On the edge of Regional Climate Models

Evaluation of interpolation methods and boundary data interval resolution for the Lateral Boundary Conditions of Regional Climate Models

by

M.E. Kootte

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday September 8, 2017 at 10:00 AM.

Student number: 4129474
Project duration: February 5, 2017 – September 8, 2017
Thesis examinor: Dr. M. Hanke, KTH Stockholm, supervisor

TU Delft

KTH Engineering Sciences

On the edge of Regional Climate Models

Evaluation of interpolation methods and boundary data interval resolution for the Lateral Boundary Conditions of Regional Climate Models

MARIA KOOTTE

Master's Thesis at School of Engineering Sciences
Advisor: M. Kupiainen
Examiner: M. Hanke

TRITA xxx yyyy-nn

Abstract

A Regional Climate Model (RCM) is a comprehensive tool to simulate high-resolution climatic factors. A RCM is driven by low resolution data obtained from a Global Climate Model (GCM). In the one-way nested method, is the GCM data fed into the RCM as a Lateral Boundary Condition (LBC) in certain updates in time, the boundary data interval resolution. The necessary information in between these updates is obtained by using linear interpolation techniques. The ability to reproduce high-resolution RCM output with low-resolution GCM data depends on the accuracy of this LBC. This thesis investigates whether third order interpolation methods lead to a more accurate approximation than the linear method. This is investigated in combination with lowering the boundary data interval resolution.

The conclusion is that a third order interpolation method does not lead to a more accurate approximation for a high boundary data interval resolution. But when the resolution is lowered, the linear interpolation method looses its accuracy earlier than the third order method. This results in a boundary data interval resolution of 1.5 hours for the linear method compared to 7.5 hours for the third order method. Implementing a lower boundary data interval resolution in combination with the third order method lead to significant gain in computational time and storage for the RCMs.

Referat

Utvärdering av interpolationsmetoder och tidsupplösning av randdata för de laterala randvillkoren för regionala klimatmodeller

En Regional Climate Modell (RCM) är ett avancerat verktyg för att simulera klimatet med hög upplösning. En RCM drivs av data från en Global Climate Modell (GCM), ofta med låg upplösning. I den så kallade envägsnästade metoden sparas GCM-data med vissa tidsintervall för att användas som ett lateralt randvillkor (LBC) till RCM. Den nödvändiga informationen mellan sparade data erhålls med användning av linjära interpoleringstekniker. Möjligheten att reproducera ett hög upplöst GCM-resultat med en RCM med liknande upplösning genom att använda envägsnästning med låg upplösning beror på exaktheten hos denna LBC. Denna avhandling undersöker huruvida tredje ordningens interpoleringsmetoder leder till en mer exakt inställning än den linjära metoden. Detta undersöks i kombination med minskningen av intervallupplösningen av randdata.

Slutsatsen är att en tredje ordningens interpolationsmetod inte leder till en mer exakt approximation för en hög dataintervallupplösning. Men när upplösningen sänks, förlorar den linjära interpolationsmetoden sin noggrannhet, tidigare än den tredje ordningens metod. Detta resulterar i en begränsad dataintervallupplösning på respektive $1,5$ timmar respektive $7,5$ timmar. Tredje orderinterpolering är således en signifikant vinst för RCM.

Acknowledgements

Exactly seven years ago my journey at the TU Delft began, which brought me to where I am today. I am blessed so many inspiring opportunities were given to me along the way. I have had the chance to see some real dutch pride at the dredgers in Singapore and the chance to stimulate the entrepreneurial environment of the TU Delft, especially the 'conservative' civil engineering sector. Furthermore, I was given the opportunity to enjoy one of the most hilarious months in my life by organizing the students introduction week and as a finishing touch, I got the chance to study in Stockholm. A city that has conquered my heart and made me realize that you have to overcome several ups and downs before you can proudly finish in a filled Olympic Stadion. I am more than grateful for all these moments and people who believed in me. Now I am almost graduated as a Mathematical Engineer, specialized in Computer Simulations. Could you imagine?

I certainly had no idea when I found my way through University College in Middelburg, along Civil Engineering in Delft, to finally graduate in the beautiful city of Stockholm. The past years have been a true adventure. Lots of memories were created and true friendships have emerged. I would like to take some time to thank some of the people who really helped me grow the past few years.

First of all, I would like to thank my supervisor Michael Hanke. My graduation year in Stockholm would not have been successful without your help. I know I have not been the most regular mathematics student as I kept on switching courses, not following the common computational track. I want to thank you for your guidance, patience and mathematical direction you have given me.
During my master's, I was convinced to use my numerical knowledge for a 'right' purpose, which I found at the Swedish Meteorological and Hydrological Institute. I want to thank Marco Kupiainen for this opportunity. Your kindness, 'lectures' and most of all open attitude made me feel very comfortable from day one. Your guidance has been very inspiring and made me enjoy my graduation internship a lot.

Another person who has been very important for me last year is Patricia, a fellow student. I am sure I would not have succeeded my year without you. You have been there as a sparring partner during my thesis, you gave me the final push to ask for graduation opportunities and you were the one I could share my struggles with. But most of all you have been a true friend, with who I have created memorable moments until the last week in Stockholm.
Without the help and perseverance of Martin van Gijzen, I could never have started my master in the first place. I would like to thank you from the bottom of my heart

for your help and your trust in me, which helped me to regain confidence.

There are some other people who have inspired me during my way. First of all Paul Reijn, who already knew I would be at a better place in the environment of Delft and on top of that made sure that I could blend in immediately. Next, Paulien, my inspirator, who I could always call for advice and who literally and figuratively had to wake me up on some crucial moments. My Topgroep, with whom I have enjoyed so many hilarious moments on Friday afternoons when we had to make concrete for a civil engineering assignment. I would like to thank Eva and Gigi for their faith in me, all my dear friends for all the beautiful moments we have shared, and lastly my dearest family. Luuc, your attitude of 'don't worry baby, everything's gonna be all-right' has been annoying but probably because I knew you were right. Thank you for your unconditional support. Bob, thanks for your sharp tongue and preparation for my professional life. Bart and Luc, for being fantastic hosts, cooks and being my cute little brothers. Dear mom and dad, thank you for all the support recent years, you have given me the freedom and the means to make me who I am today.

Marieke Kootte
Breda, 21-08-2017

Contents

Acronyms

BB Big Brother. 6, 33–36, 38–40, 42–44, 57, 58, 61, 66

BB-LBE Big and Little Brother experiment. 33–35, 40, 43

GCM Global Climate Model. 3–10, 23, 24, 26, 27, 42, 43

GRF Grid Refinement Factor. 38, 39

LB Little Brother. 33–40, 43, 44, 47, 56–58, 61, 63, 66, 70

LBC Lateral Boundary Condition. 3–5, 8–10, 23–25, 33, 34, 36, 43, 44, 55, 56, 61, 63

RCM Regional Climate Model. 3–6, 8–10, 23–25, 33, 42, 43

SBP-SAT Summation-by-Parts–Simultaneous-Approximation-Term. 9, 19, 20, 22, 35, 37, 56

SRF Spatial Resolution Factor. 10, 39, 42, 43, 47, 70, 86

Glossary

p-value The rate of convergence of the L_2-norm of the error of the numerical approximation with $SBP - SAT$ operators. 22, 23, 34, 35, 37–39, 47, 48, 56–59, 86, 99

boundary data interval resolution The frequency how often the lateral boundary condition is updated.. i, 4–6, 10, 30, 34, 36, 38–40, 42, 47–49, 55–60, 69, 99

design accuracy The convergence rate according to the SBP-SAT operators . 22, 23, 34, 35, 37, 38, 48

grid refinement factor A factor of size 2 between the spatial resolution of two Big Brother or two Little Brother models. The refinement is necessary to compare the output of one LB model to its refined equivalent . 39

lateral boundary condition Boundary condition of the regional climate model. 3–5, 11, 25, 47, 61, 70

one-way nested A technique in climate modeling to couple the regional model to the global climate model. 3, 4, 9, 33

spatial resolution difference A factor that declares how many times the spatial resolution of the LB model is refined in relation to the BB model. The factor differs between 1 and 10 . 9, 10, 39, 57, 60, 99

Nomenclature

α	Wave speed
α^*	Non-dimensional parameter describing the flow properties
$\bar{H}$	Mean depth of the fluid
ΔT	Temporal stepsize of the BB model
Δt	Temporal stepsize of the LB model
κ	Thermal conductivity
λ	Wavelength
$\mathbb{S}_i$	Penalty term introduced by the SAT method
ω	Wave frequency
ρ	Specific mass of a fluid
τ	Shear stress tensor
$\vec{f_e}$	External forces
$\vec{U}$	Vector representing the flow velocity, energy and density of a fluid in the Navier-Stokes equations
$\vec{v}$	Vector that represents the flow velocity in three dimension
E	Total energy per unit mass
f_o	Coriolis parameter
g	Gravitational acceleration
H	Spatial gridsize of the BB model
h	Spatial gridsize of the LB model
H^*	Spatial gridsize of the GCM model
h^*	Spatial gridsize of the RCM model
H_e	Total enthalpy
I	Unit tensor

J	Jacobian
k	Wavenumber
k_z	Zonal wavenumber
L	Earth's perimeter
L_1	Horizontal length of the BB domain
L_1^*	Horizontal length of the GCM domain
L_2	Length of the LB domain
L_2^*	Length of the RCM domain
M	Number of spatial gridpoints of the LB model
M_T	Number of temporal steps of the LB model
N	Number of spatial gridpoints of the BB model
N_T	Number of temporal steps of the BB model
p	Pressure
Q	Vector representing the energy sources and external forces in the Navier-Stokes equations
q_H	Energy sources
R	Earth's radius
R_o	Rossby radius of deformation
T	Simulation time of the BB-LB experiment
$Temp$	Temperature
U	Physical quantity of the flow velocity
W_f	Work of the external forces
x_1	Left boundary point of the BB model, $x_1 = 0$
x_b	The BB gridpoint that corresponds with the boundary gridpoint of the LB model
x_N	End of the BB domain. $x_N = 1$
z_1	Left boundary point of the LB model
z_M	End of the LB domain

"Let us introduce the refinement and rigor of mathematics into all sciences as far as this is at all possible, not in the faith that this will lead us to know things but in order to determine our human relation to things. Mathematics is merely the means for general and ultimate knowledge of man."

The Gay Science, Friedrich Nietzsche, 1882

Chapter 1

Introduction

1.1 Global Climate Models

Climate models are numerical models that simulate physical and atmospheric processes around the globe in order to understand the current state of global climate and climate change. By applying the basic natural laws, conservation of mass, momentum and energy, researchers are able to use these models to forecast possible future climate scenarios.
Climate models have been developed since 1956 [20]. Early climate models simulated atmospheric processes which resulted in long-period numerical forecasts of the global climate. It was not until the late 1960's that the first general circulation, or as it is now called Global Climate Model (GCM), was developed [28]. This GCM was the first model to combine the interaction between the ocean and the atmosphere.
As of today, global climate models are coupled systems of the models that define the interaction between the atmosphere, ocean, land surface, sea ice and dynamical vegetation. They are applied to predict how shifts in climatic variables, such as temperature, ocean and atmospheric current, impact the climate. The Earth's climate is a highly complex system, still climate models are significantly helpful in predicting changes in a span of a century [15].

1.2 Regional Climate Models with one-sided nesting

Regional Climate Model (RCM)s were introduced to simulate atmospheric processes with higher resolution than the global models on a regional computational domain [5] [10]. By doing so, more detailed processes could be generated on a regional scale while not requiring as high computational costs as GCMs. The RCM is embedded in the GCM with a resolution up to 10 times finer than the GCM [5]. There are several ways to embed the RCM into the GCM. One popular technique is the one-way nested method [11]. The output of the GCM is used to drive the lateral boundary conditions (LBC) of the RCM [5]. Lateral boundary conditions are prescribed conditions on the most outer part of the domain, the laterals. The prescribed conditions are time-varying values of the dependent variables in the model, which are prescribed by the data obtained from one period simulation over the GCM. The output of the RCM is not fed back into the the GCM, this would be done in *two-way* nesting. The development of RCMs continued due to a growing interest in regional and high-resolution effects of climate change and computational limitations of the coarse GCMs [9]. To produce high-resolution data

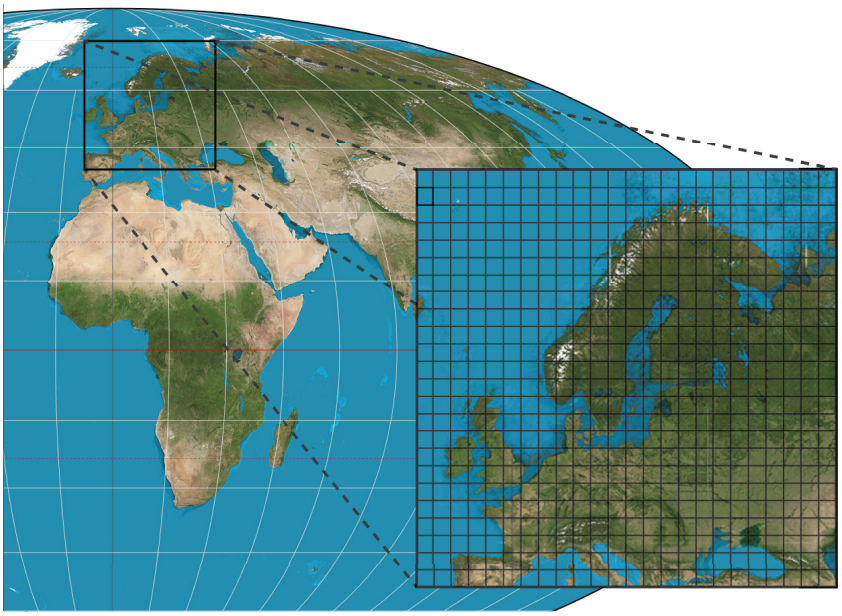

Figure 1.1: An insight in how the regional model is related to the global model. The regional domain of interest is embedded in the global domain and has a higher resolution. The figure is an adapted version of two pictures [24] [25] from Wikimedia Commons, adapted with permission.

with GCMs would be computationally too expensive. The advantage of the one-way nested technique is that it is a relatively simple method, which produces both basic climate features as well as detailed regional phenomena [5]. There are also some issues involved within this technique. These issues concern physical and numerical aspects [3][11].

1.3 Introduction of the problem

The one-way nested Regional Climate Model is the current standard for climate modeling. Nevertheless, this method is susceptible to several issues [11]. These issues are mentioned in section 2.2. Several reasons can be attributed to the emergence of these issues. One of the issues the climate researchers are acquainted with is how often the lateral boundary condition (LBC) is updated. Normally, the LBC are stored in certain intervals in time, but are required at each time step of the regional model. Interpolation in time is needed to create the LBC from the GCM output. Currently, a storage resolution around 3 and 12 hours is used [11]. The storage resolution or boundary data interval resolution is not based on extensive research but more on common sense among climate researchers, who have agreed on a certain rule of thumb. This rule prescribes that the update resolution depends on physical properties of the waves as described by the model [3]. The purpose of this work is to investigate the update of LBC intervals on computational grounds instead of physical grounds and analyze the impact it could have.
The choice for interpolation method (linear) in time for the boundary data is also not well motivated by current research. Therefore, other interpolation methods are scrutinized to better motivate the adequate choice for the LBCs.

1.4 Scope of project

The scope of this thesis is to investigate different boundary data interval resolution sizes and other interpolation methods for the LBC, by evaluating the performance of the RCMs. The boundary data interval resolution describes how often the LBC is updated in time by the output of the global model. A boundary data interval resolution of x hours means that the LBC is updated every x hours by the GCM output. The RCM timesteps between these two updates require interpolation. The objective is to see how these changes impact the lateral boundary condition, while making computational experiments varying the interpolation method and the boundary data interval resolution. Higher order interpolation methods will be investigated in combination with higher or lower boundary data resolution. Verification based on computational grounds lead to a certain compromise between the resolution and interpolation method resulting in better accuracy of the regional model. Research will be performed with a simplified model that still reflects the required properties of the full model. The simplified model will be more accessible to work with and test the different methods concerning the LBC. The simplified model gives less complex insight in the properties of the model and shows the effects of the change in method in a more straight forward matter. The model will be numerically and physically similar to the RCM and GCM, which should make the experiments scientifically qualified. If the outcome of these experiments shows improvement to the results, it can be used in the full climate models. The following research question reflects the focus point of this thesis:

What is the best compromise between boundary data interval resolution and interpolation method that leads to the desired result for the regional model output?

The question has the following research objectives that should be investigated:

- How can the model be simplified in order to investigate the involved issues while in the meantime being simple enough to work with?
- What are the physical conditions that should be captured by the full model?
- How can the argumentation of the desired result be based on computational properties instead of physical properties?
- How does altering boundary data resolution of the model (time interval of update) impact the final solution of the regional model?
- Until what extend does changing interpolation techniques impact the solution of the regional model?
- What is the combination of boundary data interval resolution and interpolation method that leads to the desired result?

1.5 Methodology

To investigate how the choice of interpolation method and boundary data interval resolution impact the output of the regional model, a modeling study is performed which will be equivalent to the set-up of the common global-regional climate model system.
An investigation on how the regional and global model interact with each other in current models has to be performed (section 2.1), before this model can be build. The current regional and climate model set-up will be discussed extensively, as well as the issues involved within this set-up (section 2.2). The full model will be simplified to get a clear insight of the boundary properties (section 3.1). This will be done in such a way that the simplified model mimics the properties of the full model, both physically and computationally. The advantage of the simplified model is that it is easier manageable and gives a transparent perception of the alterations. A study of atmospheric phenomena will be conducted (section 3.2) to reproduce the model as good as possible. Next to that, the discretization techniques used in this thesis will be explained and applied to the model problem (section 3.3). As soon as the theoretical part of the model is done, several interpolation methods will be explored and the most suitable method will be chosen (section 3.4.3).
First the model will be set up. The Big Brother experiment technique will be applied, which is a comprehensible tool to test issues involved in GCM and RCMs (section 4.1). Thereafter, the domain is determined (section 4.2) and the research focus is discussed and restricted (section 4.3). Then the arguments for the desired results are determined and explained (section 4.4) as well as the composition of several experiments (section 4.5). Parameters and variables of the model will be chosen in such a way that they are scaled versions of the full model (section 4.6).
Finally, the simulations will be run in order to draw conclusions from the results (chapter C), i.e. how the boundary procedures impact the solution. It should now be possible to draw a hypothesis for the full RCM (chapter 6) which can be investigated in the full model in subsequent research.

Chapter 2

Climate Computations

2.1 Climate model description

Climate models are three-dimensional mathematical representations of the physical process of the Earth's Climate. To simulate the Earth's climate, a Global Climate Model (GCM) is build-up by dividing the globe into a three dimensional grid (see figure 2.1) on which the basic physical equations are applied. The equations are implemented numerically and the model is run over several 'model years'. The produced results are evaluated and used for climate forecasting.

A full climate model is a coupled set of several models that describe the atmosphere, ocean, sea ice, land vegetation and biochemistry of the earth. Every model describe a certain part and use different equations and parameterizations. The physics of the

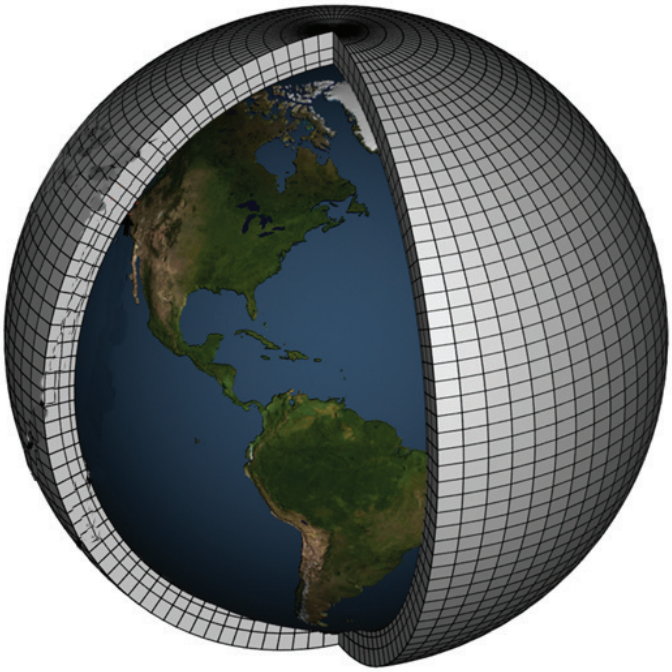

Figure 2.1: Division of the earth into a three dimensional grid. The figure displays the model domain of a global climate model. The figure is an adapted version of a figure obtained from [15].

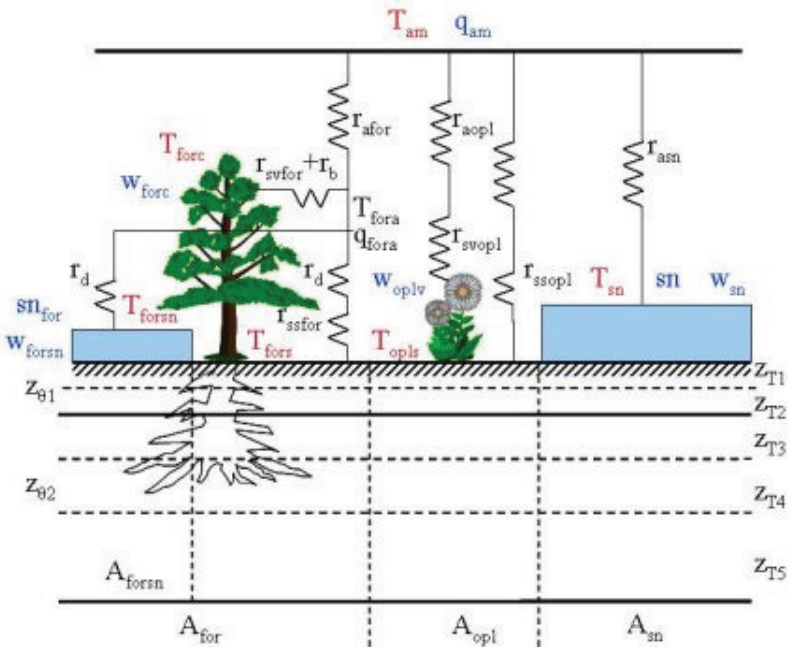

Figure 2.2: Mathematical representation of atmospheric processes. It is a schematic version of the processes to give an understanding of the physics. It is not an exact description of the model used in this thesis. The figure is property of SMHI.

coupled ocean and atmosphere models are described by the Navier-Stokes equations. A sea ice model describes the development of the sea ice. It covers the thermodynamic growth or decay of the ice, which depend on the exchanges with the atmosphere and the ocean and it covers the heat balance at the surface of the ice which tells how the ice and snow is melting. A land vegetation model simulates the water content of the soil, the soil temperature, a description of the vegetation cover and of the interactions between plants, soil and atmosphere. The biochemistry model covers the simulation of the concentration of various chemical species [1]. All these models are coupled and form together an advanced climate model.

As described in the introduction, current climate models are divided into global and regional models, where the LBC of the RCM depends on the output of the GCM. The handling of the LBC concerns mainly the aspects of the atmosphere and ocean models. The Navier-Stokes equations in these models describe the fluid as a continuum and how pressure, temperature, density and velocity of a moving fluid are related. The Navier-Stokes equations form a coupled system of equations which consist of an equation for the conservation of mass (line 1), momentum (line 2) and energy (line 3):

$$\frac{\partial}{\partial t}\underbrace{\begin{bmatrix}\rho\\ \rho\vec{v}\\ \rho E\end{bmatrix}}_{\vec{U}} + \nabla\cdot\begin{bmatrix}\rho\vec{v}\\ \rho\vec{v}\otimes\vec{v}+pI-\tau\\ \rho\vec{v}H_e-\tau\cdot\vec{v}-\vec{\nabla}Temp\end{bmatrix} = \begin{bmatrix}0\\ \rho\vec{f_e}\\ W_f+q_H\end{bmatrix}, \tag{2.1}$$

where ρ is the specific mass of the fluid, v is the flow velocity, E is the total energy per unit mass. The flow velocity $\vec{v}$ is in three dimensions, $\vec{v} = v(x, y, z)$ [**?**].
Other quantities represented by the Navier-Stokes equations are the pressure p, the

unit tensor I, the shear stress tensor τ, the total enthalpy H_e, the thermal conductivity coefficient κ, the temperature $Temp$, the external forces $\vec{f_e}$, the work of the external volume forces $W_f = \rho f_e \cdot \vec{v}$ and q_H contains the energy sources.
Both GCMs and RCMs represent the same physical system [5] and thus use the same equations.
The equations are discretized in space according the Summation-by-Parts–Simultaneous-Approximation-Term (SBP-SAT) method (the method is explained in section 3.3) and by fourth-order Runge-Kutta in time. To start the simulation, the GCM runs on a certain domain for a given simulation period. A RCM is embedded on a certain part of the global domain and requires boundary and initial data from the GCM. These data is fed into the RCM according the one-way nested method (explained in section 1.3). Thus the GCM output is stored in certain intervals in time and then interpolated to form the LBC of the RCM.

2.2 Issues within one-way nested regional climate modeling

The one-way nested technique is susceptible to several issues. They concern both physical and numerical issues. Although they are already explained in [3] and some of them already discussed there, they are once more mentioned here. The issues, as explained in [3], are as follows:

1. Numerical nesting: mathematical formulation and strategy
2. Spatial resolution difference between the driving data and the nested model
3. Spin-up
4. Update frequency of the lateral boundary conditions (LBC)
5. Physical parametrisation consistencies
6. Horizontal and vertical (spatial) interpolations errors
7. Domain size
8. Quality of the driving data
9. Climate drift or systematic errors

The research focus is based on issue 4, the update frequency of the lateral boundary condition, and combined with the investigation of several temporal interpolation methods to accomplish 'continuous' boundary conditions . Some of the other issues are briefly discussed and investigated in [3]. Issue 1, numerical nesting, concerns the numerical implementation of the Navier-Stokes equations and the way the numerical nesting technique is implemented. The numerical implementation of the Navier-Stokes equation is done by simplifying the system of equations to an equivalent workable version, the transport equation, this is explained in section 3.1 and according the same discretization techniques as the GCM and RCM implemented numerically. This similar implementation should exclude the problems due to issue 1. The implementation of the nesting strategy is motivated in section 4.1 and mimics the method from [3]. Thus the numerical nesting should not cause any problems in these experiments.
Issue 2 concerns the spatial resolution difference between the RCM and GCM, the

factor between the resolution of the GCM and the RCM (the spatial resolution difference factor (SRF)). More accurate results from the RCM are produced, when the resolution is finer than the GCM. When the SRF is too big, bridging the difference between the models cannot be realized and will result in useless output of the RCM. To find the optimal difference factor, several options of the spatial resolution difference are investigated. They will form a basis for the experiments explained in section 4.5.
Issue 3, spin-up, should not be an issue in climate simulations, as the running time is sufficiently long and the outcome is mostly determined by the boundary conditions an not by the initial conditions [5]. Furthermore, the initial condition is well known and thus the simulations are not dependent of wrong initial data, this will be come clear later in the text.
Issue 5, physical parametrisation consistency, is carefully regarded in the simplification of the Navier-Stokes equations to the transport equation, executed in section 3.1. Issue 6, horizontal and vertical interpolation errors (or shortly spatial interpolation errors as the model is one-dimensional), can lead to problems when the boundary of the regional domain is not matching a gridpoint of the global domain. Spatial interpolation is needed between two adjacent global gridpoints to form the data that updates the LBC. To investigate whether spatial interpolation influences the outcome negatively, a simulation scenario is build where both the gridpoints are matching and a scenario where the RCM boundary is not matching. This is explained in section 4.5.
Issue 7, the domain size, is already investigated in [11]. The choice for the domain size is based on this result.
Issue 8, quality of the driving data, is dealt with by validating the GCM output before implementing it as the Lateral Boundary Condition. This is evaluated in section 3.3.
Issue 9, climate drift, will not be visible in this research as the model is too simple to give rise to these errors.
The main focus of this thesis is to find the optimal relation between the boundary data interval resolution and interpolation method to obtain the most reasonable compromise.
Current models use a 'rule of thumb' to determine the boundary data interval resolution according to atmospheric properties of the model. The following is a citation from [3]: *"As a rule of thumb, the update period should be smaller than one quarter of the ratio of the length scale to the phase speed of the meteorological phenomena that we want to get correctly in the LAM domain. For instance, a typical synoptic system having a horizontal size of 1000 km and a phase speed of 50 km/h would require an updating frequency of at least 5 h."*
This means that the wavelength and phase speed of the atmospheric waves describe how often the boundary data should be updated. Scientific experiments should lead to a different motivation for the rule of thumb, based on computational properties. This shall be the outcome of the thesis.

Chapter 3

Project Description

3.1 The continuous model

The Navier-Stokes equations describe the fluid and gas flow in atmospheric and oceanic climate models. Assumptions and simplifications of the Navier-Stokes equations lead to the transport equation which is shown in the coming subsection. The transport equation has the same properties with respect to the boundary time interpolation as the full model, but it is much simpler to work with and the solution is available in closed form. Variation of parameters, to check different properties of the model with respect to the lateral boundary condition, can be easily performed.
Below it is shown how the same atmospheric characteristic waves are related to the simplified problem.

3.1.1 From Navier-Stokes to the Transport equation

The full model is represented by the full Navier-Stokes equations, which is a coupled system of equations:

$$\frac{\partial}{\partial t}\underbrace{\begin{bmatrix}\rho\\ \rho\vec{v}\\ \rho E\end{bmatrix}}_{\vec{U}} + \nabla\cdot\begin{bmatrix}\rho\vec{v}\\ \rho\vec{v}\otimes\vec{v}+pI-\tau\\ \rho\vec{v}H_e-\tau\cdot\vec{v}-\vec{\nabla}Temp\end{bmatrix} = \begin{bmatrix}0\\ \rho\vec{f_e}\\ W_f+q_H\end{bmatrix}. \tag{3.1}$$

ρ is the specific mass of the fluid, v is the flow velocity and E is the total energy per unit mass. The flow velocity $\vec{v}$ is in three dimensions, $\vec{v} = v(x, y, z)$.
Other quantities represented by the Navier-Stokes equations are the pressure p, the unit tensor I, the shear stress tensor τ, the total enthalpy H_e, the thermal conductivity coefficient κ, the temperature $Temp$, the external forces $\vec{f_e}$, the work of the external volume forces $W_f = \rho f_e \cdot \vec{v}$ and q_H contains the energy sources [**?**].
In case of non-viscous and non-heat conducting flows, the shear stresses and heat conduction terms can be neglected. This lead to the set of Euler equations. In conservation form, they read:

$$\frac{\partial\vec{U}}{\partial t} + \nabla\cdot F = Q, \tag{3.2}$$

where $\vec{U} = \begin{bmatrix} \rho \\ \rho\vec{v} \\ \rho E \end{bmatrix}$, $F = \begin{bmatrix} \rho\vec{v} \\ \rho\vec{v} \otimes \vec{v} + pI \\ \rho\vec{v}H_e \end{bmatrix}$ and $Q = \begin{bmatrix} 0 \\ \rho\vec{f_e} \\ W_f \end{bmatrix}$.

In other notation:

$$\vec{U}_t + \nabla \cdot F(\vec{U}) = Q \tag{3.3}$$

Sources and external forces are neglected in this simplification and thus $Q = 0$.
The solution of the equation depends on its initial and boundary conditions. Boundary conditions must be set normal to the boundary, this means that every boundary can be regarded separately and the equation can be regarded in one direction only. Considering the Euler equation in the x-direction only, gives enough insight of the boundary conditions. The Euler equations in one spatial dimension are:

$$u_t + F(u)_x = 0 \tag{3.4}$$

Note that $u = \begin{bmatrix} \rho \\ \rho v(x) \\ \rho E \end{bmatrix}$. Introduction of the Jacobian $J(u) = \frac{\partial F}{\partial u}$ of the flux F with respect to the quantity u, leads to the equation in quasi-linear form:

$$u_t + J(u)u_x = 0 \tag{3.5}$$

This system is completely hyperbolic since $J(U)$ has three real eigenvalues and can be diagonalized to:

$$A(u) = PJ(u)P^{-1} \tag{3.6}$$

where A is the diagonal matrix containing the eigenvalues and P formed by the eigenvectors. By substitution of $w = P^{-1}u$ the equation becomes:

$$w_t + Aw_x = 0 \tag{3.7}$$

Hyperbolity asssures that all the solutions are linearly independent. Only incoming waves have to be regarded and thus A can be simplified to a scalar α. Which lead to the final transport equation [**?**]:

$$u_t + \alpha u_x = 0 \tag{3.8}$$

The physical quantity of α is the wave speed, w has been switched to u for simplicity reasons.

The transport equation need initial and boundary conditions to get a well-posed problem: an initial boundary value problem. The well-posedness is shown in section 3.3. The solution of the one-dimensional transport equation is given in the domain $x = [0,\ L]$ (with L the length of the domain) and has to be determined by all values of time $t > 0$. The initial condition $u(x, 0)$ is imposed at time $t = 0$ along all values of x and the boundary condition $u(0, t)$ has to be imposed at all values of t. The initial condition $u(x, 0) = \phi(x)$ defines the exact solution $u(x, t)$ of the transport equation in the following way: $u(x, t) = \phi(x - \alpha t)$. An initial condition that represents the wave

characteristics of the atmospheric wave is $\phi(x) = cos(k2\pi x)$ [19] , k is the wavenumber. This leads to the following exact solution: $u(x,t) = cos(2k\pi(x - \alpha t)$. The corresponding boundary condition is $u(0,t) = cos(-k2\pi\alpha t)$.
The following set of equations represents the well-posed transport equation:

$$\begin{cases} u_t + \alpha u_x = 0 & t > 0 \quad x = [0, L] \\ u(x,0) = cos(k2\pi x) \\ u(0,t) = cos(-k2\pi\alpha t) \end{cases} \tag{3.9}$$

With exact solution:

$$u(x,t) = \cos(k2\pi(x - \alpha t)) \qquad t > 0 \quad x = [0, L] \tag{3.10}$$

3.1.2 Non-Dimensionalization

The transport equation can be non-dimensionalized, to make it easier implementable in Matlab. Hence, reference parameters are necessary for the space coordinate x, the time scale t, the velocity scale u and the wave speed scale α. The following reference parameters are introduced (do not mistaken the reference parameter U for the vector $\vec{U}$ in (3.1.1)): The following equation is the non-dimensionalized version of the transport

Table 3.1: Reference parameters

	Model variable	**Reference parameter**	**Dimensionless variable**
Velocity	u	U	$u^* = \frac{u}{U}$
Time	t	T	$t^* = \frac{t}{T}$
Length	x	L	$x^* = \frac{x}{L}$

equation

$$\frac{\partial u^*}{\partial t^*}\frac{U}{T} + \alpha\frac{\partial u^*}{\partial x^*}\frac{U}{L} = 0 \tag{3.11}$$

Multiplication of this equation by T/U leads to:

$$\frac{\partial u^*}{\partial t^*} + \frac{\alpha T}{L}\frac{\partial u^*}{\partial x^*} = 0 \tag{3.12}$$

With $\frac{\alpha T}{L} = \alpha^*$ the dimensionless number, which represents the flow properties.
To simplify the numerical implementation, the reference parameter T is chosen to be $T = \frac{L}{\alpha}$, what leads to $\alpha^* = 1$. L is chosen to be the physical length-scale on which the wave propagates, which is the earth's perimter: $L = 2\pi R$ where R is the earth's radius. Omit the $*$ and get the following dimensionless transport equation:

$$u_t + \alpha u_x = 0 \tag{3.13}$$

Where $\alpha = 1$, by the choice of the reference parameter T.

The exact solution, boundary condition and initial condition have to be non- dimensionalized as well. The non-dimensionalized version of the initial condition is:

$$u^*(x^*, 0) = \frac{1}{U}cos(2\pi kLx^*) \tag{3.14}$$

Where $kL = k_z$, the nondimensional parameter that reflects the zonal wavenumber. The zonal wavenumber is given as: $k_z = \frac{2\pi R cos(\phi)}{\lambda}$. λ is the wavelength, which is $\lambda = \frac{1}{k}$, ϕ is the latitude the wave is propagating, and can be chosen to be 0. By the choice of $L = 2\pi R$, the zonal wavenumber is correctly presented. It represents the number of wavelengths around the perimeter of the globe on a certain latitude. The dimensionless solution is then:

$$u^*(x^*, t^*) = \frac{1}{U} cos(2\pi kL(x^* - \underbrace{\alpha^*}_{=1} t^*)). \tag{3.15}$$

And after disregarding the asterisk:

$$u(x,t) \;=\; cos(2\pi k_z(x-t)) \tag{3.16}$$

The factor $\frac{1}{U}$ in front of the equation is neglected, as the amplitude won't influence the properties that have to be investigated in this research. The non-dimensionless version of the boundary condition is:

$$BC : u(0,t) = cos(-2\pi k_z t)$$

The final non-dimensional system of equations is:

$$\begin{cases} u_t + u_x = 0 & t > 0 \quad x = [0,1] \\ u(x,0) = cos(k_z 2\pi x) \\ u(0,t) = cos(-k_z 2\pi t) \end{cases} \tag{3.17}$$

With exact solution:

$$u(x,t) = cos(2\pi k_z(x-t)) \qquad t > 0 \quad x = [0,1] \tag{3.18}$$

3.2 Wave characteristics

It is important to get a clear understanding of what is going on physically. The simplification of the Navier-Stokes equations has resulted in a one dimensional transport equation, which requires the physical values of the wave speed and zonal wavenumber of waves propagating in the atmosphere to correctly describe these waves. With these wave characteristics, the simulations can be based truly on the actual climate models and are conform to the limitations and restrictions caused by the wave characteristics.

There are several types of waves propagating in the atmosphere. The first considered wave is the sound wave. This one is not included in climate models because of its high frequency [16]. There are also internal and surface gravity waves. These waves don't influence the climate and are thus not important in the climate models. The surface gravity waves can be seen as the waves in the ocean that you can observe with your eyes [16]. Internal gravity waves are similar, but instead of oscillating on the surface, they are oscillating within a fluid medium [17].
The following waves are important in climate modeling [19][18][7][8]:

- Poincaré waves
- Kelvin waves
- Rossby waves
- Topographic waves

} = Inertia waves

These four types of waves are inertia or geostrophic waves. This means that they are caused by the force produced by the rotation of the earth [2] . A Poincaré wave is a transverse wave in the ocean or atmosphere which is affected by gravity and the earth's rotation [19][4]. The wave is not acting on the surface, but in the middle of ocean/atmosphere. A Kelvin wave is also a wave that is affected by the Earth's rotation, but needs the support of a boundary [19][4]. Hence, it propagates mostly along coastlines. There are also Kelvin waves trapped along the equator. The equator acts as a boundary for these so-called Equatorial Kelvin waves. Another inertia wave is the Rossby wave. The Rossby wave is relatively slow compared to Kelvin and Poincaré waves. Rossby waves are so long that the variation of the Coriolis force with latitude are influencing them [4][18]. There are atmospheric and oceanic Rossby waves. The oceanic waves move along the boundary between the warm upper layer and the cold deeper part of the ocean. The atmospheric waves occur in high-altitude winds. Topographic waves are, just as Rossby waves, long and slowly moving waves, but affected by bottom irregularities. These waves can be treated analogously to Rossby waves [4].

The waves are modeled by the system of equations of (3.17). The characteristic zonal wave number and wave speed are the only parameters that describe which wave is regarded. Inertia waves are described by its frequency and wave speed in the dispersion relation. This relation is usually given as

$$\omega^2 = gk \tanh k\bar{H}, \tag{3.19}$$

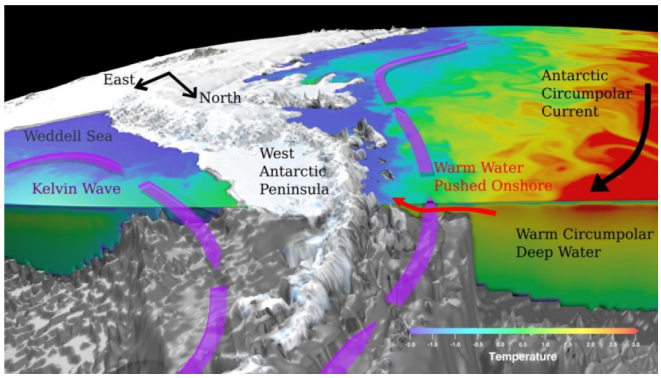

Figure 3.1: A path of the Kelvin wave interacting with Antarctic current. *Credit: Ryan Holmes / NCI [23]*

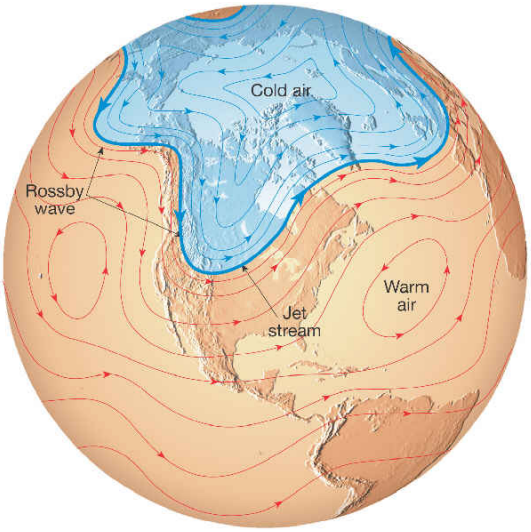

Figure 3.2: Path of a Rossby wave, acting in the Northern hemisphere. [27]

where ω is the frequency, g the gravitational acceleration, k the wavenumber and $\bar{H}$ the mean depth of the fluid. The Kelvin, Poincaré and Rossby waves are affected by the earth's rotation which modifies the dispersion relation (3.19) into the following:

$$\omega^2 = g\lambda \tanh \lambda \bar{H}, \tag{3.20}$$

where λ is related to the wavenumber k by

$$k^2 = \lambda^2 - R_o^{-2}\lambda \bar{H} \coth \lambda \bar{H}. \tag{3.21}$$

R_o is the Rossby radius of deformation, which is defined by $R_o = g\bar{H}/f_o^2$ [2]. f_o is the Coriolis parameter, the radius of deformation R_o is a length scale that relates the fluid's tendency for gravity and earth's rotation together along the earth's rotation axis [16].

Poincaré waves

Poincaré waves are surface gravity waves, affected by the Earth's rotation. The frequency is low enough to feel the effect of the Earth's rotation. This affect is visible

in the dispersion relation by the Coriolis parameter f_o. The dispersion relation (3.20) reduces to:

$$\omega^2 = f_o^2 + \alpha^2(k^2 + \pi^2 n^2/L^2), \tag{3.22}$$

where L is the earth's length along which the wave moves, n can be any positive integer and α is the wave speed defined by $\alpha = (g\bar{H})^{1/2}$ [6].

Kelvin waves

Kelvin waves are also gravity driven and similar to the Poincaré wave. The difference is that they are non-dispersive, what means that its dispersion relation simplifies to

$$\omega = \alpha k \tag{3.23}$$

Kelvin waves are trapped to a vertical boundary (the coastline) or the equator. Equatorial trapped Kelvin waves move with a wave speed of around 2.8 $m\ s^{-1}$ [7]. Its zonal wavenumbers can differ between 0.5 and 3.
There are two types of coastal Kelvin waves: coastal surface and coastal internal waves. Internal waves can reach very high phase speeds, similar to Equatorial waves. Coastal waves are much slower, they move along with a speed between 0.297 $m\ s^{-1}$ and 1.0 $m\ s^{-1}$. The wavenumbers are similar.

Poincaré waves have similar characteristics as the Kelvin waves. Our model is build to investigate how the boundary treatment influences the regional model output. It is important to relate the model to physical properties, but it isn't important to see how the waves are exactly captured by the model. Because the zonal wavenumber and phase speed of the Poincaré waves are similar to the properties of the Kelvin waves, only the Kelvin properties are regarded for the implementation.

Rossby waves

There are two types of Rossby waves

- Barotropic waves: short wavelength and high-frequency
- Baroclinic waves: long wavelength and low-frequency

Barotropic waves

A wave is called a short or Barotropic wave if its meridional length scale is smaller than its Rossby radius of deformation: $L \leq R_o$. Meridional length, is the length along a longitude circle. These waves are propagated by the wind. Its wavelength is between: $\lambda = 170 - 500\ km$ and have zonal wavenumbers between $1 - 3$ [8] [13].

Baroclinic waves

A wave is called a long or Baroclinic wave if its Meridional length scale is higher than its Rossby radius of deformation: $L \geq R_o$. It length is between $\lambda = 3000 - 10000\ km$ and its wave speed between 0.01 $m\ s^{-1}$ and 3 ms^{-1}. Its zonal wavenumber differs between $0.5 - 2.5$ but there also waves with a wavenumber up to $k_z = 5$ or even $k_z = 8$ [22] [30].
Table 3.2 summarizes the wave characteristics.

Table 3.2: Summary of the important wave characteristics

	Kelvin		**Rossby**	
	Equatorial	Coastal	Barotropic	Baroclinic
Wavespeed α $[ms^{-1}]$	2.8	$0.3-1$	20	$0.1-3$
Zonal wave number k_z	$1-2$	$0.5-3$	$1-3$	$0.5-3,\ 5,\ 8$

Only the most extreme conditions are important for the investigation. Computationally, the implementation of higher wavenumbers is less accurate than the lower wavenumbers. Thus, when the model is able to capture the highest wavenumbers and fastest wave speed, it is also able capture the lower wavenumbers. Therefore, the following wavenumbers are considered: $k_z = 1$ and $k_z = 3$ with maximum wave speed of $\alpha = 20\ ms^{-1}$ and wavenumbers $k_z = 5$ and $k_z = 8$ with maximum wave speed of $3\ ms^{-1}$. The lower wavenumbers are also investigated, because its wave speed can be very high. The wave speed relates the model parameter $T = 1$ to the actual model simulation time T^* of climate models by $T^* = T\frac{L}{\alpha}$. This is explained in more detail in section 4.6.

3.3 Discretization techniques

Implementation in Matlab can be done after discretization of the continuous equation. Spatial discretization is done with the Summation-by-Parts–Simultaneous-Approximation-Term (SBP-SAT) method. This method leads to a stable spatial discretization. Discretization of the time-dependent part is done by fourth order Runge-Kutta.

3.3.1 The Simultaneous-Approximation-Term – Summation By Parts method

The Summation-By-Parts method is a high-order finite-difference scheme beneficial for wave propagation problems. The following information is based on [26] and [12]. The Simultaneous-Approximation-Term technique imposes the boundary condition weakly by adding a penalty term, creating together with the SBP-method a stable and accurate high-order spatial discretization.
The idea of the SBP method starts by showing well-posedness of the continuous equation. This is explained in [26] as follows:
Consider a general initial value problem:

$$\begin{cases} u_t = Pu + F, & 0 \leq x \leq 1, \qquad t \geq 0, \\ u(x,0) = f(x), \\ Bu(0,t) = g_0(t), \\ Bu(1,t) = g_1(t) \end{cases} \tag{3.24}$$

where $u = (u^1, ...u^m)^T$, P a differential operator with smooth coefficients, B is a boundary operator, $g_i(t)$ are boundary functions, F is a given force function and $f(x)$ is the initial function.

Definition 3.1 *The equation is well-posed with $F = 0, g_i = 0$, if for every $f \in C^\infty$ that vanishes in a neighborhood of $x = 0, 1$, there is a unique smooth solution satisfying:*

$$||u(\cdot,t)|| \leq K \exp(\beta_c t)||f(\cdot)||. \tag{3.25}$$

K, β_c are constants independent of f. [26]

The well-posedness of the transport equation can be shown with the energy method from Definition 3.1. The transport equation in this problem is given as follows:

$$\begin{cases} u_t + \alpha u_x = 0, & 0 \leq x \leq 1, \qquad t \geq 0 \\ u(x,0) = f(x) \end{cases} \tag{3.26}$$

Mulitply the equation with u and integrate:

$$\begin{aligned} &\int_0^1 u u_t dx + \int_0^1 \alpha u u_x dx = 0 \\ &\Leftrightarrow (u, u_t) + \alpha(u, u_x) = 0 \quad \text{where } (u,u) = ||u||^2 \end{aligned}$$

Then apply integration by parts and obtain:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}||u||^2 + \alpha[u^2]_0^1 - \alpha(u_x, u) = 0 \\ \Leftrightarrow \frac{1}{2}\frac{d}{dt}||u||^2 + \frac{1}{2}\alpha[u^2]_0^1 = 0 \\ \Leftrightarrow \frac{d}{dt}||u||^2 = -\alpha[u^2]_0^1\end{aligned}$$

For well-posedness, $\frac{d}{dt}||u||^2$ cannot grow. This can be established by imposing boundary conditions.
In the case that $\alpha > 0$. Setting $u(0,t) = g_0(t)$ leads to the desired estimate:

$$\frac{d}{dt}||u||^2 = -\alpha[u^2]_0^1 \leq K\exp(\beta_c t)||f(\cdot)||$$

The bound on the discrete transport equation is similar and is done with the discrete energy method. The discrete version of the transport equation according the SBP-SAT method is [12]:

$$v_t + \alpha P^{-1}Qv = \underbrace{\sigma P^{-1}E_0(v-g)}_{\text{SAT-term } \mathbb{S}} \tag{3.27}$$

The finite difference operator P has the property that it is a symmetric positive definite matrix. The operator Q has the property that the matrix is almost skew-symmetric, meaning that $Q + Q^T = \begin{bmatrix}-1 & & \\ 0 & & \\ 0 & & 1\end{bmatrix}$, and $E_0 = \begin{bmatrix}1 & & \\ 0 & & \\ 0 & & 0\end{bmatrix}$. The vector g is a vector of same length as $u(x,t)$ containing all zeros, except for the first term. The first term contains the boundary condition at $x = 0$. The σ term is characteristic for the SAT-term. It is introduced to penalize the boundary term.
The approximation (3.27) is stable when

$$||v(t)||_h \leq K\exp(\beta_d t)||f_h||.\ [26] \tag{3.28}$$

This can be shown with the discrete energy method. The σ can be derived with this method.
The discrete energy method starts by transposing (3.27). This leads to the following:

$$u_t^T + \alpha u^T Q^T P^{-1} = \sigma(\mathbf{u}^T - \mathbf{g}^T)E_0 P^{-1} \tag{3.29}$$

Then equation (3.27) is multiplied with $u^T P$ from the left and equation (3.29) with Pu from the right. This leads to the following two equations:

$$\begin{aligned}u^T P u_t + \alpha u^T Q u = \sigma u^T E_0(u - \mathbf{g}) \\ u_t^T P u + \alpha u^T Q^T u = \sigma(u^T - \mathbf{g}^T)E_0 u\end{aligned} \tag{3.30}$$

The product of $P^{-1}P = I$ and thus vanishes in the equation. The addition of the equations in (3.30) leads to the following equation:

$$\frac{d}{dt}||u||_P^2 + \alpha u^T(Q + Q^T)u = 2\sigma u_0(u_0 - g_0) \tag{3.31}$$

It is sufficient to show it for $g_0 = 0$, then it can be written as follows:

$$\frac{d}{dt}||u||_P^2 = (2\sigma + \alpha)u_0^2 - \alpha u_N^2 \tag{3.32}$$

According to (3.28), the following should hold:

$$\frac{d}{dt}||u||_P^2 \leq K \exp(\beta_d t)||f_h||$$

This is established when the RHS is completely negative, thus when $(2\sigma + \alpha)u_0^2 \leq 0$. This leads to a bound of:

$$(2\sigma + \alpha) \leq 0 \Leftrightarrow \sigma \leq \frac{-\alpha}{2}. \tag{3.33}$$

Thus when $\sigma \leq \frac{-\alpha}{2}$, a stable solution is obtained for (3.27).

The α in our modeling problem is positive, because it was scaled to $\alpha = 1$.

On the order of accuracy and the convergence rate

The discrete approximation v of the exact solution $u(x,t)$ is determined and the truncation error can be defined.
Consider a domain of length L from $x = [0,1]$ of N gridpoints as in figure 3.3. Define u_j as the projection of u on this grid. This means that $u_i = u(x_i, t)$, $i = 1, 2, .., N$.

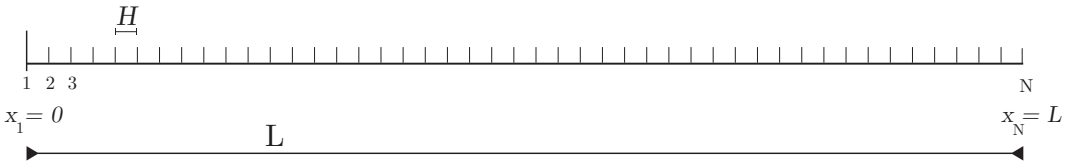

Figure 3.3: One dimensional representation of the computational domain.

The local truncation error $\mathbb{T}_i$ is defined by:

$$(u_i)_t = \alpha P^{-1} Q(u_i)_x + \mathbb{S}_i + \mathbb{T}_i, \quad i = 0, ..., N, \quad t \geq 0 \tag{3.34}$$

$\mathbb{S}$ is the SAT-term. The truncation error has the following form:

$$\mathbb{T} = (\mathcal{O}(h^r), ..., \mathcal{O}(h^r), \mathcal{O}(h^p), ..., \mathcal{O}(h^p), \mathcal{O}(h^r), ..., \mathcal{O}(h^r))^T \tag{3.35}$$

With $r < p$. These terms with r as exponent correspond the points near the boundary that have a lower accuracy than the interior, the terms with p as exponent [26]. *The order of accuracy* is the exponent in the truncation error. Thus, the order of accuracy is (r, p).
The convergence rate q, another important term is defined by the solution error:

$$\epsilon_i(t) = v_i - u_i \tag{3.36}$$

The convergence rate q comes from $||\epsilon||_2 = \mathcal{O}(h^q)$, where $||\cdot||_2$ is the discrete L_2-norm. The choice of the differential operators define the orders of accuracy (r, p) and so the convergence rate q. There are several SBP operators established to form the differential

operators leading to different rates of convergence. The four established operators lead to orders of accuracy [12]

$$(r, p) = (1, 2), \quad (2, 4), \quad (3, 6), \quad (4, 8).$$

Leading to convergence rates of

$$q = 2, \quad 3, \quad 4, \quad 5, \quad \text{respectively.}$$

Change of terminology

The terminology of order of accuracy and convergence rate can lead to quick mistakes once the implementation is started. To clarify things, the terminology is slightly changed. Once the results of the numerical implementation are present, it is expected that the $||\epsilon||_2$ will converge according the convergence rates of the SBP-SAT operators. The term *p-value* is introduced to indicate the convergence rate of the L_2-norm of the error of the numerical approximation ($||\epsilon||_2 = ||u^{exact} - u^{approx}||_2$). It is expected that the p-value of order of accuracy $(1, 2)$ is around 2, but this is not 100% certain! A new term, the *design accuracy*, is introduced to indicate the convergence rate according the SBP-SAT operators.
In summary, the order of accuracy $(1, 2)$ means a design accuracy of $p_{da} = 2$ and a p-value of $p \approx 2$.

Numerical implementation

To be sure that the implementation of the SBP-SAT operators is done correctly in Matlab, the p-values of the numerical approximations are calculated and plotted logarithmic in figure 3.4. The implementation is done for the transport equation (3.27) for increasing number of gridpoints. The figure shows that the numerical implementation of the SBP-SAT-method is done correctly and that the different SBP-SAT operators are converging according its design accuracy.

3.3.2 Fourth order Runge-Kutta

The fourth-order Runge Kutta method should be well known to the reader [31].

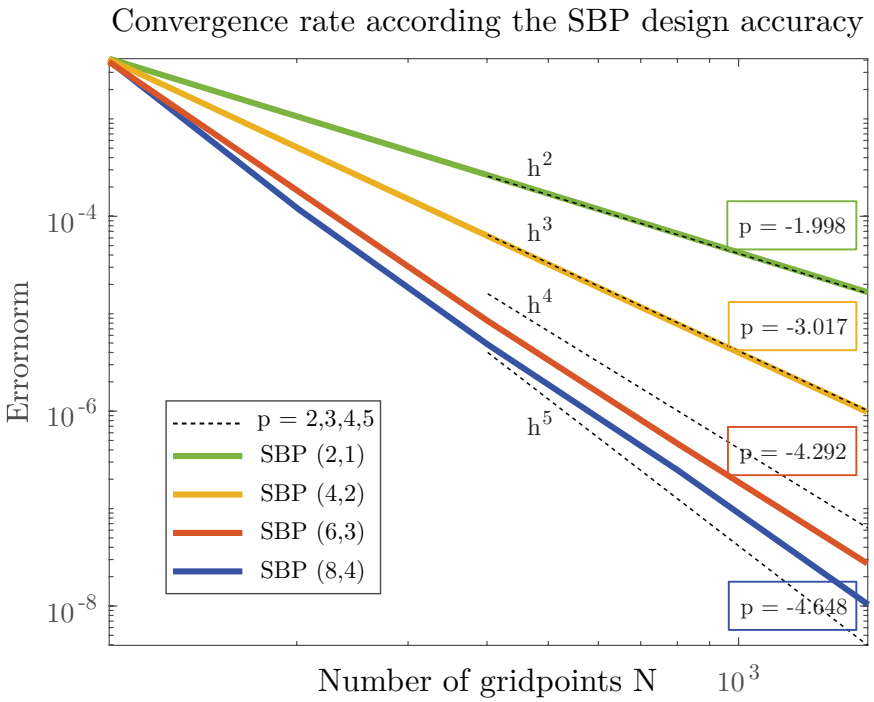

Figure 3.4: p-values of the discrete transport equation according the design accuracy of the SBP-SAT method. The coloured lines show how the numerical approximation of the transport equation converges for different SBP design accuracy. The text in the box next to its coloured lines show the slope of the power of the L_2-norm of the error (which is the p-value). The dashed lines show how the slope should converge (the design accuracy).

3.4 Interpolation methods

The Lateral Boundary Condition (LBC) is needed at each timestep of the regional model and thus interpolation is required between the boundary data updates of the global model. This is done by storing the output of the GCM in space and time. Storage in space is done only on the spatial gridpoint that corresponds to the lateral boundary of the regional domain: x_b. Storage in space is done on all the temporal steps of the GCM. Hence, the GCM output is stored in a vector $u_{GCM}(x_b, T_i)$, $i = 0, 1, 2, ..., N_T$ which has the necessary information to update the LBC. Interpolation is needed to create the required boundary information on the temporal points of the RCM. For simplicity, the vector $u_{gcm}(x_b, T_i)$ is now called $\tilde{g}(T_i)$.

$$u_{GCM}(x_b, T_i) = \tilde{g}(T_i) = \begin{bmatrix} \tilde{g}(0) \\ \tilde{g}(T_1) \\ \vdots \\ \tilde{g}(T_{N_T}) \end{bmatrix} \qquad i = 0, 1, .., N_T \tag{3.37}$$

$$\begin{cases} u_{GCM}(x_b, T_i) & \text{stored solution of the GCM on } x_b \text{ on every timestep } T_i \\ \tilde{g}(T_i) & \text{lateral boundary condition before interpolation} \\ x_b & \text{the GCM gridpoint that correponds to the RCM boundary} \\ T_i & \text{a certain GCM temporal gridpoint} \\ N_T & \text{total number of GCM timesteps} \end{cases}$$

The simulation time of the GCM is $T = 1$. After N_T steps of size ΔT, the final time is reached.
The simulation time of the RCM runs is also $T = 1$. There are in total M_T number of timesteps of size Δt and $N_T \leq M_T$ and $\Delta t \leq \Delta T$.
The created boundary data $\tilde{g}(T_i)$ can only be used as LBC after interpolation. The current RCMs use linear interpolation techniques [11]. Higher order polynomial methods can lead to more accurate solutions and have to be evaluated. In the following section several third order polynomial methods are explained and evaluated.

3.4.1 Conditions on the interpolation method

The vector $\tilde{g}(T_i)$ that needs to be interpolated is divided on an equidistant temporal grid with information on every $T_i,\ i = 0, 1, 2, ..., N_T$. Interpolation in time proceeds similar as interpolation in space, as interpolation is used to fill in the missing gridpoints and it doesn't matter whether this is a spatial or temporal point.
In case of an equidistant grid, it is not recommended to choose a higher order **global** polynomial interpolation method, this is illustrated in [21] with the Runga example ($g(x) := 1/(1 + 25x^2)$). It can be reduced by choosing a piecewise polynomial interpolation method. This method only requires information of its surrounding data.

3.4.2 Piecewise polynomial approximation

In piecewise polynomial approximation, an interval $[0, T]$ is partitioned into N_T intervals I_i as follows:

$$\underbrace{[T_0,\ T_1]}_{I_1},\ \underbrace{[T_1,\ T_2]}_{I_2},\ ...,\ \underbrace{[T_{N_T-1},\ T_{N_T}]}_{I_{N_T}} \qquad 0 \ = \ T_0 < T_1 < ... < T_{N_T} \ = \ 1.$$

The idea of piecewise approximation is that the polynomial is approximated per interval. The polynomial approximation $p_i(t_j)$ is computed on all the desired points t_j that lie in the interval $[T_i,\ T_{i+1}]$. This means that N_T polynomial approximations are created. At the end they are 'glued' to each other to form one polynomial $p(t_j)$ that can be used as boundary condition. In summary this is going as follows:

1: Partition the domain $[0,\ T]$ into N_T intervals

2: Compute the polynomial $p_i(t_j)$ piecewise for all $t_j \in [T_i,\ T_{i+1}]$, $j = 0, .., M_T$ per interval $[T_i,\ T_{i+1}]$ and $i = 0, 1, .., N_T$.

3: Glue all p_i of the intervals $[T_i,\ T_{i+1}]$ to each other to end with the final polynomial $p(t)$

Piecewise polynomial approximation can be done with first order, third order of even higher order methods. First order and third order methods are evaluated here [21] [29].

3.4.3 Piecewise linear interpolation

Piecewise linear interpolation can be seen as connecting the information $g_i = u_{GCM}(T_i)$ on all the breakpoints T_i with straight lines to each other. This can be done by local linear interpolants:

$$p_i(t) = g(T_i) + \frac{t - T_i}{T_{i+1} - T_i}(g_i - g_{i+1}) \quad \forall\ t \in [T_i,\ T_{i+1}] \tag{3.38}$$

and has the following property: $p_i(T_i) = g(T_i) \quad \forall T_i,\ i = 0, 1, ..., N_T$. There are now $N_T - 1$ polynomial approximation on $N_T - 1$ intervals that are glued to each other at the end to get the final approximant $p(t)$, which is the LBC of the RCM domain. The total interpolant is defined as follows [21]:

$$p(t) = \begin{cases} p_1(t) & \text{if} \quad 0 \leq t < T_1 \\ p_2(t) & \text{if} \quad T_1 \leq t < T_2 \\ & \vdots \\ p_{N_T}(t) & \text{if} \quad T_{N_{T-1}} \leq t \leq T_{N_T} \end{cases}$$

The linear interpolant $p(t_j)$ is used as the lateral boundary condition of the RCM domain. It is implemented for the $\vec{g}$ in equation 3.27.

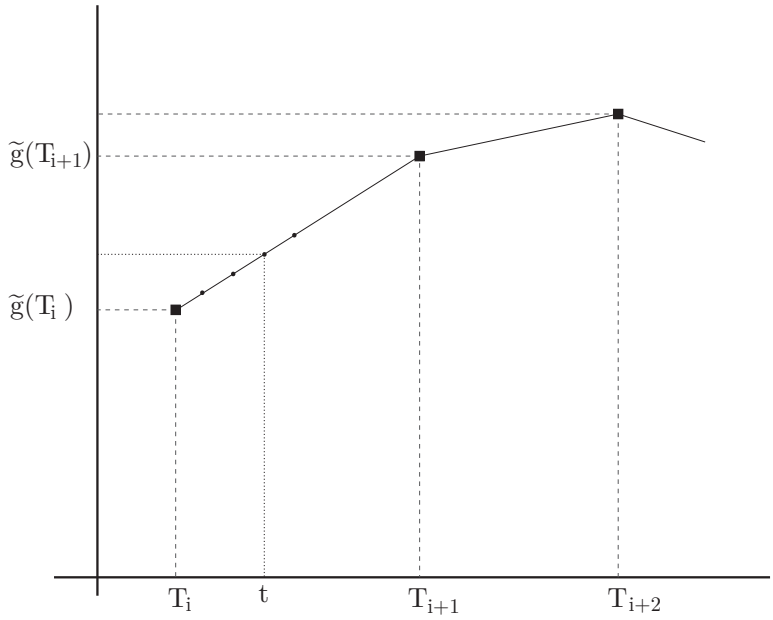

Figure 3.5: Linear interpolation of information $\tilde{g}$ on t between two adjacent points T_i and T_{i+1}.

Evaluation of the linear interpolation method

To define the accuracy of this method, the following theorem is needed [21].

Theorem 3.1 *Suppose $p_{n-1}(x)$ interpolates a function $f(x)$ at distinct points $x_1, .., x_n$. If $f \in C^n$ on an interval I containing the x_i, then for any $x \in I$*

$$f(x) = p_{n-1}(x) + \frac{f^{(n)}(\eta)}{n!}(x - x_1) \dots (x - x_n).$$

If this continuously differentiable f is present, then the error between the approximation and the function can be derived as follows: for $t \in [T_i,\ T_{i+1}]$,

$$f(t) - p(t) = \frac{f^{(\prime\prime)}(\eta)}{2!}(t - T_i)(t - T_{i+1}) \qquad \eta \in [T_i,\ T_{i+1}] \tag{3.39}$$

In this case no function is interpolated, but only data gained from the GCM domain. This means, there is no continuously differentiable function f present. The error between the interpolation and data have to be determined differently. This can be done by using the exact solution $u^*(t)$. The information g_i can be seen as the exact solution plus a certain error term ψ:

$$g_i = u^* + \psi \tag{3.40}$$

The exact solution u^* is continuously differentiable. If you fill equation (3.40) in (3.39) then you get the following error term between the linear interpolant and the function:

$$\xi(t)^* = u^*(t) + \psi - p(t) = \frac{u^*(\prime\prime)(\eta)}{2!}(t - T_i)(t - T_{i+1}) + \psi^* \qquad \eta \in [T_i,\ T_{i+1}] \tag{3.41}$$

where ψ^* is the result of the error term ψ after interpolation.
The error between the interpolation method and the data is estimated to compare the different interpolation methods. The error term ψ as a result of the discrete approximation of the exact solution is similar for all interpolation methods and only its result after interpolation, the term $\psi^*(t)$, is different per interpolation method. In order to find an estimation for the error term, this term is neglected, because its influence will be negligibly small.
This results in the following error estimate of the linear interpolation method:

$$\xi(t) = u^*(t) - p(t) \approx \frac{u^{(\prime\prime)}(\eta)}{2!}(t - T_i)(t - T_{i+1}) \qquad \eta \in [T_i,\ T_{i+1}] \tag{3.42}$$

The exact solution at the boundary of the regional model is $u^*(x_b, t) = cos(k(x_b - t))$. The second order derivative between $T = [0,\ 1]$ is bounded by

$$|u^*(\prime\prime)(\eta)| \leq k^2. \tag{3.43}$$

Because all the intervals $[T_i,\ T_{i+1}]$ are of equal size ΔT and the maximum is on the midpoint of the interval, $t_{mid} = T_i + \Delta T/2$, the final error estimate of linear interpolation:

$$|\xi(t)| \leq \frac{k^2}{2!}\left(\frac{\Delta T}{2}\right)^2 = \frac{k^2}{8}(\Delta T)^2. \tag{3.44}$$

3.4.4 Third order polynomial approximations

Third order polynomial approximations can be done in several ways, of which Cubic Spline and Cubic Hermite approximations are the most common. The most suitable third order interpolation method depends on the function that has to be approximated. Simulations of both methods shall result in the choice for the best third order method. All of the following information is based on the text in [29] and [14].

3.4.5 Piecewise Cubic Hermite interpolation

The cubic Hermite spline has to satisfy the following properties:

$$\begin{cases} p_i(t) = at^3 + bt^2 + ct + d, \quad a, b, c, d \in \mathbb{R} \\ p(T_i) = \tilde{g}(T_i), \qquad \forall i = 0, 1, ..., N_T \\ p(T_{i+1}) = \tilde{g}(T_{i+1}), \qquad \forall i = 0, 1, ..., N_T \\ p'(T_i) = \tilde{g}'(T_i), \qquad \forall i = 0, 1, ..., N_T \\ p'(T_{i+1}) = \tilde{g}'(T_{i+1}), \qquad \forall i = 0, 1, ..., N_T \end{cases}$$

Where $\tilde{g}(T_i)$ is known: these are the boundary data obtained from the GCM simulation. Normally in Cubic Hermite splines, the derivative $g'(T_i)$ is present. In this case, only the data points $\tilde{g}(T_i)$, $\tilde{g}(T_{i+1})$, *etc.* are present and no derivative. There are several ways to obtain the derivative. In this case it is done with the Catmull-Rom method. In Catmull-Rom, the derivative is approximated as follows:

$$\bar{g}'(T_i) = \frac{\tilde{g}(T_{i+1}) - \tilde{g}(T_{i-1})}{T_{i+1} - T_{i-1}} \tag{3.45}$$

For the endpoints, one-sided differences are used.

With the two values $\tilde{g}(T_i)$ and $\tilde{g}(T_{i+1})$ and it's approximated derivatives $\bar{g}'(T_i)$ and $\bar{g}'(T_{i+1})$ the parameters a_i, b_i, c_i and d_i can be derived as follows:

$$a_i = \tilde{g}(T_i)$$

$$b_i = \bar{g}'(T_i)$$

$$c_i = \frac{((\tilde{g}(T_{i+1}) - \tilde{g}(T_i))/\Delta T) - \bar{g}'(T_i)}{\Delta T}$$

$$d_i = \frac{\bar{g}'(T_{i+1}) + \bar{g}'(T_i) - 2((\tilde{g}(T_{i+1}) - \tilde{g}(T_i))/\Delta T)}{\Delta T^2}$$

3.4.6 Evaluation of the cubic Hermite spline interpolation method

The error of cubic spline is determined by using similar assumptions about the discrete error term ψ as in the previous section, explained in equation (3.40) till (3.42). Theorem 3.1, the equidistant grid with stepsize ΔT and the bound on the fourth order derivative of the exact solution lead to the following estimate of the error:

$$|\xi(t)| \leq \frac{k^4}{4!}|(t - T_i)(t - T_{i+1})(t - T_i)(t - T_{i+1})| \qquad \forall t \in [T_i, T_{i+1}] \tag{3.46}$$

The maximum value of the local polynomial is on the midpoint of its interval, $t_{mid} = T_i + \Delta T/2$, and thus

$$|\xi(t)| \leq \frac{k^4}{24}\left(\frac{\Delta T}{2}\right)^4 = \frac{k^4}{384}(\Delta T^4) \qquad \forall t \in [T_i, T_{i+1}] \tag{3.47}$$

Cubic Hermite spline interpolation is a local piecewise approximation method. The interpolant $p_i(t)$ uses only information of its surrounded breakpoints $[T_i,\ T_{i+1}]$ and T_{i-1}, T_{i+2} for the derivative approximation. Normally local polynomial approximations would be a good choice for interpolation, but in this case the derivative $\bar{g}(T_i)$ is approximated. This can result in accuracy loss, especially at the boundary points where one-sided differences are used.

3.4.7 Cubic spline interpolation

Cubic spline interpolation is another third order method. The total polynomial $p(t)$ is build from the local cubic polynomials $p_i(t)$ and has additional constraints on the first and second order derivatives, the left and right derivatives must be continuous at interior points: $p_i'(T_{i+1}) = p_{i+1}'(T_{i+1})$ and $p_i''(T_{i+1}) = p_{i+1}''(T_{i+1})$ for $i = 1, 2, ..., N_T - 2$. There are different options for the end conditions of the second order derivatives. Two options for the end conditions are considered here: the Natural spline conditions and the Not-a-Knot conditions [29][14].

The Natural spline

In case of the natural spline, the end conditions of the second order derivative satisfy: $p_0''(T_0) = p_{N_T-1}''(T_{N_T}) = 0$. In sum, the local cubic natural splines have to satisfy the following properties:

$$\begin{cases} \text{I. } p \text{ is a cubic polynomial } p_i \text{ on each subinterval } [T_i, T_{i+1}],\ i = 0, 1, .., N_T - 1 \\ \text{II. } p(T_i) = \tilde{g}(T_i), \qquad \forall i = 0, 1, ..., N_T \\ \text{III. } p_i(T_{i+1}) = p_{i+1}(T_{i+1}), \quad \forall i = 0, 1, ...N_T - 2 \\ \text{IV. } p_i'(T_{i+1}) = p_{i+1}'(T_{i+1}), \quad \forall i = 0, 1, ...N_T - 2 \\ \text{V. } p_i''(T_{i+1}) = p_{i+1}''(T_{i+1}), \quad \forall i = 0, 1, ...N_T - 2 \\ \text{VI. } p_0''(T_0) = p_{N_T-1}''(T_{N_T}) = 0 \end{cases} \tag{3.48}$$

With these properties, the values for $a_i,\ b_i,\ c_i$ and d_i can be solved.
Condition I determines the local cubic spline on each interval $[T_i, T_{i+1}]$ and its derivatives result here-from:

$$p_i(t) = a_i(t - T_i)^3 + b_i(t - T_i)^2 + c_i(t - T_i) + d_i, \quad a_i, b_i, c_i, d_i \in \mathbb{R},\ i = 0, 1, .., N_T - 1 \tag{3.49}$$

$$p_i'(t) = 3a_i(t - T_i)^2 + 2b_i(t - T_i) + c_i \tag{3.50}$$

$$p_i''(t) = 6a_i(t - T_i) + 2b_i \tag{3.51}$$

Condition II results in

$$d_i = \tilde{g}(T_i). \tag{3.52}$$

Condition III in the determination of c_i:

$$\begin{aligned} & p_i(T_{i+1}) = p_{i+1}(T_{i+1}) \\ \Leftrightarrow\ & a_i \underbrace{(T_{i+1} - T_i)^3}_{\Delta T^3} + b_i(T_{i+1} - T_i)^2 + c_i(T_{i+1} - T_i) + d_i = \tilde{g}(T_{i+1}) \\ & \Rightarrow c_i = \frac{\tilde{g}(T_{i+1}) - \tilde{g}(T_i)}{\Delta T} - \Delta T \frac{2b_i + b_{i+1}}{3} \end{aligned} \tag{3.53}$$

Condition V leads to the determination of a_i.

$$\begin{aligned} & p_i''(T_{i+1}) = p_{i+1}''(T_{i+1}) \Leftrightarrow 2b_{i+1} = 6a_i\Delta T + 2b_i \\ & \Rightarrow a_i = \frac{b_{i+1} - b_i}{3\Delta T} \end{aligned} \tag{3.54}$$

Condition IV result in a linear system of equations that have to be solved to determine b_i, followed by c_i and a_i. Condition IV is:

$$\begin{aligned} p_i'(T_{i+1}) &= p_{i+1}'(T_{i+1}) \\ \Leftrightarrow c_{i+1} &= 3a_i\Delta T^2 + 2b_i\Delta T + c_i \\ \Rightarrow b_i\Delta T + 2b_{i+1}(\Delta T + \Delta T + b_{i+2}\Delta T &= 3\frac{\tilde{g}(T_{i+2}) - \tilde{g}(T_{i+1})}{\Delta T} - \frac{\tilde{g}(T_{i+1}) - \tilde{g}(T_i)}{\Delta T} \end{aligned} \tag{3.55}$$

Condition VI is the natural end condition, resulting in

$$\begin{aligned} b_0 &= 0 \\ b_{N_T} &= 0 \end{aligned} \tag{3.56}$$

Solving the linear system of equations (3.55) and (3.54) to (3.52) from bottom-to top, results in the determination of a_i, b_i, c_i and d_i, which can be implemented in the local spline 3.49. The total spline is made up from the local pieces.

Not-a-Knot spline

The Not-a-Knot condition ensures that the *third* derivative is continuous at the points T_1 and T_{N_T-1}:

$$p_i'''(T_{i+1}) = p_{i+1}'''(T_{i+1}) \tag{3.57}$$

The third order derivative of the local spline is:

$$p_i'''(T_{i+1}) = 6a_i \tag{3.58}$$

a_i is according to 3.54 equal to: $a_i = \frac{b_{i+1}-b_i}{3\Delta T}$ and b_i can still be solved by solving the system of 3.55. It results in the following value for b_0:

$$\begin{aligned} p_0'''(T_1) &= p_1'''(T_1) \\ &\Leftrightarrow 6a_1 = 6a_2 \\ &\Leftrightarrow 6\frac{b_1 - b_0}{3\Delta T} = 6\frac{b_2 - b_1}{3\Delta T} \\ &\Rightarrow b_0 = b_2 - 2b_1 \end{aligned} \tag{3.59}$$

b_{N_T} is determined similarly:

$$\begin{aligned} p_{N_T-2}'''(T_{N_T-1}) &= p_{N_T-1}'''(T_{N_T-1}) \\ &\Leftrightarrow 6a_{N_T-1} = 6a_2 \\ &\Leftrightarrow 6\frac{b_{N_T-1} - b_{N_T-2}}{3\Delta T} = 6\frac{b_{N_T} - b_{N_T-1}}{3\Delta T} \\ &\Rightarrow b_{N_T} = 2b_{N_T-1} - b_{N_T-2} \end{aligned} \tag{3.60}$$

The determination of all the coefficients and the implementation in the final spline is similar to that of the Natural spline.

3.4.8 Evaluation of the cubic spline methods

The choice between cubic splines and cubic Hermite interpolation and its derivative conditions depends on the properties of the data that have to be interpolated. The Hermite spline has the advantage that it is a local approximation. This means that it only needs values around the point where you want to approximate. No end conditions are needed. The cubic spline does require end-conditions for the second order derivative, which can influence the approximation negatively when these conditions are chosen badly. An advantage of the cubic spline is that the second order derivative is continuous (we have set it this way). Numerical implementation helps to find out which method of the two is most suitable. The implementation is done in the same way as the implementation of the final experiment. This means that the transport equation (3.17) is implemented according the SBP-SAT method. The discrete version is (3.27). The temporal discretization is done according to the fourth order Runge Kutta method. The final numerical approximation $v(x,t_i)$, $i = 0,1,...,N_T$ is a vector of the same length in time, which is a length of $N_T + 1$ points, as the exact solution computed every numerical timestep $u(x,t_i)$. The key is to find how the output is performing after lowering the boundary data interval resolution. The lowering can be simulated here, by storing the output $v(x,t_i)$ every timestep and then removing data from this vector. Then use interpolation to approximate the removed data and compare this with the exact solution $u(x,t_i)$.
The interpolation methods can be compared when the numerical approximation $v(x,t_i)$ is stored in half the timesteps of t_i, ie $t_0, t_2, t_4, etc.$, or with a quart of the timesteps, ie $t_0, t_4, etc.$ or even with a eight of the timesteps, ie $t_0, t_8, etc.$. Then the removed data is interpolated back to a vector of length $N_T + 1$. The interpolated solution $\tilde{v}(x,t_i)$ is compared with $u(x,t_i)$ in the L_2-norm. This is going as follows:

$$\text{1. Compute } v(x,t_i). \Rightarrow \quad v(x,t_i) = \begin{bmatrix} v(x,t_0) \\ v(x,t_1) \\ v(x,t_2) \\ \vdots \\ v(x,t_{N_T}) \end{bmatrix} \tag{3.61}$$

$$\text{2. Store in half the timesteps . } \Rightarrow \quad \begin{bmatrix} v(x,t_0) \\ \bullet \\ v(x,t_2) \\ \bullet \\ \vdots \\ v(x,t_{N_T}) \end{bmatrix} \tag{3.62}$$

$$\text{3. Interpolate to } \tilde{v}(x,t_i). \Rightarrow \tilde{v}(x,t_i) = \begin{bmatrix} \tilde{v}(x,t_0) \\ \tilde{v}(x,t_1) \\ \tilde{v}(x,t_2) \\ \vdots \\ \tilde{v}(x,t_{N_T}) \end{bmatrix} \tag{3.63}$$

$$\text{4. Take the } L_2\text{-norm. } \Rightarrow \quad \epsilon_1 = ||\tilde{v}(x,t_i) - u(x,t_i)|| \tag{3.64}$$

These four steps of (3.61) - (3.64) are also executed for storing $v(x,t_i)$ one quart, one eighth and one sixteenth of all the timesteps. This changes the matrix in (3.62) to

$$\begin{bmatrix} v(x,t_0) \\ \bullet \\ \bullet \\ \bullet \\ v(x,t_4) \\ \vdots \\ v(x,t_{N_T}) \end{bmatrix}, \quad \begin{bmatrix} v(x,t_0) \\ \bullet \\ \vdots \\ \bullet \\ v(x,t_8) \\ \vdots \\ v(x,t_{N_T}) \end{bmatrix}, \quad \begin{bmatrix} v(x,t_0) \\ \bullet \\ \vdots \\ \bullet \\ v(x,t_{16}) \\ \vdots \\ v(x,t_{N_T}) \end{bmatrix}, \text{ respectively.}$$

The L_2-norm of the errors (ϵ_2, ϵ_4, ϵ_8 and ϵ_{16}) are computed and shown in table 3.3. The different numerical approximations are plotted for different storages. These are visible in figure 3.6.

Conclusion of the interpolation methods

From table 3.3 it can be concluded that the Not-a-Knot cubic spline interpolation method is in the beginning most suitable for the numerical approximation of the transport equation with exact solution:

$$u(x,t) = cos(2\pi k_z(x-t)) \qquad t > 0 \quad x = [0,1] \tag{3.65}$$

But the error ϵ_{16} is lower for the Natural spline than for the Not-a-Knot spline. Furthermore, the errors are of comparable size. It probably doesn't really matter in this case whether Not-a-Knot or Natural splines are used. It is important to test this for every single equation that has to be approximated. Cubic splines are often unfavourable due to the global approximation, but for this equation it is not the case. Because the Natural spline has a lower error for a lower storage, the Natural Cubic spline is chosen as the third order interpolation method that is going to be implemented in the final set-up of this research.

Table 3.3: The L_2-norms of the error between $u(x,t_i)$ and $\tilde{v}(x,t_i)$ for different storages

	Hermite spline	**Cubic spline**	
	Finite differences	Natural	Not-a-Knot
ϵ_2	1.62e-04	1.11e-04	1.51e-05
ϵ_4	9.92e-04	6.36e-04	2.39e-04
ϵ_8	6.95e-03	3.81e-03	3.62e-03
ϵ_{16}	5.95e-02	2.86e-02	5.65e-02

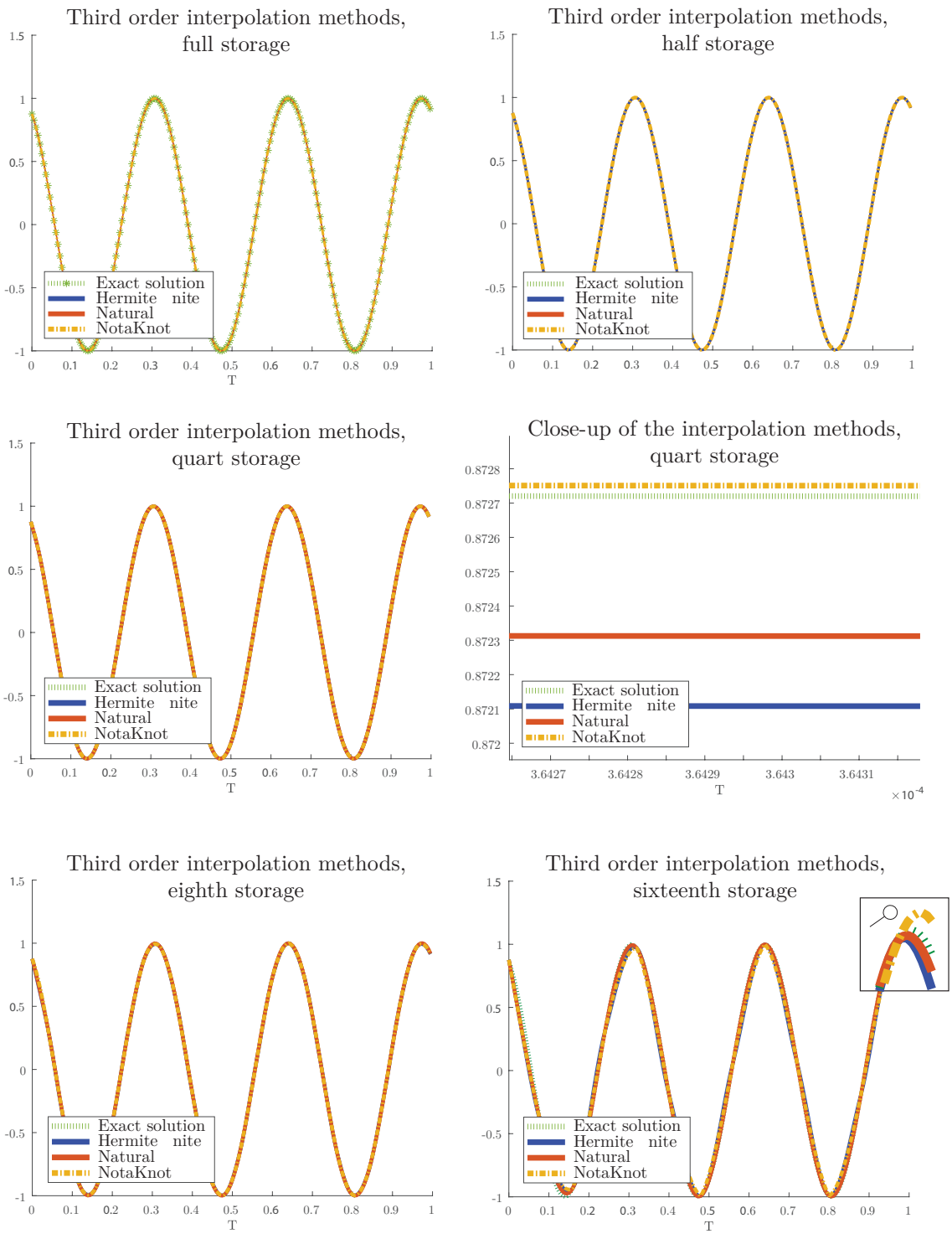

Figure 3.6: Third order interpolation methods compared for different storages. The close-up (midright) shows which method is most accurate. From the last picture it is visible that the Not-a-Knot conditions are getting worse and that the Natural spline is more reliable.

Chapter 4

Computational experiments

4.1 Big and Little Brother experiment

The experiments are performed according the Big and Little Brother experiment (BB-LBE), introduced by [3] to test the down-scaling ability of one-way nested Regional Climate Models. In the set-up of [3], the Little Brother (LB) model represents the regional model scaled to a smaller, workable version and has a Big Brother (BB) of the same resolution as the regional model, but its domain is larger. The numerical methods of the BB and LB models are the same as in the global and regional models. The simulation start by running the BB model. The model produces a dataset with large, medium and small scale features. The small scale features are filtered from the dataset, and used later for reference, such that it only contains the large and medium scale features. These are implemented at the boundary of the LB and interpolated to form the LBC. The simulation continues with running the LB model with the interpolated large and medium scale dataset as boundary data. The LB should be able to reproduce the small scale features of the model. The outcome of the LB is validated against the small scale features of the BB. The BB-LBE is introduced, because it keeps the same resolution, physics, dynamics and numerical method as the global-regional models, but it is an efficient method to test the issues concerning the LBC.

Adapted Big and Little Brother experiment for this thesis

The experiments in this thesis use a slightly different set-up of the BB-LBE then the one explained in [3]. In this case, the BB model is the equivalent of the global model and the LB of the regional model, still both scaled to a smaller, workable version. The adapted BB is of lower resolution than the BB in the original case. The LB is similar as the original one. The numerical method and resolution of the adapted BB-LBE are the same as the global-regional models. The reason that an adapted version is introduced, is because the original exact solution is present here and can be used as reference set instead of the small scale features of the BB output. The adapted BB model produces a large scale dataset due to its lower resolution and the small scale solution of the LB can now be validated against the exact solution. This makes the set-up even a better representation of the global and regional models.
The errors that are present in the adapted BB-LBE set-up are now purely due to the nesting strategy and the set-up is therefore well suitable for the validation of the nine issues, introduced in section 2.2.

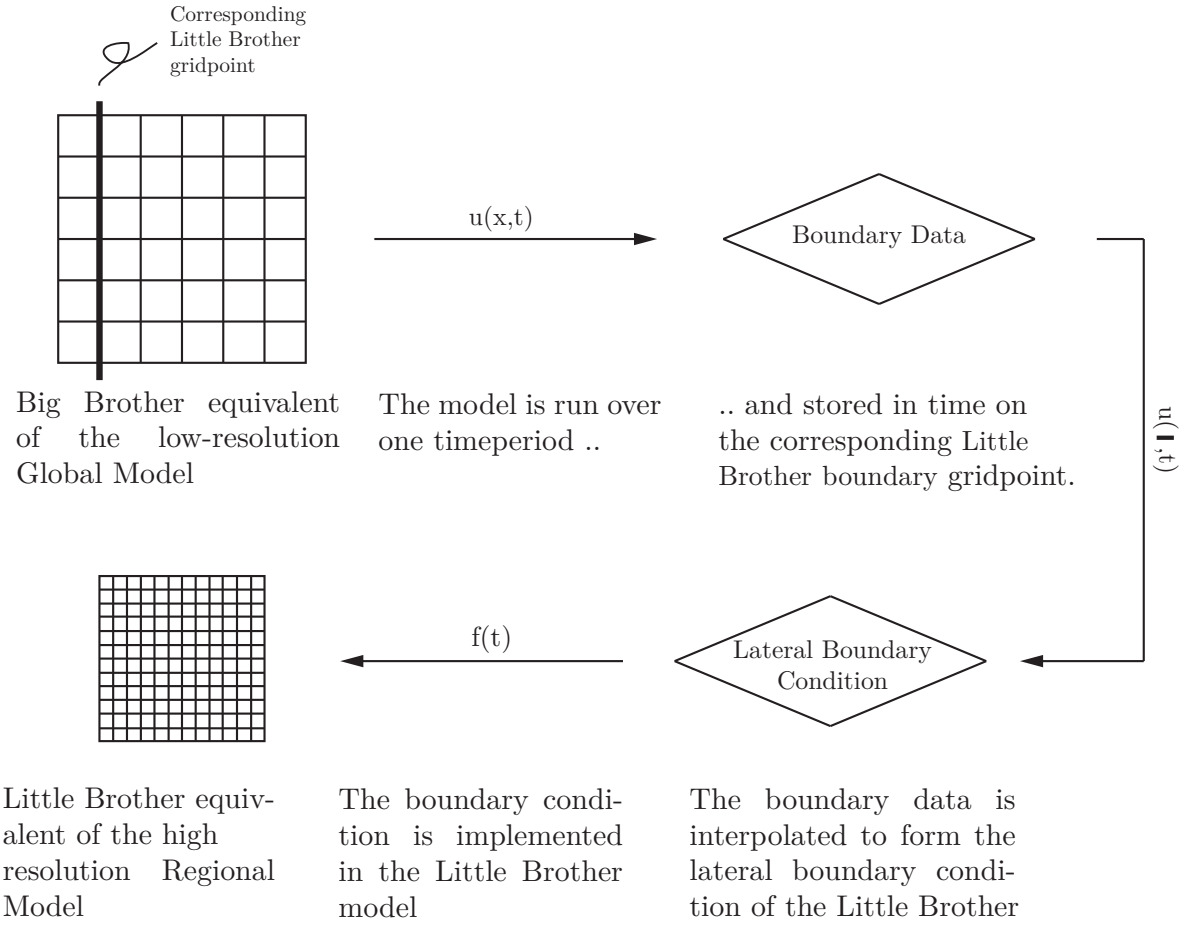

Figure 4.1: Visual graphic of the Big-Little Brother experiment.

The main focus of this thesis is to test the boundary data interval resolution and interpolation methods. With this adapted BB-LBE, the large-scale dataset of the BB contains the LBC for the LB. Focus is to investigate to what extend lowering the boundary data interval resolution, influences the outcome of the LB model and whether this can be compensated by increasing the order of interpolation.
From here on, the term 'adapted Big Brother model' is for convenience changed to 'Big Brother model'. Thus, don't mistake the Big Brother model from now on for the original Big Brother model as explained in [3]. Figure 4.1 is a schematic overview of the BB-LBE.

Remember the definitions of p-value and design accuracy. The p-value is the rate of convergence of the L_2-norm of the error of the numerical approximation with $SBP-SAT$ operators, design accuracy is the convergence rate according to the SBP-SAT operators.

4.1.1 Implementation of the BB-LBE

The BB-LBE is made in several steps to be sure that the final implementation is done correctly. These steps are as follows:

Step 1 Verification of the convergence rate according the SBP-SAT method: The model runs over a domain of size [0, 1]. Its resolution varies between $N = 51,\ 101$ to 201. With N the number of gridpoints. By taking the L_2-norm of the error between the BB solution and the exact solution: $E = ||u_{BB} - u^*||$ with u^* the exact solution, the convergence rate can be verified as follows: $p_2 = log_2(\frac{E_{201}}{E_{101}})$ and $p_1 = log_2(\frac{E_{101}}{E_{51}})$. Where p_1 and p_2 should converge according to the the design accuracy that is used, explained in section 3.3. The used order of accuracy is (42) and thus the p-value should be around $p_1, p_2 = 3$ if everything is done correctly.

Step 2 Implementation of the LB model: the LB model is driven by lateral boundary data obtained from its BB. To check the correctness of the implementation, the boundary condition is the exact solution run in time over the boundary of the LB. In this case, the boundary condition doesn't require any interpolation and thus the LB implementation can be tested. The $L2$-norm of the error between the LB solution and exact solution is now computed $E_2 = ||u_{LB} - u^*||$ and the verification is equivalent to step 1: it is tested if it converges according its design accuracy.

Step 3 Implementation of the BB solution as boundary data of the LB model. In step 3, the model is implemented such that the boundary data doesn't require any interpolation methods yet. The performance of the LB solution depends now solely on the quality of the BB solution.

Step 4 Test of interpolation methods, defined in section 3.4.3. This final step forms the basis for the analysis.

4.2 Model domain

The model is a 1D representation of the BB-LBE. The BB domain has length $L_1 = 1$ and starts at $x = 0$. The LB domain of length $L_2 = 0.2$ covers a certain part of the BB domain. The LB domain starts at z_1, where $0 = x_1 < z_1 < \underbrace{z_1 + L_2}_{z_M} < x_N = 1$. In figure 4.2 the domain is displayed.

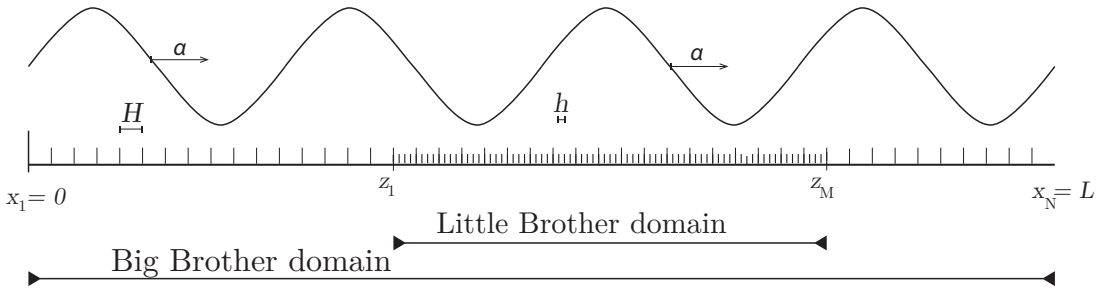

Figure 4.2: Model domain, parameters are explained in the text

4.2.1 Big Brother domain

The BB domain $x = [0\ 1]$ has N number of gridpoints. In the experiment N differs from $N = 51,\ 101,\ 201$ to 401. The gridsize is $H = 1/(N-1)$ and the timestep depends on this gridsize by the CFL-condition: $C\frac{\Delta T}{H} \leq 1$. The CFL-condition is a necessary condition for convergence of numerical solutions to partial differential equations, approximated by finite differences. This condition requires the timestep to be of a certain size or smaller than the size. If it exceeds the limit, the approximation will not converge.
The final time T is 1, so the number of timesteps N_T depends on the ΔT by $N_T = \frac{T}{\Delta T}$.

4.2.2 Little Brother domain

The LB domain covers 20% of the BB domain, hence the length of the LB domain is $L_2 = 0.2$. The gridsize h is also equivalent to the Global-Regional model relation, which differs operationally between a factor $1 - 10$ smaller than the BB gridsize. The total number of gridpoints is $M = 0.2/h$.
The LB model runs on the same time-period as the BB model, the final time T is thus 1. The number of timesteps depends again on the CFL-condition, which now relates to h instead of H, thus $C\frac{\Delta t}{h} \leq 1$.

4.3 Object of investigation

A satisfactory compromise has to be found between the boundary data interval resolution and interpolation method. The boundary interval resolution is as its finest when it is equal to the number of temporal gridpoints of the BB domain; this is all the information that can possibly be available by the model. A finer resolution would require the BB model to be run with a ΔT smaller than the one according to the CFL criterion. In this case, more datapoints are available.
The fine resolution can be coarsened until there are only two datapoints left; $T_0 = 0$ and $T_{N_T} = 1$. Every interpolation method would give a straight line between these points and is therefore not relevant for investigation, but somewhere between these options lies the desired result. The desired result can be found by coarsening the boundary data interval resolution. Coarsening is done by decreasing the resolution with a factor 2 every time. This means that in the first time that the resolution is coarsened, it is coarsened from N_T number of intervals to $N_T/2$ number of intervals. Next, the resolution is coarsened from $N_T/2$ number of intervals to $N_T/4$ number of intervals and so on.
The desired data acquired after every step is interpolated. The chosen interpolation methods interpolate the data to the necessary LBC.
The hypothesis is that the third order interpolation method can obtain more accurate results with a lower boundary data interval resolution compared to the first order method. The point of investigation is to find the best compromise between the interpolation method and the interval resolution by coarsening the resolution until the result is satisfying. The next section explains how the best compromise is chosen.

4.4 Verification

Motivation for the compromise between the boundary interval resolution and interpolation method is based on the design accuracy of the SBP-SAT method. The SBP-SAT method leads to a certain rate of convergence, described in section 3.3. With the chosen SBP-SAT method the rate of convergence should be around $p = 3$. This can be verified by refining the grid and evaluating the errors.
The error is obtained by comparing the LB solution u_{LB} with the available exact solution u^* in the L_2-norm:

$$||\epsilon||_2 = ||u^* - u_{LB}||_2 \tag{4.1}$$

The evaluation of the rate of convergence is done the following way:
The error is evaluated on three different spatial resolutions, all deviating with a factor 2 from each other. This means that the grids have a stepsize of:

$$h = \frac{0.2}{M-1},\ h_2 = \frac{1}{2}h \text{ and } h_3 = \frac{1}{2}h_2 = h/4$$

where M is the number of LB gridpoints of the first simulation. The error is related to the rate of convergence p (see equation (3.36) in section 3.3) as follows:

$$||\epsilon^h||_2 = ||u^* - u^h_{LB}||_2 \approx Ch^p \tag{4.2}$$

When the grid is refined, the error for $h_2 = h/2$ and $h_3 = h/4$ is respectively:

$$||\epsilon^{h/2}||_2 = ||u^* - u^{(h/2)}_{LB}||_2 \approx C(\frac{h}{2})^p \tag{4.3}$$

$$||\epsilon^{h/4}||_2 = ||u^* - u^{(h/4)}_{LB}||_2 \approx C(\frac{h}{4})^p \tag{4.4}$$

Then the rate of convergence p is found by introducing an auxiliary variable q_h:

$$q^h = \frac{||\epsilon^h||_2}{||\epsilon^{h/2}||_2} = \frac{h^p 2^p}{h^p} = 2^p \tag{4.5}$$

The rate of convergence p is:

$$p = log_2(q^h) \tag{4.6}$$

This is going similarly for $q_h = \frac{\epsilon^{h/2}}{\epsilon^{h/4}}$ and lead to the same p. According to the terminology in section 3.3 this p is thus the p-value of the LB approximation.

When the exact solution is not present, the same method can be applied with a slight modification of $\epsilon^{h/2}$ and $\epsilon^{h/4}$:

$$||\epsilon^{h/2}||_2 = ||u^{(h)}_{LB} - u^{(h/2)}_{LB}||_2 \approx C(\frac{h}{2})^p \tag{4.7}$$

$$||\epsilon^{h/4}||_2 = ||u^{(h/2)}_{LB} - u^{(h/4)}_{LB}||_2 \approx C(\frac{h}{4})^p \tag{4.8}$$

4.4.1 Boundary time interval resolution

It is expected that at a certain point of coarsening the boundary interval resolution, the p-value of the LB approximation is not following its design accuracy anymore but drops to lower values. For the approximation this means that the boundary data error will dominate the total error.
The p-value should be around $p = 3$ for this LB approximation. Interpolation methods should hold the p-value as long as possible around $p = 3$ while in the meantime the boundary data interval resolution is lowered.

The p-value can be determined by two different methods. To avoid confusion, the methods are explained carefully in the following subsections. The reason there are two methods, is because the p-value can be obtained in different ways. The p-value is obtained by comparing the errors of increasing gridpoints according equation 4.2 till 4.6. But increasing the spatial gridpoints, leads to an increase of temporal gridpoints as well due to the CFL-condition. The boundary data interval resolution depends on the number of temporal gridpoints, thus is influenced as well.
The size of the update interval is the boundary data interval resolution and this depends again on the number of temporal gridpoints of the BB model.

To begin with, the size of the boundary data intervals have a certain value I_1. The first boundary data interval resolution is equal to the temporal gridsize of the BB approximation, thus $\Delta T = I_1^h$. The simulation is executed for three different spatial resolutions, all refined with a Grid Refinement Factor (GRF) 2: $h, h/2$ and $h/4$. Spatial grid refinement leads to temporal grid refinement as well, due to the CFL-condition. Thus spatial resolutions $h, h/2$ and $h/4$ lead to temporal gridsizes of $\Delta T, \Delta T/2$ and $\Delta T/4$ respectively. Thus the simulations of the different spatial resolutions begin with a boundary data interval resolution of $I_1^h = \Delta T,\ I_1^{h/2} = \Delta T/2$ and $I_1^{h/4} = \Delta T/4$.
Then the boundary data interval resolution is coarsened with a factor 2 as well, that means that the new size of the update intervals become respectively: $I_2^h = 2\Delta T,\ I_2^{h/2} = 2\Delta T/2 = \Delta T$ and $I_2^{h/4} = 2\Delta T/4 = \Delta T/2$. Coarsening the boundary data interval resolution leads to the following sizes of update intervals:

$$
\begin{gathered}
I_1^h = \Delta T,\ I_1^{h/2} = \Delta T/2 \text{ and } I_1^{h/4} = \Delta T/4 \\
I_2^h = 2\Delta T,\ I_2^{h/2} = 2\Delta T/2 = \Delta T \text{ and } I_2^{h/4} = 2\Delta T/4 = \Delta T/2 \\
I_4^h = 4\Delta T,\ I_4^{h/2} = 4\Delta T/2 = 2\Delta T \text{ and } I_4^{h/4} = 4\Delta T/4 = \Delta T \\
\vdots \\
\vdots \\
I_{N_T}^h = N_T \Delta T = |T_{N_T} - T_0| = |1 - 0|,\ I_{N_T}^{h/2} = 1/2 N_T \Delta T \text{ and } I_{N_T}^{h/4} = 1/2 N_T \Delta T
\end{gathered}
$$

Two methods are designed to deal with the dependence of the boundary data interval resolution. The first method, *constant time interval size*, keeps the size of the update interval of the values that are compared equal. The second method, *constant ratio*, keeps the ratio between the size of the update intervals of the values that are compared constant.

Constant time interval size

The rate of convergence requires both the error of resolution h and $h/2$ for its computation, see equation 4.5. In the constant interval case, the absolute value of the time interval size is kept in the choice for the errors ϵ^h and $\epsilon^{h/2}$.
This means that ϵ^h is obtained for boundary data interval resolution of size:

$$I_1^h = |\Delta T|,\ I_2^h = 2|\Delta T|, I_3^h = 4|\Delta T|,\ ...,\ I_{N_T}^h = |T_{N_T} - T_0| = |1 - 0|,$$

ϵ^h is compared with $\epsilon^{h/2}$, obtained from the simulations with a boundary data interval resolution of same absolute size. This means that the p-value is obtained by comparing the errors ϵ^h vs. $\epsilon^{h/2}$ on

$$I_1^h \text{ vs. } I_2^{h/2}, \quad I_2^h \text{ vs. } I_3^{h/2}, \quad ..., \quad I_{N_T-1}^h \text{ vs. } I_{N_T}^{h/2}.$$

At a certain point of interval resolution refinement, the design order of accuracy is lost. This method shows at what size of the interval this happens.

Constant ratio between time interval resolution

In this method the intervals are compared if the ratio of both intervals is equal to the grid refinement factor. The Grid Refinement Factor (GRF) in this case is: $grf = 2$ $(h,\ h_2 = h/2,\ h_4 = h_2/2)$.
Now the intervals which ratio is equal to GRF are compared, thus I^h vs. $I^{h/2}$ when $I^h/I^{h/2} = 2$. This means that the errors ϵ^h and $\epsilon^{h/2}$ are compared for the intervals:

$$I_1^h \text{ vs. } I_1^{h/2}\ , I_2^h \text{ vs. } I_2^{h/2}, \quad \ldots, \quad I_{N_T}^h \text{ vs. } I_{N_T}^{h/2}.$$

4.5 Experiments

Several experiments are drawn to investigate the main research question. These experiments are based on the issues involved within one-way nested regional climate modeling and properties of the atmospheric waves, because the current boundary data interval resolution is based on these properties [3]. The following issues and properties form the basis of the several experiments:

1. Spatial resolution difference between the driving data and the nested model
2. Spatial interpolation errors
3. Zonal wavenumbers

4.5.1 Spatial resolution difference

The spatial resolution difference is the factor between the size of the BB gridsize and LB gridsize, the Spatial Resolution Factor (SRF). The low resolution data of the BB model is injected into the high resolution LB model. The higher resolution should not differ too much from the BB solution, but the resolutions should also not be equal as it will not lead to any significant difference between the outcome of the two models. In current regional models, the SRF varies between $1 - 10$. At a certain factor, the refinement won't influence the LB solution in a positive way anymore. To test which

factor is this limit, the first experiments are based on the choice of the current models. Not all factors are implemented, because it is not interesting to find an exact limit of the the factor, but rather gain some insight of the limit-size. Spatial resolution factor 1, 2, 4, 5 and 10 are used as the first experiments.

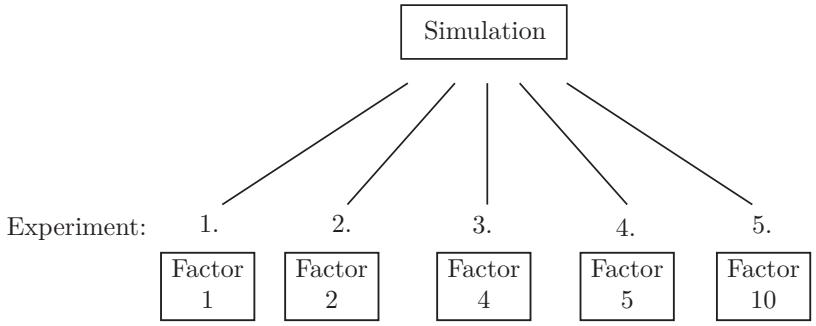

Figure 4.3: The first experiments are based on the foregoing factors between the spatial resolution of both models.

4.5.2 Spatial interpolation errors

Spatial interpolation errors can be present when the gridpoints of the BB domain are not matching with the boundary gridpoint of the LB domain. They arise when the boundary of the LB domain does not coincide with a gridpoint of its BB. Spatial interpolation is then needed to inject the boundary data into the LB. To test whether spatial interpolation influences the final solution significantly, two scenario options are created: one where the gridpoints are matching and no spatial interpolation is needed and one where the gridpoints are not matching. In the latter case, linear interpolation is used to approximate the solution at the boundary gridpoint z_1 from its adjacent BB gridpoints. The choice for linear interpolation is made, because this issue is not the main focus of the thesis and linear interpolation gives enough insight in the behaviour of the solution and is also the method used in operational climate modelling.

4.5.3 Zonal wavenumbers

The last experiment options come from the current rule of thumb that is used now as bound for the boundary data interval resolution. This depends on the zonal wavenumbers of the waves. To see whether physical properties of the wave really influence the solution, four scenarios are based on the zonal wavenumbers $k_z = 1,\ 3,\ 5$ and 8, explained in section 3.2.

4.6 GCM and RCM equivalence

The BB and LB models are equivalent to the Global and Regional Climate Models and thus so are its parameters. The wave characteristic parameters, the zonal wavenumber k_z and wavespeed α, are variables that are similar in both Global and Regional Models as in the Big and Little Brother experiment (BB-LBE). Other parameters are related to the domain of GCM and RCM.
The values of the BB-LBE simulations are scaled down to a workable value for small

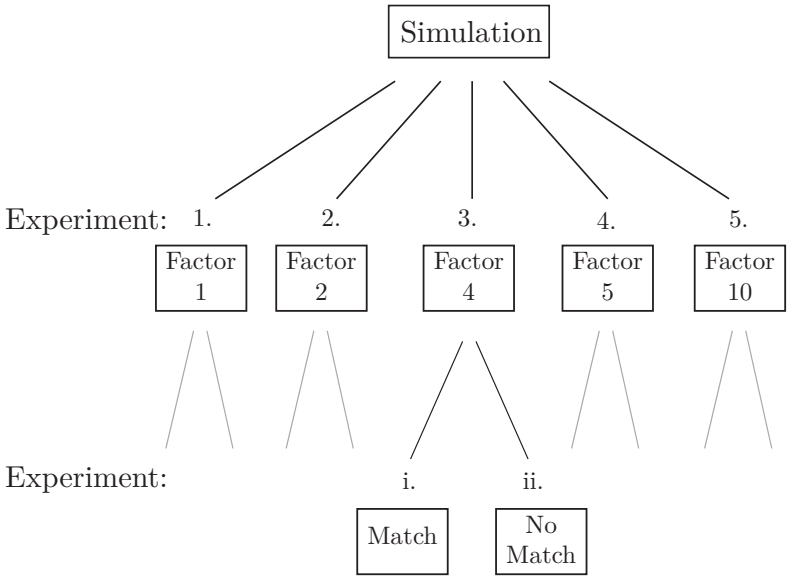

Figure 4.4: The first experiments branch into two options: experiments where the boundary matches the grid of the Big Brother domain and non-matching experiments where spatial interpolation is necessary.

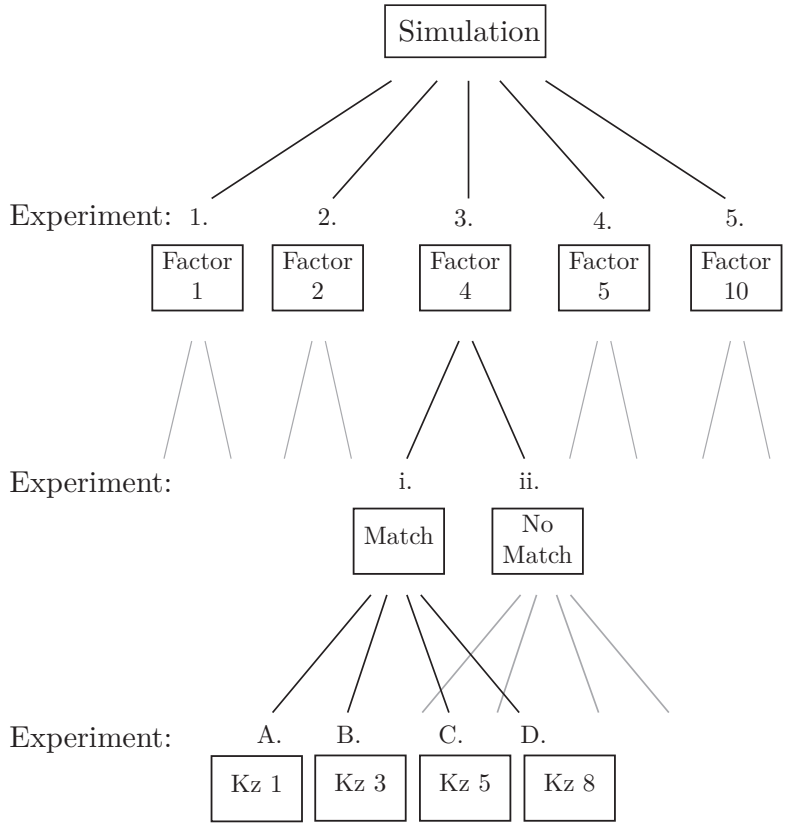

Figure 4.5: The second line of scenarios branches into four options: the zonal wavenumbers. Together this leads to 40 experiment options.

scale simulations, but still equivalent to the Global and Regional models. The parameters with an asterisk are the parameters of the GCM and RCM, the parameters without asterisk used in the simplified model here.

4.6.1 Global and Regional model parameters

The following parameters are used in the Global and Regional models:

- H^*: the gridsize of the GCM domain. It is related to N^*, the number of gridpoints: $N^* = L_1^*/H^* + 1$
- h^*: the gridsize of the RCM domain. Which is related to H^* by the Spatial Resolution Factor (SRF). Again H^* determines the the number of RCM gridpoints M^*. $M^* = L_2^*/h^* + 1$.
- L_1^*: the length of the GCM domain. Which is the perimeter of the earth.
- L_2^*: the length of the RCM domain. Which is an area of a certain part of the earth, eg Western-Europe, Scandinavia or Southern Africa.
- T^*: the simulation time.
- ΔT^*: the stepsize between two GCM timepoints
- Δt^*: the stepsize between two RCM timepoints
- the boundary data interval resolution. This is the resolution that is used now to update the RCM boundary condition in time by data of the GCM.

These parameters all have certain values that are used in the model. The size of H^* differs between 200 and 500 km. The length of the GCM domain L_1^* is the length of the earth's perimeter, which is $L_1^* = 40.000\ km$. This lead to N^* between 201 and 81. The RCM domain gridsize is related to the GCM size by the Spatial Resolution Factor (SRF). The SRF differs mainly from a factor $2 - 10$. The RCM gridsize is mostly around $h^* = 20\ km$, $h^* = 40\ km$ or $h^* = 50\ km$. The size of the RCM domain covers an area of size 5000 x 5000 km and in one dimension $L_2^* = 5000\ km$. This lead to a total number of gridpoints M^* between 101 and 251. The final simulation time T^* is related to the model simulation time T by $T^* = T\frac{L}{\alpha}$. The several waves thus relate the model simulation time to the physical simulation time. Hence, this is different per experiment. The same holds for the temporal stepsize.
The boundary data interval resolution that is used now is based on a physical rule of thumb, where the interval update is around 6 hours [3]. The boundary data interval resolution in this model is object of investigation and will be a varying variable in these experiments.

4.6.2 Big and Little Brother model parameters

The downscaled BB model has equivalent values to the GCM parameters. To get similar simulations, the number of gridpoints N have to be similar to the Global number of gridpoints N^*. Furthermore, the scaling in the nondimensional analyis has lead to a length of the BB domain of $L_1 = 1$. This lead to a BB gridsize of $H = 1/(N-1)$. The simulations are made for a number of gridpoints that differ with

a factor 2, so the size of N deviates between N is 51, 101, 201 and 401 gridpoints. The same Spatial Resolution Factor (SRF) from the global and regional models is used in the BB-LBE simulations. But not all factors are considered in the simulation, only the scenarios where SRF is 1, 2, 5 and 10, see section 4.5. This factor determines the LB gridsize: $h = H/SRF$. The length of the LB domain covers the same part of the BB domain as the RCM covers the GCM domain, which is around 1/8 of the GCM domain. In the LB experiment, the length of the LB domain is $L_2 = 0.2$. The simulation time is scaled to be $T = 1$ and the timestep is determined by the CFL condition. Thus should be at most the size of the spatial gridsize, thus $\Delta T \leq H$. In the simulations the temporal stepsize is reduced with a factor 80 to get better insight in the behaviour of the LBC. Thus $\Delta T = 1/80H$.

4.7 Set-up of the experiment

Figure 4.6 gives an overview of how the experiment is performed in one simulation. The characters A,B,C and D correspond to the characters in the figure.
A: The solution is computed on a large domain with low-resolution; the BB domain. The visible wave in A is a wave with zonal wavenumber 4 in its initial state.
B: A dataset is generated with boundary data for the LB domain. The 9 shots are snapshots of the BB solution on the LB domain, moving in time. The first shot is the initial condition, the second shot is after a certain time interval and the last shot is after one time period. The small square on the left boundary shows how the solution is stored in time. The dataset contains the solution of the BB on all its timesteps. The dataset can be filtered by removing data at certain intervals. This can be repeated until you end with a very coarse dataset. At a certain point in this repetitive process, the dataset contains too few information to generate an accurate LB solution. The optimal filter can be investigated in this step.
C: The most left graphic in C is a representation of the data set containing the boundary data in time for the LB solution. The data set has information about the boundary on all the intervals, but need information on all the LB timesteps. Interpolation methods are used to generate the full time-dependent Lateral Boundary Condition (LBC). Both first and third order interpolation methods are considered to see what generates the best results.
D: The LB solution is generated by the boundary data. Hypothesis is that the high-resolution boundary data will lead to a more accurate result then the low-resolution boundary data and that third order interpolation is able to hold more accurate approximations than linear interpolation.

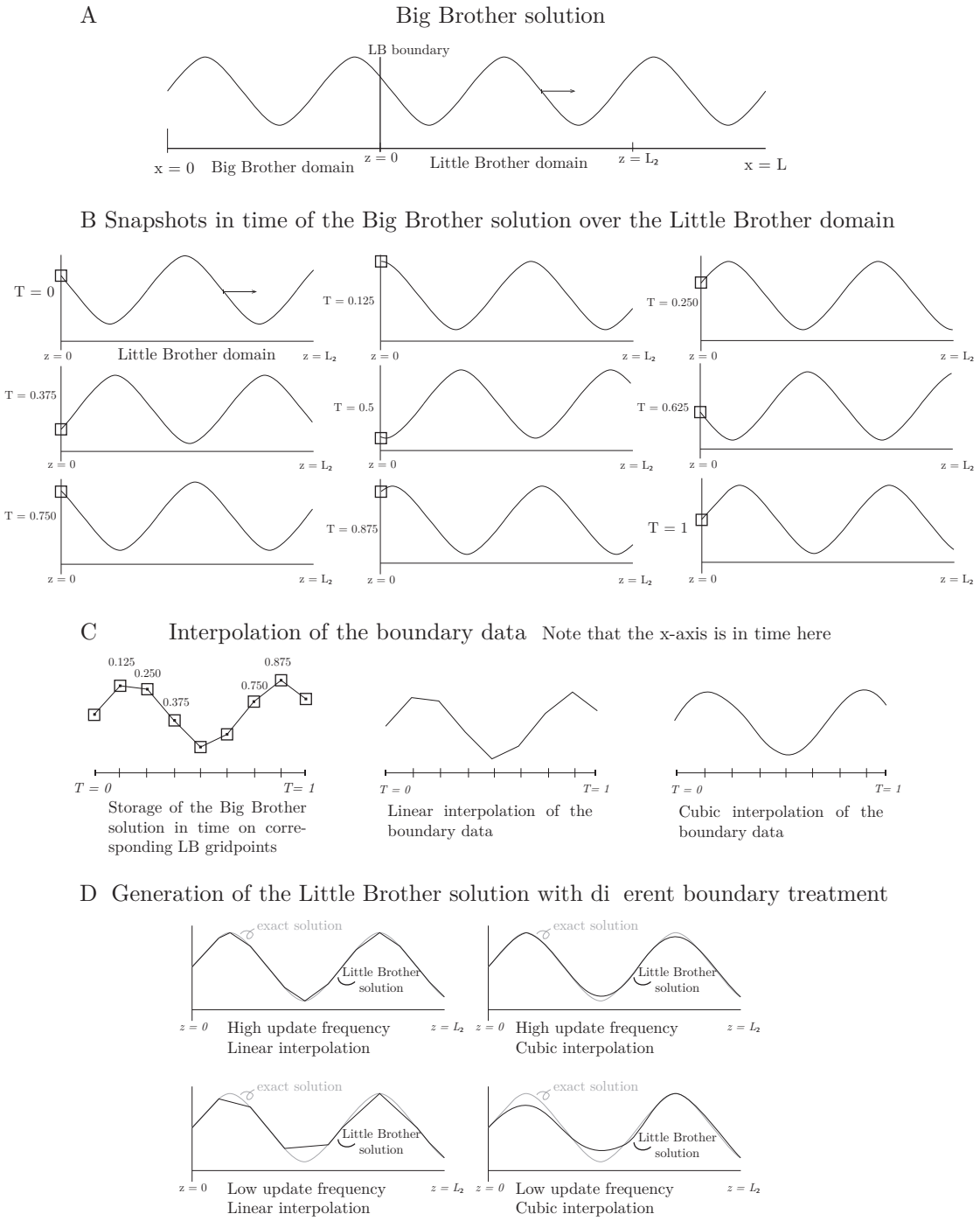

Figure 4.6: Visual of the experiment set-up:
A: The solution is computed on the Big Brother domain
B: The solution is stored in time on the boundary of the LB domain
C: Several interpolation techniques are used to form the LBC
D: Generation of the Little Brother solution

Chapter 5

Results

This section shows a part of the results of the performed experiments. Only a couple of experiments are shown, which are representative for the conclusion that can be drawn. The figures form a summary of the tables in Appendix C.

The following subsections show the L_2-norms for the several experiment options, sketched in section 4.5. Four scenarios are lighted :

- Scenario I is the simulation of $k_z = 3$, a SRF of 5 and matching gridpoints.
- Scenario II is the simulation $k_z = 1$, a SRF of 1 and matching gridpoints.
- Scenario III is the simulation $k_z = 8$, a SRF of 2 and matching gridpoints.
- Scenario IV is the simulation $k_z = 5$, a SRF of 2 and non-matching gridpoints.

These four scenarios give an indication of how the lateral boundary condition is influencing the final solutions. More experiments are performed to give an insight on how the zonal wavenumbers, spatial interpolation and spatial resolution difference are influencing the LB outcome. These can be found in Appendix C.
The results of the four scenarios are summarized in figures 5.1 till 5.5. Figure 5.1 is a summary of the L_2-norms of the regarded zonal wavenumbers with highest boundary data interval resolution possible. The columns show the absolute value of the L_2-norms for both linear and third order interpolation. The absolute value of the L_2-norm is shown to compare the accuracy of the LB outcome between linear and third order methods. Every wavenumber has the same accuracy for linear and third order interpolation, when the boundary data interval resolution is very high.

Figures 5.2 and 5.3 are a summary of tables C.11, C.12, C.13, C.14 and C.15. Figure 5.2 is a summary of the p-values of the linear results and figure 5.3 is a summary of the p-values of the third order results. The coloured lines represent the different wavenumbers. Every wavenumber is represented by two lines, because the L_2-norm is computed for $N = 51,\ 101$ and 201 and thus the p-value is computed between $N = 51$ vs. 101 and $N = 101$ vs. 201. The vertical axis represents the p-values, the horizontal axis represents the boundary data interval resolution. The lines are decaying as the boundary data interval resolution is lowered. The p-value is computed according to the comparison method in which the **ratio** between the intervals is kept constant.

The black dashed line shows the critical point where the p-value is not moving according its design accuracy anymore. For linear interpolation this happens at a boundary data interval resolution of $25\Delta T$, for third order interpolation this happens at a boundary data interval resolution of $100\Delta T$.
The grey line is a less critical margin, which is the lowest boundary data interval resolution that can be used to still obtain reasonable results.

Figures 5.4 and 5.5 are a summary of tables C.6, C.7, C.8, C.9 and C.10. Figure 5.4 is a summary of the p-values of the linear results and figure 5.5 is a summary of the p-values of the third order results. The coloured lines represent the different wavenumbers. Here again, the wavenumbers are represented by two lines. The vertical axis represents the p-values, the horizontal axis represents the boundary data interval resolution. In this case, the p-value is computed according to the comparison method in which the **interval size** is kept constant.
This comparison method leads to stricter bounds on the boundary data interval resolution. The black dashed line of linear interpolation shows that a boundary data interval resolution of $10\Delta T$ is the critical bound and for third order this is a boundary data interval resolution of $50\Delta T$. The grey dashed line leads to a less critical bound of a boundary data interval resolution of $20\Delta T$ and $100\Delta T$ respectively.

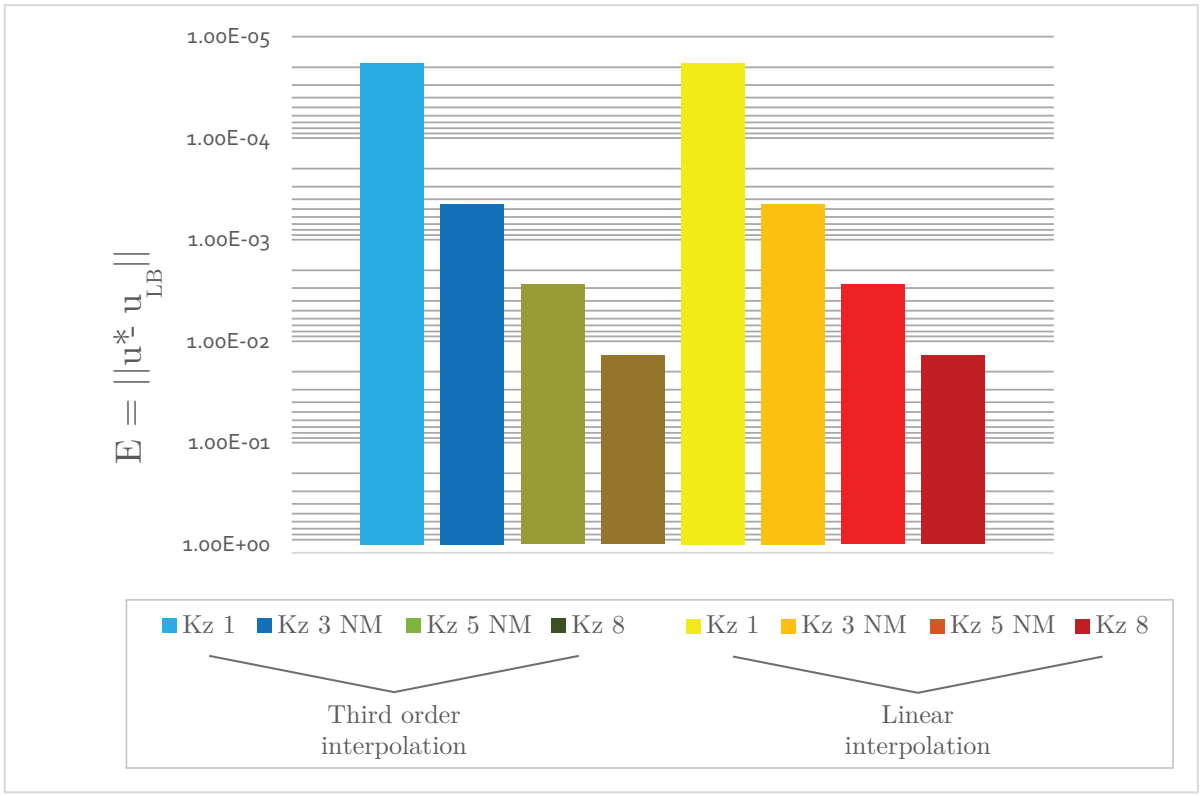

Figure 5.1: This graph is a summary of the norm tables from section C.1. It is an overview for the L_2-norms of the error between the LB approximation and the exact solution for the highest boundary data interval resolution, a boundary data interval resolution of $1\Delta T$. The norms are shown for $N = 201$ and zonal wavenumbers 1, 3, 5 and 8. (The NM means a non-matching scenario).

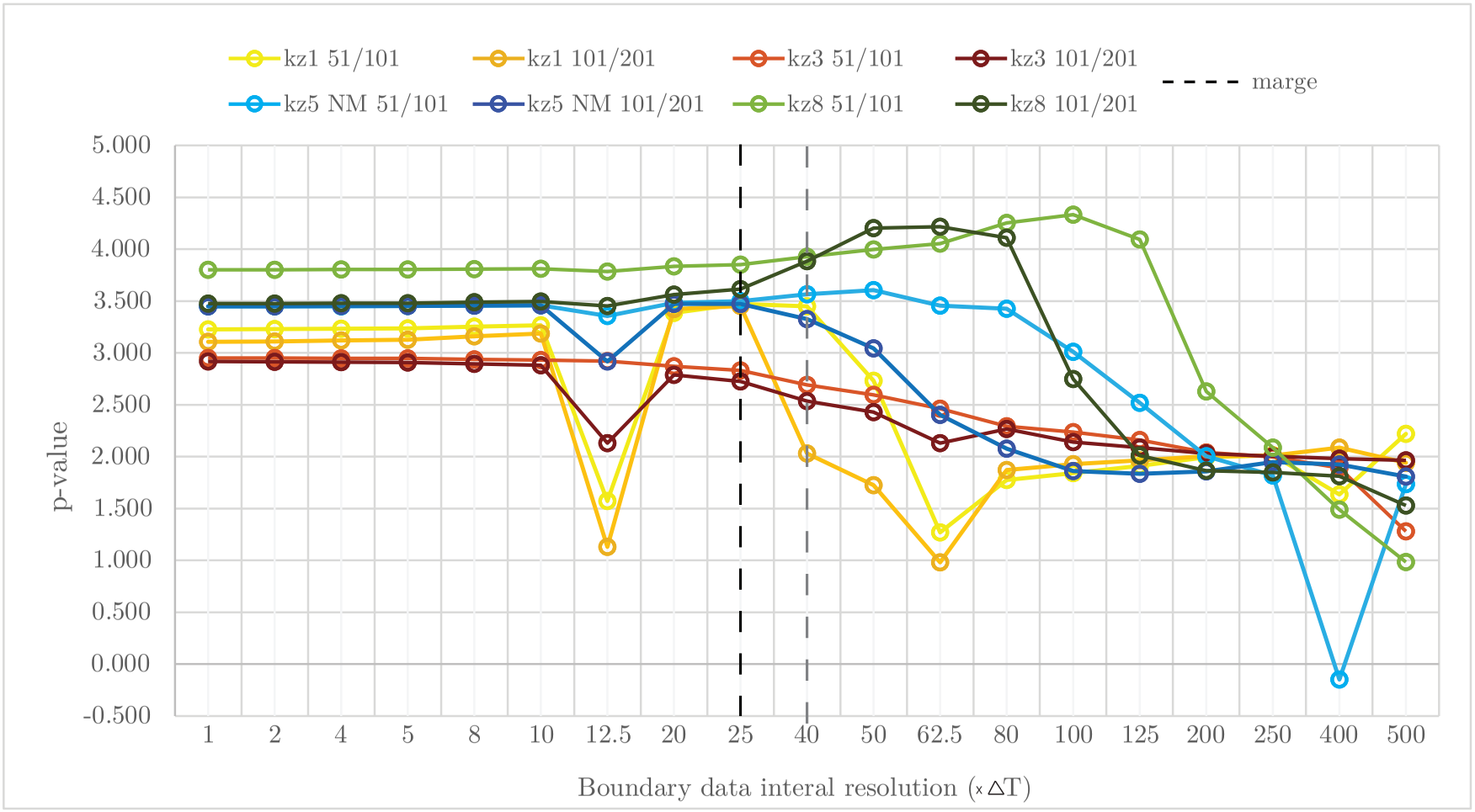

Figure 5.2: p-values of the four wavenumbers (kz1 means zonal wavenumber 1) compared for constant ratio, as in section C.2. Linear interpolation is used. The black dashed line shows the optimal combination.

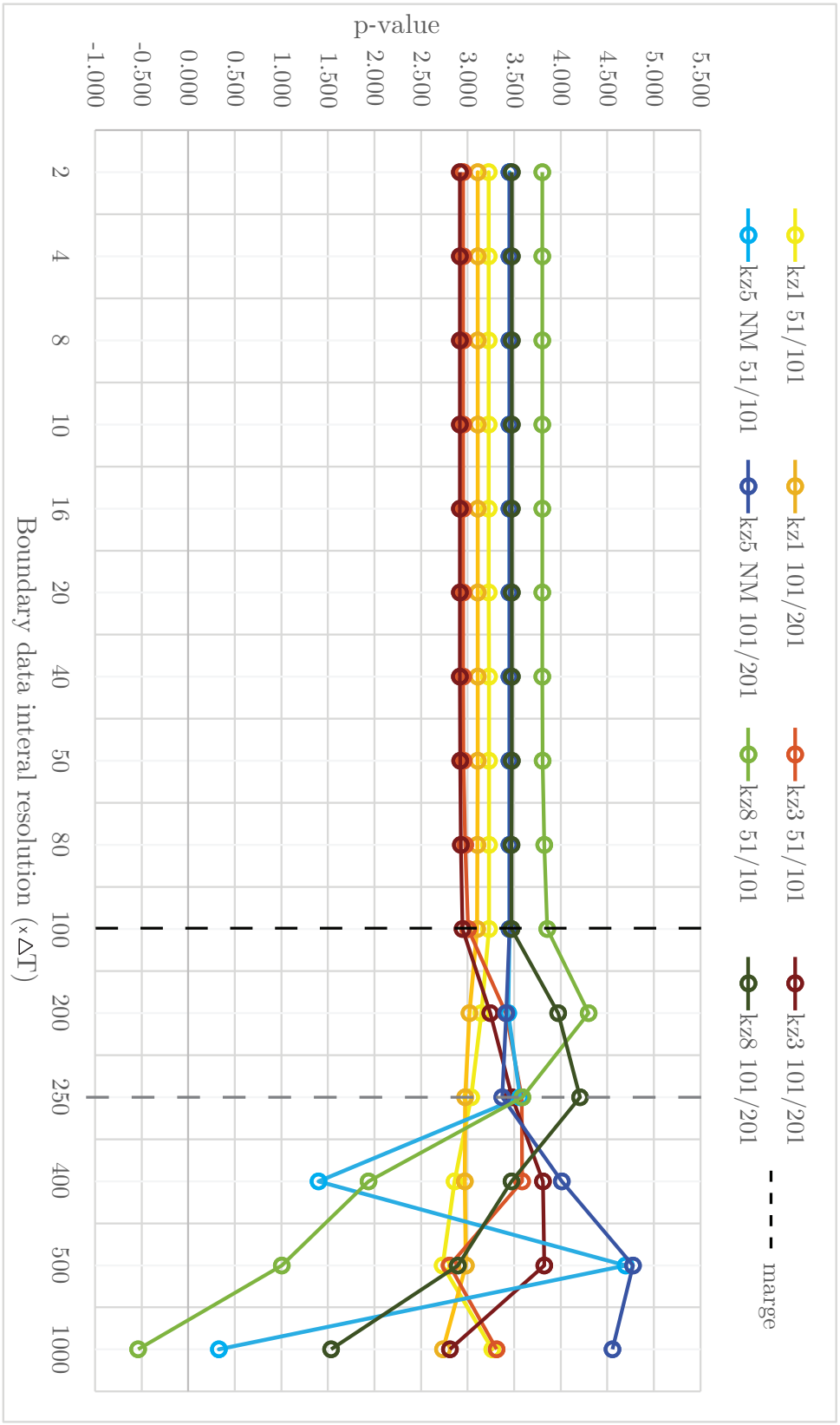

Figure 5.3: p-values of the four wavenumbers compared for constant ratio, as in section C.2. Third order interpolation is used. The black dashed line shows the optimal combination.

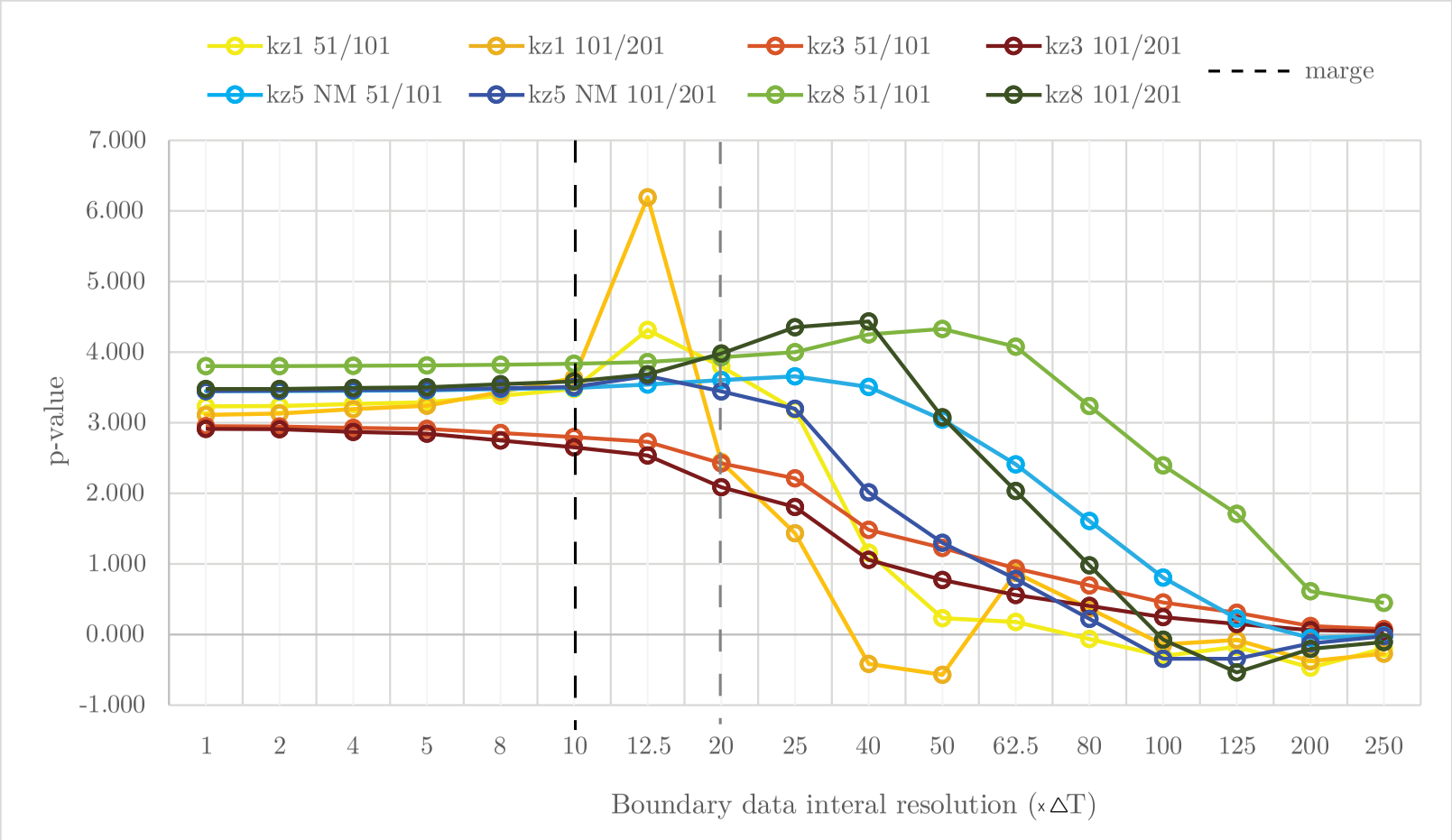

Figure 5.4: Summary of the p-values of the four wavenumbers compared for constant interval, as in section C.3. Linear interpolation is used. The black dashed line shows the optimal combination.

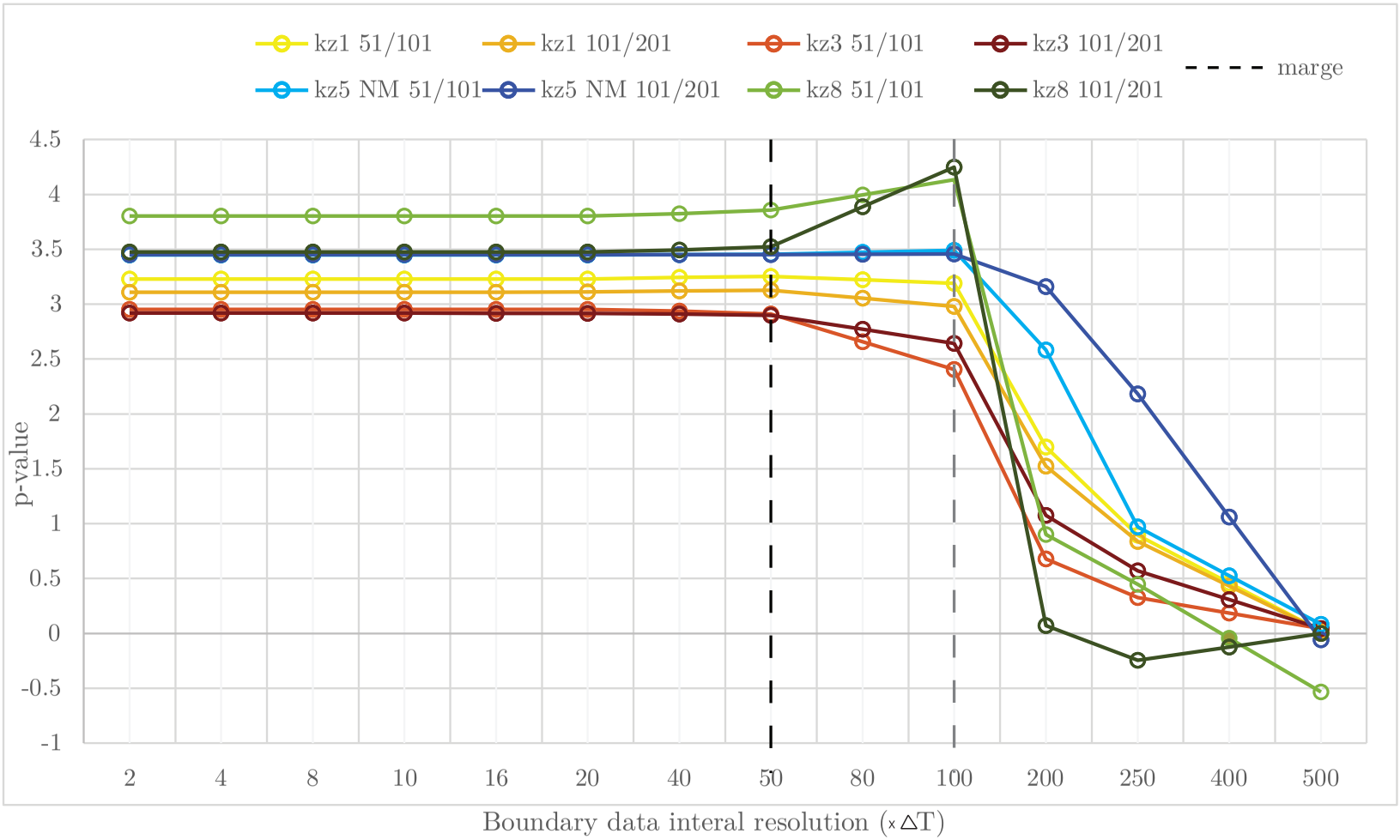

Figure 5.5: p-values of the four wavenumbers compared for constant interval, as in section C.3. Third order interpolation is used. The black dashed line shows the optimal combination.

Chapter 6

Conclusion

The goal of the thesis was to investigate the influence of the interpolation method and boundary data interval resolution of the Lateral Boundary Condition. To find an answer to it, the main research question has been divided in six sub-questions. The research and sub-questions were as follows:
What is the best compromise between boundary data interval resolution and interpolation method that leads to the desired result for the regional model output?

- How can the model be simplified in order to investigate the involved issues while in the meantime being simple enough to work with?
- What are the physical conditions that should be captured by the full model?
- How can the argumentation of the desired result be based on computational properties instead of physical properties?
- How does altering boundary data resolution of the model (time interval of update) impact the final solution of the regional model?
- Until what extend does changing interpolation techniques impact the solution of the regional model?
- What is the combination of boundary data interval resolution and interpolation method that leads to the desired result?

Sub-question 1 is answered in section 3.1 when the Navier-Stokes equations were simplified to the one-dimensional Transport Equation. Sub-question 2 is answered in section 3.2, which describes the wave characteristics of the atmospheric waves. Sub-question 3 is answered in section4.4, which describes two methods to validate the outcome. The other questions will be answered here.

6.1 Evaluation of the boundary data interval resolution

From figures 5.2-5.5 it can be concluded that on a certain point the resolution of the boundary data update becomes critical. The approximation looses its accuracy

such that its p-value doesn't coincide anymore with the expected rate of convergence according the SBP-SAT-operators. The p-values are moving horizontally, until at a certain point, the critical point, the boundary data interval resolution becomes too low. After this point, the values are acting strangely. They are moving up and down, finally resulting in a p-value around $p = 0$. After this critical point, the LB simulation is not able to correct for the temporal interpolation of boundary data. The errors caused by the inaccurate LBC are dominating the final outcome, resulting in a total loss of accuracy.
Before this critical point, the boundary data interval resolution is not influencing the outcome. This is visible in figure 5.1. The values of the same-coloured columns in this figure are equal for a boundary data interval resolution of $1\Delta T$. This means that the L_2 norm of the error between the exact solution and the LB approximation is the same.
The assessment of the boundary data interval resolution on its own is not important. It is the combination with the interpolation method that shows till what extend the resolution can be lowered.

6.2 Interpolation method influence

Figures 5.2-5.5 show that the third order interpolation method is able to gain the expected rate of convergence much longer along a lower boundary data interval resolution than linear interpolation methods. The p-values of third order interpolation remain on the same horizontal line for a lower boundary data interval resolution than for linear interpolation. Although this is visible for both the assessment methods of constant ratio and constant interval, there is a difference in at what points the p-values drop.
Assessment according a constant ratio (visible in figures 5.2 and 5.3) leads to a boundary resolution for respectively linear and third order interpolation of:

$$25\Delta T \text{ vs } 100\Delta T$$

The less critical bound (the grey dotted lines) result in :

$$40\Delta T \text{ vs } 250\Delta T$$

Assessment according a constant interval leads (visible in figures 5.4 and 5.5) to a boundary resolution for respectively linear and third order interpolation of:

$$10\Delta T \text{ vs } 50\Delta T$$

The results of the less critical bounds are:

$$20\Delta T \text{ vs } 100\Delta T$$

It also visible in figure 5.1 that interpolation on its own does not lead to a more accurate LB approximation! The norms in figure 5.1 show that linear and third order interpolation are equal, the only difference is that third order interpolation is able to hold the accurate LB outcome for a lower boundary data interval resolution.

6.3 Examination of the several experiments

Several experiments are performed to test whether the issues from section 2.2 would be of great importance. These issues gave rise to a set of experiments, but not all the drawn experiments are performed. Only a sample of experiments is performed, to give an idea of their influence. The p-values of the drawn experiments are given in tables 6.1 and 6.2.

Table 6.1: Overview of the p-values for several different experiments scenarios. The shown boundary data interval resolution is $1\Delta T$. These p-values are for the linear interpolation method. The values in the table can read as follows: the value 3.2 is the p-value for $k_z = 1$ a spatial resolution difference of $SRF = 1$ and matching boundary points. The most right column are the values of the BB simulation, which can be used as reference value.

		Match				No-match				
	SRF	1	2	5	10	1	2	5	10	BB
k_z	1	3.2		3.0						3.0
	3			2.9	2.9				2.8	3.1
	5		3.4				3.4	3.4		3.4
	8	3.8	3.8							3.8

Table 6.2: Overview of the p-values for several different experiments scenarios. These p-values are for the third order interpolation method.

		Match				No-match			
	SRF	1	2	5	10	1	2	5	10
k_z	1	3.2		2.9					
	3			2.9	2.9				2.8
	5		3.4				3.4	3.4	
	8	3.8	3.8						

6.3.1 Spatial resolution difference

The spatial resolution difference, in the tables indicated by SRF 1, 2, 5 and 10, has almost no influence on the p-values. Only in zonal wavenumber 1 is the p-value different for a spatial resolution factor of 1 and 5. For the other zonal wavenumbers the p-value is the same for all the factors.

6.3.2 Horizontal interpolation errors

The horizontal interpolation errors that can be caused in case the gridpoint of the LB boundary is not matching with a gridpoint of the BB model (indicated by match and no-match) are not influencing the p-value. The p-value is exactly the same for the both a matching and non-matching set-up. The L_2-norm of the error does change in case the gridpoints are not maching. The non-matching outcome is slightly more

accurate. In the case of zonal wavenumber 5 and spatial resolution factor of 2 has the matching experiment a norm of $E_2 = 3.19E-04$ for $N = 201$ gridpoints and a norm of $E_2 = 2.52E-04$ for the non-matching experiment.

6.3.3 Zonal wavenumbers

The zonal wavenumbers all have a different output of the p-value. The highest zonal wavenumber has a p-value of 3.8 in relation to a p-value of 2.9 for $k_z = 3$. I don't know due to which grounds this can be explained.
The zonal wavenumbers have impact on the accuracy in absolute norm. Figure 5.1 shows that the norms all differ with a factor 10 from each other. That means that the output of $k_z = 8$ is 1000 times less accurate than the output of $k_z = 1$! This is significant. This can easily be explained by the fact that the it is much harder to approximate 8 ridges and troughs in one domain of $k_z = 8$, than the one ridge of $k_z = 1$.

Most importantly is the question whether the zonal wavenumber influences the retaining of the p-value. Figures 5.2-5.5 show that this is not the case. Although the p-values differ for different wavenumbers, they retain all until the exact same boundary data interval resolution.

6.4 Final Conclusion

The conclusion from this experimental set-up should be based on the performance of the p-values. The final conclusion is thus not based on the L_2-norms of the errors between the exact solution and the LB output, but only on how the p-values are holding up. It is convenient for the reader to take a look again at figures 5.2-5.5.
From these figures it can be concluded that in case of linear interpolation, the lowest boundary data interval resolution can be between $10\Delta T$ and $25\Delta T$.
In case of third order interpolation the lowest resolution can be between $50\Delta T$ and $100\Delta T$.
This holds for all the wavenumbers, spatial resolution differences and matching and non-matching boundary points. A used boundary data interval resolution between these values should lead to the best obtainable accuracy, while using less computational time.

The boundary data interval resolution depends on the chosen temporal resolution of the BB model. The chosen temporal resolution is $\Delta T = 1/80h$. But a lower temporal resolution doesn't lead to less accurate approximations. It is thus also possible to choose a BB temporal resolution of $\Delta T = 1/8h$ and a boundary data interval resolution of $1\Delta T$. This saves even more computation time. In the case of third order interpolation, the BB temporal resolution can even be lowered to: $\Delta T = 5/8h$ and a boundary data interval resolution of $1\Delta T$.

6.4.1 What does this mean for the full Climate Models

To relate the conclusion to the final model, the boundary data interval resolution needs to be translated to the physical workable values, explained in section 4.6 and finally to the full model, the Navier-Stokes equations.
The meaning of a boundary data interval resolution of $10 - 25\Delta T$ and $50 - 100\Delta T$

Table 6.3: Final conclusion. The values in the table stand for the boundary data interval resolution.

	Constant interval	Constant ratio
Linear	10	25
Third order	50	100

have to be translated to the workable values. The model time period T was in the non-dimensional analysis scaled to $T = L/\alpha/2\pi$.

The values for α depend on the wavenumbers. $k_z = 1$ and $k_z = 3$ are investigated for $\alpha = 20\ ms^{-1}$ and $k_z = 5$ and $k_z = 8$ are investigated for $\alpha = 3\ ms^{-1}$.

The physical value of parameter L is $L = 40000000\ m$. A wavespeed of $\alpha = 3\ ms^{-1}$ and a temporal stepsize of ΔT of $\Delta T = 1/N_T = 1/4000$ lead to a physical temporal resolution of

$$\Delta T^* = 1/4000 * 40000000/3/2\pi \approx 500\ s \text{ or around } 8.5 \text{ minutes.}$$

A boundary data interval resolution of $10\Delta T$, means a resolution of $1h30$. This is significantly less than the current rule of thumb of 6 hours. A less strict bound (the grey dashed line) is a boundary data interval resolution of $25\Delta T$ or $3h45$.
In case of third order interpolation it could be increased to $50\Delta T$, which should lead to a resolution of almost $7h30$ hours or with the less strict bound to even $100\Delta T$ 15 hours.
A wave speed of $\alpha = 20\ ms^{-1}$ means a physical resolution of $\Delta T^* = 1/4000 * 40000000/20/2\pi \approx 80\ s$, which is around 1.5 minutes. The physical boundary data interval resolution would change to a value between $0h15$ and $0h40$ for linear interpolation. For third order interpolation it would change to a value between $1h15$ and $2h30$.
There is thus a significant difference in the choice of the physical boundary data interval resolution which depends on the wavespeed.
The nowadays used rule of thumb to update the boundary condition was around 6 hours. The combination of this rule of thumb with linear interpolation would lead to a decrease of accuracy of the final outcome for all the wavenumbers. A combination with third order interpolation would gain the desired accuracy in the case of the low wavespeed $\alpha = 3\ ms^{-1}$. On top I would like to add that a higher resolution than the here stated bound, would not lead to more accurate approximations. Thus a boundary data interval resolution around the stated bound is desirable to avoid useless computational time and storage.
Furthermore, the experiments with several characteristic wave properties have shown that the $BB - LBE$ simulations are less accurate as the wavenumbers increases, but that the p-values are acting similar for all the wavenumbers. This means that it is not a wise idea to choose a higher boundary data interval resolution for a higher wavenumber, because the approximation cannot exceed a certain accuracy.

6.5 Notes on the conclusion

The research is conducted for a simplified problem of the global and regional climate models. This simplified problem shows how the boundary data interval resolution and interpolation methods are influencing the solutions. If it is already visible in this simplified problem, then it is definitely visible in the full model. But a note should be made that this simplified model cannot be used to determine the exact value of the boundary data interval resolution in combination with the interpolation method. This research should be used to confirm the hypothesis that third order interpolation methods are able to longer retain an accurate solution. Another conclusion is that third order interpolation does not lead to more accurate solutions on its own! Linear and third order interpolation method lead to the same accuracy as long as the boundary data interval resolution is high enough.
To determine the exact value of the boundary data interval resolution the thesis can be repeated for the full model. The several experiment options have shown that they do not influence the final approximation. Redoing the experiments for a matching and non-matching scenarios or for several spatial resolution differences is thus unnecessary.

Another note I would like to make is that the determined wave characteristics are a result of a literature study. It was hard to determine which waves are exactly of importance and which characteristics are used in the full climate models. Especially the wave speed is important in the translation of the model value to the physical value. The most extreme conditions are chosen, but if they are too harsh this would mean that the physical update resolution is less strict than stated now. It is important to conduct more research on this topic.

Appendix A

Spatial interpolation errors

In case of a Little Brother boundary that is not matching with Big Brother gridpoints, spatial interpolation is necessary to transfer the output of the BB model to the LB model to form the lateral boundary condition. Spatial interpolation is done with a linear method, because it is not the main focus to investigate whether this causes any harm to the solution, it is more to give an insight of what would happen if they are not matching.
Linear interpolation transfers the BB output on two adjacent points (x_{LB1} and x_{LB2}) of the LB boundary to the gridpoint z_0 of the LB boundary with the following formula:

$$g_{LB}(z_0) = u_{BB}(x_{LB1}) + \frac{u_{BB}(x_{LB2}) - u_{BB}(x_{LB1})}{x_{LB2} - x_{LB1}} \tag{A.1}$$

The vector $g_{LB}(z_0)$ is not yet the lateral boundary condition! This vector contains only temporal information on the update intervals and still requires temporal interpolation before it can be used as the LBC of the Little Brother model.
The adjacent points x_{LB1} and x_{LB2} of the boundary z_0 can be found with the following algorithm:

Algorithm 1 Finding the adjacent points of the LB boundary

```
N ← number of spatial BB gridpoints
Δx ← gridsize of the BB domain
for i = 1 : N do
    if iΔx > z_0 − Δx  &  iΔx < z_0 + Δx  then
        x_LB1 ← iΔx
        x_LB2 ← (i + 1)Δx
    end if
end for
```

Appendix B

Estimation of the error

In case the exact solution is not present to use for validation, an estimate has to be made on the LB output to investigate the LBC properties. This estimation should estimate the error between the numerical approximation from the LB model and the exact solution. This can be done by unraveling what exactly is going on in the Runge-Kutta solver.
To investigate the error-norm, several terms have to be taken into account to get a clear idea how much the boundary term is influencing the error. Therefore we have precisely evaluated the equation the following way:

$$u_t + \alpha P^{-1} Q u = \sigma P^{-1} E_0 (u - g) \tag{B.1}$$

$$\Leftrightarrow u_t = \underbrace{(-\alpha P^{-1} Q + \sigma P^{-1} E_0)}_{A} u - \underbrace{\sigma P^{-1} E_0 g}_{b} \tag{B.2}$$

Thus the spatial discretization can be written as

$$u_t = Au - b(t). \tag{B.3}$$

The notation $u^n = u(t_n)$ and $b^n = b(t_n)$ and $b^{n+\frac{1}{2}} = b(t_n + \frac{1}{2}\Delta t)$ is introduced. If the whole numerical method is considered, it can be written as

$$\begin{aligned} u^{n+1} = &\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4 \right) u^n \\ &- \left(\frac{1}{6}\Delta t I + \frac{1}{6}\Delta t^2 A + \frac{1}{12}\Delta t^3 A^2 + \frac{1}{24}\Delta t^4 A^3 \right) b^n \\ &- \left(\frac{2}{3}\Delta t I + \frac{1}{3}\Delta t^2 A + \frac{1}{12}\Delta t^3 A^2 \right) b^{n+1/2} - \Delta t \frac{1}{6} b^{n+1} \end{aligned}$$

The notation $A' = I + \Delta t A + \frac{1}{2}\Delta t^2 A^2$ is introduced for convenience:

$$u^{n+1} = \left(\underbrace{I + \Delta t A + \frac{1}{2}\Delta t^2 A^2}_{A'} + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4 \right) u^n$$
$$- \frac{\Delta t}{6}\left(\underbrace{I + \Delta t A + \frac{1}{2}\Delta t^2 A^2}_{A'} + \frac{1}{4}\Delta t^3 A^3 \right) b^n$$
$$- \frac{2}{3}\Delta t \left(\underbrace{I + \Delta t A + \frac{1}{2}\Delta t^2 A^2}_{A'} - \frac{1}{2}\Delta t A - \frac{3}{8}\Delta t^2 A^2 \right) b^{n+1/2} - \Delta t \frac{1}{6} b^{n+1}$$

And even more compact by introduction of A, B, C and D:

$$u^{n+1} = \underbrace{\left(A' + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4 \right)}_{C} u^n - \underbrace{\frac{\Delta t}{6}\left(A' + \frac{1}{4}\Delta t^3 A^3 \right)}_{D} b^n$$
$$- \underbrace{\frac{2}{3}\Delta t \left(A' - \frac{1}{2}\Delta t A - \frac{3}{8}\Delta t^2 A^2 \right)}_{E} b^{n+1/2} - \underbrace{\Delta t \frac{1}{6} I}_{F} b^{n+1}$$

Thus in short:

$$u^{n+1} = Cu^n - Db^n + Eb^{n+1/2} + Fb^{n+1} \tag{B.4}$$

This equation can be seperated by evaluating the temporal steps seperately. That means that it can be written as follows: You can also write it as follows: For $n = 0, \ldots, N_T$:

$$\begin{aligned}
u^{n+1} &= Cu^n - Db^n + Eb^{n+1/2} + Fb^{n+1} \\
\Leftrightarrow u^{n+1} &= C^2 u^{n-1} - CDb^{n-1} + (CF - D)b^n + Fb^{n+1} + CEb^{n-1+1/2} + Eb^{n+1/2} \\
u^{n+1} &= C^3 u^{n-2} - C^2 Db^{n-2} - CDb^{n-1} + (CF - D)b^n + C^2 Fb^{n-2+1} \\
&\quad + Fb^{n+1} + CEb^{n-1+1/2} + C^2 Eb^{n-2+1/2} + Eb^{n+1/2} \\
u^{n+1} &= C^4 u^{n-3} + Fb^{n+1} + (CF - D)b^n + C(CF - D)b^{n-1} + C^2(CF - D)b^{n-2} - C^3 Db^{n-3} \\
&\quad + Eb^{n+1/2} + CEb^{n-1+1/2} + C^2 Eb^{n-2+1/2} + C^3 Eb^{n-3+1/2} \\
u^{n+1} &= C^4 u^{n-3} + C^0 Fb^{n+1} + CFb^n + C^2 Fb^{n-1} + C^3 Fb^{n-2} - C^0 Db^n - CDb^{n-1} - C^2 Db^{n-2} \\
&\quad - C^3 Db^{n-3} + C^0 Eb^{n+1/2} + CEb^{n-1+1/2} + C^2 Eb^{n-2+1/2} + C^3 Eb^{n-3+1/2} \\
&\quad \ldots
\end{aligned}$$

Until the equation can be regarded as the following sum:

$$u^{n+1} = C^{n+1} u^0 + \sum_{i=0}^{n} C^i F b^{n+1-i} + \sum_{i=0}^{n} C^i D b^{n-i} + \sum_{i=0}^{n} C^i E b^{n-i+1/2} \tag{B.5}$$

If characters A, B, C and D are replaced back by their original values, you end up with the following very long equation:

$$
\begin{aligned}
u^{n+1} = &\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4\right)^{n+1} u^0 \\
&+ \frac{\Delta t}{6}\sum_{i=0}^{n}\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4\right)^{i} b^{n+1-i} \\
&+ \frac{\Delta t}{6}\sum_{i=0}^{n}\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4\right)^{i} \\
&\qquad \left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{4}\Delta t^3 A^3\right) b^{n-i} \\
&+ \frac{2\Delta t}{3}\sum_{i=0}^{n}\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + \frac{1}{6}\Delta t^3 A^3 + \frac{1}{24}\Delta t^4 A^4\right)^{i} \\
&\qquad \left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 - \frac{3}{8}\Delta t^2 A^2\right) b^{n-i+1/2}
\end{aligned}
$$

If then A and b are replaced by their equation, then it becomes even longer:

$$
\begin{aligned}
u^{n+1} = &\left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) - \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad - \frac{1}{6}\Delta t^3\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^3 \\
&\qquad \left. - \frac{1}{24}\Delta t^4\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^4\right)^{n+1} u^0 \\
&- \frac{\Delta t}{6}\sum_{i=0}^{n}\left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) + \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad + \frac{1}{6}\Delta t^3\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^3 \\
&\qquad \left. + \frac{1}{24}\Delta t^4\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^4\right)^{i} \left(\sigma P^{-1}E_0 g\right)^{n+1-i} \\
&- \frac{\Delta t}{6}\sum_{i=0}^{n}\left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) + \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad \left. + \frac{1}{6}\Delta t^3\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^3 + \frac{1}{24}\Delta t^4\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^4\right)^{i} \\
&\qquad \left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) + \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad\qquad \left. + \frac{1}{4}\Delta t^3\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^3\right)\left(\sigma P^{-1}E_0 g\right)^{n-i} \\
&- \frac{2\Delta t}{3}\sum_{i=0}^{n}\left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) + \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad \left. + \frac{1}{6}\Delta t^3\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^3 + \frac{1}{24}\Delta t^4\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^4\right)^{i} \\
&\qquad \left(I + \Delta t\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right) + \frac{1}{2}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right. \\
&\qquad\qquad \left. - \frac{3}{8}\Delta t^2\left(-\alpha P^{-1}Q + \sigma P^{-1}E_0\right)^2\right)\left(\sigma P^{-1}E_0 g\right)^{n-i+1/2}
\end{aligned}
$$

Now we have separated the norm in several terms, we can verify our solution by this approximation.

B.1 Finding a bound on the error

To find a bound on the error, we take a closer look at the compact version:

$$u^{n+1} = C^{n+1}u^0 - \sum_{i=0}^{n} C^i D b^{n-i} - \sum_{i=0}^{n} C^i E b^{n-i+1/2} - \sum_{i=0}^{n} C^i F b^{n+1-i} \tag{B.6}$$

With n the considered timestep, u^0 the initial condition and b is the boundary term. Take a closer look at this term:

$$b = \sigma P^{-1} E_0 g \tag{B.7}$$

with g a vector with boundary data on the first row of the vector.

An estimation has to be found for the error between the exact solution (which is not present) and the LB output.
Firstly, the term that corresponds to the initial condition can be neglected, because after a certain time, the initial condition shouldn't influence the final solution anymore. The error between the exact solution and the Little Brother solution can be written as follows then:

$$\begin{aligned} ||u(t_{n+1}) - u^{n+1}|| = &- \sum_{i=0}^{n} C^i D ||b^{n-i} - b_{BB}^{n-i}|| \\ &- \sum_{i=0}^{n} C^i E ||b^{n-i+1/2} - b_{BB}^{n-i+1/2}|| \\ &- \sum_{i=0}^{n} C^i F ||b^{n+1-i} - b_{BB}^{n-i+1}|| \end{aligned} \tag{B.8}$$

The term between norm brackets is the error between the BB boundary data and the time interpolated solution. Now an estimation of this term can be found.

B.1.1 Estimation of the bound due to linear interpolation

The error due to linear interpolation is explained in section 3.4.3 and is estimated by (3.44), which is:

$$\xi(t) \leq \frac{k^2}{2!}\left(\frac{\Delta T}{2}\right)^2 = \frac{k^2}{8}(\Delta T)^2. \tag{B.9}$$

With k_z the zonal wavenumber and ΔT the BB timestep. This term is a bound on the boundary term g in (B.7) in case of linear interpolation. The first term on the RHS of (B.8) can then be bounded by:

$$\sum_{i=0}^{n} C^i D ||b^{n-i} - b_{BB}^{n-i}|| \;\leq\; D\frac{k^2}{8}\sigma P^{-1} E_0 (\Delta T)^2 \sum_{i=0}^{n} C^i \tag{B.10}$$

The $D\frac{k^2}{8}\sigma P^{-1}E_0(\Delta T)^2$ stands now before the sum, because we are intersted in an estimation of the error, so the maximum term.
Due to the simplifications made by estimating the second derivative of the cosinus and using constant timesteps, the same errorbounds are eligible for the second and third

term of the RHS.
This leads to a total estimate of:

$$||u(t_{n+1}) - u^{n+1}|| \leq -(D + E + F)\frac{k^2}{8}\sigma P^{-1} E_0 (\Delta T)^2 \sum_{i=0}^{n} C^i \tag{B.11}$$

Where $D + E + F = \Delta t C$. This is an analytic bound, which tells us how much the error due to interpolation is influencing the final solution.

B.1.2 Estimation of the bound due to third order interpolation

A similar bound can be accomplished for the natural cubic spline interpolation. This becomes

$$||u(t_{n+1}) - u^{n+1}|| \leq -(D + E + F)\frac{k^4}{384}\sigma P^{-1} E_0 (\Delta T)^4 \sum_{i=0}^{n} C^i \tag{B.12}$$

Appendix C

Results

This section shows a part of the results of the performed experiments. Only a couple of scenarios are shown, which are representative for the conclusion that can be drawn. The tables show the L_2 norm of the error between the exact solution and Little Brother solution for three different resolutions (columns from left to right) and different boundary interval resolution (rows). The most left column shows the size of

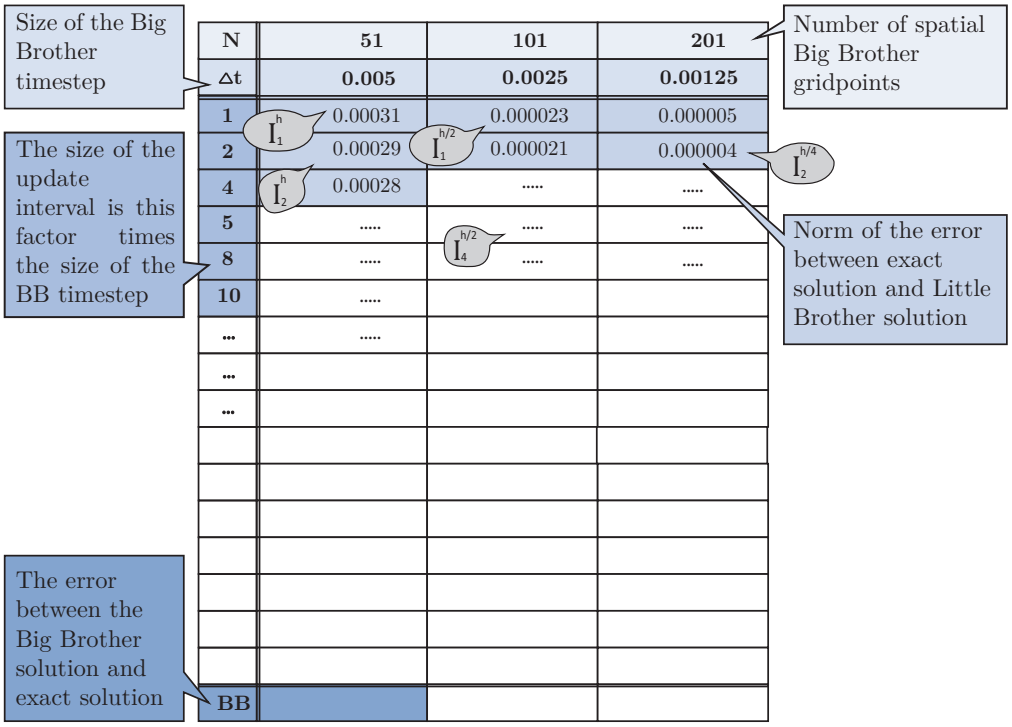

Figure C.1: This is an example of the results table. The text-clouds explain the meaning of the values in the table.

the boundary interval. The 1 means that the boundary data interval resolution has the size of one Big Brother timestep ΔT. The 2 means that the size is 2 times the Big Brother timestep, which is $2\Delta T$. This means that that the Big Brother solution is stored in half the number of points as in the first row. If there are in total 100 Big Brother timepoints, then 50 would mean that the solution is only stored at $T = 0.5$ and at $T = 1$ (the initial condition at $T = 0$ is always present). Figure C.1 explains what you see in the tables. The round clouds representing I_2^h refer to the interval

resolution. This indication is necessary for the validation section, in which two options are sketched to compare the solution. The indication corresponds to the explanation in section 4.4.

C.1 Norms

The following subsections show the L_2-norms for the several experiment options, sketched in section 4.5. Five scenarios are lighted out in the:

- Scenario I is the simulation of $k_z = 3$, a SRF of 5 and matching gridpoints.
- Scenario II is the simulation $k_z = 1$, a SRF of 1 and matching gridpoints.
- Scenario III is the simulation $k_z = 8$, a SRF of 2 and matching gridpoints.
- Scenario IV is the simulation $k_z = 3$, a SRF of 10 and non-matching gridpoints.
- Scenario V is the simulation $k_z = 5$, a SRF of 2 and non-matching gridpoints.

These five scenarios give an indication of how the lateral boundary condition is influencing the fnial solutions. More experiments are performed to give an insight on how the zonal wavenumbers, spatial interpolation and spatial resolution difference are influencing the LB outcome. These results are summarized in figures 5.2 till 5.5.

C.1.1 Scenario I: Zonal wavenumber 3, factor 5, Matching gridpoints

Table C.1: This table shows all the L_2 norms from Scenario I. The table is explained in the picture at the beginning of this chapter. The table on the left is linear interpolation. The table on the right is cubic interpolation.

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.05e-03	3.94e-04	5.22e-05
2	3.05e-03	3.95e-04	5.23e-05
4	3.06e-03	3.96e-04	5.26e-05
5	3.07e-03	3.97e-04	5.29e-05
8	3.09e-03	4.03e-04	5.41e-05
10	3.11e-03	4.07e-04	5.52e-05
12.5	3.17e-03	4.18e-04	9.54e-05
20	3.28e-03	4.47e-04	6.48e-05
25	3.40e-03	4.77e-04	7.21e-05
40	3.94e-03	6.10e-04	1.05e-04
50	4.44e-03	7.35e-04	1.36e-04
80	6.91e-03	1.41e-03	2.93e-04
100	8.95e-03	1.90e-03	4.30e-04
125	1.23e-02	2.75e-03	6.48e-04
200	2.69e-02	6.54e-03	1.60e-03
250	3.98e-02	9.95e-03	2.48e-03
400	9.18e-02	2.47e-02	6.26e-03
500	9.15e-02	3.77e-02	9.67e-03
BB	1.04e-02	1.12e-03	1.34e-04

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.05e-03	3.94e-04	5.22e-05
2	3.05e-03	3.94e-04	5.22e-05
4	3.05e-03	3.94e-04	5.22e-05
5	3.05e-03	3.94e-04	5.22e-05
8	3.05e-03	3.94e-04	5.22e-05
10	3.05e-03	3.94e-04	5.22e-05
20	3.06e-03	3.95e-04	5.22e-05
25	3.07e-03	3.95e-04	5.22e-05
40	3.14e-03	4.00e-04	5.25e-05
50	3.27e-03	4.07e-04	5.29e-05
100	6.57e-03	6.18e-04	6.52e-05
125	1.14e-02	9.56e-04	8.57e-05
200	4.93e-02	4.10e-03	2.93e-04
250	6.39e-02	9.11e-03	6.44e-04
500	6.15e-01	6.20e-02	8.83e-03
BB	1.04e-02	1.12e-03	1.34e-04

C.1.2 Scenario II: Zonal wavenumber 1, factor 1, Matching gridpoints

Table C.2: The norms of Scenario II. Left is linear interpolation, on the right is cubic interpolation.

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	1.69e-04	1.80e-05	2.09e-06
2	1.69e-04	1.80e-05	2.08e-06
4	1.68e-04	1.79e-05	2.06e-06
5	1.68e-04	1.78e-05	2.04e-06
8	1.67e-04	1.75e-05	1.96e-06
10	1.66e-04	1.72e-05	1.89e-06
12.5	2.66e-04	8.93e-05	4.08e-05
20	1.55e-04	1.49e-05	1.39e-06
25	1.48e-04	1.33e-05	1.22e-06
40	1.22e-04	1.12e-05	2.73e-06
50	1.08e-04	1.63e-05	4.93e-06
62.5	1.71e-04	7.10e-05	3.60e-05
80	1.87e-04	5.46e-05	1.49e-05
100	3.30e-04	9.22e-05	2.42e-05
125	5.70e-04	1.51e-04	3.88e-05
200	1.63e-03	4.08e-04	1.02e-04
250	2.60e-03	6.44e-04	1.60e-04
400	7.00e-03	2.25e-03	5.30e-04
500	1.39e-02	2.98e-03	7.76e-04
BB	2.91e-04	3.62e-05	4.53e-06

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	1.69e-04	1.80e-05	2.09e-06
2	1.69e-04	1.80e-05	2.09e-06
4	1.69e-04	1.80e-05	2.09e-06
5	1.69e-04	1.80e-05	2.09e-06
8	1.69e-04	1.80e-05	2.09e-06
10	1.69e-04	1.80e-05	2.09e-06
20	1.69e-04	1.80e-05	2.09e-06
25	1.69e-04	1.80e-05	2.08e-06
40	1.67e-04	1.78e-05	2.07e-06
50	1.66e-04	1.77e-05	2.06e-06
100	1.61e-04	1.82e-05	2.24e-06
125	1.71e-04	2.07e-05	2.64e-06
200	3.60e-04	4.96e-05	6.32e-06
250	6.10e-04	9.17e-05	1.16e-05
500	5.78e-03	6.00e-04	9.02e-05
BB	2.91e-04	3.62e-05	4.53e-06

C.1.3 Scenario III: Zonal wavenumber 8, factor 2, Matching gridpoints

Table C.3: The norms of Scenario III. Left is linear interpolation, on the right is cubic interpolation.

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	1.94e-01	1.39e-02	1.25e-03
2	1.94e-01	1.39e-02	1.25e-03
4	1.94e-01	1.39e-02	1.25e-03
5	1.94e-01	1.39e-02	1.24e-03
8	1.94e-01	1.39e-02	1.23e-03
10	1.94e-01	1.38e-02	1.22e-03
12.5	1.95e-01	1.41e-02	1.29e-03
20	1.94e-01	1.36e-02	1.15e-03
25	1.94e-01	1.34e-02	1.10e-03
40	1.94e-01	1.27e-02	8.61e-04
50	1.94e-01	1.21e-02	6.58e-04
62.5	1.95e-01	1.17e-02	6.29e-04
80	1.94e-01	1.02e-02	5.89e-04
100	1.95e-01	9.64e-03	1.43e-03
125	1.96e-01	1.15e-02	2.86e-03
200	2.29e-01	3.70e-02	1.01e-02
250	2.56e-01	6.01e-02	1.67e-02
400	4.19e-01	1.49e-01	4.26e-02
500	3.70e-01	1.87e-01	6.48e-02
BB	1.04e-02	1.12e-03	1.34e-04

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	1.94e-01	1.39E-02	1.25e-03
2	1.94e-01	1.39e-02	1.25e-03
4	1.94e-01	1.39e-02	1.25e-03
5	1.94e-01	1.39e-02	1.25e-03
8	1.94e-01	1.39e-02	1.25e-03
10	1.94e-01	1.39e-02	1.25e-03
20	1.94e-01	1.39e-02	1.25e-03
25	1.94e-01	1.39e-02	1.25e-03
40	1.94e-01	1.37e-02	1.23e-03
50	1.94e-01	1.34e-02	1.21e-03
100	2.18e-01	1.10e-02	7.03e-04
125	2.51e-01	2.08e-02	1.13e-03
200	4.46e-01	1.17e-01	1.05e-02
250	3.69e-01	1.84e-01	2.47e-02
500	3.69e-01	5.34e-01	1.84e-01
BB	7.10e-01	4.90e-02	3.75e-03

C.1.4 Scenario IV: Zonal wavenumber 3, factor 10, Non-Matching

Table C.4: The norms of Scenario IV. Left is linear interpolation, on the right is cubic interpolation.

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.13e-03	4.48e-04	8.04e-05
2	3.14e-03	4.48e-04	8.05e-05
4	3.14e-03	4.50e-04	8.09e-05
5	3.15e-03	4.51e-04	8.12e-05
8	3.17e-03	4.56e-04	8.25e-05
10	3.19e-03	4.61e-04	8.37e-05
12.5	3.25e-03	4.75e-04	1.19e-04
20	3.36e-03	5.01e-04	9.38e-05
25	3.48e-03	5.32e-04	1.01e-04
40	4.06e-03	6.73e-04	1.40e-04
50	4.57e-03	8.05e-04	1.71e-04
62.5	5.43e-03	1.04e-03	2.54e-04
80	6.90e-03	1.40e-03	3.22e-04
100	9.05e-03	1.95e-03	4.60e-04
125	1.24e-02	2.83e-03	6.80e-04
200	2.70e-02	6.59e-03	1.63e-03
250	3.98e-02	1.00e-02	2.51e-03
400	9.18e-02	2.48e-02	6.28e-03
500	9.11e-02	3.78e-02	9.69e-03
BB	1.04e-02	1.12e-03	1.34e-04

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.13e-03	4.47e-04	8.03e-05
2	3.13e-03	4.47e-04	8.03e-05
4	3.13e-03	4.47e-04	8.03e-05
5	3.13e-03	4.47e-04	8.03e-05
8	3.13e-03	4.47e-04	8.03e-05
10	3.13e-03	4.47e-04	8.03e-05
20	3.14e-03	4.48e-04	8.03e-05
25	3.14e-03	4.48e-04	8.04e-05
40	3.15e-03	4.49e-04	8.05e-05
50	3.18e-03	4.51e-04	8.07e-05
100	3.60e-03	4.92e-04	8.52e-05
125	4.25e-03	5.48e-04	9.20e-05
200	1.63e-02	9.51e-04	1.48e-04
250	3.11e-02	1.74e-03	2.19e-04
500	6.48e-01	2.99e-02	1.44e-03
BB	1.04e-02	1.12e-03	1.34e-04

C.1.5 Scenario V, Zonal wavenumber 5, factor 2, Non-Matching

Table C.5: The norms of Scenario V. Left is linear interpolation, on the right is cubic interpolation.

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.03e-02	2.77e-03	2.53e-04
2	3.03e-02	2.77e-03	2.53e-04
4	3.02e-02	2.77e-03	2.53e-04
5	3.02e-02	2.76e-03	2.52e-04
8	3.02e-02	2.75e-03	2.51e-04
10	3.02e-02	2.75e-03	2.50e-04
12.5	3.07e-02	3.00e-03	3.97e-04
20	3.00e-02	2.68e-03	2.41e-04
25	2.98e-02	2.63e-03	2.37e-04
40	2.92e-02	2.47e-03	2.46e-04
50	2.88e-02	2.36e-03	2.87e-04
62.5	2.88e-02	2.62e-03	4.96e-04
80	2.77e-02	2.57e-03	6.09e-04
100	2.80e-02	3.48e-03	9.59e-04
125	3.11e-02	5.42e-03	1.52e-03
200	6.43e-02	1.60e-02	4.42e-03
250	9.38e-02	2.66e-02	6.89e-03
400	5.99e-02	6.64e-02	1.75e-02
500	3.15e-01	9.46e-02	2.70e-02
BB	8.11e-02	6.54e-03	6.63e-04

N	**51**	**101**	**201**
Δt	**2.50e-04**	**1.25e-04**	**6.25e-05**
1	3.03e-02	2.77e-03	2.53e-04
2	3.03e-02	2.77e-03	2.53e-04
4	3.03e-02	2.77e-03	2.53e-04
5	3.03e-02	2.77e-03	2.53e-04
8	3.03e-02	2.77e-03	2.53e-04
10	3.03e-02	2.77e-03	2.53e-04
20	3.02e-02	2.77e-03	2.53e-04
25	3.02e-02	2.77e-03	2.53e-04
40	3.02e-02	2.76e-03	2.53e-04
50	3.02e-02	2.76e-03	2.53e-04
100	2.90e-02	2.68e-03	2.51e-04
125	3.06e-02	2.59e-03	2.50e-04
200	1.28e-02	4.84e-03	3.00e-04
250	4.05e-01	1.56e-02	5.70e-04
500	4.81e-01	3.82e-01	1.63e-02
BB	8.11e-02	6.54e-03	6.63e-04

C.2 Constant ratio between intervals to compare

There are two options to compare the norms logarithmically, as explained in section 4.4. Tables C.6 till C.10 show the p-values of the comparison of the norms that are related by a constant ratio. Figure C.2 show how the norms are compared in this case.

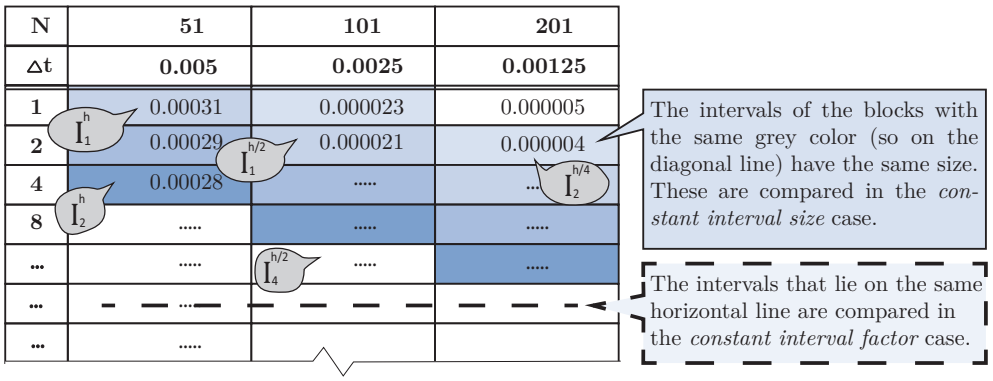

Figure C.2: This is an example of the p-value table. The text-clouds explain the meaning of the values in the table.

In the case of constant ratio, the norms from tables C.1 till C.5 are compared that lie on the same horizontal line. That results in the p-values that are showed in the coming subsections.

C.2.1 Scenario I: Zonal wavenumber 3, factor 5, Matching gridpoints

Table C.6: This table contains the p-value of the intervals compared with constant ratio of Scenario II. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	2.953	2.917
2	2.952	2.916
4	2.950	2.912
5	2.948	2.909
8	2.948	2.909
10	2.932	2.883
12.5	2.932	2.883
20	2.873	2.788
25	2.833	2.727
40	2.692	2.537
50	2.596	2.429
80	2.596	2.429
100	2.294	2.267
125	2.236	2.142
200	2.161	2.086
250	2.040	2.031
400	1.998	2.005
500	1.893	1.983
BB	3.209	3.061

N	**51-101**	**101-201**
1	2.953	2.918
2	2.953	2.918
4	2.953	2.918
5	2.953	2.918
8	2.953	2.918
10	2.953	2.918
20	2.954	2.919
25	2.957	2.920
40	2.976	2.929
50	3.005	2.945
100	3.410	3.244
125	3.581	3.479
200	3.586	3.808
250	2.809	3.824
500	3.311	2.811
BB	3.209	3.061

C.2.2 Scenario II: Zonal wavenumber 1, factor 1, Matching gridpoints

Table C.7: This table contains the p-value of the intervals compared with constant ratio of Scenario II. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.228	3.109
2	3.230	3.111
4	3.234	3.121
5	3.238	3.128
8	3.253	3.159
10	3.267	3.189
12.5	1.572	1.131
20	3.387	3.424
25	3.474	3.454
40	3.450	2.033
50	2.731	1.726
62.5	1.269	0.980
80	1.775	1.870
100	1.840	1.929
125	1.913	1.965
200	1.996	2.003
250	2.013	2.011
400	1.637	2.087
500	2.221	1.941
BB	3.007	3.002

N	**51-101**	**101-201**
1	3.228	3.108
2	3.228	3.108
4	3.228	3.108
5	3.228	3.108
8	3.228	3.108
10	3.228	3.108
20	3.229	3.108
25	3.229	3.108
40	3.229	3.106
50	3.229	3.103
100	3.148	3.019
125	3.040	2.973
200	2.860	2.972
250	2.733	2.983
500	3.266	2.734
BB	3.007	3.002

C.2.3 Scenario III: Zonal wavenumber 8, factor 2, Matching gridpoints

Table C.8: This table contains the p-value of the intervals compared with constant ratio of Scenario III. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.803	3.476
2	3.803	3.476
4	3.804	3.479
5	3.805	3.481
8	3.808	3.489
10	3.811	3.496
12.5	3.786	3.453
20	3.834	3.563
25	3.852	3.615
40	3.929	3.885
50	3.999	4.204
62.5	4.055	4.218
80	4.254	4.111
100	4.334	2.749
125	4.094	2.011
200	2.630	1.866
250	2.088	1.848
400	1.489	1.810
500	0.984	1.531
BB	3.856	3.710

N	**51-101**	**101-201**
1	3.803	3.476
2	3.803	3.476
4	3.803	3.476
5	3.803	3.476
8	3.803	3.475
10	3.803	3.475
20	3.804	3.475
25	3.806	3.475
40	3.825	3.474
50	3.856	3.472
100	4.300	3.973
125	3.590	4.207
200	1.936	3.472
250	1.004	2.898
500	-0.536	1.536
BB	3.856	3.710

C.2.4 Scenario IV: Zonal wavenumber 3, factor 10, Non-Matching gridpoints

Table C.9: This table contains the p-value of the intervals compared with constant ratio of Scenario IV. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	2.808	2.477
2	2.808	2.477
4	2.806	2.475
5	2.804	2.473
8	2.797	2.467
10	2.791	2.461
12.5	2.774	1.999
20	2.744	2.417
25	2.711	2.389
40	2.594	2.261
50	2.505	2.236
62.5	2.383	2.037
80	2.302	2.119
100	2.213	2.084
125	2.134	2.055
200	2.033	2.016
250	1.994	1.996
400	1.891	1.979
500	1.271	1.962

N	**51-101**	**101-201**
1	2.808	2.478
2	2.808	2.478
4	2.808	2.478
5	2.808	2.478
8	2.808	2.478
10	2.808	2.478
20	2.809	2.478
25	2.809	2.478
40	2.812	2.480
50	2.816	2.482
100	2.870	2.529
125	2.955	2.573
200	4.099	2.689
250	4.165	2.989
500	4.437	4.379

C.2.5 Scenario V: Zonal wavenumber 8, factor 2, Non-Matching gridpoints

Table C.10: This table contains the p-value of the intervals compared with constant ratio of Scenario V. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.450	3.451
2	3.450	3.451
4	3.451	3.452
5	3.452	3.453
8	3.455	3.457
10	3.458	3.460
12.5	3.356	2.917
20	3.483	3.475
25	3.501	3.473
40	3.567	3.326
50	3.605	3.044
62.5	3.456	2.402
80	3.426	2.078
100	3.010	1.861
125	2.520	1.834
200	2.003	1.860
250	1.820	1.948
400	-0.150	1.925
500	1.736	1.808

N	**51-101**	**101-201**
1	3.450	3.451
2	3.450	3.451
4	3.450	3.451
5	3.450	3.451
8	3.450	3.451
10	3.450	3.451
20	3.450	3.451
25	3.450	3.451
40	3.451	3.450
50	3.451	3.449
100	3.436	3.416
125	3.564	3.372
200	1.400	4.013
250	4.697	4.776
500	0.331	4.556

C.3 Constant interval size

The tables C.11 till C.15 in this section show the p-values when the intervals are compared that lie on the same diagonal line. See figure C.2 for the explanation.

C.3.1 Scenario I: Zonal wavenumber 3, factor 5, Matching gridpoints

Table C.11: This table contains the p-value of the intervals compared with constant interval of Scenario I. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	2.951	2.915
2	2.946	2.906
4	2.927	2.872
5	2.912	2.847
10	2.797	2.653
12.5	2.729	2.535
20	2.424	2.088
25	2.211	1.807
40	1.484	1.059
50	1.226	0.772
62.5	0.936	0.560
100	0.453	0.248
125	0.307	0.151
200	0.121	0.063
250	0.076	0.042

N	**51-101**	**101-201**
1	2.953	2.918
2	2.953	2.918
4	2.953	2.918
5	2.953	2.918
10	2.952	2.917
20	2.936	2.911
25	2.912	2.900
50	2.404	2.643
100	0.679	1.077
125	0.327	0.570
250	0.043	0.045

C.3.2 Scenario II: Zonal wavenumber 1, factor 1, Matching gridpoints

Table C.12: This table contains the p-value of the intervals compared with constant interval of Scenario II. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.231	3.113
2	3.238	3.129
4	3.267	3.192
5	3.289	3.240
10	3.477	3.634
12.5	4.314	6.196
20	3.798	2.444
25	3.185	1.437
40	1.161	-0.418
50	0.231	-0.571
62.5	0.176	0.872
100	-0.304	-0.141
125	-0.176	-0.078
200	-0.470	-0.379
250	-0.198	-0.270

N	**51-101**	**101-201**
1	3.228	3.108
2	3.228	3.108
4	3.228	3.108
5	3.228	3.109
10	3.23	3.11
20	3.243	3.121
25	3.254	3.128
50	3.19	2.98
100	1.699	1.523
125	0.895	0.838
250	0.022	0.024

C.3.3 Scenario III: Zonal wavenumber 8, factor 2, Matching gridpoints

Table C.13: This table contains the p-value of the intervals compared with constant interval of Scenario III. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.803	3.477
2	3.804	3.480
4	3.808	3.493
5	3.811	3.502
10	3.835	3.587
12.5	3.859	3.689
20	3.930	3.980
25	4.000	4.351
40	4.253	4.435
50	4.329	3.079
62.5	4.080	2.036
100	2.396	-0.073
125	1.709	-0.538
200	0.615	-0.204
250	0.448	-0.109

N	**51-101**	**101-201**
1	3.803	3.476
2	3.803	3.476
4	3.803	3.476
5	3.803	3.476
10	3.804	3.477
20	3.826	3.495
25	3.858	3.524
50	4.133	4.250
100	0.900	0.072
125	0.447	-0.245
250	-0.534	-0.002

C.3.4 Scenario IV: Zonal wavenumber 3, factor 10, Non-Matching gridpoints

Table C.14: This table contains the p-value of the intervals compared with constant interval of Scenario IV. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	2.807	2.476
2	2.803	2.470
4	2.785	2.447
5	2.772	2.429
10	2.670	2.296
12.5	2.610	2.225
20	2.318	1.835
25	2.113	1.637
40	1.539	1.064
50	1.227	0.806
62.5	0.942	0.614
100	0.458	0.261
125	0.311	0.173
200	0.124	0.069
250	0.077	0.045

N	**51-101**	**101-201**
1	2.808	2.478
2	2.808	2.478
4	2.808	2.478
5	2.808	2.478
10	2.808	2.477
20	2.804	2.475
25	2.799	2.472
50	2.690	2.403
100	1.919	1.738
125	1.291	1.324
250	0.057	0.271

C.3.5 Scenario V: Zonal wavenumber 5, factor 2, Non-Matching gridpoints

Table C.15: This table contains the p-value of the intervals compared with constant interval of Scenario V. Left table is due to linear interpolation. Right table is due to cubic interpolation.

N	**51-101**	**101-201**
1	3.450	3.451
2	3.451	3.454
4	3.457	3.462
5	3.461	3.468
10	3.493	3.509
12.5	3.544	3.661
20	3.603	3.446
25	3.658	3.201
40	3.506	2.017
50	3.046	1.302
62.5	2.409	0.787
100	0.807	-0.343
125	0.227	-0.345
200	-0.048	-0.126
250	-0.013	-0.025

N	**51-101**	**101-201**
1	3.450	3.451
2	3.450	3.451
4	3.450	3.451
5	3.450	3.451
10	3.450	3.451
20	3.452	3.453
25	3.454	3.453
50	3.492	3.458
100	2.584	3.161
125	0.970	2.182
250	0.083	-0.058

C.4 More scenarios

More simulations are performed to gain better insight in the impact of the wavenumber, the Spatial Resolution Factor (SRF) and the spatial interpolation errors. Their influence is described in the conclusion and the p-values are summarized in tables 6.1 and 6.2. The following four scenarios are simulated:

- Scenario Ia is the simulation of $k_z = 1$, a SRF of 1 and matching gridpoints.
- Scenario IIa is the simulation $k_z = 3$, a SRF of 10 and non-matching gridpoints.
- Scenario IIIa is the simulation $k_z = 5$, a SRF of 2 and non-matching gridpoints.
- Scenario IVa is the simulation $k_z = 8$, a SRF of 5 and non-matching gridpoints.

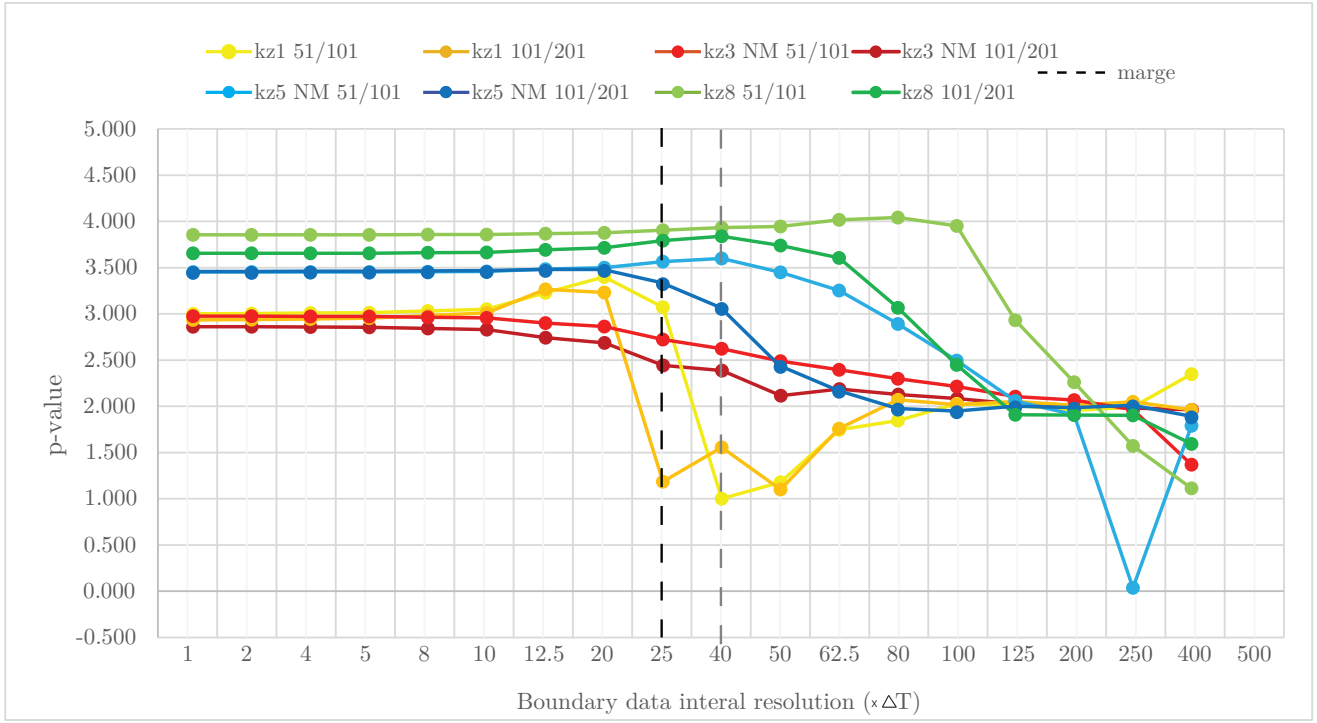

Figure C.3: p-values of the four wavenumbers (kz1 means zonal wavenumber 1) compared for constant ratio, as in section C.2. Linear interpolation is used. The black dashed line shows the optimal combination.

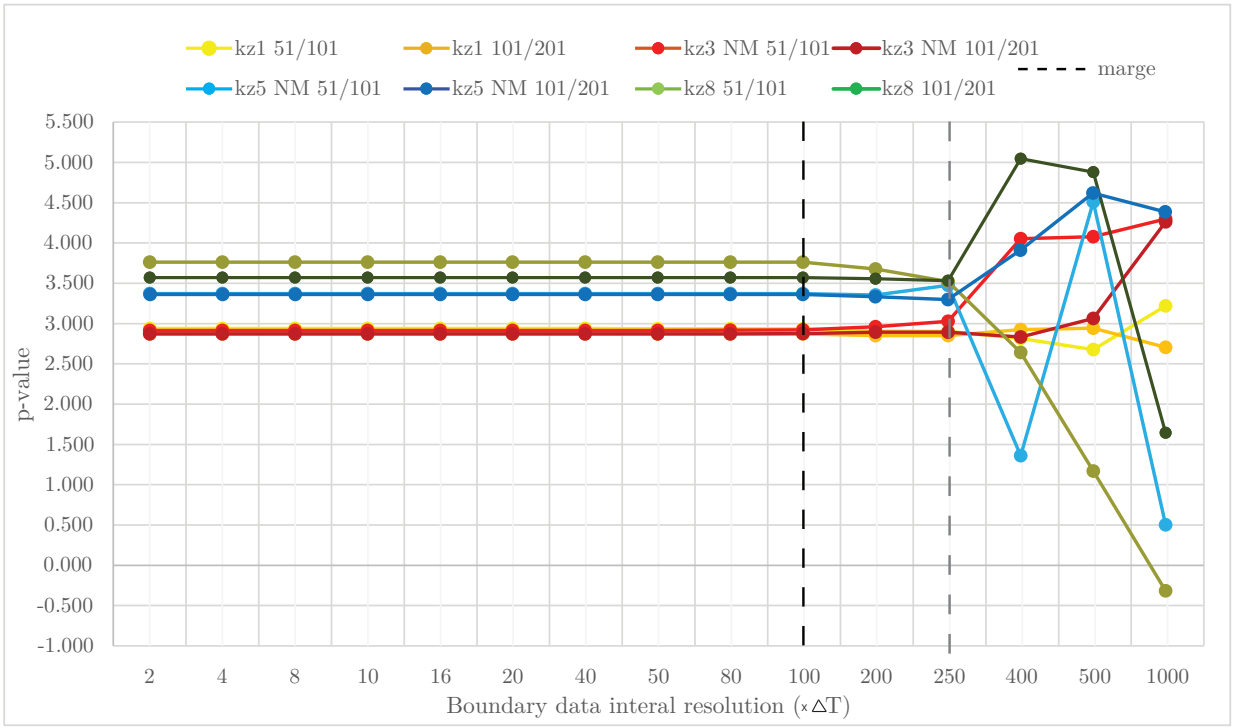

Figure C.4: p-values of the four wavenumbers (kz1 means zonal wavenumber 1) compared for constant ratio, as in section C.2. Third order interpolation is used. The black dashed line shows the optimal combination.

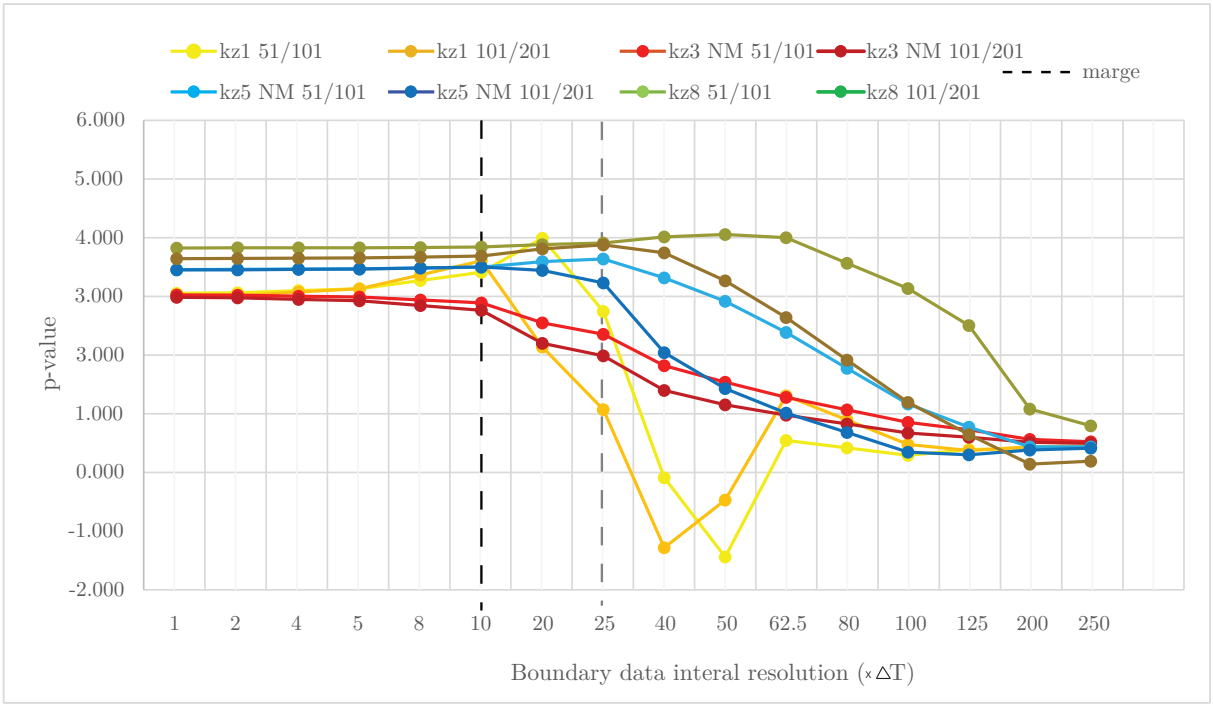

Figure C.5: Summary of the p-values of the four wavenumbers compared for constant interval, as in section C.3. Linear interpolation is used. The black dashed line shows the optimal combination.

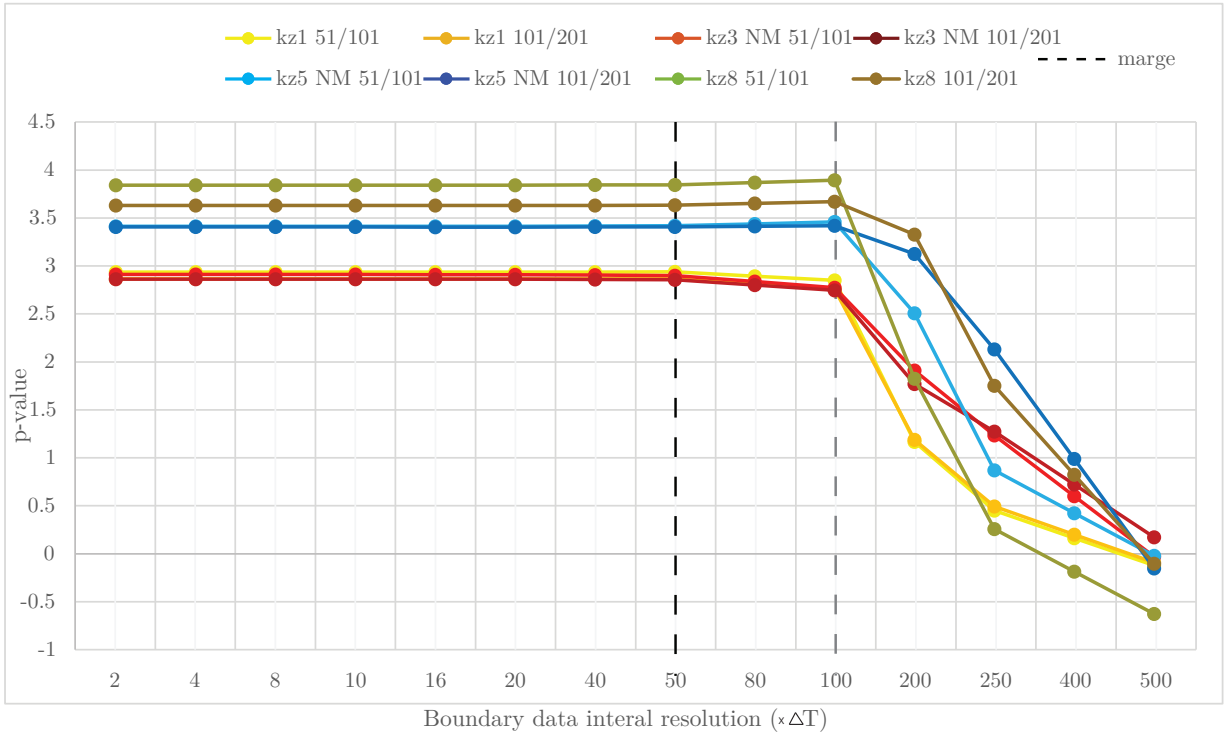

Figure C.6: p-values of the four wavenumbers compared for constant interval, as in section C.3. Third order interpolation is used. The black dashed line shows the optimal combination.

Bibliography

[1] Introduction to climate dynamics and climate modelling - Components of a climate model.

[2] geostrophic. (n.d.) American Heritage® Dictionary of the English Language, Fifth Edition. (2011). Retrieved July 19 2017 from http://www.thefreedictionary.com/geostrophic, 2011.

[3] Denis B., Laprise R., Caya D., and Cote J. Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Climate Dynamics*, 18(8):627–646, apr 2002.

[4] Benoit Cushman-Roisin and Jean-Marie Beckers. Barotropic Waves. In *IIntroduction to Geophysical Fluid Dynamics, Volume 101*, pages 271–315. 2 edition, 2011.

[5] RobertE. Dickinson, RonaldM. Errico, Filippo Giorgi, and GaryT. Bates. A regional climate model for the western United States. *Climatic Change*, 15(3):383–422, dec 1989.

[6] Adrian E. Gill. Atmosphere-ocean dynamics. In *Atmposphere Ocean Dynamics*, chapter 7, pages 204–208. Academic Press, 1982.

[7] Adrian E. Gill. Chapter Eight - Gravity Waves in a Rotating Fluid. In *Atmosphere Ocean Dynamics*, volume 30, chapter 8, pages 247–315. 1982.

[8] Adrian E. Gill. Chapter Eleven - The Tropics. In *Atmosphere Ocean Dynamics*, volume 30, chapter 11, pages 429–491. 1982.

[9] Filippo Giorgi. Simulation of Regional Climate Using a Limited Area Model Nested in a General Circulation Model. *Journal of Climate*, 3(9):941–963, sep 1990.

[10] Filippo Giorgi and Gary T. Bates. The Climatological Skill of a Regional Model over Complex Terrain. *Monthly Weather Review*, 117(11):2325–2347, nov 1989.

[11] Filippo Giorgi and Linda O. Mearns. Introduction to special section: Regional Climate Modeling Revisited. *Journal of Geophysical Research: Atmospheres*, 104(D6):6335–6352, mar 1999.

[12] Bertil Gustafsson. Boundary treatment. In *High Order Difference Methods for Time Dependent PDE*, chapter 7, pages 127–156. 2007.

[13] James R. Holton, Gregory J. Hakim, James R. Holton, and Gregory J. Hakim. Chapter 11 - Tropical Dynamics. In *An Introduction to Dynamic Meteorology*, pages 377–411. Academic Press, fifth edition, 2013.

[14] M Moller. Numerical Methods for Ordinary Differential Equations Course WI3079TU and WI3097MINOR. Technical report, Delft University of Technology, 2013.

[15] National Center for Atmospheric Research (NCAR). Climate Modeling | UCAR Center for Science Education, 2011.

[16] Joseph Pedlosky. Equations of Motion; Surface Gravity Waves. In *Waves in the Ocean and Atmosphere*, chapter 3, pages 19–31. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[17] Joseph Pedlosky. Internal Gravity Waves. In *Waves in the Ocean and Atmosphere*, pages 59–66. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[18] Joseph Pedlosky. Rossby Waves. In *Waves in the Ocean and Atmosphere*, pages 149–158. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[19] Joseph Pedlosky. Shallow Water Waves in a Rotating Fluid; Poincaré and Kelvin Waves. In *Waves in the Ocean and Atmosphere*, pages 133–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[20] Norman A. Phillips. The general circulation of the atmosphere: A numerical experiment. *Quarterly Journal of the Royal Meteorological Society*, 82(352):123–164, apr 1956.

[21] J. R. and Carl de Boor. Piecewise Linear Approximation. In JE Marsen and L Sirovich, editors, *A Practical Guide to Splines, Applied Mathematical Science*, volume 27, chapter 3, pages 31–40. jan 2001.

[22] A. J. Simmons. The Meridional Scale of Baroclinic Waves. *Journal of the Atmospheric Sciences*, 31(6):1515–1525, sep 1974.

[23] Paul Spence, Ryan M. Holmes, Andrew McC. Hogg, Stephen M. Griffies, Kial D. Stewart, and Matthew H. England. Localized rapid warming of West Antarctic subsurface waters by remote winds. *Nature Climate Change*, 7(8):595–603, jul 2017.

[24] Strebe. The world on Mercator projection between 82Â°S and 82Â°N. 15Â° graticule. Imagery is a derivative of NASA's Blue Marble summer month composite with oceans lightened to enhance legibility and contrast. Image created with the Geo, 2011.

[25] Strebe. The world on Winkel tripel projection. 15Â° graticule. Imagery is a derivative of NASA's Blue Marble summer month composite with oceans lightened to enhance legibility and contrast. Image created with the Geocart map projection software., 2011.

[26] Magnus Svärd and Jan Nordström. Review of summation-by-parts schemes for initial and boundary-value problems. *Journal of Computational Physics*, 268:17–38, 2013.

[27] Thomas Walter. rossby.waves, 2017.

[28] National Oceanic US Department of Commerce and Atmospheric Administration. Breakthrough article on the First Climate Model, 2007.

[29] Charles F. Van Loan. Piecewise polynomial interpolation. In *Introduction to scientific computing : a matrix-vector approach using MATLAB*, chapter 3, page 367. Prentice Hall, 2000.

[30] David E. Venne. Normal-Mode Rossby Waves Observed in the Wavenumber 1-5 Geopotential Fields of the Stratosphere and Troposphere. *Journal of the Atmospheric Sciences*, 46(7):1042–1056, apr 1989.

[31] C. (Cornelis) Vuik, F.J. (Fredericus Johannes) Vermolen, M.B. van (Martin Bastiaan) Gijzen, and M.J. Vuik. *Numerical methods for ordinary differential equations*. DAP, Delft Academic Press, 2015.

List of Figures

List of Tables

List of Tables