu^{-1}

Table 2: Selected parameters of Weijer and Dijkstra's model [13]. The table lists parameters to describe the domain ($\bar{D}$, ϕ_N), eddy diffusivities (K_H, K_V), atmospheric relaxation timescale τ, reference values ($\hat{\rho}, \hat{T}, \hat{S}$) and thermal expansion and haline contraction coefficients (α_T, α_S).

Temperature and salinity in this model are transported by horizontal and vertical advection, horizontal and vertical diffusion and convective adjustment (CA). The CA term is introduced as an enhanced vertical diffusion process, applied when the water column becomes statically unstable. For a tracer C_* (temperature or salinity), the convective adjustment is parameterized as:

$$\mathrm{CA}(C_*) = K_V F_0 \frac{\partial}{\partial z_*} \left(\mathcal{F} \left(\frac{\bar{D}}{\rho_0} \frac{\partial \rho_*}{\partial z_*} \right) \frac{\partial C_*}{\partial z_*} \right), \tag{10}$$

where F_0 is the efficiency of convection and $\mathcal{F}(x)$ is a continuous approximation of the Heaviside step function:

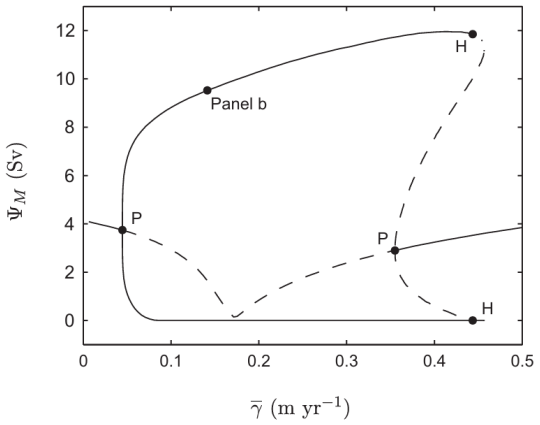
$$\mathcal{F}(x) = \max(0, \tanh[(\epsilon x)^3]). \tag{11}$$

This ensures that enhanced vertical mixing only occurs under unstable stratification. An alternative approximation of this step function, similar to Equation (4), reads:

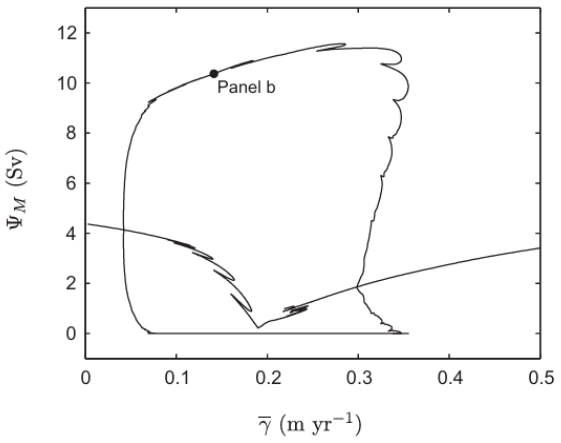
$$\mathcal{G}(x) = \frac{1}{2}(1 + \tanh[\epsilon x]). \tag{12}$$

The model is run with and without CA. The resulting bifurcation diagrams are shown in Figure 8. They show the maximum of the meridional overturning streamfunction Ψ_M over freshwater forcing strength $\bar{\gamma}$. In the solution without CA a structure with two stable states (solid lines) for the same freshwater parameter $\bar{\gamma}$ is visible. This occurs due to the salt-advection feedback [11], explained in Box 2. When CA is included, a large number of small folds appear on the branches, each corresponding to an additional saddle-node bifurcation.

Den Toom argues that this multistability is due to spurious events introduced by the CA parametrization. To show this, Den Toom reduced this model to a one-dimensional column model, isolating the convection process.



(a) Bifurcation diagram without CA. Solid lines represent stable states, dashed lines represent unstable states.



(b) Bifurcation diagram with CA. The stability of states is not indicated.

Figure 8: Comparison of bifurcation diagrams for the two-dimensional model of Weijer and Dijkstra [15] with and without CA. Freshwater forcing strength $\bar{\gamma}$ as control parameter. The plots show the maximum of the meridional overturning streamfunction Ψ_M in Sverdrups [13].

Box 2: Salt-Advection Feedback

The salt-advection feedback has first been investigated by Stommel [11]. The experimental set up shown in Figure 9 consists of two vessels, both relaxed towards a different salinity and temperature. Vessel 1 is the warm, saltier box, vessel 2 the cold, fresher box. These vessels are connected by two tubes; one at the top (overflow), one at the bottom (capillary). The rate of flow in these tube is represented with q. Stommel found that two stable states can occur:

- **Temperature Dominated State (Strong Overturning)** When the temperature differences dominate, cold water in vessel 2 is denser than in vessel 1. This induces a flow $(-q)$ in the capillary from vessel 2 to vessel 1, which also means there is a flow in the overflow from vessel 1 to vessel 2, which enhances the density in vessel 2 by the higher salinity, reinforcing the flow.
- **Salinity Dominated State (Reversed or Weak Overturning)** When the salinity differences dominate, the freshened water (due to for example ice melt) in vessel 2 is less dense than in vessel 1. Then the flow (q) reverses.

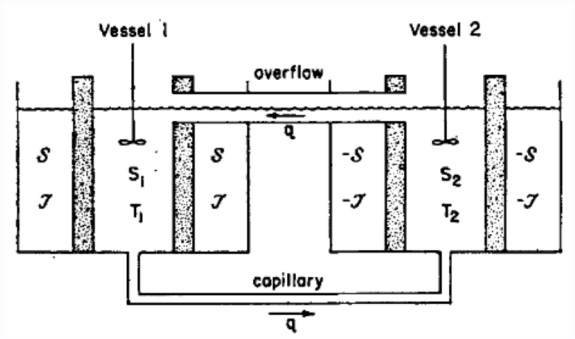


Figure 9: Model of the salt-advection feedback by Stommel [11]. Two vessels are connected by a surface overflow and a bottom capillary, with flows indicated by q. The vessels are well stirred and connected to an outside vessel each, where the temperature and salinity are maintained at constant values.

2.4 Den Toom: Spurious Multistability

To better understand the role of convective adjustment in generating multiple equilibria, Den Toom [13] reduces the two-dimensional model of Weijer and Dijkstra [15] to a one-dimensional vertical column model. The resulting model equations read:

$$\frac{\partial T_*}{\partial t_*} = K_V \frac{\partial}{\partial z_*}\left([1 + F_0 \mathcal{K}_{T_*}]\frac{\partial T_*}{\partial z_*}\right) - \kappa(i_T^{\text{res}} T_* - \tilde{T}_*(z_*)), \tag{13a}$$

$$\frac{\partial S_*}{\partial t_*} = K_V \frac{\partial}{\partial z_*}\left([1 + F_0 \mathcal{K}_{S_*}]\frac{\partial S_*}{\partial z_*}\right) - \kappa(i_S^{\text{res}} S_* - \tilde{S}_*(z_*)). \tag{13b}$$

The first term on the right-hand side represents vertical diffusion. This consists of background mixing due to eddies and the enhanced diffusion representing convective adjustment similar to equation 10. Here, $\mathcal{K}_{T_*}$ and $\mathcal{K}_{S_*}$ are the convective adjustment functions,

which are typically functions $\mathcal{F}(\frac{\partial \rho}{\partial z})$ or $\mathcal{G}(\frac{\partial \rho}{\partial z})$ as introduced earlier in Equation (11) and (12).

The second term represents the forcing of the system. This forcing happens over the entire length of the column and can be seen as the interaction between the SPG and the boundary current [18]. The properties of this boundary current are described by $\tilde{T}_*(z_*)$ and $\tilde{S}_*(z_*)$. The type of forcing can be controlled with i^{res}: if $i^{\text{res}} = 1$ the tracer is relaxed to a prescribed profile at rate κ, if $i^{\text{res}} = 0$ the term represents a fixed flux.

The model domain extends vertically from $z_* = -\bar{D}$ (bottom) to $z_* = 0$ (surface), with no-flux boundary conditions at both ends. A sketch of the model is given in Figure 10.

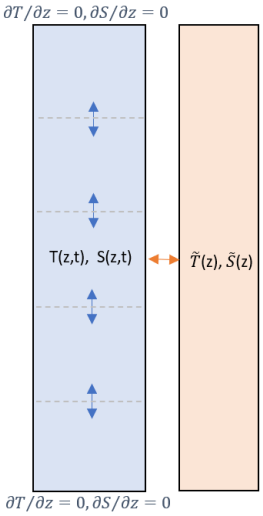


Figure 10: Schematic of Den Toom's one-dimensional vertical column model. With vertical mixing in the column due to convection and forcing from the side with a constant temperature and salinity profile $(\tilde{T}(z), \tilde{S}(z))$.

Dimensional parameters
$K_V = 1.0 \times 10^{-4}$ m^2 s^{-1}
$z_* \in [-\bar{D}, 0]$
Scaling
$\kappa^{-1} = 5$ years
$\bar{D} = 4.0 \times 10^3$ m
Nondimensional parameters
$P = 10^3$
$z \in [-1, 0]$
Modeling parameters
$i_T^{\text{res}} = 1$
$i_S^{\text{res}} = 0$
$F_0 = 100$
$\mathcal{K}_{T_*} = \mathcal{K}_{S_*} = \mathcal{K}_\rho = \mathcal{K}_\mu = \mathcal{F}(\frac{\partial \rho}{\partial z})$
$\epsilon = 10$

Table 3: Parameter values used by Den Toom [13]. The table lists the dimensional and nondimensional parameters for vertical diffusion and the domain, the used scalings, and modeling parameters for the type of forcing and convection.

This system can be nondimensionalised using Equation (3) to nondimensionalize T and S and rescaling the vertical coordinate with the column depth $\bar{D}$, $z = \bar{D}^{-1} z_*$, and of time t_* with κ, $t = \kappa t_*$. Here $\kappa^{-1} = 5$ years represents the typical timescale of interaction with the side forcing. Then, as the convective adjustment terms depend on the density, the equations are reformulated in terms of density ρ and spiciness μ using Equation (6) and (7). This leads to the following formulation of the system:

$$\frac{\partial \rho}{\partial t} = \frac{1}{P} \frac{\partial}{\partial z} \left([1 + F_0 \mathcal{K}_\rho] \frac{\partial \rho}{\partial z} \right) - \left(i_+^{\text{res}} \rho + i_-^{\text{res}} \mu - \tilde{\rho}(z) \right), \tag{14a}$$

$$\frac{\partial \mu}{\partial t} = \frac{1}{P} \frac{\partial}{\partial z} \left([1 + F_0 \mathcal{K}_\mu] \frac{\partial \mu}{\partial z} \right) - \left(i_-^{\text{res}} \rho + i_+^{\text{res}} \mu - \tilde{\mu}(z) \right), \tag{14b}$$

where

$$i_+^{\text{res}} = \frac{i_S^{\text{res}} + i_T^{\text{res}}}{2}, \quad i_-^{\text{res}} = \frac{i_S^{\text{res}} - i_T^{\text{res}}}{2}, \quad \tilde{\rho}(z) = \tilde{S}(z) - \tilde{T}(z), \quad \tilde{\mu}(z) = \tilde{S}(z) + \tilde{T}(z). \quad (15)$$

Here, $\mathcal{K}_\rho$ and $\mathcal{K}_\mu$ are convective adjustment functions similar to $\mathcal{K}_{T_*}$ and $\mathcal{K}_{S_*}$. Furthermore, $P = \frac{\bar{D}^2\kappa}{K_V}$ can be interpreted as the vertical Péclet number, representing the ratio of the timescale for vertical diffusion across the full column to the timescale of relaxation towards the prescribed side forcing. Physically, it indicates whether the column is efficiently mixed ($P \ll 1$) or whether forcing dominates and stratification can persist ($P \gg 1$). The forcings are defined as:

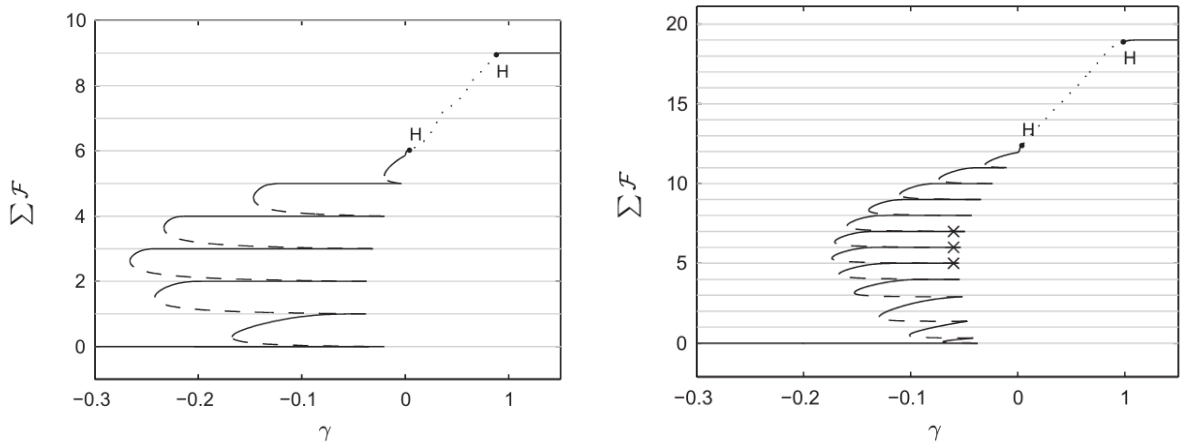
$$\tilde{T}(z) = \cos(2\pi z), \quad (16a)$$

$$\tilde{S}(z) = \gamma\cos(\pi z). \quad (16b)$$

Here, the parameter γ sets the amplitude of the freshwater flux. The shape of the salinity profile $\tilde{S}(z)$ is consistent with a typical freshwater forcing scenario, where salinity is reduced at the surface due to precipitation and ice melt.

In contrast, the temperature profile $\tilde{T}(z)$ is less physically intuitive. This corresponds to a warm-cold-warm structure over the vertical column. In reality, the boundary current may exhibit a temperature profile more like cold-warm-cold, i.e. $-\tilde{T}(z)$, where the surface is cooled by atmospheric interaction, the mid-depths are warmed by northward heat transport by the AMOC, and the deep ocean is again colder [19]. Relaxing the system towards such a profile introduces static instability and allows the model to capture convective processes. The difference between the two forcings will be discussed in Section 4.3.

Den Toom studies this model with parameters chosen to approximate realistic oceanic conditions, as listed in Table 3. The system is solved by discretizing the equations using finite differences, shown in more detail in Section 3.3. The number of vertical layers in this discretization is denoted by n_z. In Figure 11, the resulting bifurcation diagrams are shown for discretizations with $n_z = 10$ and $n_z = 20$ vertical layers. These are bifurcation diagrams with as bifurcation parameter the fresh water inflow strength γ. On the vertical axis the sum of the convective adjustment function over the interfaces is plotted: $\sum_{\text{vertical interfaces}} \mathcal{F}(\frac{\partial\rho}{\partial z})$, indicating the intensity of convection. It can be seen that the number of both equilibrium solution (stable and unstable states) depends on n_z. This resolution dependency indicates that the observed multistability is not physical, but rather a numerical artifact. The multistability is therefore labeled spurious.



(a) $n_z = 10$.

(b) $n_z = 20$.

Figure 11: Bifurcation diagram for the one dimensional model of Den Toom, using the parameters noted in Table 3 and $n_z = 10$ or $n_z = 20$ [13]. On the vertical axis $\sum \mathcal{G}$ indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ.

2.5 Comparing Bastiaansen and Den Toom

In the previous sections, the models of Bastiaansen and Den Toom have been introduced. These models form the basis for answering subquestions 1 and 2 regarding their respective multistability.

In this section, the differences in model choises and assumptions between the models of Bastiaansen and Den Toom are listed. Also, the convection terms of both models are compared. This way the parameters used in Bastiaansen are linked to those used by Den Toom. This section creates an overview of the theoretical background which will form the foundation of creating an overarching model, described in Chapter 3. This overarching model provides the foundation for answering subquestion 3 on multistability.

The differences in choises and model assumptions between the models of Bastiaansen and Den Toom are depicted in Figure 12. The main differences are:

- **Horizontal direction** Bastiaansen's model includes dynamics in the horizontal direction whereas Den Toom's model does not include a horizontal direction.

- **Vertical direction** The model by Bastiaansen is a box model, modeling the ocean as two boxes: a variable surface ocean box and a static deep ocean box. Therefore this equals $n_z = 2$ when using a vertical discretization. In contrast, the model by Den Toom focusses on dynamics in the vertical direction.

- **Bottom boundary condition** In Bastiaansen's model the deep ocean box represents a Dirichlet boundary condition: $\rho_0 = 0$. This assumes that the deep ocean has a static density. In Den Toom's model however it is assumed that at the bottom boundary there is no flux: $\frac{\partial \rho_0}{\partial z} = 0$. This reflects the assumption that the ocean bottom acts as a closed boundary with no exchange across it.

- **Forcing** In Bastiaansen's model the forcing is an atmospheric forcing only. The density of the ocean is relaxed towards an atmospheric density (temperature and salinity). In Den Toom's model the dynamcis are only forced from the side, this

can be visualized as the boundary current passing by the SPG. The temperature is relaxed towards a background profile, whereas salinity is forced by a constant flux.

- **Coupling between ρ and μ** The forcing in the model by Bastiaansen is assumed to be identical for temperature and salinity. Therefore the equation of ρ does not depend on μ, and the system is thus decoupled. In contrast, since in Den Toom there is a difference in forcing between temperature and salinity, this is a coupled system.

- **Convective mixing threshold $\Delta\rho_{\mathbf{ref}}$** This parameter determines the onset of vertical mixing. Bastiaansen uses a negative value, allowing mixing even when the upper layer is still lighter. Den Toom sets $\Delta\rho_{\text{ref}} = 0$, triggering mixing as soon as the upper layer is denser.

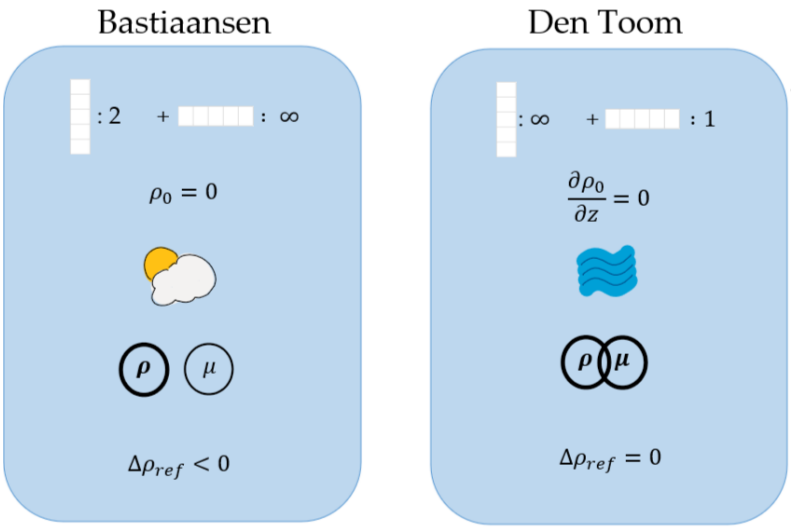


Figure 12: Overview of the key modeling elements in Bastiaansen's and Den Toom's models. Differences include dimensionality, bottom boundary conditions, types of forcing, coupling of ρ and μ, and the convective mixing threshold $\Delta\rho_{\text{ref}}$.

2.5.1 Comparison of Convection Parameterizations

The convection term is an important modeling element in both models. In this section the discretized version of the convection term used in the two models will be compared.

Den Toom The convection term in the one-dimensional column model of Den Toom, extracted from Equation (14), is

$$\frac{1}{P}\frac{\partial}{\partial z}\left(\left[1 + F_0\mathcal{K}_\rho\right]\frac{\partial \rho}{\partial z}\right), \tag{17}$$

where $\mathcal{K}_\rho = \mathcal{K}_\mu = \mathcal{G}(\frac{\partial \rho}{\partial z}) = \frac{1}{2}(1 + \tanh(\epsilon\frac{\partial \rho}{\partial z}))$. The derivatives in this equation are discretized using finite differences (with half steps):

$$\frac{\partial f_i}{\partial z} = \frac{f_{i+1/2} - f_{i-1/2}}{\Delta z}, \tag{18}$$

how these points are defined is visualized in Figure 13. Furthermore, $\Delta\rho_i = \rho_i - \rho_{i-1}$ is substituted, this yields:

$$\frac{1}{P}\frac{\left(\left[1+\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_{i+1}}{\Delta z}))\right]\frac{\Delta\rho_{i+1}}{\Delta z}\right)-\left(\left[1+\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_i}{\Delta z}))\right]\frac{\Delta\rho_i}{\Delta z}\right)}{\Delta z}. \tag{19}$$

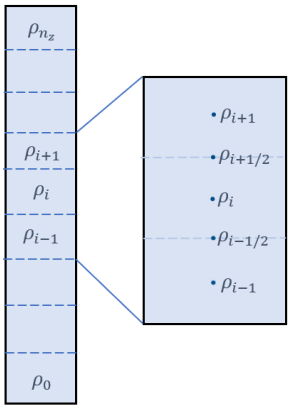


Figure 13: Visualization of the discretized column, in particular showing ρ_{i+1}, $\rho_{i+1/2}$, ρ_i, $\rho_{i-1/2}$ and ρ_{i-1}.

Bastiaansen In the extended version of Bastiaansen's model with multiple boxes in the vertical direction, the convection term becomes:

$$\kappa(\Delta\rho_{i+1})\Delta\rho_{i+1} - \kappa(\Delta\rho_i)\Delta\rho_i. \tag{20}$$

Here $\kappa(\Delta\rho) = \frac{\bar{\kappa}}{2}[1+\tanh((\Delta\rho-\Delta\rho_{\text{ref}})/\bar{\epsilon})]$ (Equation (4)) can be substituted. This results in:

$$\frac{\bar{\kappa}}{2}[1+\tanh((\Delta\rho_{i+1}-\Delta\rho_{\text{ref}})/\bar{\epsilon})]\Delta\rho_{i+1} - \frac{\bar{\kappa}}{2}[1+\tanh((\Delta\rho_i-\Delta\rho_{\text{ref}})/\bar{\epsilon})]\Delta\rho_i. \tag{21}$$

This equation can be rewritten to resemble Equation (19) by setting the parameters according to $\Delta\rho_{\text{ref}} = 0$, $\bar{\kappa} = \frac{F_0}{P\Delta z^2}$ and $\bar{\epsilon} = \frac{\Delta z}{\epsilon}$, which yields:

$$\frac{1}{P}\frac{\left(\left[\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_{i+1}}{\Delta z}))\right]\frac{\Delta\rho_{i+1}}{\Delta z}\right)-\left(\left[\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_i}{\Delta z}))\right]\frac{\Delta\rho_i}{\Delta z}\right)}{\Delta z}. \tag{22}$$

Now the similarity between the two convection terms in Equation (19) and (22) becomes clear. The only difference is that Den Toom's formulation includes an additional diffusion term:

$$\underbrace{\frac{1}{P}\frac{\left(\left[\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_{i+1}}{\Delta z}))\right]\frac{\Delta\rho_{i+1}}{\Delta z}\right)-\left(\left[\frac{F_0}{2}(1+\tanh(\epsilon\frac{\Delta\rho_i}{\Delta z}))\right]\frac{\Delta\rho_i}{\Delta z}\right)}{\Delta z}}_{\text{Bastiaansen's convective term (Equation (22))}} + \underbrace{\frac{1}{P}\frac{\frac{\Delta\rho_{i+1}}{\Delta z}-\frac{\Delta\rho_i}{\Delta z}}{\Delta z}}_{\text{extra diffusion term}}. \tag{23}$$

With the comparison of the convection terms, an expression for Bastiaansen's parameters $\bar{\kappa}$ and $\bar{\epsilon}$ can be determined in terms of Den Toom's parameters. These are:

$$\bar{\kappa} = \frac{F_0}{P}, \quad \bar{\epsilon} = \frac{1}{\epsilon}. \tag{24}$$

Substituting these into Bastiaansen's convection term (Equation (22)) recovers Den Toom's convection term (Equation (19)). The factors of Δz needed are due to making the box model continuous in the vertical direction.

In summary, this chapter has introduced the models of Bastiaansen and Den Toom, outlined their key differences in assumptions and design, and connected the underlying components of convection modeling. These findings provide the basis for the next chapter, where an overarching model is proposed.

3 Proposed Overarching Model

In the previous chapter, the key components of the models of Bastiaansen and Den Toom were outlined and compared. This chapter introduces an overarching model that integrates these components into a unified framework. Since not all features from both models are compatible, the overarching model is designed to be modular: specific elements can be turned on or off to recover either of the original models or explore intermediate configurations.

The chapter begins by presenting the structure and assumptions of the overarching model. Then, the parameter values that are used are discussed. Next, the implementation is explained, including numerical discretization and treatment of boundary conditions. Finally, the analysis method is described: bifurcation diagrams are used to investigate multistability.

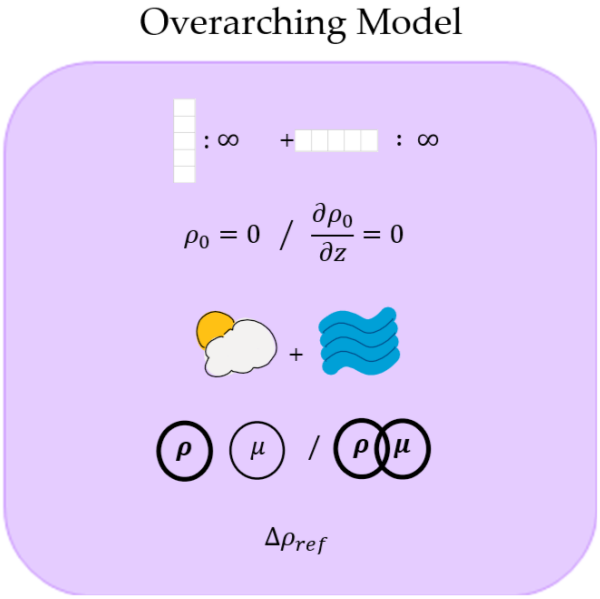


Figure 14: Overview of the key modeling elements in the overarching model. This includes dimensionality, bottom boundary conditions, types of forcing, coupling of ρ and μ, and the convective mixing threshold $\Delta\rho_{\text{ref}}$.

The components incorporated into the overarching model are summarized in Figure 14 and include:

- **Two-dimensional domain:** The model domain covers both the horizontal (x) and vertical (z) directions. This generalizes Den Toom's one-dimensional vertical column model and Bastiaansen's horizontally structured box model.
- **Bottom boundary condition, Dirichlet or Neumann:** Two types of boundary conditions can be applied at the bottom: a Dirichlet condition ($\rho_0 = 0$) sets the bottom density to a fixed value, while a Neumann (zero flux) condition ($\frac{\partial \rho_0}{\partial z} = 0$) requires the vertical density flux to vanish. These reflect the choices made in Bastiaansen and Den Toom, respectively, and affect the numerical implementation.
- **Atmospheric and/or side forcing:** Forcing can be applied at the surface and/or

from the side. This allows replication of the Bastiaansen model (atmospheric forcing) or Den Toom's setup (side forcing), or combinations thereof.

- **Coupled or uncoupled dynamics:** The interaction between temperature and salinity is determined by how forcing is applied. If both are forced identically (as in Bastiaansen's model), the system is uncoupled. If salinity and temperature receive distinct forcings (as in Den Toom's model), the system becomes coupled. This is controlled via the relaxation parameters $i_{\text{res},T}$ and $i_{\text{res},S}$.
- **Convective mixing threshold $\Delta\rho_{\text{ref}}$:** The convection term includes a reference density difference $\Delta\rho_{\text{ref}}$ that sets the stratification threshold for vertical mixing. Setting $\Delta\rho_{\text{ref}} = 0$ (Den Toom) triggers mixing only when the upper layer is denser, while a negative value (Bastiaansen) allows convection to be already triggered if the system is still stably stratified.

3.1 Mathematical Formulation

The mathematical formulation of the overarching model that describes the evolution of dimensionless density $\rho(z, x, t)$ and spiciness $\mu(z, x, t)$ reads:

$$\frac{\partial \rho}{\partial t} = \frac{1}{P}\frac{\partial}{\partial z}\Big([1 + F_0\mathcal{G}]\,\frac{\partial \rho}{\partial z}\Big) - \big(i_+^{\text{res}}\rho + i_-^{\text{res}}\mu - \tilde{\rho}(z)\big) + D\frac{\partial^2 \rho}{\partial x^2}, \tag{25a}$$

$$\frac{\partial \mu}{\partial t} = \frac{1}{P}\frac{\partial}{\partial z}\Big([1 + F_0\mathcal{G}]\,\frac{\partial \mu}{\partial z}\Big) - \big(i_-^{\text{res}}\rho + i_+^{\text{res}}\mu - \tilde{\mu}(z)\big) + D\frac{\partial^2 \mu}{\partial x^2}, \tag{25b}$$

where

$$i_+^{\text{res}} = \frac{i_S^{\text{res}} + i_T^{\text{res}}}{2}, \quad i_-^{\text{res}} = \frac{i_S^{\text{res}} - i_T^{\text{res}}}{2}, \quad \tilde{\rho}(z) = \tilde{S}(z) - \tilde{T}(z), \quad \tilde{\mu}(z) = \tilde{S}(z) + \tilde{T}(z). \tag{26}$$

The convective adjustment function is defined by:

$$\mathcal{G}(x) = \frac{1}{2}(1 + \tanh[\epsilon(x - \Delta\rho_{\text{ref}})]). \tag{27}$$

The atmospheric forcing is embedded in the boundary condition at the top. At the bottom, the model allows for either a Dirichlet boundary condition ($\rho_0 = 0$) or a zero-flux condition. No-flux boundary conditions are imposed in the horizontal direction. The resulting boundary conditions of $\rho(z, x, t)$ and $\mu(z, x, t)$ read:

$$\left[\frac{1}{P}\frac{\partial}{\partial z}\left(\left[1 + F_0\mathcal{G}(\frac{\partial \rho(0,x,t)}{\partial z})\right]\frac{\partial \rho(0,x,t)}{\partial z}\right)\right]_+ = k_S(S_A(x) - i_{A,S}^{\text{res}}S(0,x,t)) - k_T(T_A(x) - i_{A,T}^{\text{res}}T(0,x,t)), \tag{28a}$$

$$\left[\frac{1}{P}\frac{\partial}{\partial z}\left(\left[1 + F_0\mathcal{G}(\frac{\partial \rho(0,x,t)}{\partial z})\right]\frac{\partial \mu(0,x,t)}{\partial z}\right)\right]_+ = k_S(S_A(x) - i_{A,S}^{\text{res}}S(0,x,t)) + k_T(T_A(x) - i_{A,T}^{\text{res}}T(0,x,t)), \tag{28b}$$

$$\rho(-1,x,t) = \mu(-1,x,t) = 0 \quad \text{or} \quad \frac{\partial \rho(-1,x,t)}{\partial z} = \frac{\partial \rho(-1,x,t)}{\partial z} = 0, \tag{28c}$$

$$\frac{\partial \rho(z,-1,t)}{\partial x} = \frac{\partial \mu(z,-1,t)}{\partial x} = \frac{\partial \rho(z,1,t)}{\partial x} = \frac{\partial \mu(z,1,t)}{\partial x} = 0. \tag{28d}$$

Here $i^{\text{res}}_{A,S}$ and $i^{\text{res}}_{A,T}$ indicate if S and T are relaxed towards a background profile ($i^{\text{res}} = 1$) or there is a constant flux ($i^{\text{res}} = 0$). The atmospheric temperature and salinity are denoted by $T_A(x)$ and $S_A(x)$, their values will be discussed in the next section. Furthermore, to determine $T(0, x, t)$ and $S(0, x, t)$ the following transformations are applied:

$$T(0,x,t) = \frac{\mu(0,x,t) - \rho(0,x,t)}{2}, \quad S(0,x,t) = \frac{\mu(0,x,t) + \rho(0,x,t)}{2}. \tag{29}$$

3.2 Parameters

Since the parameters by Den Toom were chosen to represent a realistic ocean setting [13], these parameter values are also used as default values in the overarching model. This concerns the extend of the domain in the vertical direction, and the parameters F_0, P, ϵ, $\Delta\rho_{\text{ref}}$, i^{res}_T and i^{res}_S, see Table 4.

Dimensional parameters	
$z_* \in [-\bar{D}, 0]$	$x_* \in [-L, L]$
$K_V = 1.0 \times 10^{-4}$ m^2 s^{-1}	$K_H = 1.0 \times 10^{3}$ m^2 s^{-1}
$k_{T*} = 1/75$ days^{-1}	$\Delta\rho_{\text{ref}*} = 0$ kg m^{-3}
Scaling	
$\bar{D} = 4.0 \times 10^{3}$ m	$L = 450 \times 10^{3}$ m
$\bar{\tau} = 5$ years	$\hat{\rho} = 1.0 \times 10^{3}$ kg m^{-3}
Nondimensional parameters	
$z \in [-1, 0]$	$x \in [-1, 1]$
$P = 10^3$	$D = 0.78$
$k_T = 24.3$	$\Delta\rho_{\text{ref}} = 0$
Modeling parameters	
$i^{\text{res}}_T = 1$	$i^{\text{res}}_S = 0$
$i^{\text{res}}_{A,T} = 1$	$i^{\text{res}}_{A,S} = 0$
$F_0 = 10^2$	$\epsilon = 10$

Table 4: Parameter values used in the overarching model. The table lists dimensional and nondimensional parameters for the horizontal and vertical domain sizes, diffusion in the horizontal and vertical direction and atmospheric relaxation, the used scalings, and the modeling parameters for the type of forcing and convection.

The side forcing is adapted from Den Toom, with the temperature profile inverted to produce a cold-warm-cold structure as discussed in Section 2.4:

$$\tilde{T}(z) = -\cos(2\pi z), \tag{30a}$$

$$\tilde{S}(z) = \gamma\cos(\pi z), \tag{30b}$$

where the parameter γ sets the amplitude of the freshwater flux. However, it is important to note that translating this forcing back into dimensional variables using Equation (3) results in unphysical temperature ranges on the order of -10^4 to 10^4 °C. This indicates that Den Toom applied a rescaling, although this was not explicitly stated. A more realistic temperature variation would lie between approximately 10 and 20 °C, implying a rescaling factor of about 2000. Because the model equations are linear aside from the convection adjustment function, which resembles a step function, this rescaling can be absorbed into the nonlinearity parameter ϵ without affecting the model dynamics. A detailed derivation is provided in Appendix A.1.

The forcing from the atmosphere is determined using data from the Climate Explorer by the KNMI [20]. The average precipitation and atmospheric temperature over the SPG are shown in Figure 15 and 16. To estimate the forcing, the data within the green-highlighted region (55W to 40W at 57N) in Figure 15a is used. The length of this region is approximately 900 km. The region lies within the Labrador Sea and the data here shows a clear pattern. A cross-section in the zonal direction is selected, since the boundary current profile varies along the meridional direction [15], making the zonal direction more suitable for assuming a roughly constant side forcing. Approximating and rescaling this data the values for k_S, $S_A(x)$ and $T_A(x)$ will be estimated.

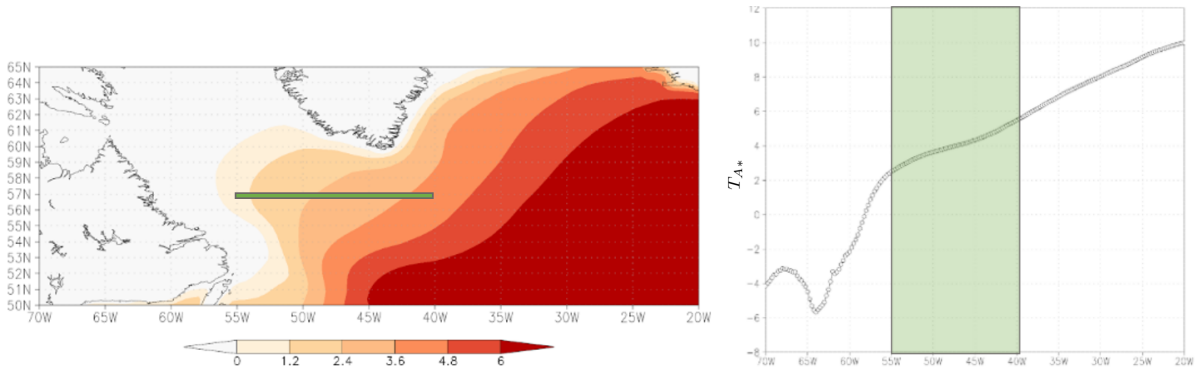
First of all, k_T is determined by rescaling the atmospheric relaxation rate k_{T*} with the timescale $\bar{\tau}$ of 5 years. Then, k_T is computed as:

$$k_T = \bar{\tau} k_{T*} \approx 24.3. \tag{31}$$

Here, the atmospheric relaxation rate $k_{T*} = 1/75$ day^{-1} is used, which follows from the timescale for atmospheric relaxation given in Table 2.

Figure 15b shows the atmospheric temperature distribution of the area marked in green, which can be approximated by a linear function ranging from 2.5 to 6.0 °C. The values are nondimensionalised using Equation (3) and rescaled by a factor of 2000, consistent with Den Toom's approach. The horizontal domain is chosen similar to Bastiaansen as $[-1, 1]$, following from the nondimensionalisation $x = \frac{x_*}{L}$ with $L = 450$ km as the typical length scale. The resulting atmospheric temperature is then given by:

$$T_A(x) = -2.15 + 0.35x. \tag{32}$$



(a) Color map of the atmospheric temperature (°C) for the entire area of the SPG.

(b) Cross-section of the atmospheric temperature (°C) at 57N.

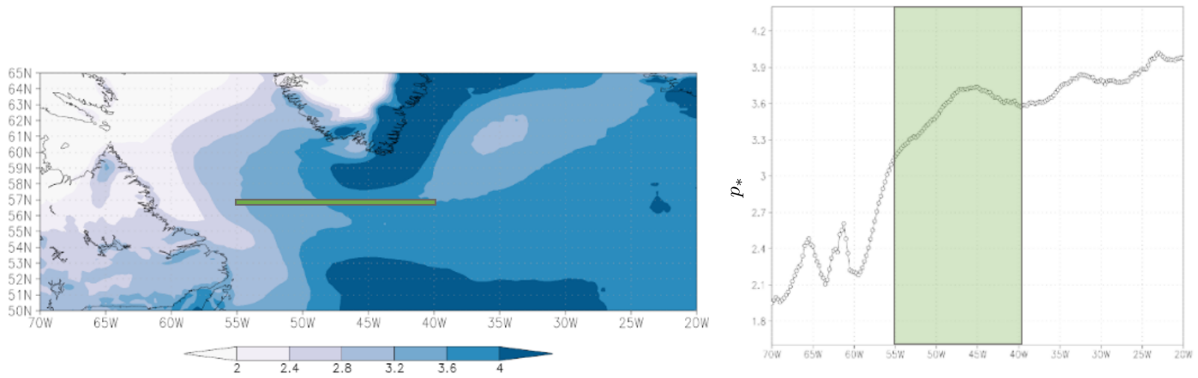
Figure 15: Atmospheric temperature in °C, averaged over 30 years (1990-2020) [20]. Based on annual mean temperature data from ERA5 [21]. The data used to approximate the atmospheric forcing is marked green (at 57N from 55W to 40W).

The precipitation distribution is shown in Figure 16b and can be approximated by $p_* = 3.4 + 0.4\cos(\frac{\pi x}{2} - 0.4)$ mm/day. This first has to be transformed in a salinity flux, using the following formula:

$$\Delta S_* = -\frac{\hat{S} p_*}{H}, \tag{33}$$

where the mixing depth $H = 50$ m and reference salinity $\hat{S} = 35$ psu are used. This is an estimate; however, since the salinity source term is multiplied by the bifurcation parameter f, this does not affect the qualitative behavior of the system. The values are nondimensionalised using Equation (3), are rescaled to a timescale of 5 years and by a factor of 2000, consistent with Den Toom's approach. Together the input for the atmospheric salinity flux is:

$$k_S S_A(x) = f(-6.6 - 0.8\cos(\frac{\pi x}{2} - 0.4)). \tag{34}$$



(a) Color map of the precipitation (mm/day) for the entire area of the SPG.

(b) Cross-section of the precipitation (mm/day) at 57N.

Figure 16: Precipitation in mm/day, averaged over 30 years (1990-2020) [20]. Based on annual mean precipitation data from ERA5 [21]. The data used to approximate the atmospheric forcing is marked green (at 57N from 55W to 40W).

Lastly, D is the horizontal diffusion parameter, corresponding to K_H which is the horizontal diffusion parameter due to eddies as noted in Table 3. This has to be rescaled using the typical scales for the length $L = 450$ km and times of $\bar{\tau} = 5$ years. With these choices, $D = \frac{K_H \bar{\tau}}{L^2} \approx 0.78$.

3.3 Implementation

The variables in the overarching model are $\rho(z, x, t)$ and $\mu(z, x, t)$. These are modeled as vectors of lexicographic order. The system of equations given in Equation (25) has to be discretized for implementation. For the first and second order derivative, finite differences are used:

$$\frac{\partial f_i}{\partial x} = \frac{f_{i+1} - f_{i-1}}{2\Delta x} \quad \text{and} \quad \frac{\partial^2 f_i}{\partial x^2} = \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2}. \tag{35}$$

First of all, the diffusion in the vertical and horizontal direction can be modeled as $A_{\text{linear}}\rho$ and $A_{\text{linear}}\mu$. A_{linear} can be formed as

$$A_{\text{linear}} = I_x \otimes A_z + A_x \otimes I_z. \tag{36}$$

Here I_x and I_z are identity matrix of the size of the horizontal and vertical grid respectively. The matrix A_x represents the diffusion in the horizontal direction. The matrix A_z represents the background diffusion in the vertical direction and can contain the atmospheric and bottom box boundary condition. Since the background diffusion is not present in Bastiaansen's model it can be turned on and off with setting `background_diffusion` to true or false.

If both boundary conditions in a dimension are no flux boundary conditions ($\frac{\partial f_i}{\partial x} = 0$), then from the first order derivatives it follows that $f_{i-1} = f_{i+1}$. Then

$$A_x = \frac{D}{\Delta x^2}\tilde{A}, \quad A_z = \frac{1}{P\Delta z^2}\tilde{A}, \tag{37}$$

with

$$\tilde{A} = \begin{bmatrix} -2 & 2 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 2 & -2 \end{bmatrix}. \tag{38}$$

The convection term is nonlinear and implemented directly using the first order derivative as in Equation (35). The appropriate boundary conditions are incorporated into this term.

The implementation of the other terms is straightforward. The numerical implementation differs slightly depending on the specific model choice within the overarching model, as discussed below.

Bottom boundary condition One of the model options is a Dirichlet boundary condition at the bottom instead of a zero flux boundary condition (`dirichlet_BC` = true). The values of the bottom box are then set to $\rho_0 = 0$ and $\mu_0 = 0$. This changes the implementation in the following way:

- First of all, the vectors describing ρ and μ reduce to vectors of size $(n_z - 1)n_x$. As the bottom value is a constant, the corresponding equation can be removed.
- For the discretization of $\frac{\partial^2 \rho}{\partial z^2}$ the discretization of ρ_1 reduces to $\frac{\partial^2 \rho_1}{\partial z^2} \approx \frac{1}{(\Delta z)^2}(\rho_0 - 2\rho_1 + \rho_2) = \frac{1}{(\Delta z)^2}(-2\rho_1 + \rho_2)$. Therefore the new discretization matrix is

$$A_z = \frac{1}{P\Delta z^2}\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 2 & -2 \end{bmatrix} \tag{39}$$

  of size $(n_z - 1, n_z - 1)$.
- Lastly, in the convection term corresponding to ρ_1, ρ_0 has to be substituted.

Coupling between ρ and μ Another choice that can be made is the coupling which depends on the relaxation parameters i_{res}. When all relaxation parameters are set to one:

$$i_T^{\text{res}} = i_S^{\text{res}} = i_{A,T}^{\text{res}} = i_{A,S}^{\text{res}} = 1, \tag{40}$$

both ρ and μ are relaxed to predefined profiles. Then the equation of ρ is independent of μ, thus the system is uncoupled.

Forcing The atmospheric forcing is directly implemented as the boundary condition in the convection term as described in Equations (28a) and (28b). The atmospheric forcing and side forcing can be turned on by setting `atmos_BC` and `side_forcing` to true respectively.

Lastly, it is important to note that for specific settings of the overarching model the solution for salinity is only determined up to a constant. This occurs when all boundary conditions for the diffusion in the domain are zero-flux and the salinity forcing consists solely of fluxes ($i_S^{\text{res}} = i_{A,S}^{\text{res}} = 0$). Then the distribution of S is free up to a constant. Therefore an extra constraint is required to ensure a unique solution:

$$\int_{-1}^{0}\int_{-1}^{1} S dx dz = 0. \tag{41}$$

This constraint is implemented using a Lagrangian multiplier.

To be consistent with this constraint, it is necessary that the total integrated salinity flux over the domain is also zero. The side forcing satisfies $\int_{-1}^{0} \tilde{S} dz = 0$ by definition. However, the atmospheric forcing must be adjusted to satisfy this condition. The salinity flux is then modeled as:

$$k_S S_A(x) = f(0.469 - 0.8\cos(\frac{\pi x}{2} - 0.4)). \tag{42}$$

Since the implementation is discrete, a small correction factor is recalculated at each run to ensure that $\int_{-1}^{1} k_S S_A(x) dx = 0$ holds numerically. The implications of this assumption will be discussed in Section 4.4.2.

3.4 Continuation

To answer the research questions the multistability of the system is analyzed using bifurcation diagrams. In this section it is introduced how such a bifurcation diagram is computed.

In a bifurcation diagram, an indicator of the system's equilibrium state is plotted against a control parameter related to freshwater forcing (either atmospheric, f, or from the side, γ). In this thesis, the vertical axis represents either the mean of ρ over the whole domain, or the sum of the convection over the vertical interfaces: $\sum \mathcal{G}$, averaged in the horizontal direction. This last value gives an indication of how much convection there is in the system. To compute such diagrams, the BifurcationKit package in Julia is used [22], specifically employing its continuation function.

This continuation function requires a sufficiently accurate initial guess of an equilibrium state, which is obtained using Newton's method. It is essential to start from a configuration where the initial state can be found, such as a state of full convection or a

scenario without spatial forcing. At higher spatial resolutions, convergence becomes more challenging. In such cases, an interpolated solution from a coarser resolution is used as the initial guess.

The continuation function uses a pseudo-arclength continuation, allowing the step size to adapt dynamically, more efficiently tracking solution branches through folds [23]. Although this pseudo-arclength continuation is used, the step size must be small enough to avoid missing bifurcations. However, using very small steps can lead to memory issues. In order to resolve this, the bifurcation diagram is divided in segments and a memory node of DelftBlue is used [24]. After each segment, continuation is restarted from the last computed solution. The direction of the restart step is determined by the recent change in the continuation parameter. The sign of this difference determines in which direction the continuation proceeds. For a more detailed discussion of the restart method see Appendix A.2.

Altogether, this chapter has introduced the overarching model, which unifies the key elements of the models by Bastiaansen and Den Toom into a single modular framework. By allowing individual components to be switched on or off, the model can replicate either of the original models or explore their combined effect in a two-dimensional setting. Furthermore, the parameter choices for this configuration have been introduced. This chapter provides a foundation for the systematic analysis of multistability, which is the focus of the next chapter.

4 Results: Multistability in Overarching Model

In this chapter the model is first validated by reproducing the known limit cases of the models of Bastiaansen and Den Toom. These two limit cases are then studied in more depth, answering the following subquestions:

1. How does multistability in Bastiaansen's model depend on its parameters? In particular, does the structure persist within a realistic parameter range?

2. How does the multistability in Den Toom's model depend on the number of vertical layers? Specifically, does the structure converge as the vertical resolution increases?

Finally, the two-dimensional overarching model is set up with parameters chosen to capture key SPG features and analyzed to investigate when multistability arises from physical mechanisms versus numerical artifacts, answering the following subquestion:

3. How does the bifurcation structure of the overarching model change under fresh water forcing (atmospheric, f, and from the side, γ)?

A visual summary of the relationships between the overarching model, the two limit cases, and the model adaptations needed to reproduce the results of Figure 6 and 11a is shown in Figure 17.

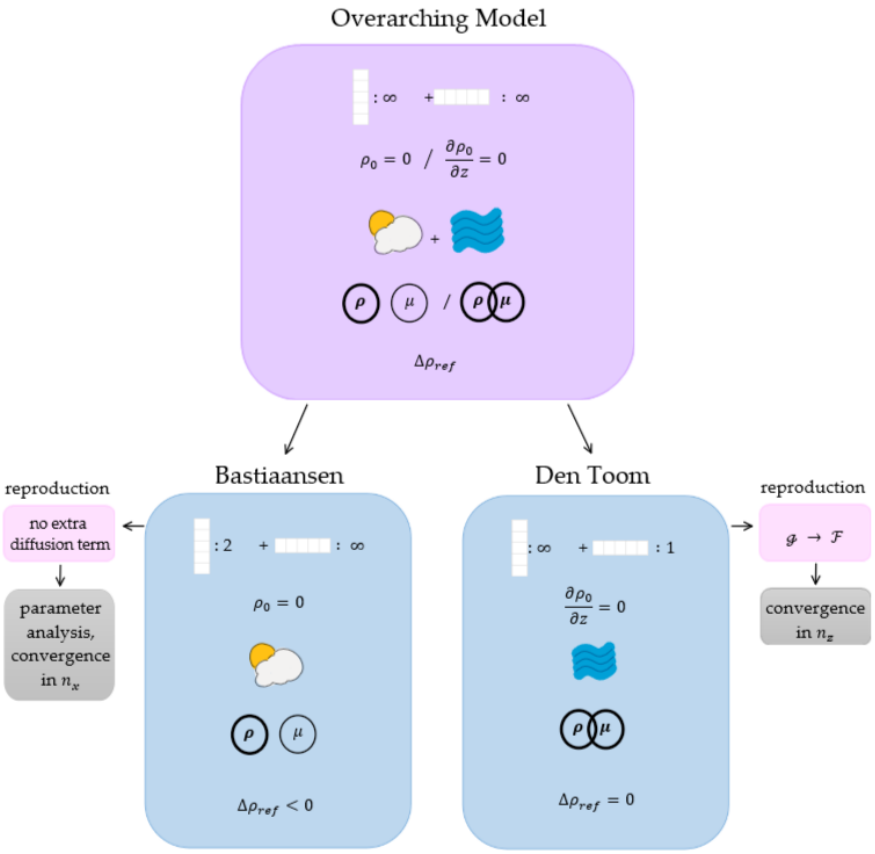


Figure 17: Overview of the key modeling elements in the overarching model and its limits (Bastiaansen's and Den Toom's model). This includes dimensionality, bottom boundary conditions, types of forcing, coupling of ρ and μ, and the convective mixing threshold $\Delta\rho_{\text{ref}}$. Also, the slight modifications for needed to reproduce the results of Figure 6 and 8b are included.

4.1 Model Validation

To validate the overarching model, it is first shown that the results of Figure 6 and 11a can be reproduced by using the settings as denoted in Table 6.

Parameter	**Bastiaansen**	**Den Toom**
`dirichlet_BC`	true	false
`atmos_BC`	true	false
`side_forcing`	false	true
`background_diffusion`	false	true
n_z	2	10
n_x	70	3
F_0	10^4	10^2
P	10^3	10^3
ϵ	1	10
$\Delta\rho_{\text{ref}}$	-1	0
convective adjustment	$\mathcal{G}$	$\mathcal{F}$
D	0.01	0.01
$i^{\text{res}}_{A,T}$	1	–
$i^{\text{res}}_{A,S}$	1	–
k_T	1	–
k_S	1	–
$T_A(x)$	-2	–
$S_A(x)$	$f(1+\cos(\frac{\pi x}{2}))$	–
i^{res}_T	–	1
i^{res}_S	–	0
$\tilde{T}(z)$	–	$\cos(2\pi z)$
$\tilde{S}(z)$	–	$\gamma\cos(\pi z)$

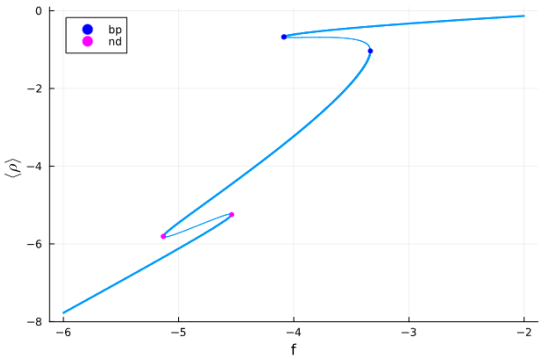
Table 5: Simulation parameters for reproducing Den Toom and Bastiaansen. The table lists boolean values which describe the model configuration, the grid resolution (n_x, n_z), the convection parameters $(F_0, P, \epsilon, \Delta\rho_{\text{ref}})$ and convective adjustment function, the horizontal diffusion D, the side forcing (i^{res}_T, i^{res}_S, $\tilde{T}(z)$, $\tilde{S}(z)$) and the atmospheric forcing ($i^{\text{res}}_{A,T}$, $i^{\text{res}}_{A,S}$, k_T, k_S, $T_A(x)$, $S_A(x)$).

In Figure 18a, the reproduction of the bifurcation diagram of Bastiaansen is plotted, with bifurcation parameter f, representing atmospheric freshwater forcing, on the horizontal axis. On the vertical axis the mean density of the solution, $\langle\rho\rangle$, is plotted. The reproduction is exactly the same as the result of Bastiaansen in Figure 6.

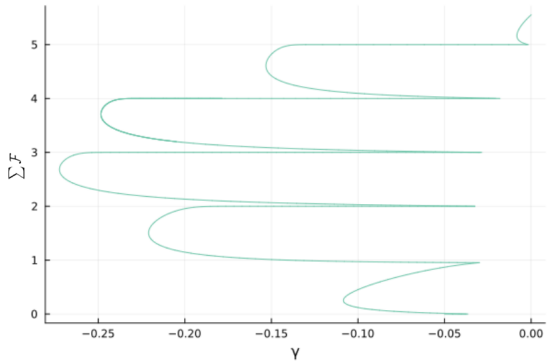
In Figure 18b, the reproduction of the bifurcation diagram of Den Toom is plotted, with bifurcation parameter γ, representing freshwater forcing from the side, on the horizontal axis. On the vertical axis $\sum\mathcal{F}$ averaged over x is plotted, which indicates the intensity of convection in the system. Den Toom has no dynamics in the horizontal direction, however in the reproduction $n_x = 3$ is used. Since there is no heterogeneity in the horizontal direction in the system, this is similar to a one dimensional vertical system. Comparing this reproduction to the result of Den Toom in Figure 11a, there are two differences. In the reproduction the stability is not included, due to computational costs.

Also, the bifurcation diagram is only shown for $\gamma < 0$, since this area is of interest as it represents situations with melting water. Apart from these differences in presenting the solution, the solutions are the same.

Together, these limit cases contain all model elements of the overarching model, thus their successful reproduction provides strong confidence in the implementation of the overarching model.



(a) Reproduction of bifurcation structure found by Bastiaansen. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f.

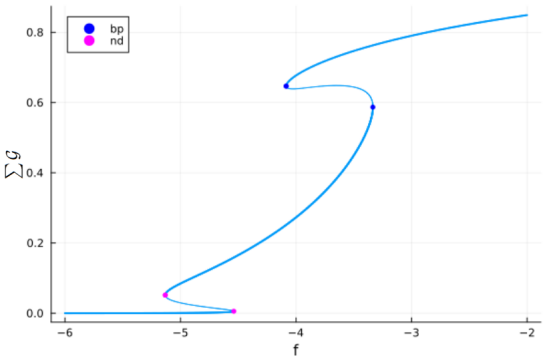


(b) Reproduction of bifurcation structure found by Den Toom. On the vertical axis $\sum \mathcal{F}$, averaged over x, indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ.

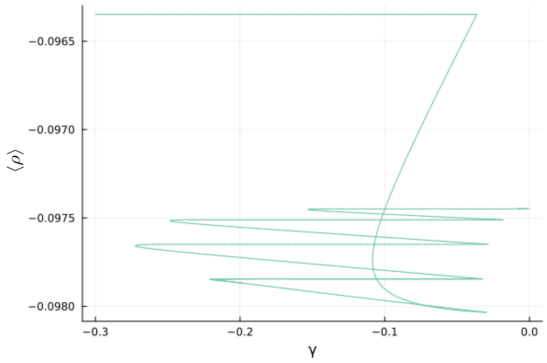
Figure 18: Reproduction of the bifurcation diagrams corresponding to the models of Den Toom and Bastiaansen as limits of the overarching model.

The two bifurcation diagrams use different axes. The control parameter on the horizontal axis is related to freshwater forcing, either atmospheric (f) or from the side (γ). The vertical axis also differs for both cases: $\sum \mathcal{F}$, averaged over x, indicates the intensity of convection in the system, whereas $\langle\rho\rangle$ measures the mean density of the solution.

For the analysis of the two dimensional overarching model it is preferable to use a vertical axis that captures the multistability arising from both mechanisms from Bastiaansen and Den Toom. Therefore, the variable on the vertical axis is changed, with $\sum \mathcal{G}$, averaged over x, in Figure 19a and $\langle\rho\rangle$ in Figure 19b. Since in Bastiaansen's and Den Toom's model different convective adjustment functions are used, $\mathcal{G}$ and $\mathcal{F}$ respectively, the measure on the vertical axis changes slightly. However, both indicate the intensity of convection in the system. Although the multiple equilibria can still be observed in both plots in Figure 19, the use of $\sum \mathcal{F}$ or $\sum \mathcal{G}$, averaged over x, offers clearer insight into the system's equilibrium state as it indicates the intensity of convection. For this reason, this metric is used while analyzing the two dimensional overarching model.



(a) bifurcation structure found by Bastiaansen. On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection. On the horizontal axis the atmospheric freshwater parameter f.



(b) Reproduction of bifurcation structure found by Den Toom. Reproduction Den Toom. On the vertical axis $\langle \rho \rangle$, the mean density of the domain. On the horizontal axis the freshwater parameter from the side, γ.

Figure 19: Reproduction of the bifurcation diagrams corresponding to the models of Den Toom and Bastiaansen with different values on the vertical axis.

4.2 Analysis of Bastiaansen's Model: Physical Meaningful Multistability

As shown in Section 4.1 the bifurcation structure in Bastiaansen's model can be successfully reproduced using the overarching model. This section presents a more detailed analysis of that structure. Specifically, it investigates its sensitivity to model parameters and examines the horizontal resolution required to ensure numerical convergence.

The bifurcation structure corresponding to Bastiaansen's model is plotted in Figure 20 as the turquoise line. Two different types of bifurcations are detected. The upper right double fold consists of two bifurcations, F1 and F2, where one eigenvalue crosses zero. The lower left double fold consists of two bifurcations, F3 and F6, where two eigenvalues cross zero. An additional solution branch, shown in green, emerges from these bifurcations and contains an extra stable solution. The equilibria along this branch are asymmetric, in contrast to those on the main branch. A more detailed discussion of this additional branch is provided in Appendix A.3. In the rest of this thesis, this additional branch is omitted, since it represents secondary multistability that can only exist in the presence of multistability in the main branch.

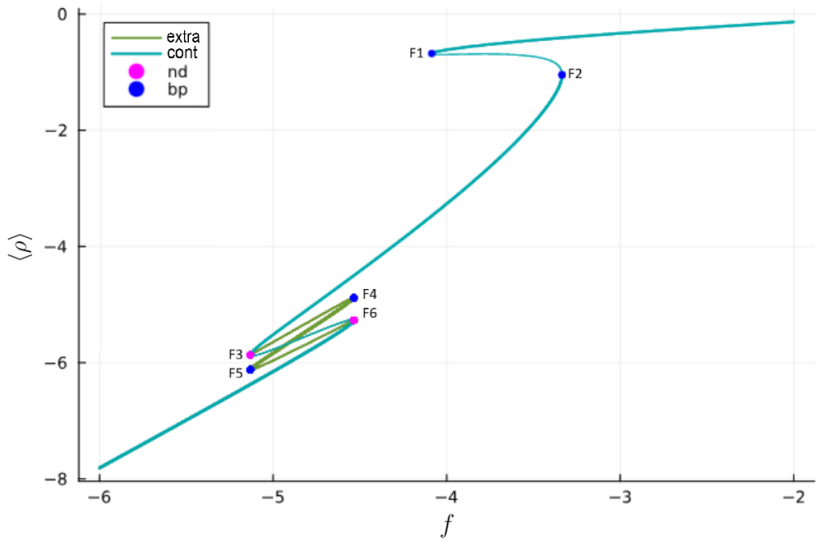


Figure 20: Bifurcation diagram using the function bifurcationdiagram from BifurcationKit. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f. The turquoise line shows the general continuation of the branch. The green line shows the extra fold. The dots represent bifurcation points. In *bp* one eigenvalue crosses zero, in *nd* two. Stable solutions are displayed by a bold line, unstable by a thinner line.

4.2.1 Influence of $\Delta\rho_{\text{ref}}$ on Bifurcation Structure

The bifurcation structure observed in Bastiaansen's model is highly sensitive to the parameters $\Delta\rho_{\text{ref}}$, k_T, and D. This subsection focuses on $\Delta\rho_{\text{ref}}$, which determines the critical value at which the convective adjustment function $\mathcal{G}$ triggers mixing.

In Bastiaansen's setup, $\Delta\rho_{\text{ref}} = -1$, which means that convection is triggered even under stable stratification (when the surface is less dense than the deep layer). This behavior is not physically realistic. However, it has been argued in the literature that using a small negative value for $\Delta\rho_{\text{ref}}$ improves numerical stability, while still preserving the essential dynamics of the system in the limit as $\Delta\rho_{\text{ref}} \to 0^-$ [14]. From a physical perspective, $\Delta\rho_{\text{ref}} = 0$ is the more appropriate choice: convection occurs only when the system becomes unstably stratified, which is consistent with the underlying physics.

To test this, in Figure 21 the bifurcation diagram is plotted for a range of $\Delta\rho_{\text{ref}}$ values between -1 and 1 with steps of 0.1. It can be seen that by increasing the value, the unstable part of the solutions decreases, until the bifurcations eventually vanish. In Figure 21, it can be seen that at $\Delta\rho_{\text{ref}} = -0.2$ the bottom left double fold has vanished. All bifurcations have disappeared at $\Delta\rho_{\text{ref}} = -0.1$. This indicates that the structure does not persist in the physically consistent limit ($\Delta\rho_{\text{ref}} = 0$), contrary to earlier assumptions.

The bifurcation point where the fold bifurcations disappear when varying $\Delta\rho_{\text{ref}}$ is called a cusp bifurcation. This is a bifurcation where two branches of saddle-node bifurcations meet tangentially [25]. The cusps are marked with red stars in Figure 21.

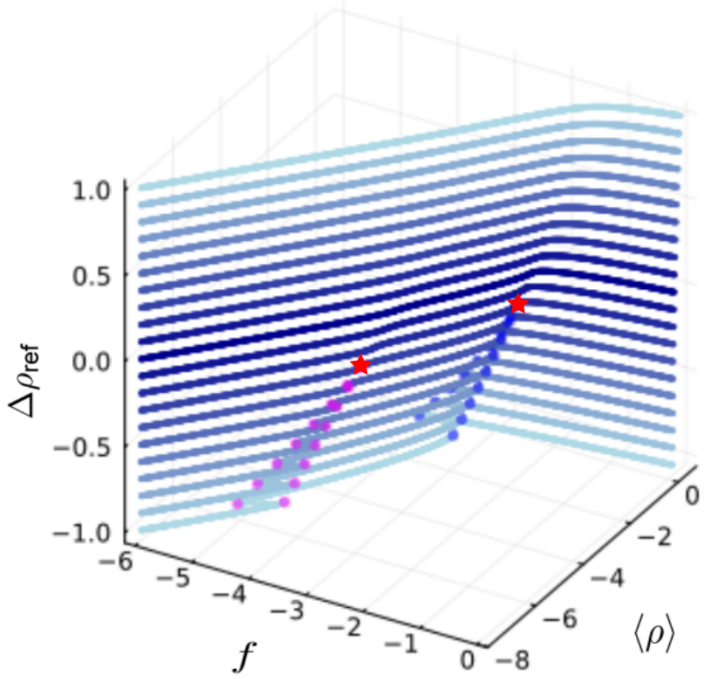


Figure 21: Three-dimensional bifurcation diagram using the parameters of Bastiaansen (table 1) and ranging $\Delta\rho_{\text{ref}} \in [-1, 1]$ with steps of 0.1 on the z-axis. The darkest line represents $\Delta\rho_{\text{ref}} = 0$. On the y-axis $\langle\rho\rangle$, the mean density of the domain. On the x-axis the atmospheric freshwater parameter f. The dots represent bifurcation points: in blue dots one eigenvalue crosses zero, in pink dots two. The red stars indicate the cusp bifurcations.

The parameters used in the overarching model are listed in Section 3.2. It is not straightforward to use these parameters in a simplified two-box model as Bastiaansen's, in which one box represents the surface ocean and the other the deep ocean. For instance, atmospheric forcing in the overarching model affects only a thin layer at the top of the surface box. The relative magnitudes of k_T, D and $\kappa = \frac{F_0}{P}$ determine which processes dominate: the atmospheric influence, the smoothing by diffusion, or vertical convection. Different values can therefore lead to qualitatively different bifurcation structures. To assess the effect of the parameters k_T and D these parameters are systematically varied.

Figure 22 summarizes the resulting bifurcation types across the $(\Delta\rho_{\text{ref}}, k_T, D)$ parameter space. The purple transparent markers correspond to bifurcation diagrams without any bifurcation points, the lighter blue markers correspond to two bifurcation points, and the dark blue markers correspond to four or more bifurcation points. Within the last category, a variety of bifurcation shapes can occur, though a detailed classification is beyond the scope of this study.

The main observation here concerns the role of $\Delta\rho_{\text{ref}}$. A bifurcation structure with multistability can be found for negative values of $\Delta\rho_{\text{ref}}$, but this multistability vanishes when $\Delta\rho_{\text{ref}}$ is set to zero. This further supports the finding that the observed multistability does not persist under physically realistic parameter choices.

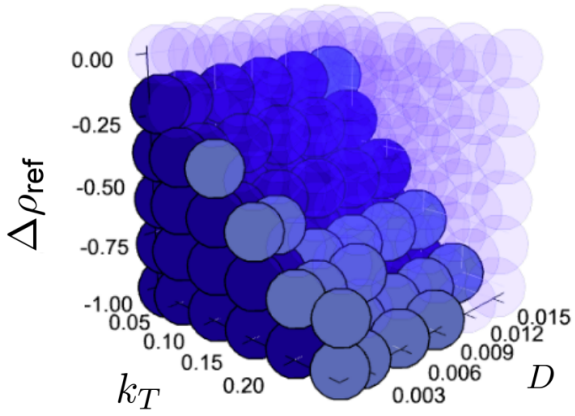
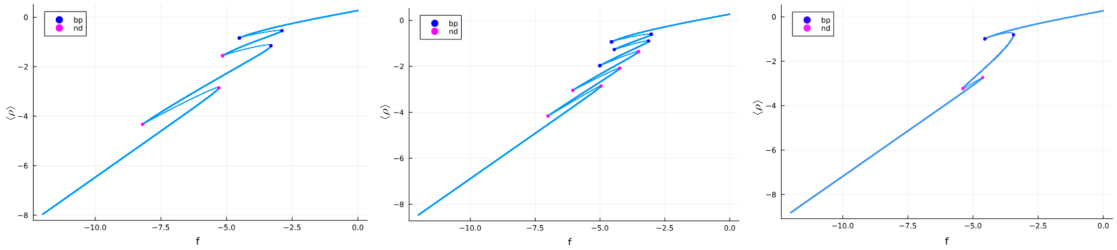


Figure 22: Indication of the bifurcation structure for different values of k_T (atmospheric exchange rate), D (diffusion in the horizontal direction) and $\Delta\rho_{\text{ref}}$ (convection threshold parameter). The purple transparent markers correspond to bifurcation diagrams without any bifurcation points, the lighter blue markers correspond to two bifurcation points, and the dark blue markers correspond to four or more bifurcation points.

4.2.2 Horizontal Resolution and Convergence

An important consideration in analyzing Bastiaansen's bifurcation structure is the choice of horizontal resolution, denoted by n_x. If n_x is too low, the solution fails to converge, and spurious bifurcations appear as numerical artifacts. This behavior is illustrated in Figures 23a and 23b, reproducing the results presented by Zanuttini [26]. In this setup, an additional variable box is introduced in the vertical direction ($n_z = 3$). At low resolution in the horizontal direction multiple double folds appear. Their number equals $\lceil \frac{n_x}{2} \rceil$, where n_x is the number of horizontal grid cells. In principle, each grid cell gives rise to a double fold. However, due to the left–right symmetry of the domain, symmetrically located cells behave identically and thus share the same double fold, halving the total count. If n_x is odd, the central grid cell adds an additional double fold, which explains the ceiling function in $\lceil \frac{n_x}{2} \rceil$. For sufficiently large n_x the solution converges. In Figure 23c the converged solution is plotted for $n_x = 70$, for larger values of n_x the bifurcation structure remains the same.



(a) $n_x = 5$. (b) $n_x = 10$. (c) $n_x = 70$.

Figure 23: Bifurcation diagrams showing the dependency on n_x. Bastiaansen's parameters (Table 1) and model with an additional variable box are used. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f. The dots represent bifurcation points: in blue dots one eigenvalue crosses zero, in pink dots two.

4.3 Analysis of Den Toom's Model: Spurious Multistability

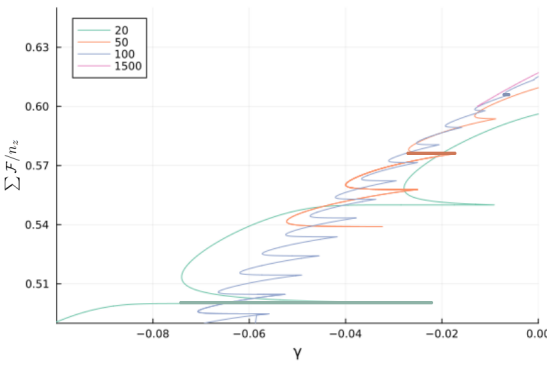
As shown in Section 4.1 Den Toom's bifurcation structure can be successfully reproduced using the overarching model. In the original analysis, bifurcation diagrams were only presented for vertical resolutions up to $n_z = 20$, leading to the suggestion that the observed bifurcations were spurious. However, it remains unclear whether this structure might converge at higher resolutions. To investigate this, a convergence check is performed.

4.3.1 Vertical Resolution and Convergence

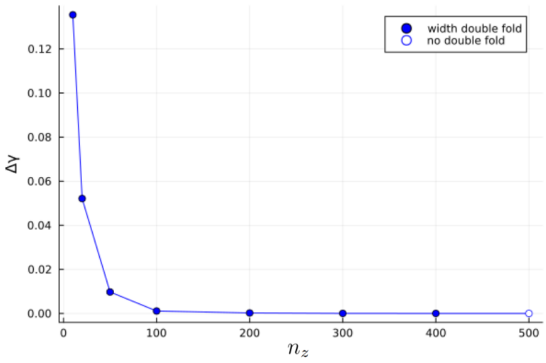
Figure 24a shows part of the bifurcation diagram for increasing values of n_z. As n_z increases, the distance between the folds decreases, eventually leading to their disappearance. This suggests that the multistability is not physical, but rather introduced by limited vertical resolution.

To quantify this behavior, Figure 24b plots the width of the second double fold starting from $\gamma = 0$ as a function of n_z. This width is indicated for $n_z = 20$, 50 and 100 with a thick, slightly darkened line in Figure 24a. Empty dots indicate that the fold has disappeared entirely, providing further evidence that the bifurcations vanish with sufficient resolution.

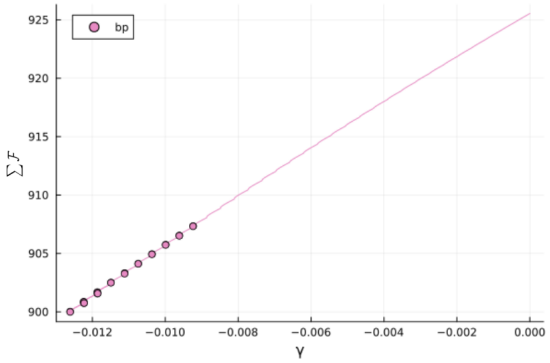
The required resolution for convergence depends on the value of the freshwater parameter γ: for more negative values, finer vertical resolution is necessary to eliminate spurious folds. This behavior is illustrated more explicitly in Figure 24c, which shows the bifurcation diagram at $n_z = 1500$, with bifurcation points highlighted. These bifurcations only appear for $\gamma < -0.009$, for larger values of γ, the diagram becomes smooth and free of bifurcations.



(a) Bifurcation structure for different values of n_z. In order to compare the bifurcation diagrams, the vertical axis is set to $\frac{\sum \mathcal{F}}{n_z}$. The thick, slightly darker lines indicate the width of the second double fold starting from $\gamma = 0$.



(b) Width ($\Delta\gamma$) of the second double fold starting from $\gamma = 0$ versus n_z. For $n_z = 20, 50, 100$ these are indicated with a thick, slightly darkened line in Figure 24a. An empty dot means no double fold.



(c) Bifurcation structure for $n_z = 1500$. The dots mark bifurcation points.

Figure 24: Convergence with n_z. (a,b) Bifurcation structure for different values of n_z. On the vertical axis $\sum \mathcal{F}$, averaged over x, indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ.

Because achieving convergence for all values of γ requires extremely high vertical resolution, the exact structure of the bifurcation diagram at full convergence remains unclear. To test the hypothesis that the bifurcations disappear for more negative γ as n_z is increased, the analysis is repeated using $P = 50$, which strengthens vertical diffusion and likely reduces the required resolution. The resulting bifurcation diagrams for varying n_z are shown in Figure 25. These results reveal the fully converged solution: one double fold remains.

For the analysis in this section, $\tilde{T}(z) = \cos(2\pi z)$ is used in line with the analysis performed by Den Toom. However, in the analysis of the overarching model $\tilde{T}(z) = -\cos(2\pi z)$ is used since this represents a cold-warm-cold configuration as discussed in Section 2.4. The results of this section still apply since the bifurcation diagrams using both temperature profiles are the same, as shown in Appendix A.4.

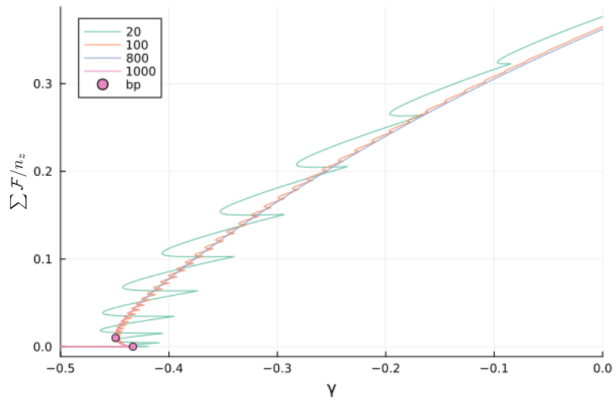


Figure 25: Bifurcation structure for different values of n_z, using stronger vertical background mixing ($P = 50$). On the vertical axis $\frac{\sum \mathcal{F}}{n_z}$, averaged over x, indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ. The dots represent the bifurcation points remaining in the converged solution.

4.4 Multistability in the Overarching Model under Freshwater Forcing

After analyzing both limit cases, this section examines the impact of fresh water forcing on the two-dimensional overarching model. The bifurcation structure under fresh water forcing, parameterized by the atmospheric forcing amplitude f and side forcing amplitude γ, is systematically studied.

First, the side forcing and atmospheric forcing are separately applied on the two-dimensional system to study their respective influence. Thereafter, both forcings are applied simultaneously and the bifurcation structure is generated on a grid of f and γ values.

Unless stated otherwise, the parameter values used are given in Table 6 (see Section 3.2 for a motivation). A zero-flux bottom boundary condition is used, consistent with the configuration in most complex climate models.

Furthermore, typical climate models have a vertical resolution of around 64 layers and a horizontal resolution of 1.0 ° (common) to 0.1 ° (high resolution). Since the modeled area in this overarching model spans 15 °, that corresponds to a horizontal resolution of $n_x = 15$ to $n_x = 150$.

The computational power is however a limiting factor for this analysis. Therefore in the analysis that follows $n_z = 10$ is chosen. This choice is further motivated in Section 4.3 where it is shown that the solution using $n_z = 64$ is qualitatively similar to the solution using $n_z = 10$. The baseline horizontal resolution is $n_x = 10$, which is varied to explore resolution effects.

Parameter	Overarching model
`dirichlet_BC`	false
`atmos_BC`	true
`side_forcing`	true
`background_diffusion`	true
n_x	10
n_z	10
F_0	10^2
P	10^3
ϵ	10
$\Delta\rho_{\text{ref}}$	0
convective adjustment	$\mathcal{G}$
D	0.78
$i^{\text{res}}_{A,T}$	1
$i^{\text{res}}_{A,S}$	0
k_T	24.3
$T_A(x)$	$-2.15 + 0.35x$
$k_S S_A(x)$	$f(0.469 - 0.8\cos(\frac{\pi x}{2} - 0.4))$
i^{res}_T	1
i^{res}_S	0
$\tilde{T}(z)$	-cos($2\pi z$)
$\tilde{S}(z)$	$\gamma\cos(\pi z)$

Table 6: Default parameter values used in the overarching model. The table lists boolean values which describe the model configuration, the grid resolution (n_x, n_z), the convection parameters $(F_0, P, \epsilon, \Delta\rho_{\text{ref}})$ and convective adjustment function, the horizontal diffusion D, the side forcing (i^{res}_T, i^{res}_S, $\tilde{T}(z)$, $\tilde{S}(z)$) and the atmospheric forcing ($i^{\text{res}}_{A,T}$, $i^{\text{res}}_{A,S}$, k_T, $T_A(x)$, $k_S S_A(x)$).

4.4.1 Dynamics under Side Forcing

First of all, the model is studied when only side forcing is present (`atmos_BC` = false). In this case the model is quite similar to the model by Den Toom. The only difference is the convective adjustment function, which is $\mathcal{G}$ instead of $\mathcal{F}$.

In Figure 26 the bifurcation diagram is shown for $n_z = 10$ with both $\mathcal{G}$ and $\mathcal{F}$. It can be seen that the bifurcation structure is qualitatively similar. With $\mathcal{G}$, the bifurcation structure appears slightly smoother, which is consistent with the fact that $\mathcal{G}$ (Equation (12)) provides a smoother approximation of the step function than $\mathcal{F}$ (Equation (11)).

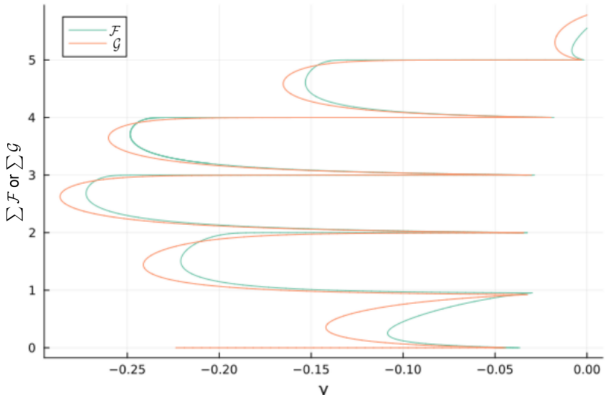


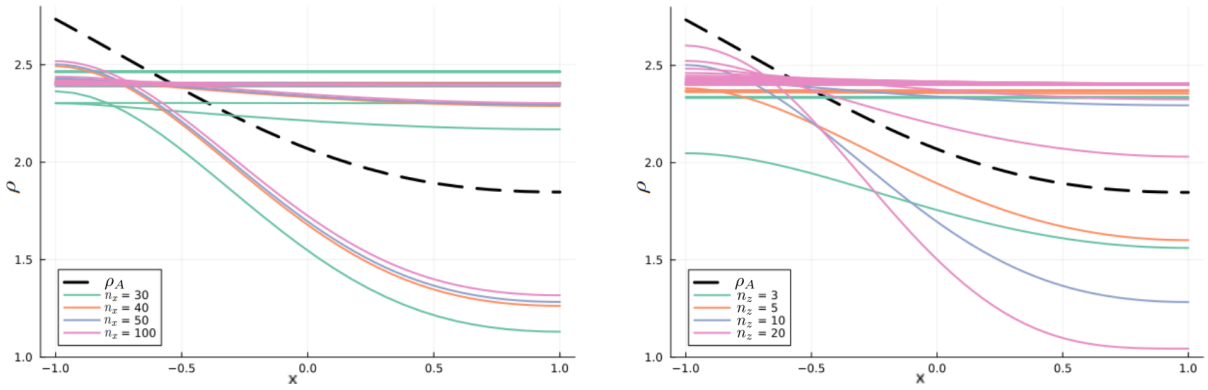
Figure 26: Bifurcation structure with only side forcing for $n_z = 10$, using $\mathcal{F}$ and $\mathcal{G}$. On the vertical axis $\sum\mathcal{F}$ or $\sum\mathcal{G}$, averaged over x, indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ.

4.4.2 Dynamics under Atmospheric Forcing

Next, the influence of the atmospheric forcing on the system is studied by neglecting the side forcing (`side_forcing` = false). Because the model behavior depends on the resolution in both the horizontal and vertical direction, first the influence of n_x and n_z is examined. Thereafter, the bifurcation diagram over fresh water parameter f is analyzed.

In Figure 27, the density profiles at $f = 0.3$ are shown for different combinations of horizontal (n_x) and vertical (n_z) resolution. For each configuration (n_x, n_z), the density profile at every discretized layer is plotted, and all curves belonging to the same configuration are shown in the same color. Thus, the number of curves corresponds directly to the chosen vertical resolution n_z. Figure 27a shows the effect of varying n_x with $n_z = 10$, while Figure 27b shows the effect of varying n_z with $n_x = 50$. The dashed black line indicates the atmospheric density ρ_A. Differences between the density profiles are most pronounced in the upper ocean layers, where atmospheric forcing has the strongest impact.

In Figure 27a, it can be seen that a sufficiently large horizontal resolution n_x is required for convergence. In this case, $n_x = 40$ is adequate, as increasing the resolution beyond this does not change the solution. However, the value of n_x needed for convergence increases with f, since then sharper gradients occur in the system. Figure 27b demonstrates the effect of vertical resolution. Higher resolution leads to thinner layers, therefore the impact of ρ_A can be seen in more top layers.



(a) Profiles varying with n_x, $n_z = 10$.

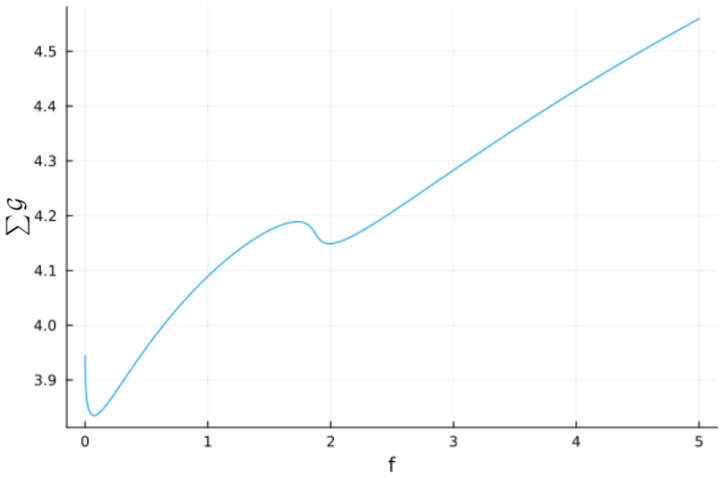
(b) Profiles varying with n_z, $n_x = 50$.

Figure 27: Density profiles ρ at $f = 0.3$ for different vertical (a) and horizontal (b) resolutions. For each configuration (n_x, n_z), the density profile at every discretized layer is plotted, and all curves belonging to the same configuration are shown in the same color. Thus, the number of curves corresponds directly to the chosen vertical resolution n_z. The dashed black line represents the atmospheric density, ρ_A.

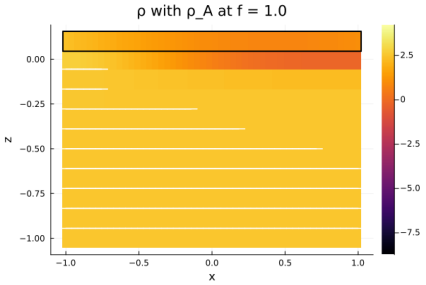
The bifurcation diagram over f is shown in Figure 28a using $n_x = 50$ and $n_z = 10$. It can be seen that there is no multistability. There are however slight changes visible in the total convection strength in the system (expressed by $\sum \mathcal{G}$).

To explain this behavior, it is important to note that the condition $\int_{-1}^{1} S_A dx = 0$ is implemented to ensure that the system is uniquely solvable. Due to this assumption the atmospheric density ρ_A is positive on the left side and negative on the right side. This results in an unstable stratification on the left, which triggers convective mixing there. The positive density spreads downward and, due to the diffusion in the horizontal direction, spreads over the entire horizontal domain. As the freshwater forcing amplitude f increases, this effect intensifies, eventually causing the negative density to spread upwards again because of background diffusion in the vertical direction. This vertical spread is possible since strong convective mixing does not occur in the upper right region, where the stratification remains stable.

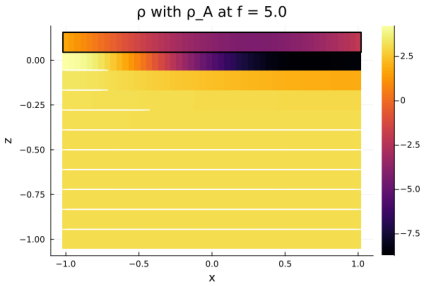
This behavior is illustrated in Figures 28b and 28c, where the solutions for different values of f are shown. Here, convective mixing across an interface is indicated by a white line. This convection pattern is mainly induced by the salinity, as shown in Appendix A.5. The turning points in the bifurcation diagram coincide with the onset of convective mixing across an entire vertical layer.



(a) Bifurcation structure. On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection. On the horizontal axis the atmospheric freshwater parameter, f.



(b) Visualization at $f = 1.0$.



(c) Visualization at $f = 5.0$.

Figure 28: (a) Bifurcation structure for $n_z = 10$ and $n_x = 50$, with only atmospheric forcing. (b,c) Visualizations of the solution for two different atmospheric freshwater parameter values. The value of ρ is displayed on the (x, z)-domain as a color plot. White lines between vertical layers indicate static instability (top layer denser than the layer underneath). The upper layer in the black box represents the atmospheric forcing.

4.4.3 Dynamics under Atmospheric and Side Forcing

In this section, the two-dimensional model is investigated under the combined influence of atmospheric and side forcing. First, the bifurcation structure is analyzed with freshwater forcing only from one side: varying the freshwater parameter from the side, γ, with no atmospheric freshwater forcing, $f = 0$, and varying the atmospheric freshwater parameter, f, with no freshwater forcing from the side, $\gamma = 0$, focusing on the dependency of the number of horizontal cells n_x. Thereafter, the bifurcation structure is computed for a grid of freshwater parameters γ and f.

Freshwater forcing from the side First of all, the bifurcation structure is studied varying the freshwater parameter from the side, γ, with no atmospheric freshwater forcing, $f = 0$, meaning that the only freshwater forcing in the system originates from the side. The bifurcation diagrams for $n_x = 10$ and $n_x = 50$ are shown in Figure 29. The qualitative structure is similar to the results obtained by Den Toom (Figure 26). However, since the surface is now cooled, making the top layer denser, a larger freshwater flux is required to have stratification. This is reflected by the larger values on the horizontal axis. Furthermore, for $n_x = 10$ extra multistability appears, whereas for $n_x = 50$ these folds vanish. This indicates that this multistability is a result of insufficient resolution.

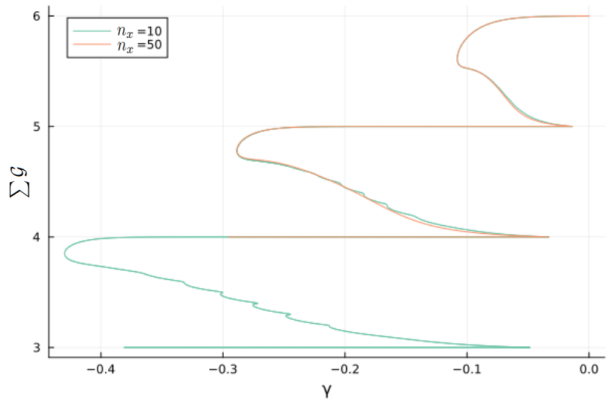
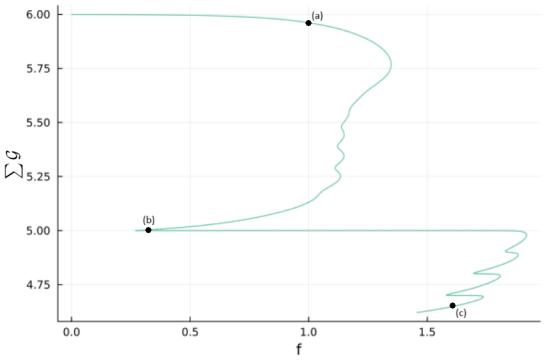
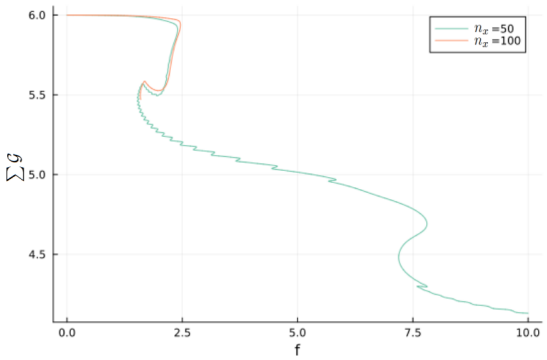


Figure 29: Bifurcation diagram for with $n_z = 10$ and $n_x = 10$ or $n_x = 50$. Only freshwater forcing from the side is applied ($f = 0$). On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection. On the horizontal axis the freshwater parameter from the side, γ.

Freshwater forcing from the atmosphere Next, the bifurcation structure is studied varying the atmospheric freshwater parameter, f, with no freshwater forcing from the side, $\gamma = 0$, meaning that the only freshwater forcing in the system originates from the atmosphere. The bifurcation diagrams for $n_x = 10$, $n_x = 50$ and $n_x = 100$ are shown in Figure 30. These plots illustrate the strong impact of the horizontal resolution: the bifurcation diagram for $n_x = 10$ (Figure 30a) is structurally very different from those for $n_x = 50$ and $n_x = 100$ (Figure 30b). These seem to converge to the same solution, although for $n_x = 50$ some additional saddle nodes appear that vanish for $n_x = 100$, indicating that these arise due to insufficient horizontal resolution.



(a) $n_x = 10$. The dots indicate the solutions shown in Figure 31.

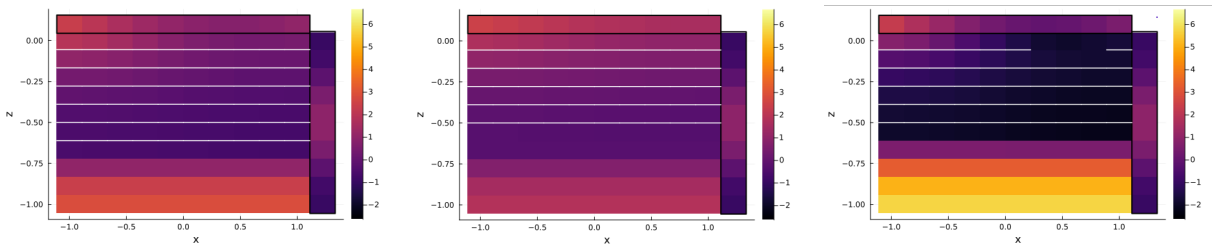


(b) $n_x = 50$ and $n_x = 100$.

Figure 30: Comparison of bifurcation diagrams for different n_x values with $n_z = 10$. Only freshwater forcing from the atmosphere is applied ($\gamma = 0$). On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection. On the horizontal axis the atmospheric freshwater parameter f.

The difference in structure seems to relate to the values of the atmospheric forcing, represented by T_A and S_A. On a coarser grid, these values are less extreme. For $n_x = 10$, the first bifurcation that occurs relates to the lowest row of the mixed layer becoming stably stratified. This transition is illustrated by the change between Figures 31a and 31b, where the density fields in the (x, z)-plane are shown together with white lines

indicating the layers between which convection occurs. In Figure 31b, the bottom white line has vanished, indicating that this layer has become stably stratified. The bifurcation at $f \approx 1.9$ corresponds to the upper layer that separates itself from the rest. This process is illustrated in Figures 31b and 31c, where the breakup of the top white line indicates detachment of the upper layer.



(a) $f = 1.0$. (b) $f = 0.29$. (c) $f = 1.64$.

Figure 31: Comparison of ρ fields at different atmospheric forcing values f, corresponding to Figure 30a, visualizing the different possible structural changes. The value of ρ is displayed on the (x, z)-domain as a color plot. White lines between vertical layers indicate static instability (top layer denser than the layer underneath). The upper layer in the black box represents the atmospheric forcing, the column in the black box represents the side forcing.

For $n_x = 50$, the bifurcation structure changes: the double folds correspond to an additional top layer separating from the rest of the column, while the lowest row of the mixed layer becomes stably stratified between these folds. Although the structure appears converged with respect to n_x, it remains strongly dependent on n_z.

Grid Finally the bifurcation structure can be computed on a grid of freshwater parameters f and γ. In Figure 32 ($n_x = 10$) and 33 ($n_x = 50$) these are shown only looking at $f > 0$ and $\gamma < 0$, representing the situations with meltwater. The bifurcation structure for $n_x = 50$ is much smoother than that for $n_x = 10$. Especially in the bifurcation diagrams over γ the spurious bifurcations due to horizontal resolution have disappeared. The spurious bifurcations due to vertical resolution remain.

Altogether, these experiments highlight that the bifurcation structure of the overarching model is controlled primarily by the side-forcing mechanism, producing spurious multistability as found in Den Toom's model. Horizontal heterogeneity introduces additional folds, but these disappear once horizontal resolution is sufficient.

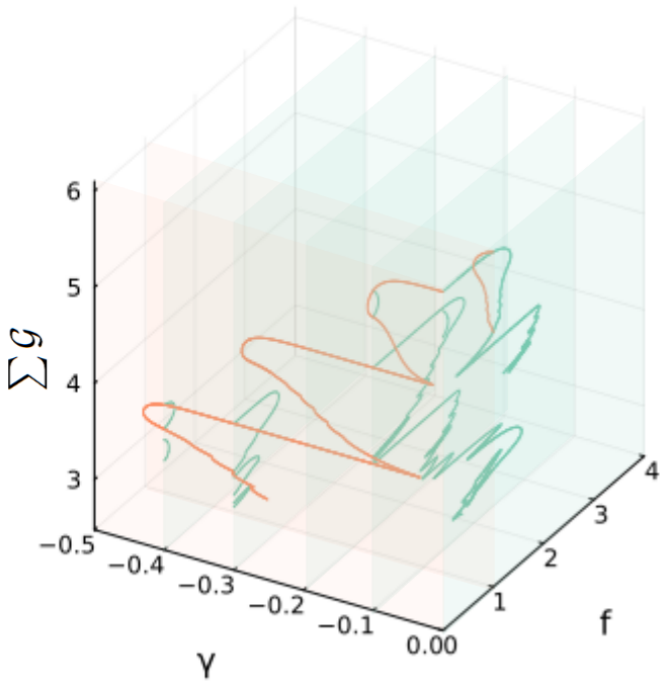


Figure 32: Bifurcation structure on a (f,γ)-grid, representing freshwater forcing from the atmosphere (f) and from the side (γ), using $n_x = 10$ and $n_z = 10$. The orange (green) lines show bifurcation structures over γ (f) for a constant f (γ). On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection.

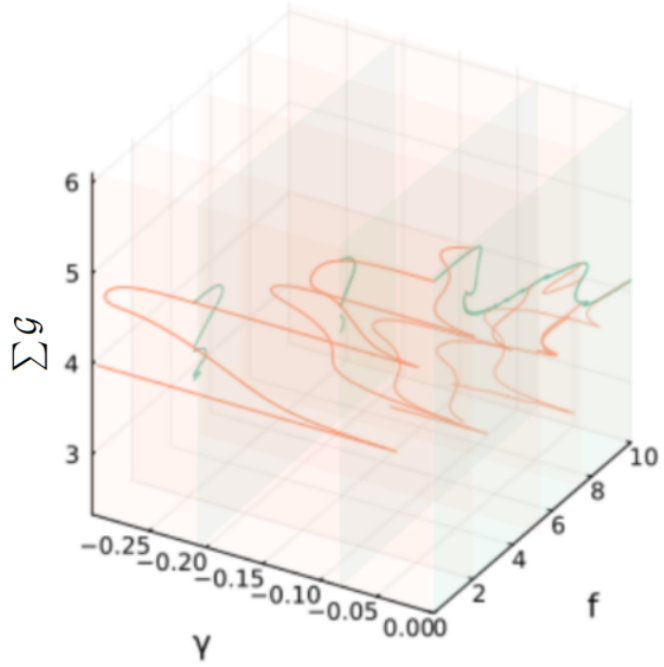


Figure 33: Bifurcation structure on a (f,γ)-grid, representing freshwater forcing from the atmosphere (f) and from the side (γ), using $n_x = 50$ and $n_z = 10$. The orange (green) lines show bifurcation structures over γ (f) for a constant f (γ). On the vertical axis $\sum \mathcal{G}$, averaged over x, indicating the intensity of convection.

In summary, the analysis in this chapter shows that multistability in Bastiaansen's model depends critically on parameter choices, disappearing when physically realistic parameters are used. In Den Toom's model, multistability vanishes with increased vertical resolution, confirming that it is a numerical artifact. Since climate models lack sufficient vertical resolution, this artifact is relevant in practice. The overarching model demonstrates that a combination of both mechanics can introduce additional multistability on top of spurious multistability similar to Den Toom, but these features vanish with higher horizontal resolution.

5 Discussion and Further Research

This chapter reflects on the main findings of this thesis, highlighting the assumptions and limitations of the modeling approach. Furthermore, the results are discussed in the context of idealized ocean models and their role in climate research. Meanwhile, potential directions for future work are suggested.

5.1 Limitations of Modeling Approach

The model developed in this thesis is intentionally idealized to enable systematic analysis of convection dynamics under controlled conditions. While this approach provides clear insights into the influence of specific processes, it also relies on assumptions and simplifications that constrain its realism and scope. This section examines four such aspects in detail: the spatial representation of the system, the extra constraint imposed for solvability, the treatment of freshwater forcing, and the effects of numerical resolution. For each, the implications for interpretation are discussed and possible directions for future improvement are outlined.

5.1.1 Spatial Assumptions for Idealized Model

The idealized model relies on several simplifying spatial assumptions in space that can influence the results. The two-dimensional setup of the model limits the representation of oceanic processes, which are three-dimensional. The horizontal domain is chosen as a zonal line on the area where deep convection takes place. The longitudinal dynamics are represented by the side forcing, which mimic the boundary current.

In reality, the gyre is an area enclosed by a boundary current, see Figure 34 [18]. To be able to realistically model the spatial heterogeneity within the gyre it would therefore be necessary to make a three-dimensional model. This also raises questions about the assumption that the side forcing acts uniformly on all grid cells, as the boundary current interacts locally with the neighbouring water directly, not uniformly with the entire gyre. This side forcing is also set to a very simplified constant, but in reality the boundary current and the gyre influence each other. It could therefore be better to implement a variable boundary current, similar to the model by Born which is introduced in Box 3.

This also brings attention to the choice of a constant horizontal diffusion coefficient D, which implies uniform eddy mixing across the domain. This contradicts observational evidence for strong spatial variability [27].

Box 3: Variable Boundary Current

Born [18] introduced a model with a variable boundary current, which better captures the interaction between the boundary current and the convective basin. The model consists of four boxes: the boundary current and convective basin contain an upper and lower layer each.
The cylindrical interior region models the convective basin. The upper and lower region are instantly mixed by convection when stratification becomes unstable. Note that this differs from the modeling of convection in this thesis. The upper layer is forced by temperature and freshwater flux.
The boundary current encircles the convective basin. The layers of the boundary current are in contact with the convective basin through parameterized eddy fluxes of heat and salt. The fluxes depend on the differences in temperature and salinity between the two regions and are scaled by an exchange coefficient, which is influenced by the velocity of the boundary current.

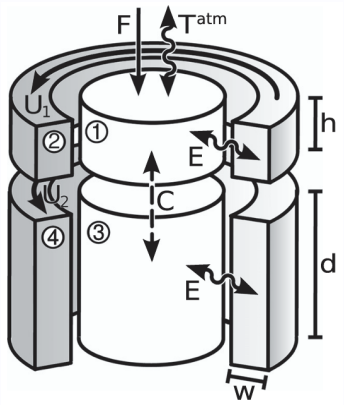


Figure 34: Schematic of the model by Born with a variable boundary current [18]. The convective basin is represented as a cylindrical interior region with an upper layer of height h and lower layer of height d, which are mixed by convection (C) when unstable stratification occurs. Surrounding it is the boundary current, also split into an upper and lower layer with velocities U_1 and U_2, respectively. Heat and salt exchange between the basin and current occurs through parameterized eddy fluxes (E). The upper layer of the interior region is forced by a fresh water flux F and relaxed towards the atmospheric temperature T^{atm}.

Further research could address these limitations by expanding the model by Born. This model could be expanded to a vertical continuous column to further study the interaction between the boundary current and the gyre.

Another topic for further research could be to implement a different type of spatial structure, considering multiple sites of deep convection. Connecting these sites could also induce different patterns [28], where certain sites are off while others show deep convection. This could maybe explain the different convection patterns in the multistability observed in climate models [7].

5.1.2 Extra Constraint for Solvability

Another modeling choice is the extra constraint on the salinity: $\int_{-1}^{0}\int_{-1}^{1} S dx dz = 0$. This constraint serves to determine a unique solution to a system that is determined up to a constant due to Neumann boundary conditions at all boundaries. This choice can be justified as the convection depends on the gradient, which is the same for all constants.

However, to be able to find a solution, the forcings in salinity also have to be net zero. For the atmospheric forcing this means: $\int_{-1}^{1} S_A dx = 0$. This means that at a part of the surface the water is modeled to become saltier instead of fresher. This contrasts with reality, where ice melt and precipitation lead to freshening across the whole surface. In Section 4.4.2 the impact of this assumption became clear. An artifact that has to be taken into consideration while analyzing the results of the overarching model.

5.1.3 Fresh Water Forcing in Overarching Model

The study of freshwater forcing via parameters f (atmospheric forcing amplitude) and γ (side forcing amplitude) provides insight into their separate and combined effects on the bifurcation structure. But the bifurcation structures have only been studied for one of these forcing amplitudes set to a constant. The freshwater input from the atmosphere and the boundary current are however linked.

Also, since the freshwater forcing is modeled as a constant salinity flux, it does not take into account that in time the surface water freshens. However, the difference in salinity does depend on the current salinity and amount of precipitation, as given by Equation (33).

Future work should explore reducing the freshwater forcing parameters f and γ to a single parameter. This could simplify interpretation within real oceanographic contexts and highlight joint effects.

5.1.4 Numerical Resolution and Convergence

In Section 4.2.2 and 4.3.1 the importance of using sufficiently large n_x and n_z for convergence is highlighted. However, the analysis method is limited in computational efficiency, although a pseudo-arclength algorithm and a memory node of DelftBlue are used. This limits the possibilities for analyzing the two-dimensional model where increasing the resolution comes with great computational cost. The resolution does however alter the results, as shown in Section 4.4.3. Caution is needed while analyzing these results, since solutions are analyzed that have not converged.

Further research is therefore needed to develop a more efficient analysis method. There is mainly improvement possible in varying the step size within the pseudo-arclength algorithm, as the super small step size is only necessary for the sharp turns in the bifurcation structures. The other parts could do with bigger steps.

Another way to improve numerical efficiency is to vary the thickness of the vertical boxes with depth, using thinner boxes near the surface where more dynamic processes occur.

5.2 Role in Climate Research

While the immediate findings of this study are important, the modeling framework also fits within a broader scope of using idealized ocean models to advance climate science.

This section considers two aspects in particular: the relationship between bifurcation analyses and transient behavior in a changing climate, and relating climate model closure schemes to effective vertical mixing parameters.

5.2.1 Bifurcation Analysis vs Transient Analysis

Throughout this paper, bifurcation analysis has been used to systematically explore the multiple equilibria of the idealized ocean model. In contrast, most climate models are typically run in a transient mode, evolving forward in time until an equilibrium is reached or conditions change. Despite this difference in approach, knowledge of multiple steady states remains crucial: even in transient simulations, the presence of alternative equilibria can manifest through phenomena such as rate-induced tipping or noise-induced tipping, which may cause the system to shift abruptly between states under gradual or stochastic forcing. Understanding the bifurcation structure thus provides a framework for interpreting potential transitions in time-dependent climate simulations.

5.2.2 Relating Closure Schemes to Effective Mixing Parameters

Vertical mixing is a key process in ocean models, but its parameterization varies across climate modeling frameworks [29]. Common schemes include enhanced vertical diffusion (used in this thesis) and turbulence closure schemes such as Gaspar [30] and KPP [31]. To connect the results of this thesis to climate models using such closure schemes, it is important to relate them to the enhanced vertical diffusion by estimating effective mixing parameters.

Future work could develop a method for estimating effective mixing parameters from turbulence closure schemes used in climate models. A suggested approach is:

1. Initialize a vertical column with a linearly stable stratification due to temperature alone.
2. Apply a constant surface cooling until convection occurs.
3. Keep forcing constant and monitor the evolution of surface temperature, mixed-layer depth, and vertical mixing coefficient.
4. Compare the timescales of these changes using one of the turbulence closure schemes used in climate models with those from a reference enhanced-diffusion simulation to estimate equivalent parameters ϵ and $\Delta\rho_{\text{ref}}$.

This procedure translates mixing in turbulence closure schemes into equivalent enhanced vertical diffusion parameters, enabling comparison between idealized models and climate models.

6 Conclusion

This thesis set out to investigate whether the multistability observed in complex ocean models represents genuine physical processes or is instead an artifact of the modeling techniques used. This question arises from the two different explanations for multistability in the Subpolar Gyre (SPG) presented in literature: the box model with continuous horizontal direction by Bastiaansen suggests physically meaningful multistability, whereas the one-dimensional column model by Den Toom suggests spurious multistability. In this thesis an overarching model combining elements from both approaches is presented, designed to reproduce each limit case and explore the interplay of mechanisms in a two-dimensional model. This model is analyzed by answering the following subquestions:

1. How does multistability in Bastiaansen's model depend on its parameters? In particular, does the structure persist within a realistic parameter range?
The multistability in Bastiaansen's model is highly sensitive to the chosen parameters, in particular the convection threshold parameter $\Delta\rho_{\text{ref}}$. In the original results, $\Delta\rho_{\text{ref}} < 0$ allows convection to occur even under stable stratification. However, adopting the physically realistic value $\Delta\rho_{\text{ref}} = 0$, where mixing is triggered only when the upper layer becomes denser than the deep layer, eliminates all multistability. This shows that the observed structure is not robust under realistic parameter values.

2. How does the multistability in Den Toom's model depend on the number of vertical layers? Specifically, does the structure converge as the vertical resolution increases?
In Den Toom's model, the multistability vanishes as the number of vertical layers increases. The spacing between bifurcation folds decreases with resolution and eventually disappears, confirming that the original multistability is a numerical artifact caused by coarse vertical discretization. However, the limited vertical resolution typically used in climate models is insufficient to eliminate these spurious bifurcations.

3. How does the bifurcation structure of the overarching model change under fresh water forcing?
When both dynamics of the models are combined in the overarching model, side forcing dominates, producing multistability similar to in Den Toom's model. Additional multistability also arises, but diminishes with higher horizontal resolution, again pointing to a resolution-related origin.

Based on these results, the main research question, **Is the multistability observed in complex ocean models a true reflection of physical processes, or an artifact of modeling techniques?**, can be answered.

The evidence strongly suggests that the multistability in the models studied in this thesis is spurious. In Bastiaansen's case, it depends on unrealistic parameters and in Den Toom's case, it is eliminated with sufficient resolution. Since climate models lack sufficient vertical resolution, this artifact is relevant in practice. In the overarching model the bifurcation structure corresponding to Den Toom's model dominates, showing extra bifurcation points due to the resolution in the horizontal direction. However, to be able to fully link this research to complex ocean models, the respective vertical mixing parameters should be extracted from these models.

These findings underline the importance of physically justified parameters and sufficient numerical resolution when interpreting multistability in climate models. Without these, bifurcation structures may reflect model artifacts rather than real climate tipping behavior.

References

[1] R. M. van Westen and M. L. Baatsen, "European temperature extremes under different amoc scenarios in the community earth system model," *Geophysical Research Letters*, vol. 52, no. 12, p. e2025GL114611, 2025.

[2] R. Seager, D. S. Battisti, J. Yin *et al.*, "Is the gulf stream responsible for europe's mild winters?" *Quarterly Journal of the Royal Meteorological Society*, vol. 128, no. 586, pp. 2563–2586, 2002.

[3] T. M. Lenton, D. I. A. Mckay, S. Loriani, J. F. Abrams, S. J. Lade, J. F. Donges, J. E. Buxton, M. Milkoreit, T. Powell, S. R. Smith *et al.*, "The global tipping points report 2023," Ph.D. dissertation, University of Exeter, 2023.

[4] P. T. Spooner, D. J. Thornalley, D. W. Oppo, A. D. Fox, S. Radionovskaya, N. L. Rose, R. Mallett, E. Cooper, and J. M. Roberts, "Exceptional 20th century ocean circulation in the northeast atlantic," *Geophysical Research Letters*, vol. 47, no. 10, p. e2020GL087577, 2020.

[5] C. Li and A. Born, "Coupled atmosphere-ice-ocean dynamics in dansgaard-oeschger events," *Quaternary Science Reviews*, vol. 203, pp. 1–20, 2019.

[6] A. Levermann and A. Born, "Bistability of the atlantic subpolar gyre in a coarse-resolution climate model," *Geophysical Research Letters*, vol. 34, no. 24, 2007.

[7] J. Lohmann, H. A. Dijkstra, M. Jochum, V. Lucarini, and P. D. Ditlevsen, "Multistability and intermediate tipping of the atlantic ocean circulation," *Science advances*, vol. 10, no. 12, p. eadi4253, 2024.

[8] D. Swingedouw, A. Bily, C. Esquerdo, L. F. Borchert, G. Sgubin, J. Mignot, and M. Menary, "On the risk of abrupt changes in the north atlantic subpolar gyre in cmip6 models," *Annals of the New York Academy of Sciences*, vol. 1504, no. 1, pp. 187–201, 2021.

[9] D. I. Armstrong McKay, A. Staal, J. F. Abrams, R. Winkelmann, B. Sakschewski, S. Loriani, I. Fetzer, S. E. Cornell, J. Rockström, and T. M. Lenton, "Exceeding 1.5 c global warming could trigger multiple climate tipping points," *Science*, vol. 377, no. 6611, p. eabn7950, 2022.

[10] D. Häfner, R. L. Jacobsen, C. Eden, M. R. B. Kristensen, M. Jochum, R. Nuterman, and B. Vinter, "Veros v0.1 – a fast and versatile ocean simulator in pure Python," *Geoscientific Model Development*, vol. 11, no. 8, pp. 3299–3312, Aug. 2018. [Online]. Available: https://gmd.copernicus.org/articles/11/3299/2018/

[11] H. Stommel, "Thermohaline convection with two stable regimes of flow," *Tellus*, vol. 13, no. 2, pp. 224–230, 1961.

[12] R. Bastiaansen, H. A. Dijkstra, and A. S. von der Heydt, "Fragmented tipping in a spatially heterogeneous world," *Environmental Research Letters*, vol. 17, no. 4, p. 045006, 2022.

[13] M. den Toom, H. A. Dijkstra, and F. W. Wubs, "Spurious multiple equilibria introduced by convective adjustment," *Ocean Modelling*, vol. 38, no. 1-2, pp. 126–137, 2011.

[14] P. Welander, "A simple heat-salt oscillator," *Dynamics of Atmospheres and Oceans*, vol. 6, no. 4, pp. 233–242, 1982.

[15] W. Weijer and H. A. Dijkstra, "A bifurcation study of the three-dimensional thermohaline ocean circulation: The double hemispheric case," 2001.

[16] L. Yu, "Global air–sea fluxes of heat, fresh water, and momentum: Energy budget closure and unanswered questions," *Annual review of marine science*, vol. 11, no. 1, pp. 227–248, 2019.

[17] J. Pedlosky, "A study of the time dependent ocean circulation," *Journal of Atmospheric Sciences*, vol. 22, no. 3, pp. 267–272, 1965.

[18] A. Born and T. F. Stocker, "Two stable equilibria of the atlantic subpolar gyre," *Journal of physical oceanography*, vol. 44, no. 1, pp. 246–264, 2014.

[19] R. S. Pickart and M. A. Spall, "Impact of labrador sea convection on the north atlantic meridional overturning circulation," *Journal of Physical Oceanography*, vol. 37, no. 9, pp. 2207–2227, 2007.

[20] V. Trouet and G. J. Van Oldenborgh, "Knmi climate explorer: a web-based research tool for high-resolution paleoclimatology," *Tree-Ring Research*, vol. 69, no. 1, pp. 3–13, 2013.

[21] H. Hersbach and et al., "The ERA5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020.

[22] R. Veltz, "BifurcationKit.jl," Jul. 2020. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02902346

[23] H. B. Keller, "Lectures on numerical methods in bifurcation problems," *Applied Mathematics*, vol. 217, p. 50, 1987.

[24] Delft High Performance Computing Centre (DHPC), "DelftBlue Supercomputer (Phase 2)," https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2, 2024.

[25] J. Guckenheimer and Y. A. Kuznetsov, "Cusp bifurcation," *Scholarpedia*, vol. 2, no. 4, p. 1852, 2007, revision #137541.

[26] Z. Aldo, "Amoc transitions: spurious or fragmented tipping?" Ph.D. dissertation, 2024.

[27] C. A. Katsman, M. A. Spall, and R. S. Pickart, "Boundary current eddies and their role in the restratification of the labrador sea," *Journal of Physical Oceanography*, vol. 34, no. 9, pp. 1967–1983, 2004.

[28] A. Neff, A. Keane, H. A. Dijkstra, and B. Krauskopf, "Bifurcation analysis of a north atlantic ocean box model with two deep-water formation sites," *Physica D: Nonlinear Phenomena*, vol. 456, p. 133907, 2023.

[29] G. Madec and the NEMO System Team, *NEMO Ocean Engine Reference Manual*, 2024. [Online]. Available: https://doi.org/10.5281/zenodo.1464816

[30] nbruegge, "pyvmix," accessed August 26, 2025. [Online]. Available: https://github.com/nbruegge/pyvmix

[31] G. L. Wagner, "Oceanturb.jl," accessed August 26, 2025. [Online]. Available: https://glwagner.github.io/OceanTurb.jl/latest/

[32] M. Vellinga, "Multiple equilibria in ocean models as a side effect of convective adjustment," *Journal of physical oceanography*, vol. 28, no. 4, pp. 621–633, 1998.

A Appendix

A.1 Implicit Rescaling by Den Toom

Studying the side forcing by Den Toom more closely, it becomes apparent that an implicit rescaling is applied to the equations. The temperature side forcing is given by

$$\tilde{T}(z) = \cos(2\pi z). \tag{43}$$

Translating this forcing back into dimensional variables using the non-dimensionalization

$$T = \alpha_T(T_* - \hat{T}) \tag{44}$$

results in unphysical temperature ranges on the order of -10^4 to 10^4 °C. This indicates that Den Toom applied a rescaling, although this was not explicitly stated:

$$T = \tilde{\alpha}\alpha_T(T_* - \hat{T}). \tag{45}$$

A more realistic temperature variation would lie between approximately 10 and 20 °C, implying a rescaling factor of about $\tilde{\alpha} = 2000$.

Because the model equations are linear aside from the convection adjustment function, which resembles a step function, this rescaling can be absorbed into the nonlinearity parameter ϵ without affecting the model dynamics. This is shown in the following derivation.

The original equation for the temperature was:

$$\frac{\partial T}{\partial t} = \frac{1}{P}\frac{\partial}{\partial z}\left(\left[1 + F_0\left(\frac{1}{2}(1 + \tanh[\epsilon\frac{\partial \rho}{\partial z}])\right)\right]\frac{\partial \rho}{\partial z}\right) - \left(i_+^{\text{res}}T - \tilde{T}(z)\right). \tag{46}$$

Rescaling the temperature (and also salinity and thus density) with $\frac{1}{\tilde{\alpha}}$ gives:

$$\frac{1}{\tilde{\alpha}}\frac{\partial T}{\partial t} = \frac{1}{P}\frac{\partial}{\partial z}\left(\left[1 + F_0\left(\frac{1}{2}(1 + \tanh[\epsilon\frac{1}{\tilde{\alpha}}\frac{\partial \rho}{\partial z}])\right)\right]\frac{1}{\tilde{\alpha}}\frac{\partial T}{\partial z}\right) - \left(i_+^{\text{res}}\frac{1}{\tilde{\alpha}}T - \frac{1}{\tilde{\alpha}}\tilde{T}(z)\right). \tag{47}$$

By multiplying this equation with $\tilde{\alpha}$, the original equation can be recovered, except for the nonlinear term $\frac{1}{2}(1 + \tanh[\epsilon\frac{1}{\tilde{\alpha}}\frac{\partial \rho}{\partial z}])$. However, ϵ is a modeling parameter determining the steepness of the step function, and is therefore up for change. Replacing ϵ by $\tilde{\epsilon} = \tilde{\alpha}\epsilon$ makes sure that the nonlinear term is also the same as before.

Thus, considering this extra rescaling factor $\tilde{\alpha} = 2000$ the side forcing temperature $\tilde{T}(z) = \cos(2\pi z)$ models a realistic temperature range of 10 and 20 °C, with the modeling equations staying intact. However, this rescaling does effect the parameter ϵ modeling the steepness of the step function.

A.2 Difficulties Restart Method

Using the restart method as described in Section 3.4 three kinds of events occur where a manual restart is needed. These three events are displayed in Figure 35 and described below.

1. **Going around the corner.** If a segment of the bifurcation diagram goes around the corner, because of the corner the automatic determined step size might be in the wrong direction. Then the bifurcation diagram will start backtracking.

2. **Unstable part.** The bifurcation diagram is not completely numerically stable. It oscillates and therefore it can be that the automatic determined step size might be in the wrong direction. To have this event occur less it is better to choose more steps in between which the direction is determined, however note that this will increase the probability of event 1 to occur.

3. **Backtracking.** Backtracking can also occur within a segment itself. Then the continuation probably gets to close to a zero eigenvalue for it to be able to handle it. This has only occurred when using $\mathcal{F}$ which contains a step function, probably leading to this numeric instability. This can be overcome by restarting from a slightly different solution, approaching the bifurcation point differently.

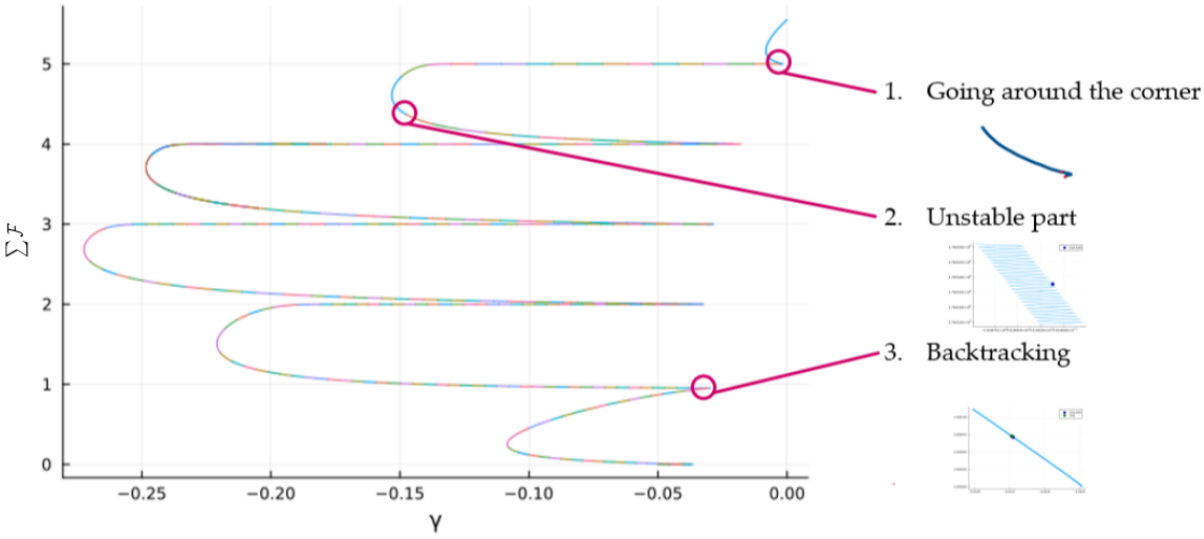


Figure 35: Difficulties in constructing a bifurcation diagram. An example of a bifurcation diagram is shown, in which all colored pieces represent a segment. Furthermore, typical error locations are indicated including a visualization of the error.

A.3 Extra Branch Bastiaansen

In Figure 36, the same as Figure 20, it can be seen that apart from the structure found by Bastiaansen in turquoise, another branch is also present in the system. In this section this extra branch is studied in more detail.

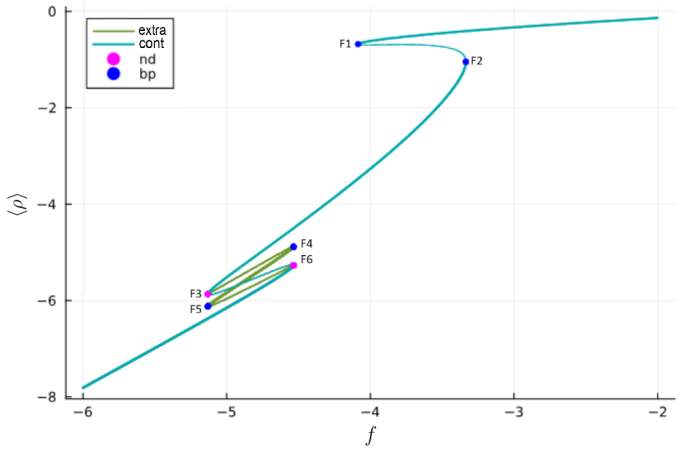


Figure 36: Bifurcation diagram using the function bifurcationdiagram from BifurcationKit. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f. The turquoise line shows the general continuation of the branch. The green line shows the extra fold. The dots represent bifurcation points. In *bp* one eigenvalue crosses zero, in *nd* two. Stable solutions are displayed by a bold line, unstable by a thinner line.

To show the difference in equilibrium solutions between the different branches, the equilibrium solutions are plotted at $f = -4.7$ (lower left fold) and $f = -3.7$ (upper right fold). At $f = -4.7$ separate plots are made for the solutions corresponding to the general continuation and the extra fold. The difference between the solutions is that those corresponding to the general continuation are symmetrical, whereas the ones corresponding to the extra fold are asymmetrical. Note that the mirrored asymmetrical solutions also exist.

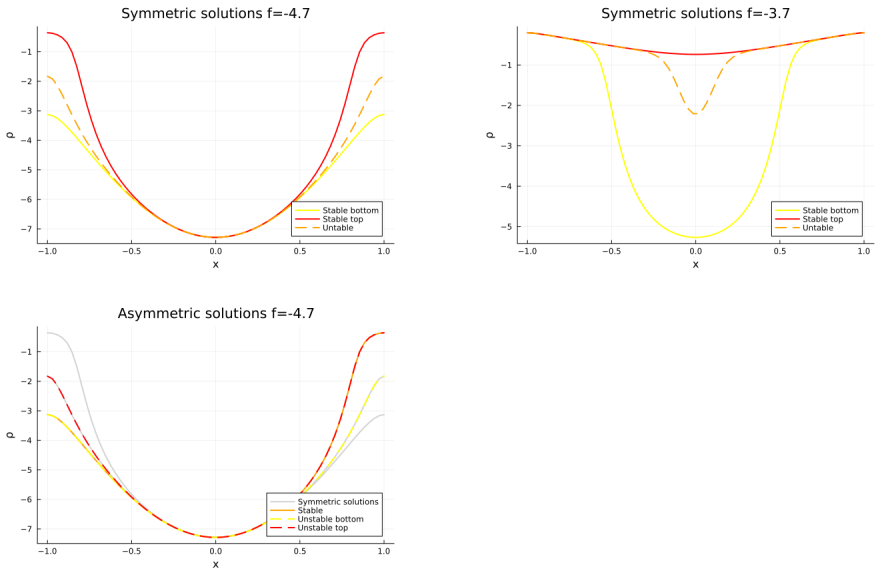
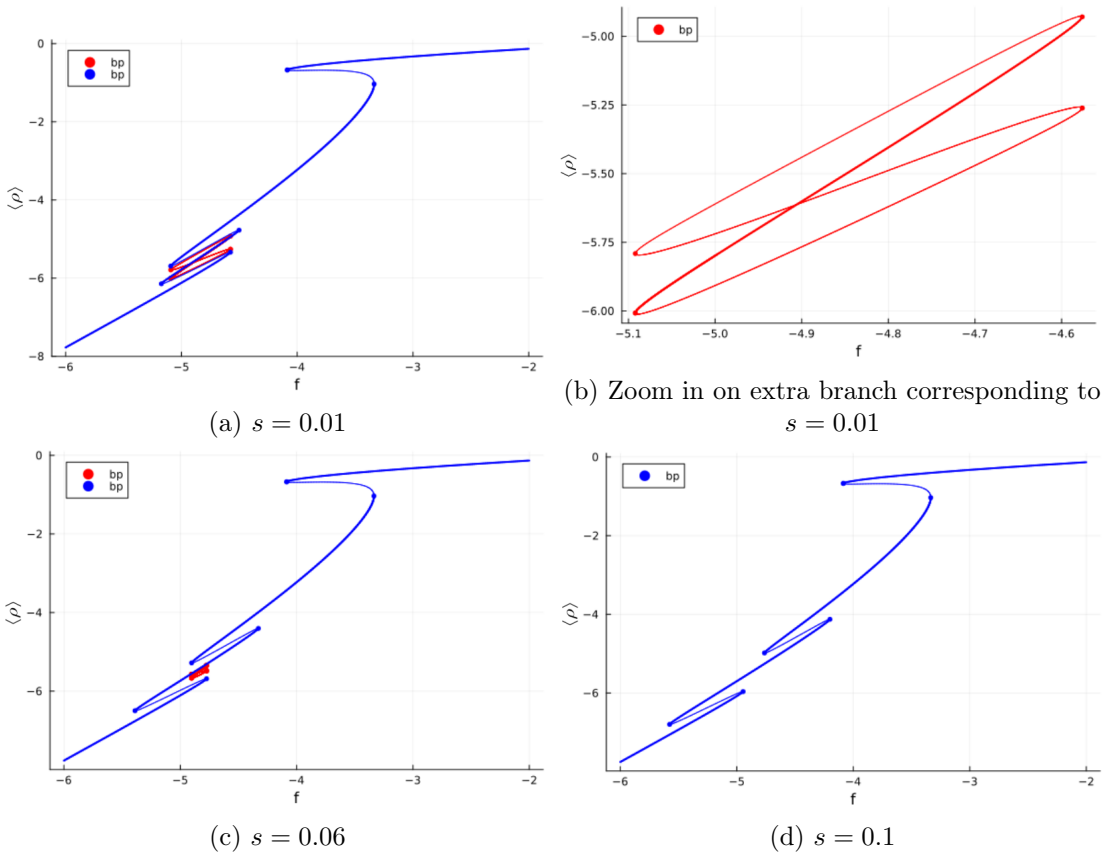


Figure 37: Equilibrium solutions at folds for $f = -4.7$ and $f = -3.7$, plotted as the density profile ρ over the horizontal domain. Yellow, orange and red lines represent bottom, middle and upper solutions. Dashed lines represent unstable solutions, full lines stable solutions.

This phenomena occurs due to the symmetry in the atmospheric density. Due to this symmetry, convection in combination with diffusion can either impact the left or right side of the steady state solution (extra fold), or both at the same time (general continuation).

This is a special case and can be altered by creating an asymmetry in the atmospheric density. In order to achieve that $\Delta\rho_A = 2 + f(1 + \cos[\frac{\pi x}{2} + s])$ is used. Here s shifts the cosine creating an asymmetric density. The bifurcation diagrams corresponding to $s = 0.01$, $s = 0.06$ and $s = 0.1$ are displayed in Figure 38. The general bifurcation structure is plotted in blue. The previous double fold connected to the extra branch gets separated in two double folds. This happens as the left and right side of the cosine are decoupled. They both get impacted by convection in combination with diffusion at a different threshold. The bigger the shift the more these double folds get separated, until they do not even overlap anymore.



(a) $s = 0.01$

(b) Zoom in on extra branch corresponding to $s = 0.01$

(c) $s = 0.06$

(d) $s = 0.1$

Figure 38: Bifurcation diagrams for different values of s using $\Delta\rho_A = 2 + f(1 + \cos[\frac{\pi x}{2} + s])$, including the extra branch (in red) if existing. On the vertical axis $\langle\rho\rangle$, the mean density of the domain. On the horizontal axis the atmospheric freshwater parameter f.

At the limit of $s = 0$, the folds are exactly on top of each other, and an extra solution connected to the branch appears, which is the unstable solution in the general continuation. The extra branch found at $s = 0$ cannot suddenly appear, thus a similar solution should also be present in the case with $s \neq 0$. By starting with a solution at the extra branch at $s = 0$ and taking this as a starting point for a continuation with $s = 0.00015$ the extra branch can be found. This process can be repeated: increasing s with very small steps (0.0002 up to 0.001) and finding the extra branch each step.

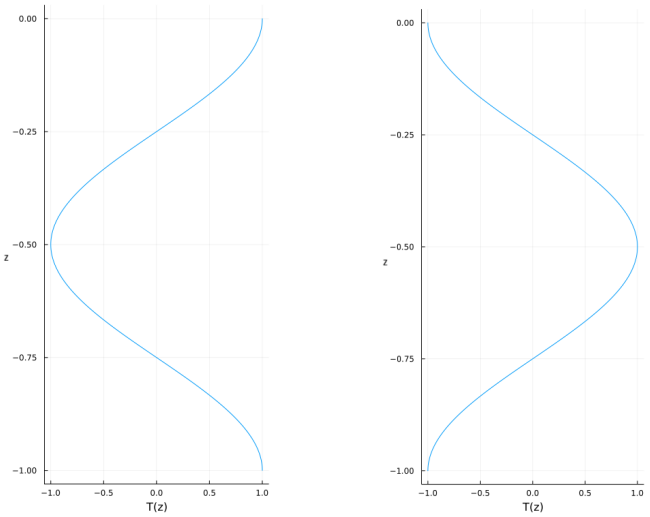
In Figure 38, the continuation (blue) with extra branch (red) is plotted for several values of s. It can be seen that the extra branch has a butterfly shape with bifurcation points on the tips. The width of this extra branch depends on the overlap between the two folds created by the asymmetry. With increasing s the branch decreases and eventually disappears.

A similar bifurcation structure due to symmetry has also been found by Neff et al [28]. Here, a simple box model with two polar boxes with deep-water formation is studied. When asymmetry is introduced in freshwater and thermal forcing, 'partial shutdown' states are possible. This supports the findings that spacial heterogeneity can give rise to multistability.

A.4 Side Forcing: $\tilde{T}(z)$

In this section the temperature to which the ocean is relaxed from the side is further investigated. In the model by Den Toom the temperature to which the ocean is relaxed is $\tilde{T}(z) = \cos(2\pi z)$. The explanation given for this function is: "which is such that it tends to stabilize (destabilize) the stratification above (below) $z = 0.5$" [13]. This is copied from Vellinga [32]. It does not seem to be based on physical reasoning.

Plotting this temperature forcing it can be seen that the profile shows warm water on the top, cold in the middle and warm at the bottom (figure 39a). However, in the Subpolar Gyre, it would be expected to have colder water at the top due to atmospheric cooling, warmer in the middle due to the AMOC and cold at the bottom. Thus exactly a reverse of $\tilde{T}(z)$: $\tilde{T}_{rev}(z) = -\cos(2\pi z)$, see figure 39b.

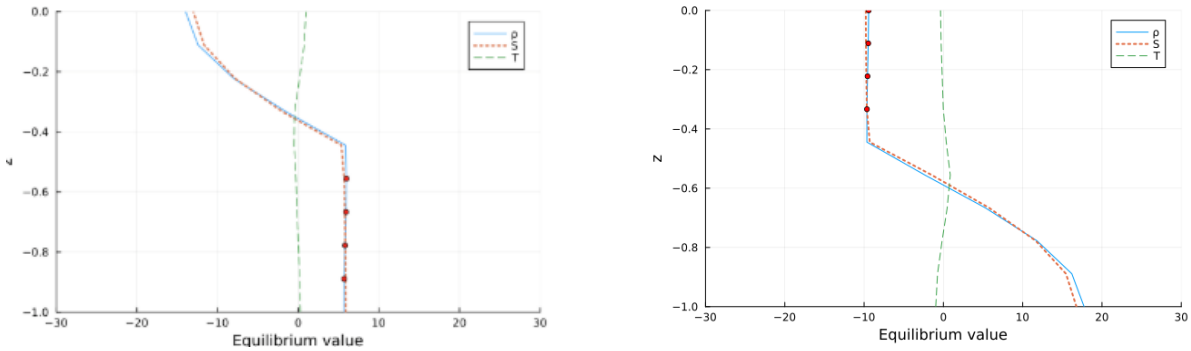


(a) $\tilde{T} = \cos(2\pi z)$: warm-cold-warm configuration.

(b) $\tilde{T}_{rev} = -\cos(2\pi z)$: cold-warm-cold configuration.

Figure 39: Plots of the temperature with which the ocean is relaxed from the side. Temperature on the x-axis and the z-domain on the y-axis.

The impact of these different side forcings is shown in the figure below. With $\tilde{T}$ convection would find place at the bottom, with $\tilde{T}_{rev}$ on the top. Thinking of mixed layer depth, the solution found using $\tilde{T}_{rev}$ is more in line with observations.



(a) $\tilde{T}$ (b) $\tilde{T}_{rev}$

Figure 40: Equilibrium solutions corresponding to $\tilde{T}$ and $\tilde{T}_{rev}$ for $\gamma \approx -0.205$. The profiles of ρ, T and S are plotted over the z-domain. Red dots indicate static instability (top layer denser than the layer underneath).

Due to symmetry of the side forcings of salinity and temperature, the bifurcation structure corresponding to Den Toom remains the same, as can be seen in figure 41. The orange line corresponds to $\tilde{T}_{rev}$, the green one to $\tilde{T}$. Thus for the bifurcation analysis of Den Toom's model this would not change anything.

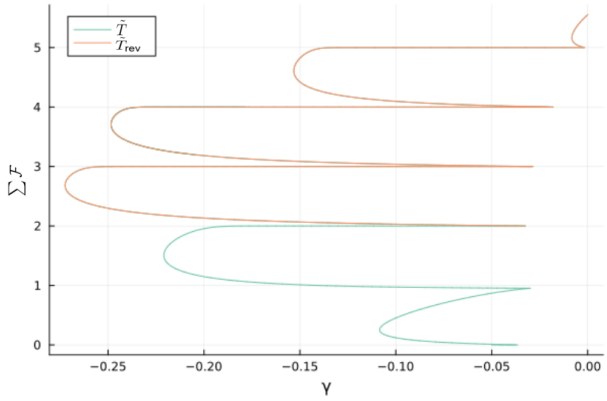
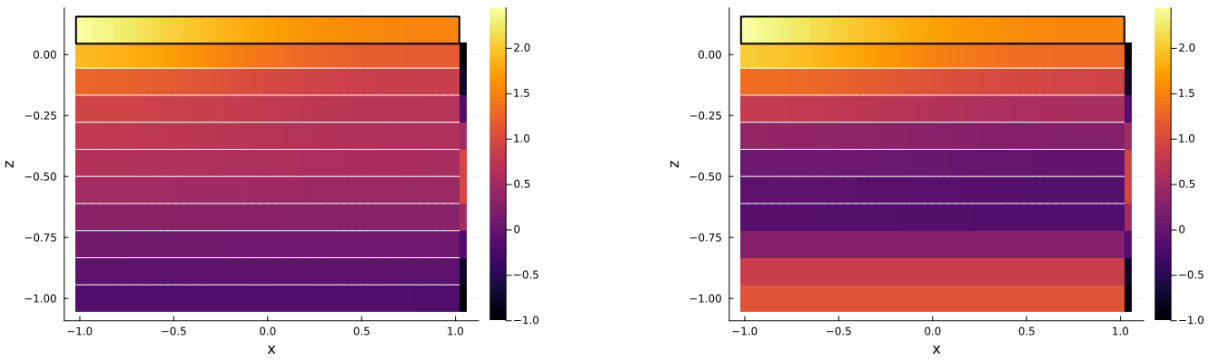


Figure 41: Comparison of bifurcation structures for different temperature profiles. On the y-axis $\sum \mathcal{F}$ indicating the intensity of convection. On the x-axis the freshwater parameter from the side, γ.

However, for the analysis of the two-dimensional overarching model this does make a great difference. The different structures are displayed below. Now the atmospheric forcing will enhance or counter the side forcing and therefore effect the convection in a different way.

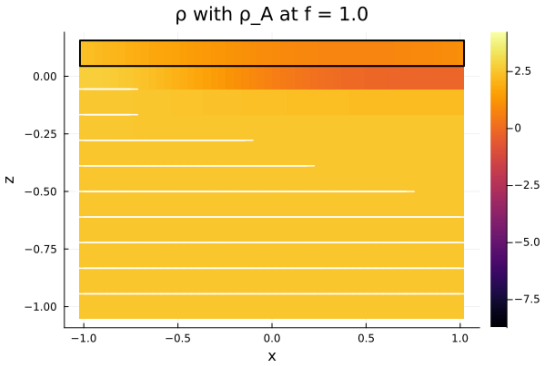


(a) $\tilde{T}$ (b) $\tilde{T}_{rev}$

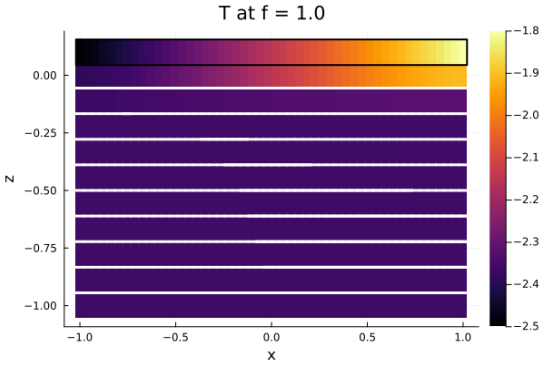
Figure 42: Solutions corresponding to $\tilde{T}$ and $\tilde{T}_{rev}$ for $\gamma = 0$ and $f \approx 0.34$. The value of ρ is displayed on the (x, z)-domain as a color plot. White lines between vertical layers indicate static instability (top layer denser than the layer underneath). The upper layer in the black box represents the atmospheric forcing, the column in the black box represents the side forcing.

A.5 Dynamics under Atmospheric Forcing

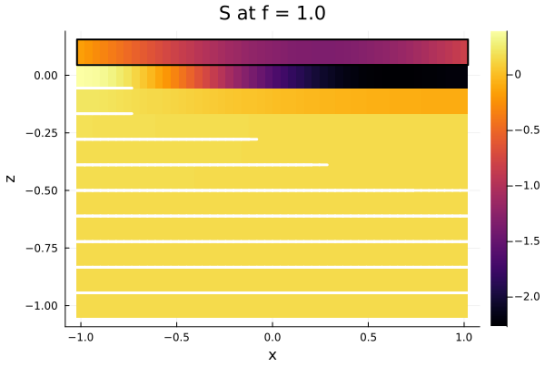
In this appendix the density profile shown in Figure 43a, the same as Figure 28b, is studied in more detail. To determine which variable induces the static instability structure visualized by the white lines in Figure 43a, the temperature and salinity profiles are separately plotted in Figure 43. In Figure 43b the white lines indicate where the upper layer has a lower temperature than the layer underneath. In Figure 43c the white lines indicate where the upper layer has a higher salinity than the layer underneath. Since the density is related to the temperature and salinity as $\rho = S - T$ (Equation (6)), these white lines give an indication of the contribution of temperature and salinity to the static instability. The pattern in Figure 43c highly resembles the pattern in Figure 43a, indicating that salinity induces the static stability in this solution.



(a) Density field. White lines between vertical layers indicate static instability (top layer denser than the layer underneath).



(b) Temperature field. White lines between vertical layers indicate a top layer with lower temperature than the layer underneath.



(c) Salinity field. White lines between vertical layers indicate a top layer with higher salinity than the layer underneath.

Figure 43: Visualizations of the solution for freshwater parameter $f = 1.0$. The value of ρ, T and S are displayed on the (x, z)-domain as a color plot. The upper layer in the black box represents the atmospheric forcing.