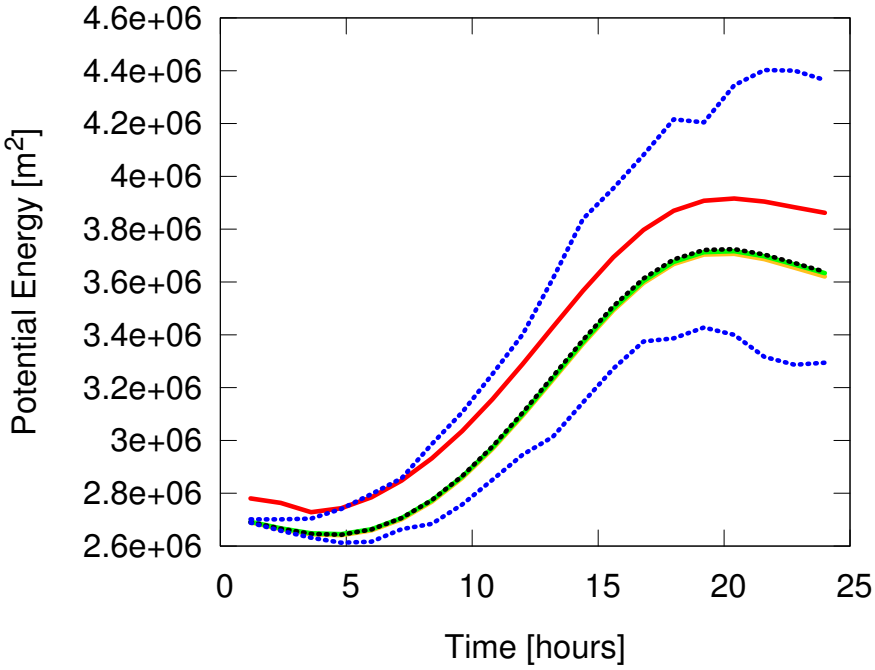
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Error Estimation Algorithms for Ocean Models



Jan Ackmann

Hamburg 2017

Berichte zur Erdsystemforschung
Reports on Earth System Science

186
2017

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

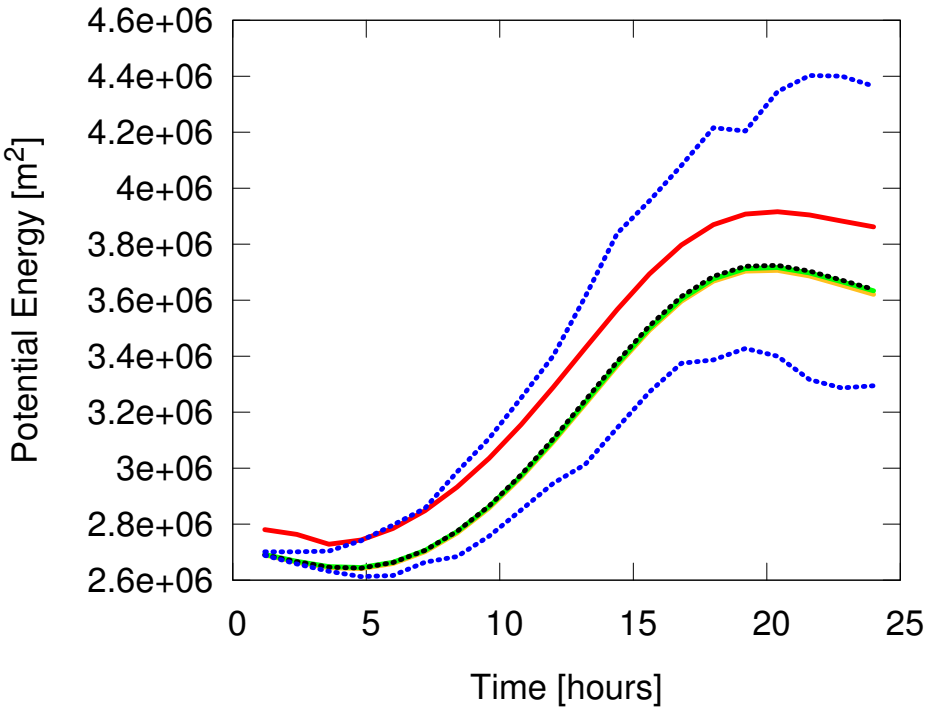
name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Error Estimation Algorithms for Ocean Models



Dissertation with the aim of achieving a doctoral degree
at the Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg

submitted by

Jan Ackmann

Hamburg 2017

Jan Ackmann

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 8.7.2016

Folgende Gutachter empfehlen die Annahme der Dissertation:

Dr. Peter Korn
Prof. Dr. Michael Hinze

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 186
Reports on Earth System Science / Max Planck Institute for Meteorology 2017

ISSN 1614-1199

Abstract

Errors due to discretization are inherent to the numerical solutions of Ocean General Circulation Models (OGCM). However, obtaining reliable estimates for these errors is a difficult undertaking.

In this thesis, we develop a stochastic dual-weighted error estimator for the estimation of the discretization error in physical quantities of interest (goals) that is applicable to ocean models. Towards this aim, we extend the dual-weighted error estimation technique by a stochastic process with memory. In this, we extend previous work on memory-less stochastic dual-weighted methods for two-dimensional wave-type flows.

The introduction of memory effects is the key new element of our extension and is shown to be crucial in the estimation of goal errors in an ocean model setting. The memory governs the temporal evolution of the stochastic process. We interpret the memory as a stochastic representation of physical constraints on the time-evolution of the essential building block of our stochastic dual-weighted error estimator – the local truncation error. The memory of the stochastic process is represented by temporal correlation coefficients directly or by their upper bound. The temporal correlation coefficients and other required stochastic quantities of the stochastic process are estimated from high-resolution model information at near-initial times. Our resulting stochastic dual-weighted approach is equivalent to a linearized stochastic-physics ensemble, but in contrast to the ensemble it only requires a single model integration and a single adjoint integration.

In order to study the applicability of our stochastic dual-weighted error estimator for OGCMs, we focus on important oceanic features: the presence of lateral boundaries with their associated boundary currents, and the phenomenon of baroclinic instability within a stratified ocean. Both phenomena are studied by means of idealized experiments, the Munk gyre and the flow against an island for lateral boundaries, and the spherical channel experiment for baroclinic instabilities. For flows with boundaries, we find that our stochastic error estimator provides meaningful error bounds for a range of physically relevant goals. For the eddying flow regime due to baroclinic instabilities, the stochastic process of our error estimator is modeled as a compound of a horizontal, a vertical, and a temporal structure. To be applicable in our stochastic dual-weighted error estimation framework, we propose a generalization of its temporal structure to include additional, possibly negative time-correlations.

Zusammenfassung

Diskretisierungsfehler sind Bestandteil der numerischen Lösungen von Ozeanzirkulationsmodellen (OGCM). Die Berechnung belastbarer Abschätzungen für diese Fehler ist allerdings ein schwieriges Unterfangen.

In der vorliegenden Arbeit entwickeln wir einen stochastischen dual-gewichteten Fehlerschätzer für die Abschätzung des Diskretisierungsfehlers in physikalischen Zielfunktionalen, welcher auf Ozeanmodelle angewendet werden kann. Dafür erweitern wir die dual-gewichtete Fehlerschätzungsmethode um einen stochastischen Prozess mit Gedächtnis. Damit erweitern wir vorangegangene Arbeiten über stochastische dual-gewichtete Fehlerschätzungsmethoden ohne Gedächtnis für zweidimensionale wellendominierte Strömungen.

Die Berücksichtigung von Gedächtniseffekten ist das Schlüsselelement unserer Erweiterung und wir zeigen, dass sie bei der Schätzung von Fehlern in Zielfunktionalen für Ozeanmodelle essentiell sind. Das Gedächtnis regelt die zeitliche Entwicklung des stochastischen Prozesses. Wir interpretieren das Gedächtnis als eine stochastische Darstellung von physikalischen Beschränkungen bezüglich der zeitlichen Entwicklung des essentiellen Bausteins unseres stochastischen dual-gewichteten Fehlerschätzers – dem lokalen Diskretisierungsfehler. Das Gedächtnis des stochastischen Prozesses ist entweder direkt durch zeitliche Korrelationskoeffizienten oder durch ihre obere Schranke repräsentiert. Die zeitlichen Korrelationskoeffizienten und andere benötigte stochastische Größen des stochastischen Prozesses werden mittels hochaufgelöster Modellinformation auf nah-initialen Zeitskalen geschätzt. Unsere resultierende stochastische dual-gewichtete Methode ist äquivalent zu einem linearisierten stochastische-Physik Ensemble, aber im Gegensatz zum Ensemble benötigt sie nur eine Integration des primalen Modells und eine Integration des adjungierten Modells.

Um die Anwendbarkeit unseres stochastischen dual-gewichteten Fehlerschätzers für OGCMs zu untersuchen, fokussieren wir uns auf wichtige ozeanographische Merkmale: die Anwesenheit von lateralen Rändern mit ihren assoziierten Randströmen, und das Phänomen der baroklinen Instabilität im stratifizierten Ozean. Beide Phänomene werden mit Hilfe idealisierter Experimente untersucht, dem Munk gyre und der Strömung um eine Insel für laterale Ränder, und dem sphärischen Kanalexperiment für barokline Instabilitäten. Für Strömungen mit Rändern erhalten wir sinnvolle Fehlerschranken für eine Reihe physikalisch relevanter Zielfunktionale für unsere stochastischen Fehlerschätzer. Für die wirbelbehaftete Strömung durch barokline Instabilitäten modellieren wir den stochastischen Prozess unseres Fehlerschätzers als Komposition einer horizontalen, einer vertikalen, und einer zeitlichen Struktur. Zur Anwendung im Rahmen unseres stochastischen dual-gewichteten Fehlerschätzers schlagen wir eine Generalisierung

der zeitlichen Struktur des stochastischen Prozesses vor, um zusätzliche, möglicherweise negative, zeitliche Korrelationen miteinzubeziehen.

Contents

Chapter 1

Introduction

In a comprehensive Earth System Model, the modeling of the ocean component is crucial for properly understanding the climate system, its variability and its response to external forcings. The governing equations that are typically chosen to represent the ocean component in an Earth System model are too complex to solve them analytically. Thus, the governing equations are instead solved approximately by a numerical model on a discrete grid with finite degrees of freedom. The grid spacing of a state-of-the-art ocean model used for the scientific basis of the IPCC report [12] is typically about 1 degree. At this resolution, important flow features are still unresolved due to the multiscale nature of the ocean dynamics that ranges from the order of kilometers down to the millimeter range. Thus, errors due to discretization of the governing equations are inevitably part of the numerical solution. However, so far, there are no methods available to reliably quantify the error due to discretization for an arbitrary numerical ocean model simulation, even though in all of quantitative science it is essential to gain and provide information about the uncertainty in model output or measurements. The purpose of this thesis is thus the derivation and implementation of an error estimation algorithm, targeted at the specifics of ocean models, for the *a posteriori* estimation of errors in physical quantities of interest.

The total error in a model can be conceptionally separated into two parts, the formulation error and the error due to discretization [36]. The first part of the total error already occurs with the formulation of the mathematical model. This is because the mathematical model can necessarily only be an approximation of the real processes in nature. Physical processes are usually too complex, not well enough understood or in part even completely unknown, and thus can only be represented in a simplified form. Apart from this fact, it is also often even desirable to work with a mathematical model that is as simple as possible and only keep the 'important' processes of the physical process one is interested in studying. In the final mathematical model, parts of the real dynamics are thus neglected which introduces the first layer of errors, the formulation error. For most practical applications, the resulting mathematical model is still too complex to be solved by analytical means. The mathematical model is then discretized

and solved approximately on a computer, which introduces the second part of model errors, which we refer to as discretization error. Under this term, all errors are collected that have their roots in the discretization of the continuous equations and boundary conditions of the mathematical model.

In ocean modeling the mathematical models are usually based on approximations of the Navier-Stokes Equations. More specifically, for Ocean General Circulation Models (OGCMs), as used in a comprehensive Earth System Model, the governing equations are the so-called hydrostatic primitive equations [21]. These equations govern the ocean dynamics, in the horizontal and in the vertical, as well as their interplay with temperature and salinity. One key assumption is that the vertical scale of the ocean is small compared to the horizontal scale. A model of lower complexity used for idealized ocean circulation studies is given by the shallow-water equations which can be understood as a primitive equation model that only has one degree of freedom in the vertical flow structure [29], i.e. only consists of one vertical ocean layer. Both models only incorporate the ocean dynamics necessary for their respective tasks in order to keep the computational costs for obtaining a numerical solution of the model to a minimum.

The discretizations of these mathematical models are mostly based on Finite-Volume methods, Finite-Difference methods or a combination thereof. When solving the discretized models, computational limitations require a trade-off between resolution, typically the grid spacing in the horizontal, and simulation time. If the required simulation times go from centuries to millennia, even state-of-the-art OGCMs are limited to a resolution of about 1 degree, which is equivalent to approximately $100km$. This leaves many flow features such as boundary currents or mesoscale eddies unresolved or highly underresolved, which is the source of the discretization error in these model simulations.

The combination of the errors due to formulation and discretization constitutes the total model error, and its quantification proves difficult for the ocean. This is because the model solution cannot easily be compared to observations, as ocean observations are sparsely available in time and in space. To fill these gaps, Ocean State Estimation (OSE) [53] aims at producing numerical solutions of high-resolution ocean models that in some sense are 'close' to observations. OSE is thus closely intertwined with data assimilation [25, 49]. These resulting numerical solutions are generally referred to as reanalyses, see for instance ORAS4 [2] or GECCO2 [26]. Comparing reanalyses to a model solution is a standard approach to gain information about the total model error. However, due to the limited amount of observations, reanalyses have large uncertainties. Reanalyses are by definition only available in hindsight and can only provide error information for the actual ocean circulation. Error information for future projections of the ocean circulation as well as for ocean-type experiments in other settings are thus not covered. Also, a clear separation of the formulation and the discretization parts of the total model error is not possible but would be highly desirable to understand the nature of model errors.

One extension towards gaining knowledge about discretization errors is then to additionally compare solutions on different model resolutions. One example is given in [22], who perform a series of transient coupled climate model simulations in which the resolution of the ocean model ranges from 1 degree to 1/10-th of a degree and compare them among each other, and to observational data whenever possible. Valuable insights about the influence of unresolved processes can be gained from these comparisons. However, these comparisons are very expensive as several high-resolution simulations need to be performed. For an ocean simulation that itself is already at the current computational limits, this approach is thus not applicable. Bypassing the computational limit by inferring the discretization error of a model simulation from a series of even coarser model solutions, such as by Richardson extrapolation [46], is not applicable because the coarser model solutions will be even more under-resolved than the model simulation itself.

Apart from the quantification of the discretization error, much current work aims at its reduction and can be divided into two general approaches. In the first approach, the work is targeted on improving the discretization itself. Much work has gone into the improvement of the numerical schemes and their implementation on High Performance Computers (HPC). With the next generation of ocean models, such as the ICON-Ocean model [27] or the MPAS-O model [44], further progress on the matter of discretization errors is expected by formulating the discretized model in a way that it scales better on future High Performance Computers which would allow higher spatial resolution, and by increasing the flexibility to adjust the model resolution for specific regions if deemed beneficial. In the second approach, the error due to discretization is reduced by employing subgrid-scale parametrizations to mimic the effects of processes that cannot be resolved. This approach can be understood as shifting parts of the discretization error into the domain of the model formulation. Standard ocean model parametrizations are an eddy-viscosity in order to parametrize the effect of eddies, especially near boundaries, enhanced vertical diffusion as a parametrization for deep convection, schemes like the KPP parametrization [28] for the vertical mixing of the upper ocean, and the Redi [43] neutral diffusion and the Gent and McWilliams [14] eddy advection parametrization to capture the missing effects of mesoscale eddies on the ocean circulation.

Although these continuous reductions of the discretization error have significantly improved the numerical solutions of ocean models, there is still a significant amount of errors due to discretization in a resulting numerical solution. This is due to the wide range of spatial and temporal scales in the ocean that will remain being unresolved in the future. One example is given in [22] for the vertical distribution of heat that changes significantly towards a net warming at around 1000 meters depth once mesoscale eddies are not explicitly resolved any more by the model but are only parametrized. As it is not foreseeable when mesoscale eddies can be resolved in climate model simulations, errors like these will be inherent to the numerical ocean solutions. Despite its undoubt-

ful importance, there are still no methods available to reliably quantify the error due to discretization for a given ocean model simulation. Therefore, we think that the development of means to quantify the uncertainty associated with these errors is key and should be pursued with great efforts.

The field of error estimation can be roughly separated into *a priori* and *a posteriori* error estimators. *A priori* error estimators are derived analytically for a specific model discretization. However, an ocean model can be considered to be too complex to obtain useful *a priori* error estimators, as even for relatively simple problems, *a priori* error bounds are difficult to derive and often not very useful as they are not tight enough. Additionally, the calculation of the bounds usually requires constants that have to be estimated. In contrast, *a posteriori* error estimators are calculated for a specific, already calculated, model solution. *A posteriori* error estimators usually perform better, i.e. yield tighter bounds and can even yield model corrections, but in contrast to *a priori* estimators, the calculated bounds are usually not guaranteed. Often the term asymptotic bound is instead used, which means that the bounds are guaranteed if the model resolution is within the asymptotic range, i.e. higher than a certain threshold grid spacing. Due to the complexity of ocean models, we focus on the field of *a posteriori* error estimators.

A simple method for *a posteriori* error estimation is the study of grid convergence, where model errors are inferred from running the same experiment on grids with different model resolution [46, 35]. The previously introduced study [22] can be counted towards this class of error estimators. The method's advantage is that it is easily accessible and non-intrusive, i.e. the method does not require changes in the model code. The disadvantage is that running an ocean model for several different model resolutions is computational expensive.

Another class of *a posteriori* error estimators aims at quantifying the uncertainty in a model solution from a series of perturbed model solutions on the same grid. For instance, methods from the field of statistical mechanics aim at understanding and modeling the terms that need to be added to a discretized model in order to obtain corrected model solutions. As these terms are often described stochastically, solving the corrected discretized model yields an ensemble of perturbed model solutions which is then interpreted as a measure for the uncertainty in the model solution. However, even for simple models, this is a difficult undertaking, as is illustrated in [8] for a low-complexity Hamiltonian system. In a GFD environment, a closely related approach has become increasingly popular that aims at the quantification of parts of the model uncertainty due to model errors by employing stochastic sub-grid scale parametrizations [7, 4, 5]. Although these approaches are promising, they are not yet at the level where they can be used to estimate the full error due to discretization in a complex ocean model.

An important and widespread class of *a posteriori* error estimators is given by the

dual-weighted methods [1, 40, 3, 15, 17, 11, 19, 51]. They are designed to estimate the error in physical quantities of interest (goals) by combining the dual solution with local error information on the grid element level, i.e. the residual. The dual solution is the sensitivity of the goal to the residual. There are many flavors of dual-weighted methods that can be applied to various tasks, such as estimating corrections for a goal, estimating error bounds or deriving error indicators to guide adaptive mesh refinement. Although the method has a strong theoretical foundation, most of the work on dual-weighted methods is done for the Finite-Element Method (FEM). Only very little work is dedicated to other discretization schemes, in which it is then often tried to relate the used discretization schemes back to Finite-Element methods [31], or to estimate residual information from a series of coarse model solutions [13]. Both approaches are not deemed to be applicable here because the model is too complex to relate it back to FEM, and the model solutions are typically too under-resolved to derive residual information from them.

Recently, an approach for the quantification of the error due to discretization in a general GFD environment has been put forward in the literature [41]. The method belongs to the *a posteriori* dual-weighted error estimation techniques and yields confidence intervals on the error in a time-dependent goal, such as energetic quantities or fluxes through a cross-section. The method is itself embedded into the general framework of the dual-weighted error estimation approach [15] and extends it to the estimation of goal errors for inviscid 2D flows on the sphere. In the extension, the local truncation error, which represents the residual in the framework of [15], is replaced by a stochastic component. The step towards a stochastic representation of the local truncation errors was motivated by insights from the field of statistical mechanics.

1.1 Thesis Objective:

In this thesis, we develop a stochastic dual-weighted error estimation algorithm to quantify the discretization error in goals that is applicable to ocean model simulations. Our *a posteriori* error estimator is conceptionally based on the idea of [41] to replace the local truncation error by a stochastic component, however, we extend it to the specifics of ocean models in various ways.

One important aspect in this extension is the inclusion of a memory effect for the stochastic error component. The memory governs how future states of this stochastic component are connected to past states. We describe the representation of this memory and show how it needs to be adapted according to the complexity of the experiment and the numerical ocean model.

We evaluate our extended stochastic dual-weighted error estimation approach by studying it on important oceanic features. In a first step, we extend our error estimation algorithm to experiments with lateral boundaries. The resulting boundary currents

bring in the concept of regionally varying rates of error production in the computational domain, and the emergence of biases in the local truncation errors that are persistent in time. The fundamentals of these flows are studied in a 2D model environment. In a second step, we discuss our error estimator for a baroclinic instability in the flow. We do this within the framework of a 3D ocean model with full dynamics. This step introduces a downward energy cascade to the problem of error estimation and a vertical ocean structure that the algorithm needs to adhere to. The overarching question is whether and how a stochastic error component can be modeled that is dominated by the effects of mesoscale eddies.

Another important aspect of this work is the analysis of the stochastic extension of the dual-weighted error estimation approach in general. It is clarified how the method is related to classical ensemble techniques, and the analysis of the method's results is extended.

1.2 Thesis Outline

The thesis is structured as follows:

- In **chapter 2**, we define the basic mathematical notation and framework that will be used throughout this thesis and derive the dual-weighted error estimation approach of [15]. We then show how the dual-weighted error estimation approach can be linked to the field of statistical mechanics and in particular the Mori-Zwanzig formalism. With this background, the concept of a stochastic dual-weighted error estimator is introduced.

- In **chapter 3**, the stochastic error estimation algorithm of [41] is extended by introducing memory effects for the stochastic component that describes the temporal evolution of the local truncation error. This step broadens the range of applicability of the algorithm to flows with lateral boundaries. We show this by investigating our error estimation algorithm with a viscous shallow-water model on two classical oceanographic experiments, the Munk gyre and the flow around an island. Also, the relation between our stochastic error estimator and classical ensemble methods is explained. This chapter has been submitted for publication in the Journal of Computational Physics in 2016 and is currently under revision.

- In **chapter 4**, we lay the groundwork for applying the stochastic error estimator to the phenomenon of baroclinic instability. The testbed is an idealized ocean channel experiment implemented in a full 3D ocean model. In the experiment, a tilted temperature field drives the formation of baroclinic instabilities in the flow. A 3D flow structure with mesoscale eddies emerges that is reflected in the local

truncation error. We attempt to model the resulting stochastic component of the error estimator and discuss its properties.

In chapter 5, we conclude and give an outlook to possible future extensions and open questions that might be of importance for the further development of a stochastic dual-weighted error estimator towards full global ocean circulation simulations.

Chapter 2

Problem Statement and Approach

We here put the previously discussed notations about the different model error sources into a rigorous mathematical framework. Towards that, we define the concept of a model, continuous as well as discretized, in a general framework. Given this framework we then define the formulation error, the discretization error and the total model error in a physical quantity of interest.

We start by defining the continuous form of a general model

$$N\left(q\left(x,t\right)\right)=0,\ \ for\ \left(x,t\right)\in\Omega\times\left[0,T\right], \tag{2.1}$$

where N contains the model equations as well as initial conditions $q(x,0)=q^0$ and boundary conditions $q(x,t)=q^b(x,t)$ on $\partial\Omega$. The function $q\left(x,t\right)$ is defined on $\Omega\times[0,T]$ and satisfies system (2.1).

For an ocean model and more generally for a model in GFD, $N\left(q\left(x,t\right)\right)$ is usually given by partial differential equations derived from the Navier-Stokes equations that are combined with subgrid-scale parametrizations. The function $q(x,t)$ then represents the variables that define the state of the fluid at position x and time t and cannot be obtained by analytical means and is thus unknown. However, even if the continuous solution to system (2.1) were known, it would already be flawed in the way that the formulation error is already inherent in it. Thus, even the unknown continuous solution $q(x,t)$ would be erroneous if compared to nature.

What we are actually able to solve is a discretization of system (2.1) which can be written as

$$N_\Delta\left(q_\Delta\right)=0,\ \ for\ \left(x,t\right)\in\Omega_\Delta\times\{t_0,\ldots,t_n\}, \tag{2.2}$$

where N_Δ contains the discretized model equations in discrete time $T_\Delta:=\{t_0,\ldots,t_n\}$ and discrete space Ω_Δ with initial conditions $q^0_\Delta=P_\Delta(q^0)$ and boundary conditions $q_\Delta=P_\Delta(q^b)$ on $\partial\Omega_\Delta$. Here, Ω_Δ denotes the discrete model domain with boundary $\partial\Omega_\Delta$. P_Δ represents a projection $P_\Delta:V(\Omega)\to V_\Delta(\Omega_\Delta)$ from the continuous function space $V(\Omega)$ on Ω to the discretized function space $V_\Delta(\Omega_\Delta)$ on Ω_Δ. The solution of system

(2.2) is denoted by $q_\Delta = \begin{pmatrix} q_\Delta^0 \\ \vdots \\ q_\Delta^n \end{pmatrix}$ and is a vector of state vectors $q_\Delta^i \in \mathbb{R}^m,\ i = 0, \ldots, n$ in discrete time T_Δ and discrete space Ω_Δ.

The discretization error in the solution of the discretized system (2.2) is then given by the difference between its discrete solution q_Δ and the continuous solution q of the continuous system (2.1)

$$e_\Delta := q_\Delta - P_{\Omega_\Delta \times T_\Delta}(q), \tag{2.3}$$

where $P_{\Omega_\Delta \times T_\Delta}$ denotes a projection of the continuous solution q from the continuous function space $V(\Omega \times [0, T])$ on $\Omega \times [0, T]$ to the discretized function space $V_\Delta(\Omega_\Delta \times T_\Delta)$ on $\Omega_\Delta \times T_\Delta$.

For the investigated physical system, we are not only interested in estimates of the discretization error e_Δ but especially in the error in physical quantities of interest (goals) that characterize said physical system. As discretization errors are inherent to the model solution q_Δ, they are consequently also in the physical quantities of interest (goals) that are derived from q_Δ. In the continuous context, a goal is defined as a functional J that is applied to the continuous solution $q\,(x, t)$. We denote its discrete approximation by $J_\Delta(q_\Delta)$, where J_Δ is the discrete version of J. The discretization error ϵ in a goal is then the difference between the true value $J(q)$ and its approximation $J_\Delta(q_\Delta)$

$$\epsilon := J_\Delta(q_\Delta) - J(q), \tag{2.4}$$

that can be separated into two parts

$$\epsilon = \left(J_\Delta(P_\Delta(q)) - J(q)\right) + \left(J_\Delta(q_\Delta) - J_\Delta(P_\Delta(q))\right). \tag{2.5}$$

The first term is the error due to the discretization of the functional, and it is usually assumed that this term is sufficiently small and can be neglected. The second term

$$\epsilon_\Delta := J_\Delta(q_\Delta) - J_\Delta(P_\Delta(q)) \tag{2.6}$$

directly contains the error in a goal due to the difference between the solution q_Δ of the discretized model (2.2) and $q\,(x, t)$, the solution of the continuous model (2.1).

The problem can be stated as follows: For a given ocean model (2.1) and its discretized version (2.2), how can we obtain an estimate for the error in a goal ϵ? In this, the possible algorithm needs to be applicable to an already existing model, and thus should not be intrusive.

2.1 Dual-weighted Error Estimation

The technique by which we approach the problem is called dual-weighted error estimation. In dual-weighted error estimation approaches, the goal error ϵ is translated into a weighted sum of local truncation errors at the grid cell level. The method belongs to the *a posteriori* error estimators that are calculated for a specific model solution. The motivation for dual-weighted error estimation is that it is usually easier to obtain information about these local truncation errors than the discretization error e_Δ directly. In this section, we derive the dual-weighted approach of [15] based on the framework and notation provided in the previous section.

First, we perform a Taylor series expansion of $N_\Delta\left(q_\Delta\right)$ around the point $P_\Delta(q)$ up to first order which yields

$$N_\Delta\left(P_\Delta(q)\right) \approx N_\Delta\left(q_\Delta\right) - \frac{\partial N_\Delta}{\partial q_\Delta} e_\Delta. \tag{2.7}$$

Because $N_\Delta\left(q_\Delta\right)$ is zero by definition, we obtain a direct relationship between the discretization error e_Δ and $N_\Delta\left(q\right)$

$$N_\Delta\left(P_\Delta(q)\right) + \frac{\partial N_\Delta}{\partial q_\Delta} e_\Delta \approx 0, \tag{2.8}$$

which is equivalent to

$$e_\Delta \approx -\left(\frac{\partial N_\Delta}{\partial q_\Delta}\right)^{-1} N_\Delta\left(P_\Delta(q)\right). \tag{2.9}$$

$N_\Delta\left(P_\Delta(q)\right)$ is known as the local truncation error which is the result of inserting the continuous solution q into the discretized model N_Δ, i.e. the error done by the discretized model within one timestep.

The relationship (2.9) still has to be connected to the goal J. This is done by performing a Taylor expansion up to first order of J_Δ around $P_\Delta(q)$

$$J_\Delta\left(P_\Delta(q)\right) \approx J_\Delta\left(q_\Delta\right) - \frac{\partial J_\Delta}{\partial q_\Delta} e_\Delta. \tag{2.10}$$

The approximation (2.10) can be reformulated in terms of ϵ_Δ defined in (2.6)

$$\begin{aligned}
\epsilon_\Delta = J_\Delta\left(q_\Delta\right) - J_\Delta\left(P_\Delta(q)\right) &\approx \frac{\partial J_\Delta}{\partial q_\Delta} e_\Delta && (2.11)\\
&= -\frac{\partial J_\Delta}{\partial q_\Delta}\left(\frac{\partial N_\Delta}{\partial q_\Delta}\right)^{-1} N_\Delta\left(P_\Delta(q)\right) && (2.12)\\
&= {q_\Delta^*}^T N_\Delta\left(P_\Delta(q)\right). && (2.13)
\end{aligned}$$

The vector q^*_Δ is the discrete adjoint solution, which is the solution to the so called adjoint flow equations

$$\left(\frac{\partial N_\Delta}{\partial q_\Delta}\right)^T q^*_\Delta + \left(\frac{\partial J_\Delta}{\partial q_\Delta}\right)^T = 0. \tag{2.14}$$

The discrete adjoint solution thus provides a translation from the local truncation error to the error in the goal J_Δ.

Equation (2.13) can be used as an error estimator for ϵ_Δ

$$\epsilon_\Delta \approx \langle q^*_\Delta, N_\Delta(q)\rangle_{\Omega_\Delta \times T_\Delta}, \tag{2.15}$$

where $\langle\cdot\rangle_{\Omega_\Delta \times T_\Delta}$ is the standard euclidean scalar product in discrete time and discrete space. If we can further assume that the first term of (2.4) is small, i.e. the goal discretization is negligible, then (2.13) is also an error estimator for the full discretization error ϵ in the goal J

$$\epsilon \approx \langle q^*_\Delta, N_\Delta(q)\rangle_{\Omega_\Delta \times T_\Delta}. \tag{2.16}$$

This error estimator requires two components, the local truncation error and the discrete adjoint solution.

The discrete adjoint solution can be obtained in several ways. One way is to formulate the continuous adjoint flow equations to the continuous model (2.1), discretize them and solve the discretized adjoint flow equations. Another way is to formulate and solve the adjoint model to the discretized model (2.2). The similarities and differences between both approaches are illustrated in [48]. For a complex ocean model environment, we chose a third approach which is referred to as Algorithmic Differentiation (AD) [9, 20]. Basically, in AD the model run is divided into its elementary operations for which their derivative is known. Using the chain rule, the AD tool now moves backwards through the computational graph of the model run and accumulates the derivatives of the elementary operations which then yields the discrete adjoint solution. Once the model code has been prepared for a specific AD tool, obtaining the discrete adjoint solution is straightforward.

The local truncation error on the other hand is difficult to obtain continuously for all timesteps due to the complexity of an ocean model and is thus the focus of this thesis. The difficulty stems from how the local truncation error, or more generally residual information, is obtained. We here mention the term residual as it is frequently encountered when discussing the dual-weighted error estimation for the Finite-Element Method [40, 3, 17, 11, 19, 51]. Generally, the evaluation of the residual requires two different functional forms of the model, one from which the model solution is obtained and one in which the residual is evaluated [46]. For practical applications, where continuous model solutions are unavailable, obtaining residual information that can be used to drive dual-weighted error estimation then comes down to two basic approaches [47]. In

the first approach, the discrete model solution is evaluated in a higher-resolved model, which yields what we here refer to as the classical residual (see [11]). In the second approach used for the framework described in [15], a higher-resolved model solution is evaluated in the discrete model which yields the local truncation error. For many applications, the described residual evaluation is simply an insertion into the respective functional form and does not involve any additional model runs.

However, for complex geophysical fluid dynamic models such as an ocean general circulation model (OGCM), we deem the continuous evaluation of residual information at every model timestep to be prohibitive. For the evaluation of the classical residual, we expect the residual evaluation to be associated with high computational costs. For instance, taking the ICON-Ocean model [27] as an example, we expect the evaluation of the classical residual to cost at least 30 percent of a respective high-resolution model simulation, because many calculations performed within the model simulation are still required in the evaluation of the residual. Thus, continuously evaluating the classical residual is impracticable. For the evaluation of the local truncation error, an often-used, straight-forward approach is based on using a smooth reconstruction of the coarse model solution as a surrogate for the unknown higher-resolved model solution. However, in case of an OGCM, the coarse model solution is typically strongly under-resolved, especially due to the strong influence of resolution-dependent subgrid-scale parametrizations. Thus, its smooth reconstruction will not be a good representation of the respective higher-resolved model solution, and consequently we do not expect the resulting local truncation error estimates to be good approximations. Techniques that try to bypass the higher resolution entirely by inferring the local truncation error from a selection of local truncation error evaluations on even coarser model resolutions, see for instance the $\tau-$estimation technique [13], are then also not deemed to be applicable due to the under-resolution of the coarse model solutions. As a consequence, for models with the specifics of an OGCM neither the classical residual nor the local truncation error can be continuously evaluated in practice.

In order to make the dual-weighted error estimation technique available for these models, [41] developed a stochastic extension in which the local truncation error is represented by a stochastic process. This bypasses the need to continuously evaluate the local truncation error, but introduces the challenge of modeling a stochastic process that is able to mimic the temporal evolution of the local truncation error. The idea to stochastically extend the dual-weighted approach was motivated by the Mori-Zwanzig Formalism from statistical mechanics.

2.2 Mori-Zwanzig Formalism: Local Truncation Errors as Local Model Uncertainty

The Mori-Zwanzig Formalism from statistical mechanics [18, 8] illustrates how the discrete model solution q_Δ would need to be corrected in order to make up for the missing influence of the unresolved scales. First, we write our continuous model (2.1) in the form

$$N(q) := \frac{dq}{dt} - h(q) = 0, \tag{2.17}$$

where the function h represents the continuous dynamics of system (2.1). Given the initial and boundary conditions of (2.1), the solution of (2.17) is that of system (2.1).

The model (2.17) can be conceptually separated into two parts, the resolved or macroscopic dynamics and the unresolved dynamics. Towards this aim, we separate the solution q into a resolved part $\tilde{q}_{\Delta,t}$ and an unresolved part $\hat{q}$. Here, we interpret the resolved part $\tilde{q}_{\Delta,t}$ as a discrete function on our discrete space Ω_Δ but still in the continuous time-domain. After also performing the same separation for our model equations, the model (2.17) is separated into a macroscopic part that governs the evolution of the resolved components

$$\frac{d\tilde{q}_{\Delta,t}}{dt} - f(\tilde{q}_{\Delta,t}, \hat{q}) = 0, \tag{2.18}$$

and a microscopic part that governs the evolution of the unresolved part of the solution

$$\frac{d\hat{q}}{dt} - g(\tilde{q}_{\Delta,t}, \hat{q}) = 0. \tag{2.19}$$

In equation (2.18), we can already see that the unresolved part influences the evolution of the resolved components, because f is not only a function of $\tilde{q}_{\Delta,t}$ but also of $\hat{q}$. However, the system we are actually able to solve is the time-discretized version of equation

$$\frac{dq_{\Delta,t}}{dt} - \bar{f}(q_{\Delta,t}) = 0, \tag{2.20}$$

where $\bar{f}$ now only depends on the resolved variables and represents our spatial discretization scheme. In the Mori-Zwanzig notation this step would be seen as the application of a suitable projection operator to f that yields the so-called Markovian term $\bar{f}$. In system (2.20) the influence of the unresolved dynamics is not captured any more, and the solution to (2.20) is then of course not $\tilde{q}_{\Delta,t}$ anymore but its approximation $q_{\Delta,t}$. If we were to additionally introduce a discretization scheme for the time-domain, we would again obtain the solution q_Δ, the solution of system (2.2).

The Mori-Zwanzig Formalism now tells us how we would need to correct (2.20) in

order to again obtain the solution $\tilde{q}_{\Delta,t}$

$$\frac{d\tilde{q}_{\Delta,t}}{dt} - \bar{f}(\tilde{q}_{\Delta,t}) = \int_0^t K(\tilde{q}_{\Delta,t}(t-s), s)ds + \eta(\tilde{q}_{\Delta,t}(0), \hat{q}(0), t). \tag{2.21}$$

Two correction terms appear on the right-hand side of equation (2.21). The first term $\int_0^t K(\tilde{q}_{\Delta,t}(t-s), s)ds$ is called the memory term and the second term $\eta(\tilde{q}_{\Delta,t}(0), \hat{q}(0), t)$ is known as the orthogonal dynamics term. Both terms together are also referred to as the 'noise' term (see for instance [8])

$$n(t) := \int_0^t K(\tilde{q}_{\Delta,t}(t-s), s)ds + \eta(\tilde{q}_{\Delta,t}(0), \hat{q}(0), t).$$

For complex systems such as the shallow-water equations in an ocean-type setting or even a full 3D ocean model, we expect a direct derivation of a computable noise term or its components to be extremely difficult if not impossible. In these settings, it is even already difficult to disentangle the dynamics of the noise and clearly attribute them to be either stochastic or deterministic. For instance in the orthogonal dynamics, the initial conditions $\hat{q}(0)$ of the unresolved variables appear. From the point of view of the resolved scales, it makes sense to interpret $\hat{q}(0)$ as a stochastic quantity. On the other hand, a large part of the memory term is usually considered to be deterministic, but with the complexity of our systems of interest its computation is anything else than clear. Therefore, we hope to instead make progress by following a data-driven approach in which we choose to interpret the entirety of the noise term as a stochastic process.

If the noise term $n(t)$ is now represented by a stochastic process, the corrected solution $\tilde{q}_{\Delta,t}$ becomes a stochastic process too, which introduces the concept of uncertainty to the model solution. This model uncertainty is the source of the uncertainty in $J_\Delta(q_\Delta)$ that we want to estimate. We can apply the same line of reasoning to our dual-weighted error estimation approach, where the corrections to the model solution which we now interpret stochastically are the local truncation errors. In contrast to (2.21) the thereby introduced model uncertainty is however not propagated through the system directly, but *a posteriori* by the discrete adjoint solution.

The Mori-Zwanzig formalism thus motivates to represent the local truncation error stochastically as local model uncertainty in the form of a time-discrete stochastic process $X(\mathbf{p})$ with parameter set $\mathbf{p}$. The local model uncertainty at every timestep i is then represented by a random vector $X_{t_i}(\mathbf{p}_i)$ with parameters $\mathbf{p}_i$ and dimension $\mathbb{R}^m$, the same dimension as the state vectors q^i_Δ. The sequence of these random vectors defines the stochastic process $X(\mathbf{p}) := \{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ that is our stochastic representation of the local truncation error.

Replacing the local truncation error by the stochastic process $X(\mathbf{p})$ changes the de-

terministic dual-weighted error estimator (2.16) for ϵ to a stochastic error estimator

$$E_{app} := \langle q^*_\Delta, X(\mathbf{p})\rangle_{\Omega_\Delta \times T_\Delta} \tag{2.22}$$

that yields a random variable E_{app}. This random variable is a measure for the uncertainty in the goal J due to discretization errors in the solution q_Δ.

2.3 Extracting Local Uncertainty Information from Local Truncation Errors at Near-initial Timescales

Given we have already obtained a discrete model solution q_Δ, the discrete goal $J_\Delta(q_\Delta)$, the discrete adjoint solution q^*_Δ, and an appropriate stochastic process $X(\mathbf{p})$ has been chosen, then a major task lies in estimating a problem-specific parameter set $\mathbf{p}$. Towards this problem, [41] developed an algorithm for wave-type flows in a shallow-water model. It is based on estimating the parameters of the stochastic process from a near-initial learning period. For this time period, approximations to the local truncation error are made available from short model simulations on high-resolution. The resulting parameters are then extrapolated from the near-initial learning period to later times. In this, it is assumed that the estimated statistical properties of the local truncation errors at near-initial times are also capable to describe the behavior of the local truncation error at later times.

The proposed algorithm we build around this idea is an extension to the original algorithm by [41] that is able to cope with the higher complexity of flows we encounter. The following algorithmic steps are part of an algorithmic cycle that is formulated as questions that the user of the algorithm needs to answer in order to obtain a valid parameter set $\mathbf{p}$. This already indicates that the algorithm is in many regards dependent on educated guesses by the user who has to make decisions about the trade-off between estimating reliable parameters and keeping the estimation process as light-weight and computationally cheap as possible. The algorithmic cycle is exited when the user is satisfied with the performance of the obtained parameter set.
The algorithmic cycle looks as follows:

- **Step 1:** Do we encounter different dynamical flow regimes in the model solution? This is important because different dynamical flow regimes might come with significantly different rates of error production. If this is the case, it might be necessary to chose the parameter set $\mathbf{p}$ to vary locally according to the respective dynamical flow regime. The user then needs to decide how the computational domain is divided into regions of different dynamical flow regimes.

- **Step 2:** How long should the near-initial learning period be? The model should not be in a state of initial shock to ensure that the near-initial information is

valid for a long time-window after the near-initial learning phase. At the same time, the learning phase needs to be kept short to keep the computational cost to a minimum.

- **Step 3:** What is a suitable probability distribution for the stochastic process $X(\mathbf{p})$? A Gaussian distribution might be tempting as it is a standard choice to represent model errors, but in principle other probability distributions are possible and might fit better for certain dynamical flow regimes.
- **Step 4:** Which statistical parameters need to be estimated? The variance is here considered to be a standard quantity to be estimated, but do we also need estimates for the mean values or can the means be neglected? Also correlations, spatial as well as temporal, might be crucial if the local truncation error follows certain patterns in space or time. The correlation coefficients then need to be either estimated or upper-bounded.

In this algorithm, we assume that the stochastic process $X(\mathbf{p})$ has already been chosen beforehand, but the choice of the stochastic process could as well be integrated into the algorithmic cycle as an additional algorithmic step between step 2 and step 3. This additional step can become important in an environment where different stochastic processes are available that each have their specific strengths and weaknesses in modeling the local truncation error but would require different stochastic quantities to be estimated. One stochastic process might for instance be based on temporal fluctuations in the local truncation error, while another one models the local truncation error directly.

Throughout this thesis, we will constantly use this algorithmic framework to obtain the problem-specific stochastic process $X(\mathbf{p})$. With the stochastic process being estimated, all components of the stochastic error estimator (2.22) are known and error estimates for the goal J can be computed.

2.4 Research Questions

Throughout this thesis, we show how the algorithm presented above needs to be adapted to the specific oceanic phenomena in order to obtain a stochastic process for the local truncation error that can be used for goal-oriented error estimation. In this, we constantly increase the complexity of the employed ocean model and the studied numerical ocean experiments throughout this thesis.

For chapter 3, we study the described stochastic error estimation approach for ocean-type experiments with lateral boundaries in a viscous 2D shallow-water model.

- What probability distribution can be used to model the local truncation error?
- What is the relation between the stochastic error estimator (2.22) and a classical stochastic-physics ensemble?

- How big is the effect of the discretization of the goal itself compared to the error in the goal ϵ?
- What part of the stochastic process represents the memory term that is predicted by the Mori-Zwanzig formalism?
- How can spatial and temporal correlations be represented in the stochastic process? And, can these correlations be estimated reliably?
- Does this approach lead us to valid error estimates for physically relevant goals?

Chapter 4 is then concerned with the concept of baroclinic instabilities in the full 3D ocean model environment. We discuss the required properties of the stochastic process and lay the groundwork for a stochastic process that is able to model the local truncation error and can be applied in our stochastic error estimation approach.

- Does the concept of the local truncation error even make sense in an environment where the local truncation error primarily results from only partly resolving the mesoscale eddy field?
- How can the vertical structure, especially stratification, of the ocean experiment be preserved under the influence of the stochastic process?
- Is it possible to use a stochastic process to model the local truncation error that is due to not fully resolving the mesoscale eddy field?

Chapter 3

Stochastic Goal-oriented Error Estimation with Memory

We propose a stochastic dual-weighted error estimator for the viscous shallow-water equation with boundaries. For this purpose, previous work on memory-less stochastic dual-weighted error estimation is extended by incorporating memory effects. The memory is introduced by describing the local truncation error as a sum of time-correlated random variables. The random variables itself represent the temporal fluctuations in local truncation errors and are determined from high-resolution information at near-initial times.

The resulting error estimator is evaluated experimentally in two classical ocean-type experiments, the Munk gyre and the flow around an island. In these experiments, the stochastic process is adapted locally to the respective dynamical flow regime. Our stochastic dual-weighted error estimator is shown to provide meaningful error bounds for a range of physically relevant goals. We prove, as well as show numerically, that our approach can be interpreted as a linearized stochastic-physics ensemble.

3.1 Introduction

Quantifying the uncertainty due to discretization errors is essential in judging the quality of a numerical model solution. For many applications we are not only interested in estimates of the discretization error in certain norms but especially in estimating the resulting error in key physical quantities of interest (goals) such as energetic quantities or volume transports. These goals are linear or non-linear functionals of the model solution. A conceptual framework for this type of error estimation is provided by dual-weighted error estimation techniques [40, 3, 15, 17, 11, 19, 51]. Dual-weighted methods are applied to a model solution *a posteriori* and combine the adjoint model solution with residual information; both are deterministic quantities. The adjoint solution is the sensitivity of the goal with respect to the residual. In the context of this study, the residual is the local truncation error that describes to which extend the exact solution to the continuous equations fails to satisfy the discrete equations. A stochastic extension

of the dual-weighted error estimation approach was developed in [41] by modeling the local truncation error as a stochastic process. The purpose of this paper is to extend the stochastic dual-weighted error estimation by including memory effects and to gain a more profound theoretical understanding of the properties of stochastic dual-weighted approaches. This is done by theoretical analysis as well as by an experimental evaluation of the new error estimation algorithm.

For practical applications, where continuous model solutions are unavailable, obtaining residual information comes down to two basic approaches [47] that each can be used to drive dual-weighted error estimation. In the first approach, the discrete model solution is inserted into a higher-resolved discrete model which yields what we refer to as the classical residual, (see [11]). In the second approach, which is for instance used in [15], a higher-resolved discrete model solution is inserted into the discrete model which yields the local truncation errors.

We propose a stochastic representation of the local truncation error because the established methods for calculating the local truncation error and the classical residual via a higher-order reconstruction are difficult to obtain within our discrete model. The shallow-water model that we use to illustrate our method uses a finite-difference-finite-volume discretization on an unstructured grid with a staggered distribution of variables (Arakawa C-staggering). For such models, a higher-order interpolation of the state vector would for instance require to interpolate the normal components of the velocity vector from an unstructured and non-orthogonal grid at a coarse resolution to the same grid type at higher resolution. This task poses its own difficulties. An additional reason is due to the specific target application for which we aim to develop an error estimation algorithm, namely global ocean modeling. For this model framework, we expect a combination of high computational costs of such a higher-order representation, significant computational costs of the residual evaluation itself, and the typical under-resolution of ocean model solutions, which would make the established methods prohibitive.

An alternative approach to continuously obtaining residual information was introduced in [41] where local truncation errors are modeled as a stochastic process. This approach was proposed and tested for wave-type flows in a shallow-water model. The replacement of deterministic local truncation errors by a stochastic process was motivated by the Mori-Zwanzig formalism [33, 32, 54, 18] from statistical mechanics. The Mori-Zwanzig formalism is used here as a conceptual picture that provides us with a guideline of how a model that acts on a limited amount of (finite) scales could be supplemented to incorporate the influence of the unresolved scales. In this, we do not aim for a systematic or rigorous implementation of the Mori-Zwanzig formalism. From the point of view of the resolved scales, the influence of the unresolved scales can then be interpreted as being stochastic. In the dual-weighted error estimation approach, the information about the unresolved dynamics inherently lies within the local trunca-

tion errors. Therefore, local truncation errors can be interpreted stochastically as local model uncertainty, and the temporal evolution of this uncertainty can be described by a stochastic process. This poses the question of how to determine a suitable stochastic process. In [41], a white-noise process was chosen whose parameters were estimated using near-initial, high-resolution information, which was in accordance to the flows under consideration.

Our work is oriented along the lines of [41], but we go beyond this work in several aspects. First, we extend the algorithm to stochastic processes with memory, a property deemed to be highly important in modeling the effect of the unresolved scales (see e.g. [33, 32, 54, 18, 8]). The white-noise process used in [41] does not provide such a memory effect as future states are independent of previous states. Our algorithm now models the local truncation error by considering temporal fluctuations in local truncation errors at all previous timesteps, and thus naturally inherits a memory effect. Second, we deepen the analysis of the algorithm and the algorithm's results, we clarify its relation to ensemble techniques, and we discuss the assumption of the dual-weighted method underlying our approach, namely that the goal discretization error is assumed to be negligible. As a consequence of our algorithmic extension we are able to consider the estimation of errors in goals for two-dimensional flows with boundaries and viscosity.

In the framework of the two-dimensional shallow-water equations, we study two ocean-type experiments, the so-called Munk gyre and the flow around an island. The investigation of viscous flows with lateral boundaries poses several challenges. First, changing the model resolution can now coincide with a change in the model parameters, possibly introducing systematic biases in the local truncation errors that are persistent in time. Thus, local truncation errors at different timesteps cannot be assumed to be uncorrelated. Also, a memory term is needed to account for these biases. Second, we encounter transient flows, such as flows being spun up from initial rest. For these flows, the local truncation errors are expected to grow in time. For these reasons, a white-noise process as it was used in [41] is not sufficient any more. The third issue concerns the presence of lateral boundaries. As the dynamical flow regime near these boundaries changes, so does the production rate of local truncation errors. Information about a change of the flow regime needs to be featured into the stochastic process. We will show how to derive a stochastic process that satisfies these new requirements. An important assumption for our choice of underlying dual-weighted approach, which is for instance derived in [15], is that the error in the goal due to discretization of the goal itself is negligible. Our numerical results indicate that this assumption does not necessarily hold for our experiments. However, we show that, under certain conditions, we can correct for this error and thus still obtain viable error estimates.

The paper is structured as follows: In section 2, we explain the basic dual-weighted error estimation approach for a general framework and connect it to the shallow-water model and our chosen discretized model. Section 3 describes how our specific dual-

weighted error estimation framework can be carried into a stochastic framework. Based on this stochastic framework, we propose an algorithm to goal-oriented error estimation in Section 4. Section 5 describes the results of our error estimation algorithm on the two experiments including a discussion of the results, and in section 6 we conclude.

3.2 The Shallow-Water Equations and Goal Errors

The model that we use to illustrate our stochastic error estimation approach are the viscous shallow-water equations on the sphere, which are here given in the vector-invariant form by

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial t} &= -\left(\xi + f\right)\mathbf{k} \times \mathbf{v} - \nabla\left(gh + K\right) + \nu\nabla^2\mathbf{v} \\ \frac{\partial h}{\partial t} &= -\nabla \cdot \left(h^*\mathbf{v}\right). \end{aligned} \tag{3.1}$$

In this, $\mathbf{v} = (u, v)$ are the horizontal velocities, ξ is the relative vorticity, f the Coriolis parameter, g is the gravitational constant and $K = \frac{1}{2}(u^2 + v^2)$ is the kinetic energy per unit mass, $h = h^* + h_s$ is the total height of the free surface, ν is the viscosity parameter, h_s the orography and h^* the fluid thickness.

Our numerical model to obtain the discrete solution to the continuous shallow-water equations (3.1) is the ICON-Shallow-Water model [6, 45]. The model is based on a mixed finite-difference-finite-volume discretization in space implemented on a triangular grid on the sphere, using a C-type staggering for the prognostic variables height and velocity. For the discretization in time, we employ a fourth-order Runge-Kutta method. The discrete heights h_Δ are located at the center of grid cells, while the normal components of the discrete velocity vector reside on the midpoints of the triangle edges. The normal velocity components are denoted by vn_Δ.

We can write the continuous shallow-water equations (3.1) in the form of a general continuous model

$$N\left(q\left(x,t\right)\right) = 0, \tag{3.2}$$

with initial conditions $q(x, 0) = q^0$ and boundary conditions $q(x, t) = q^b(x, t)$ on $\partial\Omega$. The function $q\left(x,t\right)$ is defined on $\Omega \times [0, T]$ and satisfies system (3.2). The solution q then represents the functions h and $\mathbf{v}$ of system (3.1).

The discretization of the continuous system (3.2) can then formally be written as

$$N_\Delta\left(q_\Delta\right) = 0, \tag{3.3}$$

with initial conditions $q^0_\Delta = P_\Delta(q^0)$ and boundary conditions $q_\Delta = P_\Delta(q^b)$ on $\partial\Omega_\Delta$. Here, Ω_Δ denotes the discrete model domain with boundary $\partial\Omega_\Delta$. P_Δ represents a projection $P_\Delta : V(\Omega) \to V_\Delta(\Omega_\Delta)$ from the continuous function space $V(\Omega)$ on Ω to the

discretized function space $V_\Delta(\Omega_\Delta)$ on Ω_Δ. The solution of system (3.3) is denoted by $q_\Delta = \begin{pmatrix} q_\Delta^0 \\ \vdots \\ q_\Delta^n \end{pmatrix}$ and is a vector of state vectors $q_\Delta^i \in \mathbb{R}^m,\ i = 0, \ldots, n$ in discrete time $T_\Delta := \{t_0, \ldots, t_n\}$ and discrete space Ω_Δ. In the notation of the discretized shallow-water model, the vector q_Δ then represents the vector (h_Δ, vn_Δ) .

3.2.1 Errors in physical quantities of interest

Discretization errors are inherent in the model solution q_Δ, and consequently also in the physical quantities of interest (goals) that are derived from q_Δ. In the continuous context, a goal is defined as a functional J that is applied to the continuous solution $q\,(x,t)$. We denote its discrete approximation by $J_\Delta(q_\Delta)$, where J_Δ is the discrete version of J. The total error ϵ in a goal is then the difference between the true value $J(q)$ and its approximation $J_\Delta(q_\Delta)$

$$\epsilon := J_\Delta(q_\Delta) - J(q), \tag{3.4}$$

that can be separated into two parts

$$\epsilon = (J_\Delta(P_\Delta(q)) - J(q)) + (J_\Delta(q_\Delta) - J_\Delta(P_\Delta(q)))\,. \tag{3.5}$$

The first term is the error due to the discretization of the functional, and in our error estimation approach for the ICON-Shallow-Water model we assume that this term is sufficiently small and can be neglected. The second term

$$\epsilon_\Delta := J_\Delta(q_\Delta) - J_\Delta(P_\Delta(q)) \tag{3.6}$$

is the error in the discrete goal J_Δ due to the discretization errors in the model solution. This error can be approximated via the dual-weighted error estimation approach that is illustrated in [15]

$$\epsilon_\Delta \approx \epsilon_{app} := \langle q_\Delta^*, N_\Delta(q)\rangle_{\Omega_\Delta \times T_\Delta}. \tag{3.7}$$

Here, $\langle\cdot\rangle_{\Omega_\Delta \times T_\Delta}$ denotes the standard Euclidean scalar product in discrete space and time, q_Δ^* is the discrete adjoint solution to q_Δ, and $N_\Delta(q)$ are the local truncation errors. More precisely, we note here that the definition of the local truncation error should of coarse also include an application of P_Δ and would thus be written as $N_\Delta(P_\Delta(q))$. We however choose to omit P_Δ here and in all future mentions simply to increase readability.

For our ICON-Shallow-Water model, the two components in the scalar product (3.7) can be defined as following. The vector of local truncation errors at timestep i is defined

by

$$N_\Delta^i(q) = P_\Delta\left(q(x,t_i)\right) - S_\Delta\left(P_\Delta\left(q(x,t_{i-1})\right)\right), \tag{3.8}$$

where S_Δ is one application of the discrete time-stepping operator of our model.The discrete adjoint solution q_Δ^* consists of the adjoint components for the discrete heights h_Δ and the velocities vn_Δ and is here obtained via algorithmic differentiation, which is described in detail in [42] for the ICON-Shallow-Water model.

Combining an error estimate ϵ_{app} derived from (3.7) with equation (3.6) could then be used to correct for the error in $J_\Delta(q_\Delta)$, or derive bounds $\epsilon_1, \epsilon_2 > 0$ that satisfy

$$J_\Delta(q_\Delta) - \epsilon_1 < J_\Delta(P_\Delta(q)) < J_\Delta(q_\Delta) + \epsilon_2. \tag{3.9}$$

Here our aim is to derive an error estimator similar to (3.9). To do this, we require approximations for the local truncation errors, but as already pointed out in the introductory part of this thesis the standard approaches are not applicable for our type of numerical model. Our alternative approach to approximating the local truncation errors is motivated by the Mori-Zwanzig formalism (see section 2.2).

3.3 A Stochastic Framework for Dual-weighted Error Estimation

We now formulate a stochastic framework for the dual-weighted error estimation approach. We wish to obtain a discrete-time stochastic process that mimics the temporal evolution of the local truncation errors. This stochastic process is denoted by $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ consisting of random vectors $X_{t_i} \in \mathbb{R}^m$ with parameters $\mathbf{p}_i$. The temporal character of the Euclidean scalar product in equation (3.7) becomes obvious when written as a weighted sum of local truncation errors at the timesteps $i = 1, \ldots, n$

$$\epsilon_{app} = \sum_{i=1}^{n} \langle q_\Delta^{*\,i}, N_\Delta^i(q)\rangle_{\Omega_\Delta}, \tag{3.10}$$

where $\langle\cdot\rangle_{\Omega_\Delta}$ denotes the Euclidean scalar product in discrete space Ω_Δ, $q_\Delta^{*\,i}$ is the vector of the discrete adjoint solution at timestep i, and $N_\Delta^i(q)$ is the vector of local truncation errors at timestep i. Our ansatz then follows from rewriting the local truncation errors at timestep i by the sum of the local truncation errors at the previous timestep $i-1$

plus a small fluctuation. Equation (4.7) then becomes

$$\epsilon_{app} = \sum_{i=1}^{n} \langle q_{\Delta}^{*\,i}, N_{\Delta}^{i-1}(q) + \big(N_{\Delta}^{i}(q) - N_{\Delta}^{i-1}(q)\big)\rangle_{\Omega_{\Delta}} \tag{3.11}$$

$$= \sum_{i=1}^{n} \langle q_{\Delta}^{*\,i}, \sum_{j=1}^{i} \Big(N_{\Delta}^{j}(q) - N_{\Delta}^{j-1}(q)\Big)\rangle_{\Omega_{\Delta}}, \tag{3.12}$$

with fluctuations $N_{\Delta}^{i}(q) - N_{\Delta}^{i-1}(q)$. Note that $N_{\Delta}^{0}(q)$ is simply a zero vector. We bring in stochasticity by assuming that the distribution of the vectors $N_{\Delta}^{i}(q) - N_{\Delta}^{i-1}(q)$ can be described by random variables $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$,which are scalar quantities with parameter parameter sets $\mathbf{p}_{[t_{i-1},t_i]}$ that characterize their distribution. Each entry of the vector $N_{\Delta}^{i}(q) - N_{\Delta}^{i-1}(q)$ is thus interpreted as one realization of the random variable $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$. Given that the random variables $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$ are known, the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ is defined immediately. The $k-th$ entry of its random vectors $X_{t_i}(\mathbf{p}_i)$, a scalar quantity denoted by $X_{t_i,\,k}(\mathbf{p}_i)$ with $X_{t_i}(\mathbf{p}_i) := (X_{t_i,\,k}(\mathbf{p}_i))_{k=1,\ldots,m}$, can be written as the sum

$$X_{t_i,\,k}(\mathbf{p}_i) := \sum_{j=1}^{i} Y_{[t_{j-1},t_j]}(\mathbf{p}_{[t_{j-1},t_j]}). \tag{3.13}$$

Assuming the existence of the random variable's mean $E(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})) = \mu_{[t_{i-1},t_i]}$ and its variance $Var(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})) = \sigma^2_{[t_{i-1},t_i]}$, the mean and the variance of $X_{t_i,\,k}(\mathbf{p}_i)$ at timestep i follow from the summation rule for random variables. The mean can be calculated straightforwardly as

$$E(X_{t_i,\,k}(\mathbf{p}_i)) = \sum_{j=1}^{i} E(Y_{[t_{j-1},t_j]}(\mathbf{p}_{[t_{j-1},t_j]})) = \sum_{j=1}^{i} \mu_{[t_{j-1},t_j]} \tag{3.14}$$

The variance follows from

$$\begin{aligned} Var(X_{t_i,\,k}(\mathbf{p}_i)) = \; & Var(X_{t_{i-1},\,k}(\mathbf{p}_{i-1})) + Var(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})) \\ & +2Cov(X_{t_{i-1},\,k}(\mathbf{p}_{i-1}), Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})). \end{aligned} \tag{3.15}$$

The covariance inherits the temporal correlation of the fluctuations at different timesteps. Since we want to estimate bounds for the error in a goal, we choose an upper bound for the covariance, which is given by the Cauchy-Schwarz inequality

$$\Big|Cov(X_{t_{i-1},\,k}(\mathbf{p}_{i-1}), Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]}))\Big| \leq \sqrt{Var(X_{t_{i-1},\,k}(\mathbf{p}_{i-1}))}\sqrt{Var(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]}))}. \tag{3.16}$$

The upper bound of this estimate is attained in case of maximal correlation between $X_{t_{i-1},\,k}(\mathbf{p}_{i-1})$ and $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$ and thus the maximum growth rate in variance of the stochastic process. Combining equation (3.15), the upper bound on the covariances (3.16), and the variance of the random variable $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$ then yields the variance of $X_{t_i,\,k}(\mathbf{p}_i)$

$$Var(X_{t_i,\,k}(\mathbf{p}_i)) = \sum_{j=1}^{i} \sigma^2_{[t_{j-1},t_j]} + 2\sum_{1\leq j}\sum_{<k\leq i} \sigma_{[t_{j-1},t_j]}\sigma_{[t_{k-1},t_k]}. \tag{3.17}$$

In order to further simplify these sums, we will assume certain properties for the stochastic process:

- First, we assume that $N^1_\Delta(q) - N^0_\Delta(q)$ is known. The random vector $X_{t_1}(\mathbf{p}_1)$ is then chosen to have the mean vector $N^1_\Delta(q) - N^0_\Delta(q)$ and variances $Var(X_{t_1}(\mathbf{p}_1)) = \mathbf{0}$. If there are systematic biases in the local truncation errors, the mean will act as a bias correction for the local truncation errors.
- Second, we assume that the random variables $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$ are entirely defined by their means and variances,
- and third we assume that the random variables' distributions are centered

  $$E(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})) = 0,$$

  and their variance is constant throughout all timesteps i:

  $$Var(Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})) = \sigma^2.$$

  As all random variables $Y_{[t_{i-1},t_i]}(\mathbf{p}_{[t_{i-1},t_i]})$ are the same, we denote them by $Y_{\Delta t}(\mathbf{p}_{\Delta t})$ for simplicity.

The stochastic process then belongs to the class of L2-stochastic processes that are characterized by their first two moments. The subsequent random vectors $X_{t_i}(\mathbf{p}_i)$ will then have mean vectors

$$E(X_{t_i}(\mathbf{p}_i)) = N^1_\Delta(q) - N^0_\Delta(q) \tag{3.18}$$

and variance vectors

$$Var(X_{t_i}(\mathbf{p}_i)) = (i-1)^2\sigma^2\mathbf{1}, \tag{3.19}$$

where $\mathbf{1}$ denotes the all-ones vector in $\mathbb{R}^m$. We note that in this derivation we have neglected possible spatial correlations of the temporal fluctuations in local truncation errors. Nevertheless, information about spatial correlations is contained in the mean

vector and is thus part of the stochastic process. If we can assume that the temporal fluctuations in local truncation errors are small compared to the mean vector, the spatial pattern in the mean vector dominates possible spatial correlations in the temporal fluctuation in local truncation errors. In comparison to [41], the variance of the stochastic process now grows in time and the stochastic process is not centered around zero any more.

We have not yet assumed a probability distribution for the random variables $Y_{\Delta t}(\mathbf{p}_{\Delta t})$ or discussed the probability distribution of the stochastic process. For our stochastic process, the type of probability distribution of its entries $X_{t_i,\,k}(\mathbf{p}_i)$ follows directly from the chosen probability distribution of the random variables $Y_{\Delta t}(\mathbf{p}_{\Delta t})$. The reason is that the random variables $X_{t_i,\,k}(\mathbf{p}_i)$ are defined as sums of fully-time correlated random variables $Y_{\Delta t}(\mathbf{p}_{\Delta t})$, which means that the shape of the sum's distribution is preserved and only the sum's variance grows in time. This fact makes it easy to create realizations of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$. For each entry of $X_{t_1}(\mathbf{p}_1)$, one realization is drawn from $Y_{\Delta t}(\mathbf{p}_{\Delta t})$ and added i times to its respective entry in order to obtain $X_{t_{i+1}}(\mathbf{p}_{i+1})$. The probability distribution that we use in our numerical experiments is a Laplace distribution. Our method is however not limited to these probability density functions.

3.3.1 Deriving an Error Estimator from Parametrized Local Truncation Errors

Replacing the local truncation errors in (4.7) by a stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ makes ϵ_{app} a random variable, and we denote this random variable by E_{app}

$$E_{app} = \langle q^*_\Delta, \begin{pmatrix} X_{t_1}(\mathbf{p}_1) \\ \vdots \\ X_{t_n}(\mathbf{p}_n) \end{pmatrix} \rangle_{\Omega_\Delta \times T_\Delta}, \tag{3.20}$$

where each random vector $X_{t_i}(\mathbf{p}_i) \in \mathbb{R}^m$ is multiplied by the transpose of vector ${q^*_\Delta}^i \in \mathbb{R}^m$. E_{app} then represents an estimate for the probability distribution of the error in the goal $J_\Delta(q_\Delta)$. For our experiments, it is realistic to assume that the distribution of E_{app} converges towards a Gaussian distribution. To illustrate this, we look at the contribution to E_{app} at each timestep, which can be represented as a sum of adjoint-weighted random variables. For our experiments, we know that the adjoint solution does not have any singularities. Also, we know that we can expect to repeatably find adjoint-weighted random variables with similar variances. This in turn means that the central limit theorem holds for the sum of these adjoint-weighted random variables, making the contribution of each timestep to E_{app} a Gaussian-distributed random variable.

The parameters of this Gaussian variable can be estimated by creating realizations

of it. For each realization of E_{app}, the scalar product (3.20) is calculated using a realization of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$.

Ultimately, we want to obtain a result that is similar to the information contained in equation (3.9). This brings us to confidence intervals that give us a certain probability for the error in the goal being within the given bounds

$$J_\Delta(q_\Delta) - \epsilon_1 < J(q) < J_\Delta(q_\Delta) + \epsilon_2, \tag{3.21}$$

where $\epsilon_1, \epsilon_2 \in \mathbb{R}$. Here, we define ϵ_1, ϵ_2 not to be restricted to positive numbers anymore, which means that our confidence intervals do not necessarily need to confine $J_\Delta(q_\Delta)$. This makes sense because the resulting random variable E_{app} can have a non-zero mean if our stochastic process has non-zero means.

We identify two criteria for the resulting error estimates that need to be fulfilled, in order to consider the estimate as valid:

1. The resulting confidence intervals need to confine the 'true' goals.

2. Additionally, we want the confidence intervals to be as tight around the 'true' goal as possible. To quantify this, we define an effectivity index eff as the ratio between the half confidence interval width $\frac{\epsilon_2+\epsilon_1}{2}$ and the total error in the goal ϵ

$$eff := \frac{\epsilon_2 + \epsilon_1}{2\epsilon}. \tag{3.22}$$

   For our experiment we choose a value of up to 10 to still be 'useful', which was also chosen in [16]. This value is however a subjective choice based on visual impressions and could potentially be done differently by other users of the error estimator.

There is the possibility that only one of these two criteria is fulfilled for a given confidence interval, thus we only consider an error estimate as valid if both criteria are met.

In summary, we identify three main conditions that are crucial for the resulting confidence intervals to hold. First, the assumptions on the random variable $Y_{\Delta t}(\mathbf{p}_{\Delta t})$ need to hold. Second, the discrete adjoint solution needs to be a valid indicator for the influence of model perturbations on the goal. Third, the error due to discretization of the goal has to be negligible.

3.3.2 Connection to Established Ensemble Techniques

In this section we establish a connection between our error estimator and stochastic-physics ensembles. This ensemble method is here understood such that a random forcing is applied to the current state vector q^i_Δ at every timestep. In our case, the

random forcing is provided by the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ in (3.20). We now provide theoretical evidence that the distribution of goals J_Δ, calculated from model solutions perturbed this way, is comparable to correcting $J_\Delta(q_\Delta)$ by our stochastic dual-weighted approach (3.20).

Theorem 1. *Let q be the solution to the continuous model (3.2) and $N^i_\Delta(P_\Delta q)$ the corresponding vectors of local truncation errors at timesteps i, where N_Δ denotes the discretized model (3.3). Then, consecutively adding $N^i_\Delta(P_\Delta q)$ to the state vectors q^i_Δ at all timesteps $k = 1, \dots, i$ yields a corrected state vector at timesteps i, denoted by $q^i_{\Delta,\,C}$, and the state vector at timestep n satisfies*

$$q^n_{\Delta,\,C} = P_\Delta q(x, t_n). \tag{3.23}$$

The corrected state vectors $q^i_{\Delta,\,C}$ are thus identical to the continuous solution q at all grid points for all discrete timesteps.

Proof. The proof is done by induction. Without loss of generality, we assume that the local truncation errors can be written by means of a discrete time-stepping operator S_Δ

$$N^i_\Delta(P_\Delta q) = P_\Delta\left(q(x, t_i)\right) - S_\Delta\left(P_\Delta\left(q(x, t_{i-1})\right)\right). \tag{3.24}$$

Basis:
We show that the statement holds for the first timestep $n = 1$. q^1_Δ is calculated by applying one application of the discrete time-stepping operator of our model S_Δ to q^0_Δ

$$q^1_\Delta = S_\Delta\left(q^0_\Delta\right). \tag{3.25}$$

Now, we correct q^1_Δ by the first vector of local truncation errors $N^1_\Delta(q)$

$$N^1_\Delta(P_\Delta q) = P_\Delta\left(q(x, t_1)\right) - S_\Delta\left(P_\Delta\left(q(x, t_0)\right)\right). \tag{3.26}$$

The corrected solution $q^1_{\Delta,\,C}$ is then

$$q^1_{\Delta,\,C} = q^1_\Delta + P_\Delta\left(q(x, t_1)\right) - S_\Delta\left(P_\Delta\left(q(x, t_0)\right)\right). \tag{3.27}$$

By definition of the discrete system (3.3), $P_\Delta\left(q(x, t_0)\right) \equiv q^0_\Delta$, which means that we can replace $S_\Delta\left(P_\Delta\left(q(x, t_0)\right)\right)$ by q^1_Δ. This yields

$$q^1_{\Delta,\,C} = P_\Delta q(x, t_1) \tag{3.28}$$

which proves our statement to be true for $n = 1$.
Inductive Step:
In the induction step from n to $n + 1$ we use that $q^n_{\Delta,\,C} = P_\Delta q(x, t_n)$. It follows

immediately that

$$\begin{aligned} q^{n+1}_{\Delta,\,C} &= S_\Delta\left(P_\Delta q(x,t_n)\right) + N^{n+1}_\Delta(P_\Delta q) \\ &= S_\Delta\left(P_\Delta q(x,t_n)\right) + P_\Delta\left(q(x,t_{n+1})\right) - S_\Delta\left(P_\Delta\left(q(x,t_n)\right)\right) \\ &= P_\Delta q(x,t_{n+1}), \end{aligned}$$

by using the induction statement in the second line. □

This result tells us that correcting $J_\Delta(q_\Delta)$ by the scalar product (3.7) is an approximation of directly calculating $J_\Delta(q^n_{\Delta,\,C})$ from a consecutively corrected state $q^n_{\Delta,\,C}$. Thus, we can also expect qualitatively similar results in case of replacing the local truncation errors by our stochastic process. In general, the distribution of the stochastic-physics ensemble might be different from the probability distribution resulting from the error estimator (3.20) due to the nonlinearity of the model. The advantage of using an adjoint-based approach lies in its lower computational cost.

3.4 Goal Error Ensemble Method

In this section, we propose an algorithm based on the framework described in section 2. To obtain the probability distribution of the error in the goal E_{app} from (3.20), we need the algorithm to yield estimates for the parameters of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$. The main ingredient to estimate these parameters is the knowledge about the local truncation errors $N_\Delta(q)$ and thus of the unknown $q(x,t)$. To bypass the knowledge of the analytical solution $q(x,t)$, we make use of information from higher grid resolutions. More specifically, we use information from a near-initial, high-resolution solution q_{high}. Near-initial means that we start at the same initial conditions but integrate the highly-resolved model for only a small number of timesteps m. The solution q_{high} is then interpreted as our 'true' solution, and the vectors $N^i_\Delta(q_{high}),\ i=0,\dots,m$ are used as approximations to the vectors $N^i_\Delta(q),\ i=0,\dots,m$. With this step, we make the assumption that the local truncation error characteristics from near-initial behavior can be extrapolated to a much longer time window.

In the case that $q(x,t_i)$ is replaced by a high-resolution solution q_{high}, we need to actually define a mapping P_Δ that fits to our ICON-Shallow-Water model. Our choice is motivated by the structure imposed by the grid refinement process of grids in ICON. To perform one refinement for a grid of chosen resolution, the triangle edges are simply bisected. Each triangle is thus split into four refined sub-triangles. Thus, we choose to use point-wise projection to map the heights onto the coarse resolution. As each edge is bisected into an even number of edges, the velocities are chosen to be mapped onto the coarse grid by averaging over the two innermost velocity values.

We formulate our algorithm description for a stochastic process that is fully described

by the parameters mean and variance. The algorithm is, however, not dependent on this specific choice, and other parameters may be chosen according to the probability distribution of the random variable $Y_{\Delta t}(\mathbf{p}_{\Delta t})$.

We identify three steps for our error estimation algorithm, a pre-processing, a processing and a post-processing step. The parameter estimation for the stochastic process is at the core of the method and is performed in the processing step. The general structure of our algorithm is similar to the goal error ensemble method in [41]. The differences are the separation of different flow regimes that is at the core of a now cyclic processing step, and we now estimate random variables for the temporal fluctuations of the local truncation errors $Y_{\Delta t}(\mathbf{p}_{\Delta t})$ instead of estimating the random vector $X_{t_i}(\mathbf{p}_i)$ directly.

3.4.1 Pre-processing: Calculate the approximate goal $J_\Delta(q_\Delta)$ and the corresponding adjoint solution q^*_Δ

Our error estimation algorithm is an *a posteriori* method. Therefore, the pre-processing part consists of solving the model (3.7) to calculate the discrete model solution q_Δ and finally the approximate goal $J_\Delta(q_\Delta)$. This specific $J_\Delta(q_\Delta)$ is then the initial value for the calculation of the discrete adjoint solution q^*_Δ.

3.4.2 Processing: Derive a problem-specific stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$

The following processing steps describe a cycle that needs to be run until the parameter sets of the stochastic process are satisfactory. Because there might be many possibilities for reasonable stochastic processes $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$, this calls for a subjective user decision to exit the processing step.

1. Separate different flow regimes:
   As a consequence of the multiscale nature of Geophysical Fluid Dynamics, we expect to encounter different dynamical flow regimes and thus changing error behavior throughout the space domain. Therefore we divide the flow into regions of different dynamical regimes. This means separating Ω_Δ into subsets $\Omega_{\Delta,k}, k = 1, \ldots, l$. We choose the subsets to be disjunct and their union to be again Ω_Δ. Each region $\Omega_{\Delta,k}$ then has its own parameter set for the random variable $Y_{\Delta t}(\mathbf{p}_{\Delta t})$. The regional change in parameter set is indicated by the subscript k in $\mathbf{p}_{\Delta t,k}$. The separation of flow dynamics has to be visible and derived just from the knowledge of the coarse-grid solution q_Δ. Furthermore, for every flow there might be several possibilities to separate the flow regime. In principle, if dynamical features are prominent in one field only, different variables could use different separations.

2. Choose length of near-initial learning phase m:
   For simplicity we assume that i timesteps at standard resolution are equivalent to ki timesteps, with $k \in \mathbb{N}$, at high resolution. In total, we then have to perform km timesteps on the high-resolution grid. Applying the coarse model $N_\Delta^i(q_{high})$ then yields the approximations for the local truncation errors for the first m timesteps.

3. Choose the probability distribution of the random variables $Y_{\Delta t}(\mathbf{p}_{\Delta t,k})$:
   Our experimental experience shows that for the flows considered here a Laplace distribution fits better than a Gaussian. For other models, entirely different probability distributions might have to be considered.

4. Estimate the parameter sets $\mathbf{p}_{\Delta t,k}$:
   The entries of the vectors $N_\Delta^i(q_{high}) - N_\Delta^{i-1}(q_{high})$ that correspond to subset $\Omega_{\Delta,k}$ are interpreted as realizations of the random variable $Y_{\Delta t}(\mathbf{p}_{\Delta t,k})$. The parameter estimation is performed separately for each model variable. To calculate $\mathbf{p}_{\Delta t,k}$, we first estimate the first two moments, mean and variance, of the truncation error rate of change from timestep $i-1$ to i

$$\mu_{[t_{i-1},t_i],k} := \frac{1}{\#_{\Omega_{\Delta,k}}} \sum_{j \in \Omega_{\Delta,k}} N_\Delta^i(q_{high})(j) - N_\Delta^{i-1}(q_{high})(j) \tag{3.29}$$

   and

$$\sigma_{[t_{i-1},t_i],k} := \sqrt{\frac{1}{\#_{\Omega_{\Delta,k}} - 1} \sum_{j \in \Omega_{\Delta,k}} \left(\left(N_\Delta^i(q_{high})(j) - N_\Delta^{i-1}(q_{high})(j) \right) - \mu_{[t_{i-1},t_i],k} \right)^2}. \tag{3.30}$$

   $\#_{\Omega_{\Delta,k}}$ is the number of model variables in region $\Omega_{\Delta,k}$. $N_\Delta^{i-1}(q_{high})(j)$ denotes the j-th entry of vector $N_\Delta^{i-1}(q_{high})$. By using time averages of the estimated values $\mu_{[t_{i-1},t_i],k}$ and $\sigma_{[t_{i-1},t_i],k}$, we intend to make our parameter estimations more robust

$$\mu_{\Delta t,k} := \frac{1}{m-d} \sum_{i=d}^{m} \mu_{[t_{i-1},t_i],k} \tag{3.31}$$

$$\sigma_{\Delta t,k} := \frac{1}{m-d} \sum_{i=d}^{m} \sigma_{[t_{i-1},t_i],k}, \tag{3.32}$$

   with $d < m$. d should be chosen such that the solution is not in a state of an initial shock any more and is able to represent the error characteristics of the following timesteps.

3.4.3 Post-processing: Calculate Confidence Intervals on the Error in the Goal

Our stochastic process $\{X_{t_i}(\mathbf{p}_i)\}$ is then fully described by the estimated parameter sets $\mathbf{p}_{\Delta t,k}$ and the first vector of local truncation errors $N^1_\Delta(q_{high})$. With this knowledge we can estimate the parameters μ and σ of the Gaussian random variable E_{app}. These two moments are sufficient to calculate the confidence intervals described in (3.21), where ϵ_1 and ϵ_2 are chosen as

$$\epsilon_1 = \mu - c\sigma,$$

$$\epsilon_2 = \mu + c\sigma.$$

The constant $c \in \mathbb{R}$ relates to the reliability of the error estimator, and we find a reasonable choice for our application to be $c = 3$ for 99.7% confidence intervals. These confidence intervals act as bounds for the error in the goal. Repeating the algorithm for a series of integration times t_n yields a stochastic process of random variables E_{app} that represents the temporal evolution of confidence intervals.

3.5 Numerical Results

We evaluate our error estimation algorithm in two ocean-type experiments, the Munk gyre and a flow around an island. We show how the algorithmic steps we have described lead to confidence intervals that bound the error in a goal. The Munk gyre experiment is designed to show the capabilities of our error estimator in the case of resolution-dependent viscosity parameters. In this experiment, the flow is in steady state. The flow around an island experiment is designed to show the capabilities of our error estimator for a flow that is in a transition phase. Here, the viscosity parameters do not change with resolution. The separation of these two factors, parametrization and transition, enables us to study their influence on our error estimation algorithm independently.

For both experiments, the error estimation procedure is started from a mapped-down, spun-up state of a high-resolution run. To these two experiments, we apply three physically relevant goals: the area-averaged potential energy, the area-averaged kinetic energy, and the volume flux through a cross-section. We choose the energies as goals because of their non-linearity in their respective variable, whereas the volume flux is an important quantity of interest in oceanography, where often flows through cross-sections, such as the Drake passage, are discussed.

For each combination of goal and experiment, we derive a series of confidence intervals that are distributed equidistant in time. The resulting temporal evolutions of confidence intervals are discussed and compared to reference goals calculated from high-resolution solutions. In order to illustrate that the error estimator works for different resolutions, confidence intervals are calculated for two sets of model resolutions, resolution $\Delta 3$ with

Table 3.1: Averages over the square roots of the triangle areas for employed grids

Resolution	Grid Spacing in km
$\Delta 3$	320
$\Delta 4$	158
$\Delta 5$	79
$\Delta 6$	39

$\Delta 5$ as reference resolution and resolution $\Delta 4$ with $\Delta 6$ as reference resolution. In this grid notation, the numbers refer to the number of edge bisections performed on a base grid. The corresponding grid spacings are given in table 4.1. With this specific choice of the two sets of model resolutions, the distance between the grids is kept constant in order to highlight the convergence properties of our error estimator.

To demonstrate the connection of our error estimator to the class of stochastic-physics ensembles (see 3.3.2), we also show series of error bounds derived from full stochastic-physics ensembles (SPE) that are forced by the respective stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$.

3.5.1 Goals

The three different goals are all area-averaged or length-averaged quantities, calculated from the state vector at one specific timestep. The region or cross-section that is averaged over is denoted by Ω_R, $A(\Omega_R)$ is its area or length respectively, and their discretized counterparts are $\Omega_{\Delta R}$ and $A(\Omega_{\Delta R})$. In this section, $A(\Omega_{\Delta R})$ is chosen to be equal to the size of one specific element at resolution $\Delta 2$, which coincides well with the size of the dynamic flow features encountered in the two experiments. For a chosen grid Ω_Δ, consisting of cells C^i and edges E^i, the three diagnostics can be expressed in the following way.

Area-averaged Potential Energy: The area-averaged potential energy in region $\Omega_{\Delta R}$ is defined by:

$$J_1(q_\Delta^n) := \frac{g}{2A(\Omega_{\Delta R})} \sum_{C^i \in \Omega_{\Delta R}} A(C^i) h_{\Delta i}^2. \tag{3.33}$$

For the displayed results we omit the factor $\frac{g}{2}$.

Area-averaged Kinetic Energy: For the kinetic energy, the velocities vn_Δ, residing on the edges, are first interpolated into the cell centers by radial basis function interpolation. This yields one zonal velocity component u and one meridional velocity component v per cell. With these velocities we can define the area-averaged kinetic

energy by:

$$J_2(q_\Delta^n) := \frac{1}{2A(\Omega_{\Delta R})} \sum_{C^i \in \Omega_{\Delta R}} A(C^i)(u_i^2 + v_i^2). \tag{3.34}$$

Volume Flux: For the volume flux, the heights h_Δ are first interpolated onto the edges. These height variables are denoted by $\hat{h}_\Delta$. We then define the volume flux by:

$$J_3(q_\Delta^n) := \frac{1}{A(\Omega_{\Delta R})} \sum_{E^i \in \Omega_{\Delta R}} A(E^i)(vn_{\Delta i}\hat{h}_{\Delta i}). \tag{3.35}$$

3.5.2 Experiments

Experiment 1: Munk Gyre

The first experiment is the Munk gyre experiment, a wind-forced gyre circulation in an idealized ocean basin, for which the main feature is a western boundary intensification [34], which we will refer to as "Munk layer" from now on. The basin extends in longitudinal direction from 35W to 35E and in longitudinal direction from the equator to 60N. Due to the triangular grid, the shape is not fully rectangular as in the original Munk gyre experiment. We choose the basin to have a depth of 1 km with flat bathymetry. As for the wind-forcing, the meridional wind stress τ_λ is set to zero, while the zonal wind stress τ_ϕ is prescribed as

$$\tau_\phi = \frac{-1.08 \cdot 10^{-03} cos\left(\frac{\pi\lambda}{60^\circ}\right) \frac{m^2}{s^2}}{h_\Delta},$$

where λ is the latitude and h_Δ is the height. This wind-forcing mimics the mean winds prevailing in the Atlantic basin that change from westwards winds to eastwards winds with latitude. The viscosity for the simulation on $\Delta 3$ is chosen in a way that the flow at the western boundary is resolved by at least one grid-point [50]. A criterion for the viscosities of the subsequent higher resolution runs is then given by [23], where the aim is to have the same ratio between inertial and frictional forces in the Munk layer at all resolutions. The resulting eddy viscosities and timestep lengths are given in table 4.2. From a physical point of view, decreasing viscosity when increasing resolution ensures a more realistic representation of the gyre circulation. Figure 3.1 shows the spun up state of the Munk gyre at resolution $\Delta 5$. The error estimation procedure is started after a short spin-up phase of approximately one week of integration time on the respective reference resolution. This high-resolution solution is then mapped onto the coarse-resolution grid using the operator P_Δ. The error estimation is started from

Table 3.2: Eddy viscosity parameters and timestep lengths on different resolutions for the Munk gyre experiment

Resolution	ν in $\frac{m^2}{s}$	Δt in s
$\Delta 3$	4.92E+05	600
$\Delta 4$	3.15E+05	600
$\Delta 5$	1.9E+05	600
$\Delta 6$	1.15E+05	300

this initial state and run for another 10 days of integration time.

Experiment 2: Flow Around an Island

The second experiment is a flow around an island that is located at the equator. The flow is driven by wind-forcing that has a cosine shape in latitude and has its maximum intensity at the equator. We define two setups for this experiment. Although the general setups are that of an ocean-type experiment, we here choose the parameters such that the flow velocities and heights are one to two orders of magnitude higher than typical oceanographic values to cover different value ranges for the flow variables.

In the main setup, which we refer to as "flow around an island setup", the viscosity parameter is chosen to be $\nu = 6.5 * 10^6 \frac{m^2}{s}$ for all resolutions. We prescribe a zonal wind stress of

$$\tau_\phi = \frac{1.08cos\left(\frac{\pi\lambda}{180^\circ}\right)\frac{m^2}{s^2}}{h_\Delta},$$

which is 1000 times stronger than the forcing for the Munk gyre experiment and acts in opposite direction. The timestep is $60s$ for resolutions $\Delta 3$ to $\Delta 5$ and $30s$ for resolution $\Delta 6$. Figure 3.2 shows the state of the flow around the island setup after an initial spin-up phase of 30 hours at resolution $\Delta 5$. After the initial spin-up phase at the respective reference resolutions, the typical flow velocity upstream of the island is about $80\frac{m}{s}$ for all resolutions. For the island with a width of around $1500km$ and viscosity $\nu = 6.5*10^6\frac{m^2}{s}$ the Reynolds-number is about 20, and thus the flow is still in the laminar regime. As the flow is further spun up during the error estimation phase of additional 24 hours of integration time, a pair of symmetric vortices develops in the flow behind the island.

The second setup is especially designed to show that the error estimator also works for regimes with higher Reynolds numbers, and we refer to this setup as "Kármán vortex street setup". In this setup, the wind forcing is chosen to be a factor 16 smaller than in the main flow around an island setup, and the viscosity is $\nu = 2.4*10^4\frac{m^2}{s}$. The timestep is chosen to be $60s$ for all resolutions. The resulting velocity upstream of the

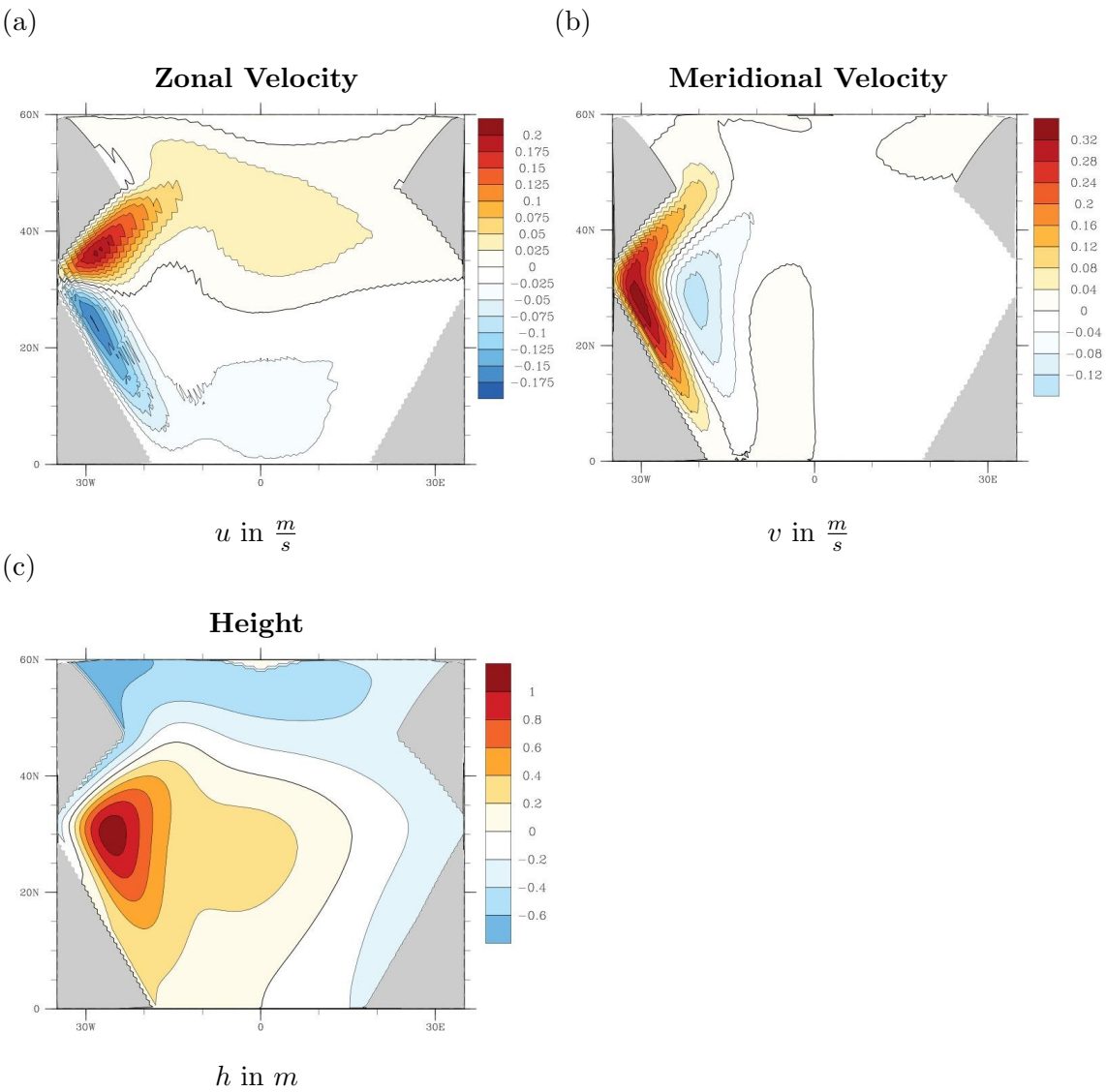


Figure 3.1: State of the flow after the short spin-up phase of one week of integration time, the starting time of the error estimation, on the reference resolution $\Delta 5$ in the Munk gyre experiment. (a) zonal velocities, (b) meridional velocities, (c) heights.

island is about $20\frac{m}{s}$. Thus, the calculated Reynolds-number is 1250 for which a Kármán vortex street downstream of the island emerges. This dynamical feature develops after an initial spin-up phase of 4 days of integration time. The error estimation phase is an additional 24 hours of integration time. The state of the flow at the end of the error estimation phase is shown in figure 3.3. The Kármán vortex street is most pronounced in the meridional velocities (Figure 3.3 b). For this setup, error estimation is only performed at resolution $\Delta 4$ with reference solution $\Delta 6$.

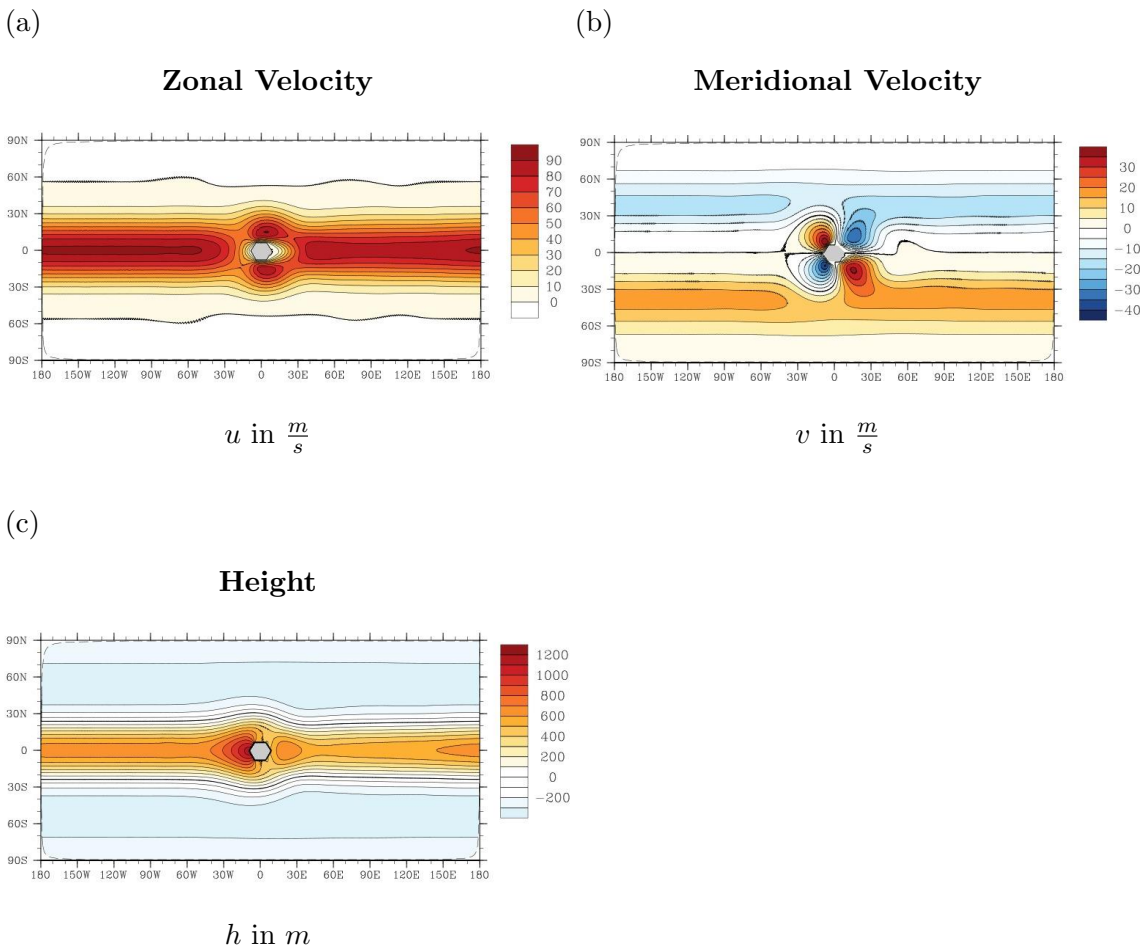


Figure 3.2: State of the flow after the spin-up phase of 30 hours, the starting time of the error estimation, on the reference resolution $\Delta 5$ for the flow around an island setup. The island is located at the center of each plot. (a) zonal velocities, (b) meridional velocities, (c) heights.

3.5.3 Estimating the Problem-specific Stochastic Processes $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$

We now present the results for the processing cycle of our algorithm that is described in section 3 and show how to obtain a problem-specific stochastic process for a given experiment. We start with a description of the chosen separation of dynamical flow regimes:

- Experiment 1: Munk gyre
  We identify two dynamical regimes. The first is the region of the western boundary current west of $15W$ longitude, with high flow velocities of up to $0.32\frac{m}{s}$ in the boundary layer and a boundary layer separation at around $45N$. Compared to

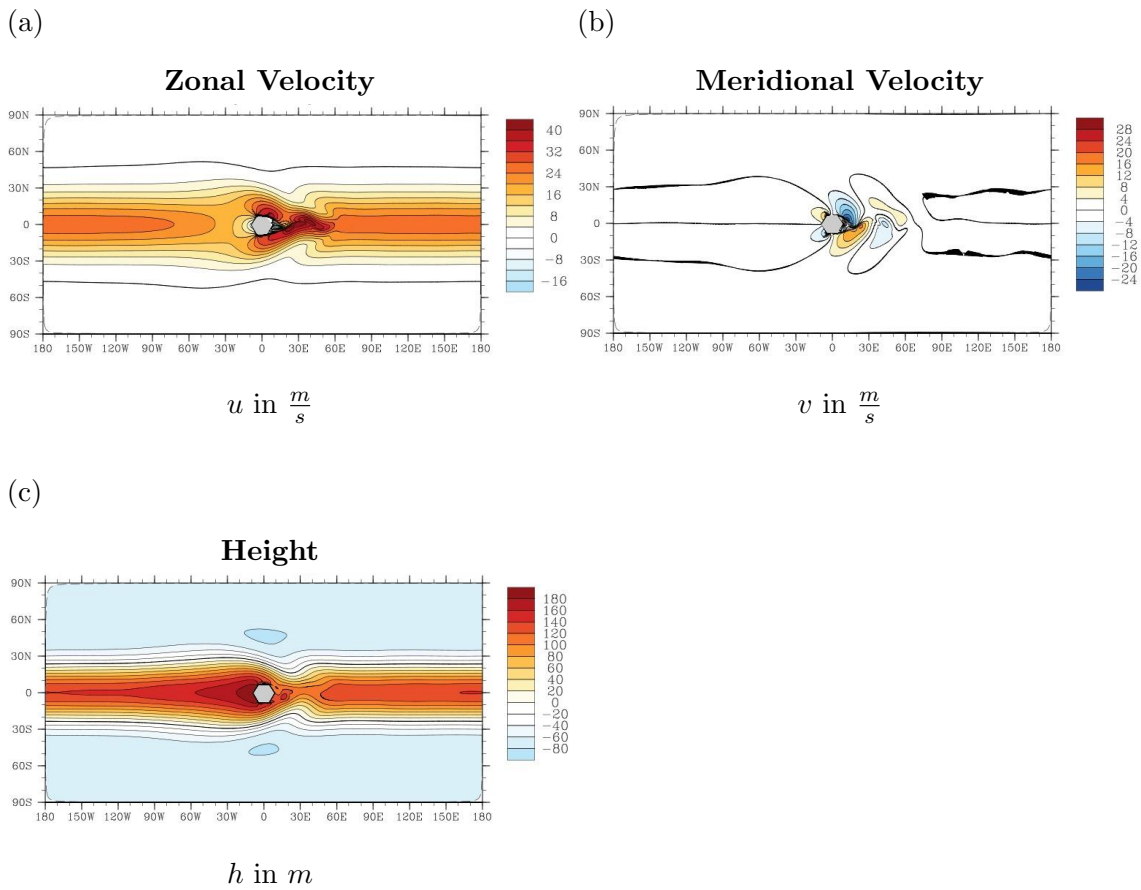


Figure 3.3: State of the flow after 5 days of integration time, at the end of the error estimation phase, on the reference resolution $\Delta 6$ for the flow around an island setup. The island is located at the center of each plot. (a) zonal velocities, (b) meridional velocities, (c) heights.

this western boundary region, the area-averaged kinetic energy in the rest of the basin is several orders of magnitude smaller. Our choice is therefore to separate the basin into the Munk layer region west of longitude $15W$, with its western boundary current, and the area east of longitude $15W$, referred to as the basin interior.

- Experiment 2: flow around an island
  Similarly to the Munk gyre experiment, we separate the computational domain into two regions.
  For the flow around an island setup, based on the initial state at the start of the error estimation, we choose to have the island's boundary layer to be confined by a circle with radius 40, with the center on the equator at $15E$. We refer to this

region as the boundary region. The area lying outside of this circle is called the interior region, to be consistent with the Munk gyre experiment. This separation ensures that all boundary effects are confined in one region.

For the Kármán vortex street setup, the changed dynamical flow regime due to the emerging vortex street in the wake of the island motivates the definition of the boundary layer region by a rectangle with the lower left corner at $10W, 10S$ and the upper right corner at $40E, 10N$. This choice ensues that the vortex shedding is enclosed within the boundary region, as this is a major source of local truncation error production. The area lying outside of the rectangle is defined as the interior region.

For both experiments and all resolutions, we choose the length of the near-initial learning phase to be $m = 20$ timesteps. As the probability distribution of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$, we choose a Laplace distribution. The corresponding probability density function is given by

$$f(x|b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right), \tag{3.36}$$

where the shape parameter b is defined by the variance, $2b^2 = \sigma^2$. Thus, the stochastic process is completely defined by its mean and variance.

For these choices, we perform the last step of the processing part, which is estimation of the parameter sets $\mathbf{p}_{\Delta t,k}$ that now only consist of the parameters $\mu_{\Delta t,k}$ and $\sigma_{\Delta t,k}$. We choose $d = 10$, thus average the parameters for 10 timesteps.

We find that we can neglect the influence of the means $\mu_{\Delta t,k}$, because they are consistently at least one order of magnitude smaller than $\sigma_{\Delta t,k}$ for all regions and timesteps. We prove that the influence of the means on the uncertainty in the goal is negligible. For the random vector $X_{t_i}(\mathbf{p}_i)$, the mean is the sum $\sum_{j=2}^{i} \mu_{\Delta t,k} = (i-1)\mu_{\Delta t,k}$, and its standard deviation is $(i-1)\sigma_{\Delta t,k}$. The difference of one order of magnitude is thus persistent throughout all timesteps. Calculating the scalar product (3.20) also does not change this difference, as it is a linear operation, and since we use c times σ intervals in the end, the influence of the means becomes even smaller. As an example, for $3\sigma = 90$, the associated influence of the means would be of the order of 3. Thus, we will only give values for the standard deviations.

The estimated parameters for all experiments are given in table 3.3. As expected, the standard deviations in the boundary regions are bigger than the ones in the interior regions for all experiments and all resolutions. For the Munk Gyre, the parameters are consistently bigger for resolution $\Delta 4$ with reference resolution $\Delta 6$. This is counterintuitive at first, because we would expect the error growth rate to decrease when increasing resolution because of convergence of the solution. However, the resolution-dependent viscosity parameters hinder a fast convergence of the solution. The error growth rate

Table 3.3: Estimated parameters at resolution $\Delta 3$ with reference $\Delta 5$ and at resolution $\Delta 4$ with reference $\Delta 6$ for all experiments

	Resolution $\Delta 3$		Resolution $\Delta 4$	
Munk gyre	$\sigma_{\Delta t,Boundary}$	$\sigma_{\Delta t,Interior}$	$\sigma_{\Delta t,Boundary}$	$\sigma_{\Delta t,Interior}$
h	2.23E-05	1.09E-05	2.62E-05	1.53E-05
vn	8.24E-07	7.15E-07	1.12E-06	8.39E-07
Flow around an island	$\sigma_{\Delta t,Boundary}$	$\sigma_{\Delta t,Interior}$	$\sigma_{\Delta t,Boundary}$	$\sigma_{\Delta t,Interior}$
h	2.78E-04	4.24E-05	1.61E-04	2.51E-05
vn	4.18E-05	1.31E-05	4.48E-05	1.30E-05
Kármán vortex street			$\sigma_{\Delta t,Boundary}$	$\sigma_{\Delta t,Interior}$
h			5.46E-04	8.97E-06
vn			1.54E-05	8.96E-07

might thus just be bigger because the differences in the flow between resolution $\Delta 4$ and $\Delta 6$ are slightly bigger than the differences between $\Delta 3$ and $\Delta 5$. Comparing the parameter values of the Kármán vortex street setup and the flow around an island setup, it is noteworthy that the parameters in the boundary region are of the same order of magnitude. As the velocity upstream of the island in the Kármán vortex street setup is however four times smaller, the relative error growth rate is much larger in the boundary region for the Kármán vortex street setup.

From these standard deviations, we derive the shape parameter b. For resolution $\Delta 3$ with its reference resolution $\Delta 5$, we show the spatial distribution of the temporal fluctuations $N_{\Delta}^{20}(q_{high}) - N_{\Delta}^{19}(q_{high})$ (blue bars) in figure 3.4 for the Munk gyre experiment, in figure 3.5 for the flow around an island setup,and in figure 3.6 for the Kármán vortex street setup. We find these distributions to be sufficiently stable for the entire time of the error estimation and all experiments. We compare the distributions to the relative frequencies given by the estimated random variables $Y_{[t_{19},t_{20}]}(\mathbf{p}_{[t_{19},t_{20}],k})$ (dashed red line), where the shape parameter b is taken to be equal to the standard deviations $\sigma_{[t_{19},t_{20}],k}$. This choice of shape parameter is basically an overestimate. However, a Laplace distribution with this choice of parameters is a reasonable fit (Figures 3.4, 3.5, and 3.6). We thus also choose the shape parameters of the random variables $Y_{\Delta t}(\mathbf{p}_{\Delta t,k})$ to be equal to the standard deviation parameters $\sigma_{\Delta t,k}$, which then fully defines our problem-specific stochastic processes $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$.

3.5.4 Error Estimates

We now study the temporal evolutions of the confidence intervals for the chosen goals. For a combination of goal and region, we calculate 20 confidence intervals equidistantly

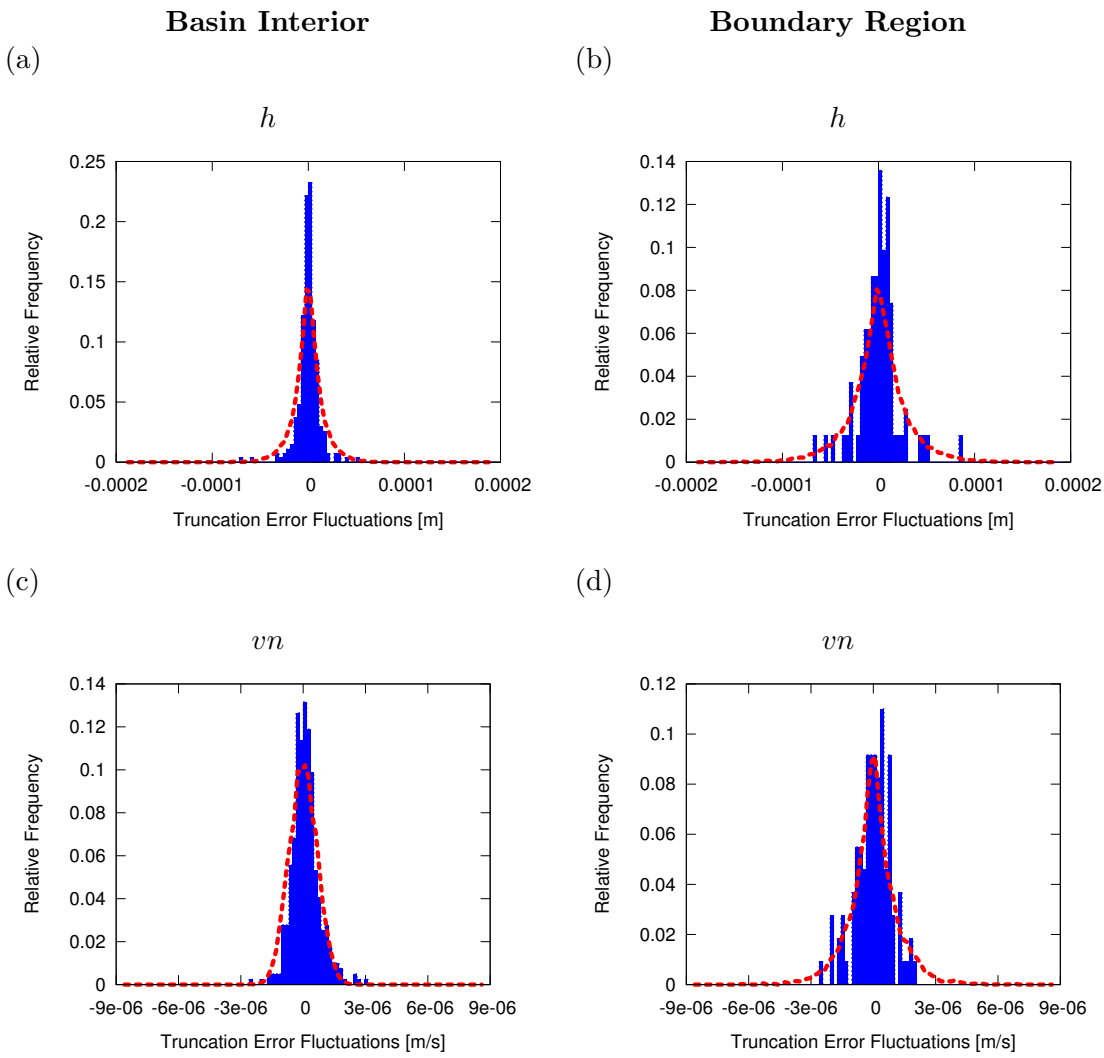


Figure 3.4: Distribution of the temporal fluctuations in local truncation errors $N_\Delta^{20}(q_{high}) - N_\Delta^{19}(q_{high})$ (blue bars) and the distribution of the estimated random variable $Y_{[t_{19},t_{20}]}(\mathbf{p}_{[t_{19},t_{20}]})$ (red dashed line) in the Munk gyre experiment. (a) h in the basin interior, (b) h in the boundary region, (c) vn in the basin interior region, (d) vn in the boundary region.

in time, covering the whole time interval of 1440 timesteps. We focus on the regions near boundaries.

In our analysis, we make the distinction between two reference goals, $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$, to be able to discuss both error terms appearing in (3.5) separately. Ideally both forms of the reference goal are bounded by our confidence intervals. In this case our main criterion for an error estimator is fulfilled. If, however, $J_{high}(q_{high})$

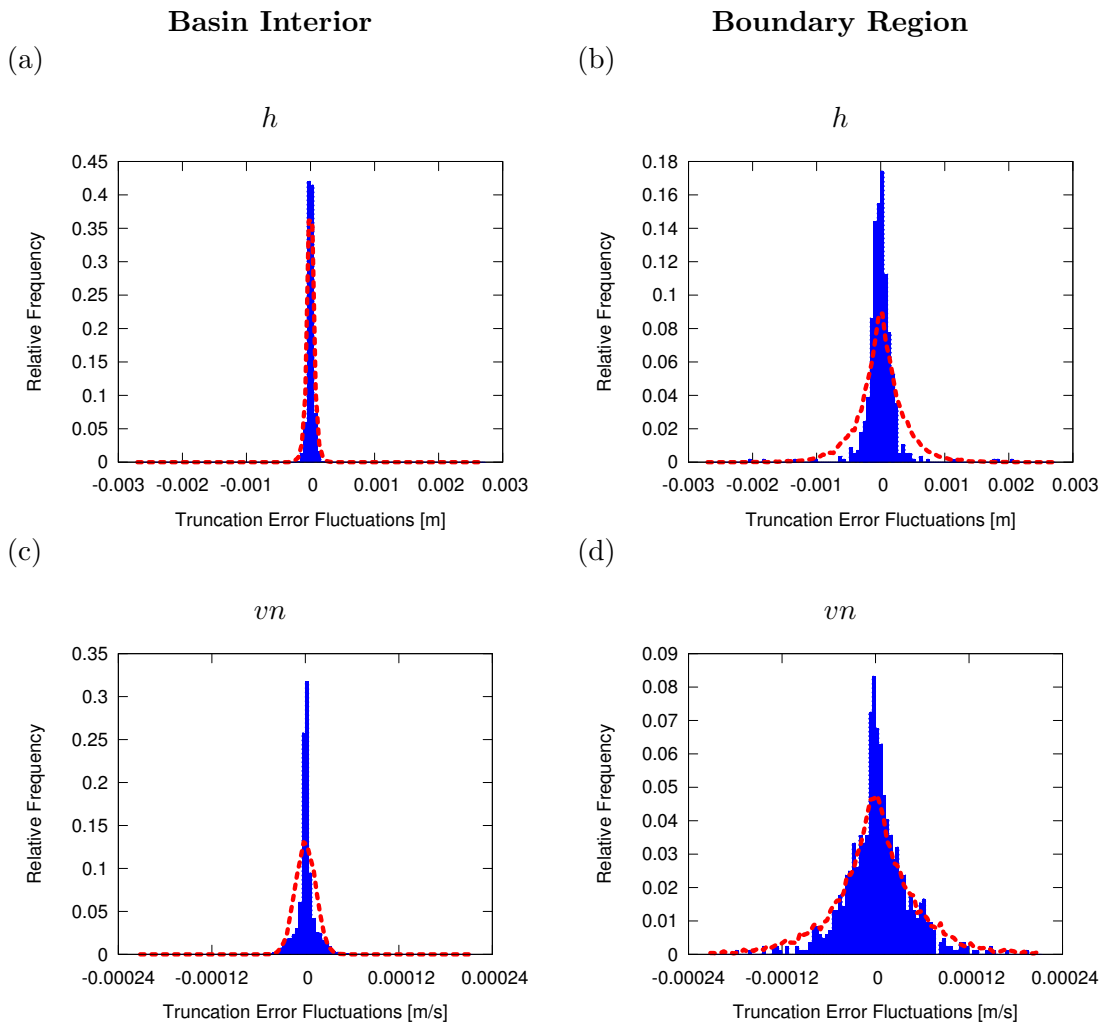


Figure 3.5: Distribution of the temporal fluctuations in local truncation errors $N_\Delta^{20}(q_{high}) - N_\Delta^{19}(q_{high})$ (blue bars) and the distribution of the estimated random variable $Y_{[t_{19},t_{20}]}(\mathbf{p}_{[t_{19},t_{20}]})$ (red dashed line) in the flow around an island setup. (a) h in the basin interior, (b) h in the boundary region, (c) vn in the basin interior region, (d) vn in the boundary region.

is not bounded, we can distinguish between two cases.

- If $J_\Delta(P_\Delta(q_{high}))$ is not bounded by the confidence interval either, our main criterion from section 3.3.1 not being fulfilled can be attributed to our error estimator, the scalar product (3.20). We can then also identify which component of the scalar product is responsible by looking at the approximated goal $J_\Delta(q_\Delta)$ that is corrected by the scalar product $\langle q_\Delta^*, N_\Delta(q_{high})\rangle$. This means, $J_\Delta(q_\Delta)$ is cor-

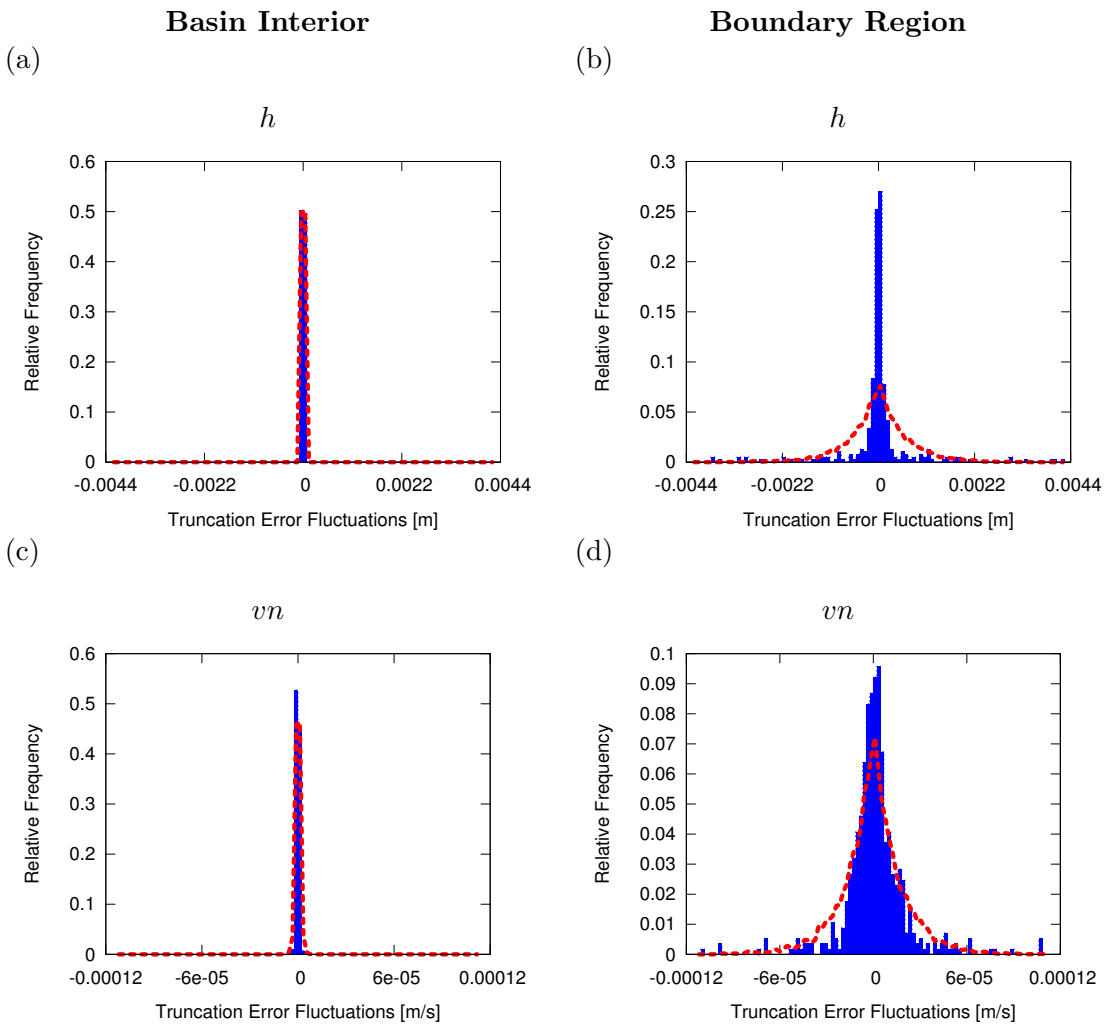


Figure 3.6: Distribution of the temporal fluctuations in local truncation errors $N_\Delta^{20}(q_{high}) - N_\Delta^{19}(q_{high})$ (blue bars) and the distribution of the estimated random variable $Y_{[t_{19},t_{20}]}(\mathbf{p}_{[t_{19},t_{20}]})$ (red dashed line) in the Kármán vortex street setup. (a) h in the basin interior, (b) h in the boundary region, (c) vn in the basin interior region, (d) vn in the boundary region.

rected with the full knowledge about the local truncation errors at all timesteps. If this corrected goal is bounded by the confidence interval, the adjoint solution q_Δ^* is responsible for the bad performance of the error estimator and is thus not a valid surrogate for the full model information. If the corrected goal is however not bounded, the only remaining error source is the chosen stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$.

- If $J_\Delta(P_\Delta(q_{high}))$ is bounded by the confidence intervals but $J_{high}(q_{high})$ lies outside, the discretization error in the goal cannot be neglected. In this case we have reached a fundamental limit of any dual-weighted error estimator based on framework (3.7).

This rationale enables us to analyze the results of our error estimator and its limitations.

Confidence Intervals at Resolution $\Delta 3$

We find $J_\Delta(P_\Delta(q_{high}))$ to be bounded by the confidence intervals not only for the following selection of results but for all possible combinations of experiment, region and goal at this resolution.

Experiment 1: Munk gyre

The results for the Munk gyre experiment are summarized in figure 3.7. For this experiment, we show confidence intervals for the Munk layer. For each goal, we show one temporal evolution of confidence intervals for one specific region, except for the kinetic energy, where we give results for two regions of the Munk Layer region. The system is in steady state, however, the steady state for the model resolution $\Delta 3$ is different from the steady state of the 'true' reference solution at resolution $\Delta 5$. In particular, the kinetic energy in Region 2 stabilizes at around $0.04\frac{m^2}{s^2}$ for resolution $\Delta 5$ compared to $0.02\frac{m^2}{s^2}$ for resolution $\Delta 3$, a relative error of 100 percent.

For the potential energy, the volume flux and the Kinetic energy in Region 1, we find our main criterion to be fulfilled; the confidence intervals confine the 'true' goal $J_{high}(q_{high})$. We now discuss whether the second criterion is fulfilled, which requires that the effectivity index eff given by the ratio between 3σ and ϵ does not exceed 10. We find that for certain phases of the error estimation the confidence intervals are even small enough that they do not confine $J_\Delta(q_\Delta)$ anymore. In these cases, eff is below 1. This occurs approximately for the first 50 hours of integration time and it tells us that for this time window we have not just derived a bound on the error in the goal but we can go beyond this and even effectively correct for this error. After this initial phase, the effectivity index grows continuously and eventually reaches 10. For the Kinetic energy in Region 1 eff exceeds a value of 10 after 225 hours of integration time. For the volume flux this threshold is crossed after 150 hours of integration time, and for the potential energy eff exceeds 10 after 120 hours of integration time. In figure 3.8, we illustrate how the confidence intervals behave in comparison to intervals that indicate when the effectivity index eff exceeds a value of 10. As an example we here choose the volume flux goal but we find the same behavior for the other goals as well. The intervals for the effectivity index increase strongly in comparison to the confidence intervals for the first 75 hours of integration time. This is because the flows have not reached their respective steady states yet, and as a result the error in the goal is strongly increasing in this phase. However, after the steady states are reached, the intervals for eff do not

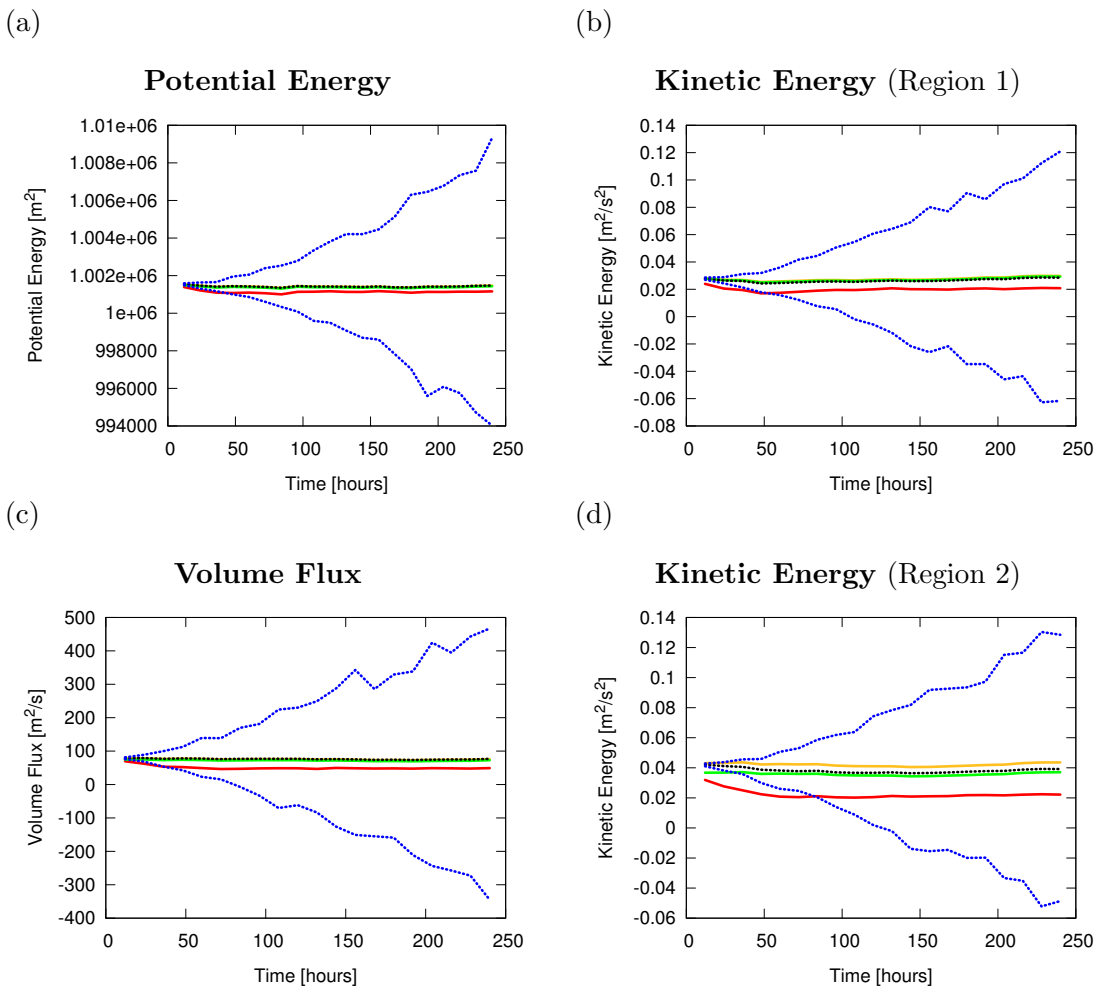


Figure 3.7: Temporal evolution of four different goals: (a) the potential energy, (b) the kinetic energy (Region 1), (c) the volume flux, and (d) the kinetic energy (Region 2) for the Munk gyre experiment at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to the discrete goal applied to a high resolution solution $J_\Delta(P_\Delta(q_{high}))$ (solid orange line) and $J_{high}(q_{high})$ (solid green line). If $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$ coincide, only $J_{high}(q_{high})$ is shown. The confidence intervals (dashed blue lines) are derived to confine $J_{high}(q_{high})$, $J_\Delta(P_\Delta(q_{high}))$, and $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line).

grow further while the confidence interval width keeps increasing at a steady pace. The skill of the error estimate is thus very high in the early phases of the error estimation but then steadily decreases in time until our second criterion is not fulfilled any more after 150 hours of integration time.

We now compare the behavior of the confidence intervals for the kinetic energy goals in Region 1 and Region 2. Both regions are located at the center of the Munk layer

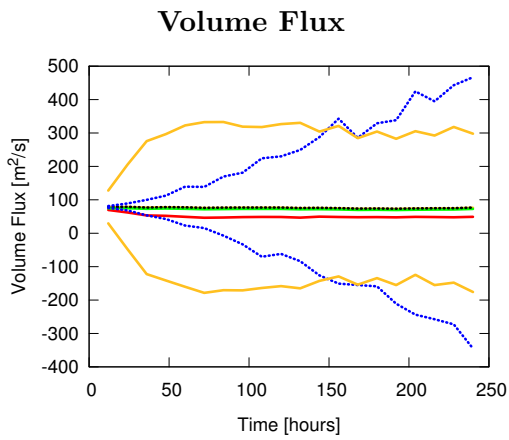


Figure 3.8: Comparison of the confidence intervals (dashed blue lines) with intervals that show when the effectivity index equals 10 (solid orange lines) for the volume flux in the Munk gyre experiment at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown, as well as $J_{high}(q_{high})$ (solid green line), and $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line).

region and overall we observe a similar behavior for the temporal evolution of the confidence intervals. A closer look to the first 25 hours of integration time however reveals differences. Where the confidence intervals also confine the 'true' goal $J_{high}(q_{high})$ for Region 1, they do not so for Region 2. Applying the reasoning we developed at the beginning of the section, we can identify that the cause is the discretization error in the goal that cannot be neglected for the total error in the goal. At later timesteps, the discretization error in the goal is again negligible compared to the uncertainty that is accumulated in the confidence intervals.

The question arises, whether this discretization error in the goal can be corrected for, given the information available. For the kinetic energy goal in Region 2, we find that the discretization error in the goal stays almost constant in time. This motivates us to use a bias correction on the confidence intervals to make them confine $J_{high}(q_{high})$. We make use of the information from the near-initial learning phase and correct all confidence intervals by adding $J_{high}(q^{20}_{high}) - J_\Delta(P_\Delta(q^{20}_{high}))$, where q^{20}_{high} denotes the high-resolution solution at timestep 20. The bias-corrected confidence intervals are shown in figure 3.9. The corrected confidence intervals now bound $J_{high}(q_{high})$ at all times which fulfills our first criterion. Our second criterion is also fulfilled as the effectivity index never exceeds a value of 7. We thus obtain a corrected error estimator that is valid for the whole integration time. As a result of the bias correction, $J_\Delta(P_\Delta(q_{high}))$ and $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ are not necessarily bounded by the confidence intervals any more, as can be seen here for the first 50 hours of integration

time.

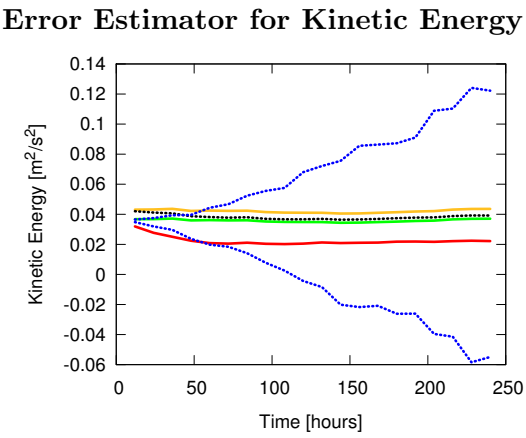


Figure 3.9: Bias-corrected confidence intervals (dashed blue lines) for the kinetic energy in Region 2 in the Munk gyre setup at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to the discrete goal applied to the high resolution solution $J_\Delta(P_\Delta(q_{high}))$ (solid orange line), $J_{high}(q_{high})$ (solid green line), and $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line).

Experiment 2: flow around an island

The results for the flow around an island setup are shown in figure 3.9. For the flow around an island setup, we are mostly interested in the regions near the island and especially focus on the area upstream of the island, where the maximum in meridional velocities is reached and the potential energy is at its peak due to the stagnation point in front of the island. For each goal we show one temporal evolution of confidence intervals for one specific region. The examples for the potential energy and the kinetic energy indicate that the flow is in a transition phase and that energy is constantly added to the system. For the potential energy, the chosen region is located around longitude $4W$ and latitude $11N$, which is close to the stagnation point. Thus, as the flow is further spun up, the potential energy in this area will be mostly increasing in time. The kinetic energy goal is from an area around the location longitude $9W$, latitude $4N$. The kinetic energy roughly doubles for the time window of the error estimation, and this increase is almost entirely due to the increase in meridional velocities. The increase in the respective energies occurs very differently at different resolutions (Figure 3.10). In contrast to these two goals, the volume flux is calculated for a cross-section downstream of the island. The chosen cross-section lies parallel to the equator at latitude $14N$ and stretches from longitude $2E$ to longitude $7E$. Due to this specific choice of cross-section, the volume flux indicates the formation of the vortex pair downstream of

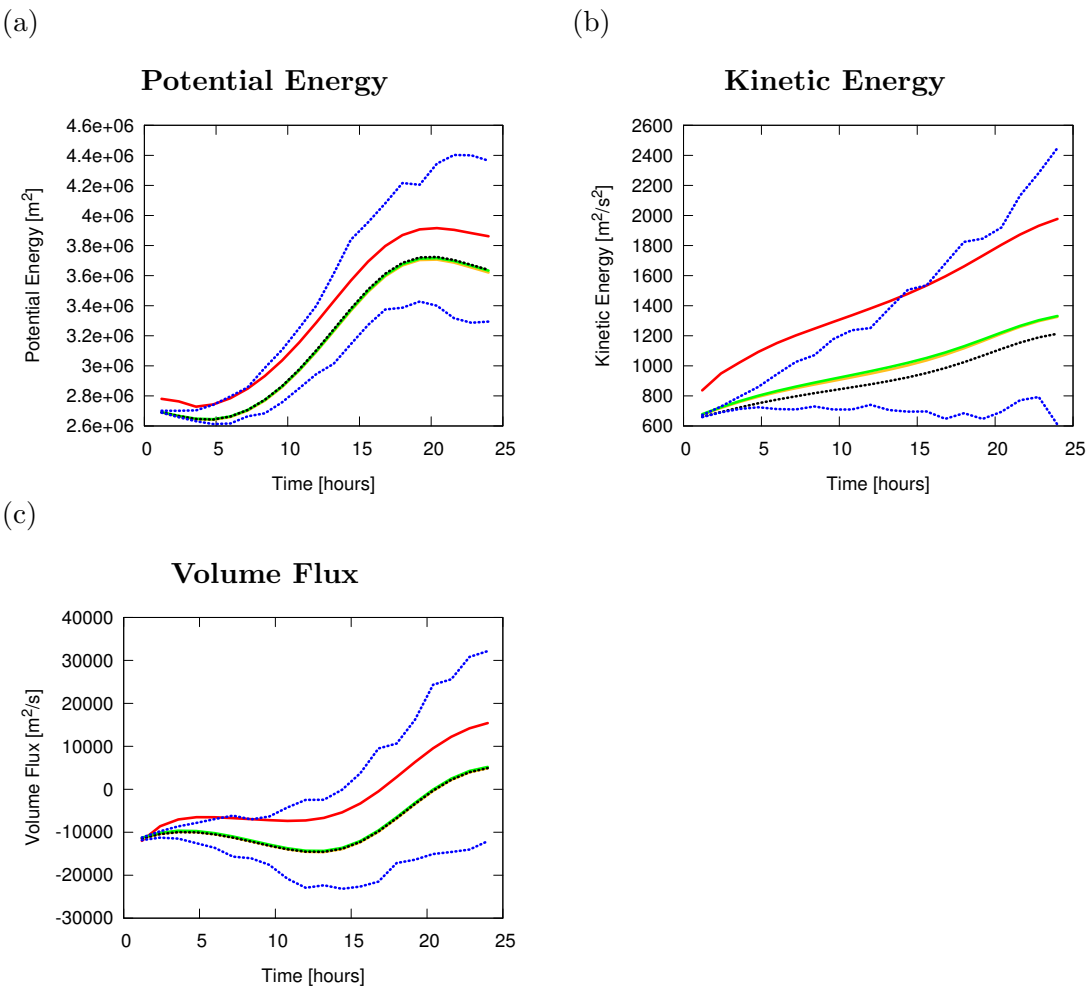


Figure 3.10: Temporal evolution of three different goals: (a) the potential energy, (b) the kinetic energy, and (c) the volume flux for the flow around an island setup at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to the discrete goal applied to a high resolution solution $J_\Delta(P_\Delta(q_{high}))$ (solid orange line) and $J_{high}(q_{high})$ (solid green line). If $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$ coincide, only $J_{high}(q_{high})$ is shown. The confidence intervals (dashed blue lines) are derived to confine $J_{high}(q_{high})$, $J_\Delta(P_\Delta(q_{high}))$, and $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line).

the island, as the volume flux changes from negative to positive values, which means a change in direction of meridional velocities from southward to northward.

As for the Munk gyre experiment, the potential energy, the volume flux, and the kinetic energy fulfill our main criterion for the error estimator. Concerning the second criterion, we find that the effectivity index is far lower than for the Munk gyre experiment. The maximum values for eff ranges from 1.6 to 2.5. The confidence intervals

for the flow around an island setup are thus tighter around the 'true' goal. At the start of the error estimation, we have an effective error correction for the error in the goal. This effect prevails for a time window of about 5 hours of integration time for the potential energy and the volume flux and up to 13 hours for the kinetic energy plot, which corresponds to approximately 800 timesteps.

Comparison to Stochastic-Physics Ensembles

We perform full stochastic-physics ensemble simulations in the manner we described in section 2.5 at resolution $\Delta 3$. The ensemble of perturbed goals calculated from the perturbed model solutions is used to derive error bounds in a similar fashion as for our adjoint-based error estimator. From our 100 ensemble members, we choose the lower error bound ϵ_1 to be the ensemble's minimum and the upper error bound ϵ_2 to be the ensemble's maximum. We compare these bounds to the bounds from the previous section.

The results are shown in figure 3.11 for the Munk gyre experiment and in figure 3.12 for the flow around an island setup. The main finding is that the bounds derived from the stochastic-physics ensemble also bound the error in the goals and are in this way qualitatively similar. For the flow around an island setup, the bounds for all regions and goals are of comparable size. For the Munk gyre experiment this holds true for the potential energy and the volume flux. For the kinetic energy, we again find that $J_{high}(q_{high})$ is not bounded in the beginning of the error estimation for Region 2 while it is bounded for Region 1. Additionally, the bounds derived from the stochastic-physics ensemble behave differently compared to our error estimator. The first difference is that the lower bounds do not cross the zero line, which is to be expected as it is the kinetic energy's lower bound. The second difference is the far stronger growth of the upper bound to $0.3\frac{m^2}{s^2}$ towards the end of the error estimation window compared to $0.14\frac{m^2}{s^2}$ in the case of our error estimator. The increased growth rate is for the most part confined in the second half of the error estimation window and thus coincides with stronger perturbations by the stochastic process. Our error estimator is an approximation to the stochastic-physics ensemble, and thus the question arises whether the adjoint solution provides a valid surrogate under these perturbations.

Confidence Intervals at Resolution $\Delta 4$

To have comparable results, we show the same combinations of regions and goals as presented for resolution $\Delta 3$. In general, the results are very similar, which is why we will only give a short description and focus on the cases where the results differ significantly.

Experiment 1: Munk gyre

The error estimates for the potential energy, the volume flux and the kinetic energy in

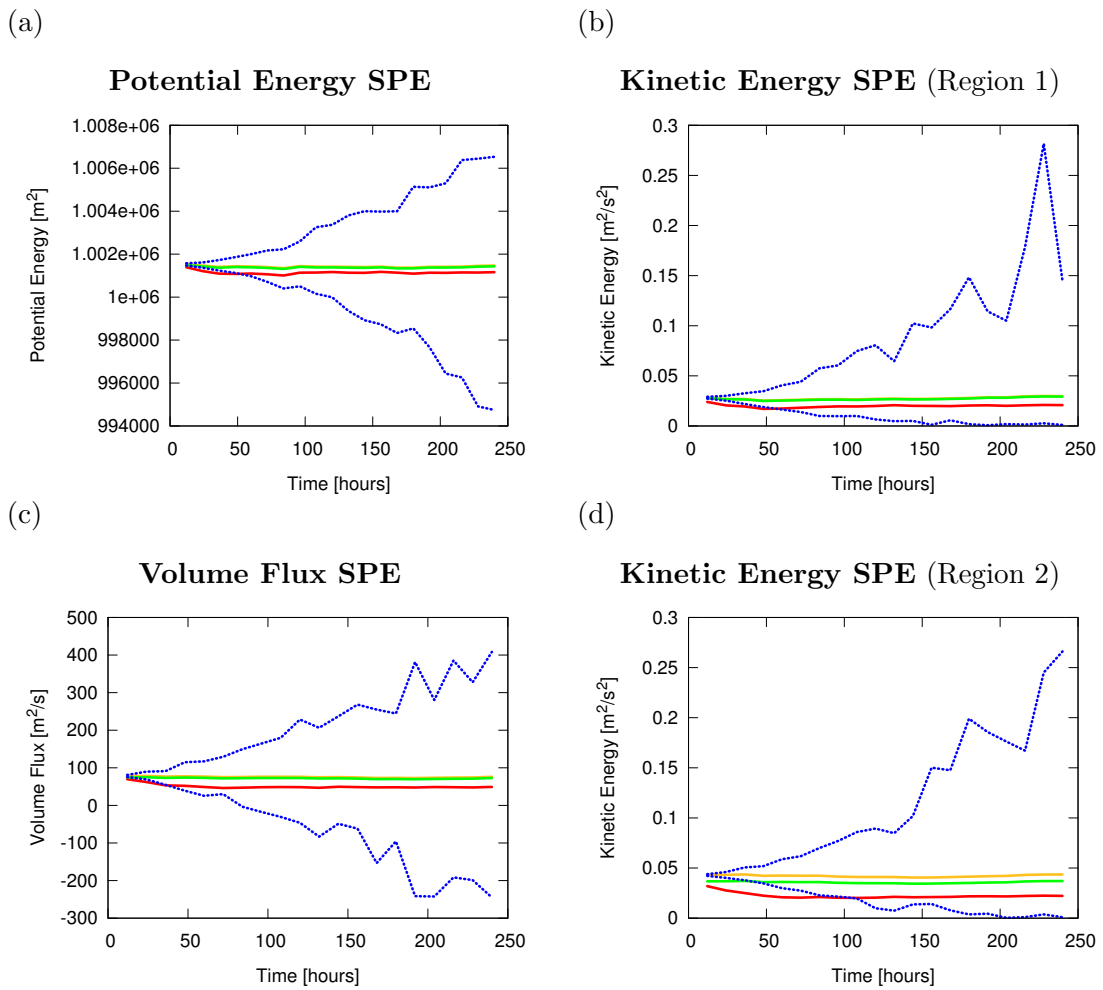


Figure 3.11: Temporal evolution of error bounds derived from stochastic-physics ensembles for four different goals: (a) the potential energy, (b) the kinetic energy (Region 1), (c) the volume flux, and (d) the kinetic energy (Region 2) for the Munk gyre experiment at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to the discrete goal applied to a high resolution solution $J_\Delta(P_\Delta(q_{high}))$ (solid orange line) and $J_{high}(q_{high})$ (solid green line). If $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$ coincide, only $J_{high}(q_{high})$ is shown. The confidence intervals (dashed blue lines) are derived to confine $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$.

Region 1 (not shown) stay qualitatively the same, and we find an effective correction of the goal for the first 50 to 75 hours of integration time. The confidence intervals for the kinetic energy in Region 2, shown in figure 3.13 (a), again do not bound $J_{high}(q_{high})$ for the first 75 hours of integration time. However, we now find that $J_\Delta(P_\Delta(q_{high}))$ is not bounded by the confidence interval either, which indicates that the problem lies within the scalar product (3.20). More precisely, the problem lies with the adjoint

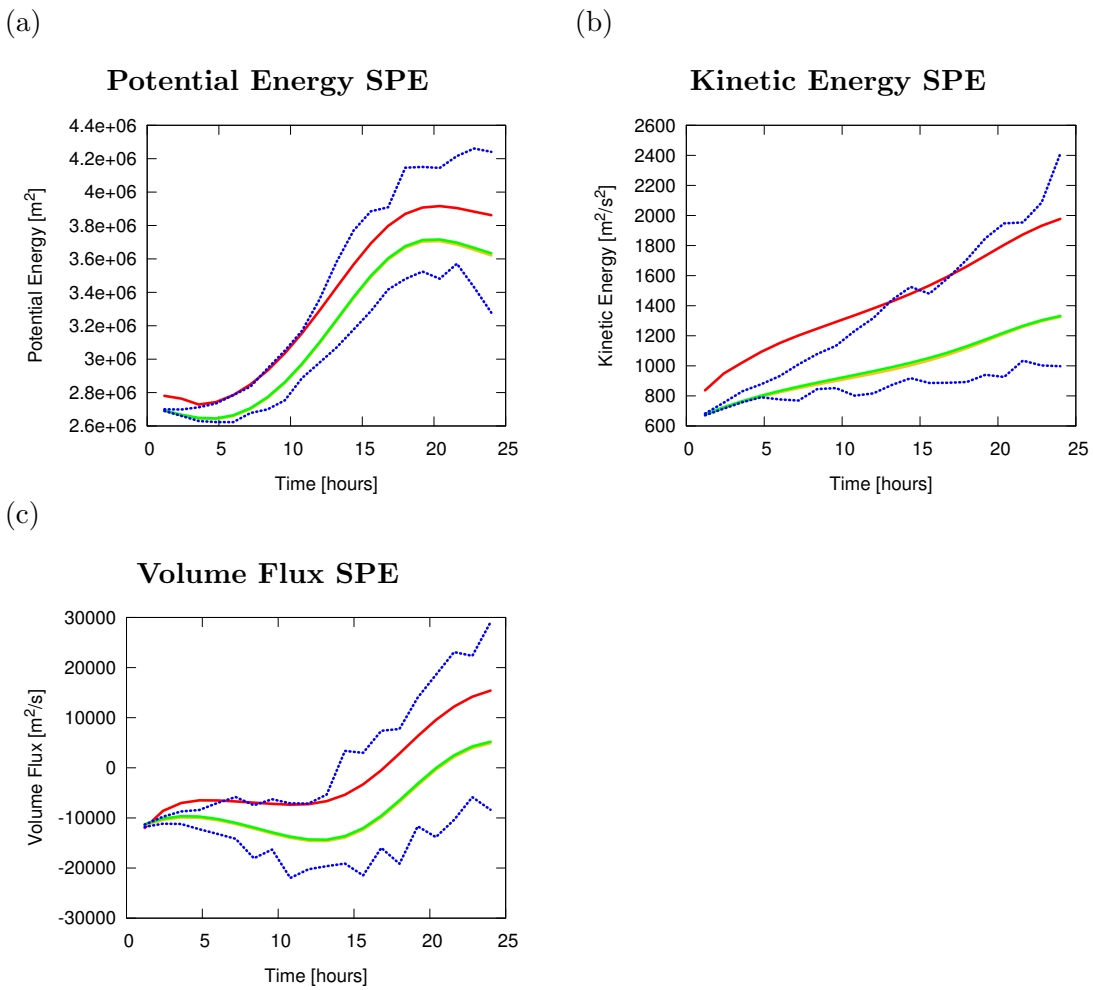


Figure 3.12: Temporal evolution of error bounds derived from stochastic-physics ensembles for three different goals: (a) the potential energy, (b) the kinetic energy, and (c) the volume flux for the flow around an island setup at resolution $\Delta 3$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to the discrete goal applied to a high resolution solution $J_\Delta(P_\Delta(q_{high}))$ (solid orange line) and $J_{high}(q_{high})$ (solid green line). If $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$ coincide, only $J_{high}(q_{high})$ is shown. The error bounds (dashed blue lines) are derived to confine $J_\Delta(P_\Delta(q_{high}))$ and $J_{high}(q_{high})$.

solution q^*_Δ, because $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ does not coincide with $J_\Delta(P_\Delta(q_{high}))$. To further strengthen this argument, we show the error bounds derived from the stochastic-physics ensemble (Figure 3.13 b). The error bounds derived in this way fully bound $J_\Delta(P_\Delta(q_{high}))$ as well as $J_{high}(q_{high})$ at all times and even yield an effective goal correction for 125 hours of integration time.

For all goals, the widths of the confidence intervals are consistently reduced by 20 to

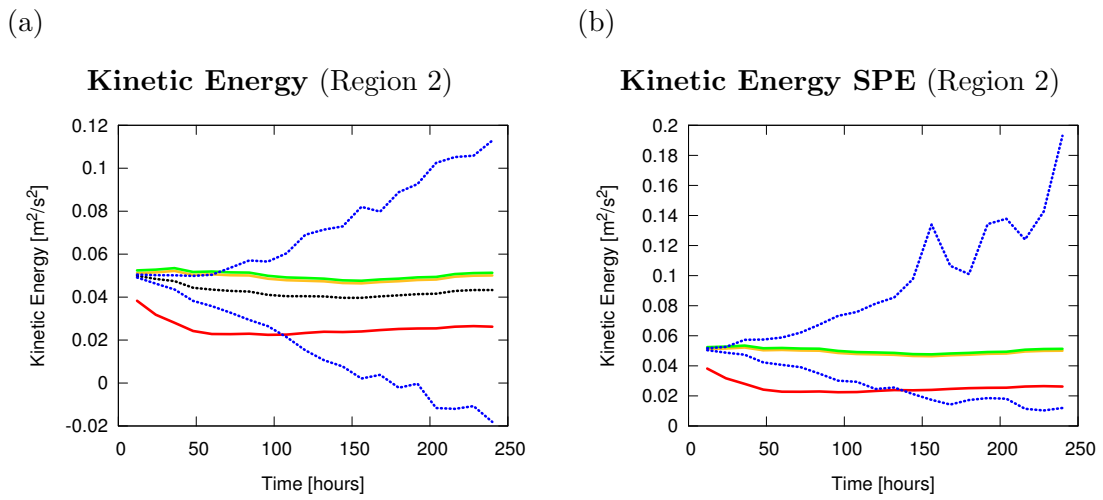


Figure 3.13: Comparison of temporal evolutions of (a) confidence intervals for the kinetic energy (Region 2) with (b) error bounds derived from a stochastic-physics ensemble for the Munk gyre experiment at resolution $\Delta 4$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to $J_{high}(q_{high})$ (solid green line) and to $J_\Delta(P_\Delta(q_{high}))$ (solid orange line). The derived confidence intervals and error bars (dashed blue lines) are given, and if applicable $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line) is shown.

40 percent compared to resolution $\Delta 3$. To further compare the results, we calculate a score for the change in confidence interval width. The score is calculated as the ratio between the effectivity index at resolution $\Delta 4$ and the effectivity index at resolution $\Delta 3$. This score indicates how fast the error in a goal decreases in relation to the decrease in confidence interval width when increasing the resolution. For the Munk gyre experiment, this score is mostly found between 1 and 4, meaning that the width of the confidence interval decreases faster than the error in the goal when increasing the resolution from $\Delta 3$ to $\Delta 4$. The confidence intervals are thus getting tighter for the goals shown.

Experiment 2: flow around an island

For the flow around an island setup, the 'true' goals $J_{high}(q_{high})$ are bounded for all plots (not shown), and we find an effective correction of the goal for the first 5 to 7.5 hours of integration time. For the kinetic energy this time window is thus roughly halved compared to resolution $\Delta 3$.

Concerning the widths of the confidence intervals, we find that they are consistently reduced by 40 to 70 percent. We observe the strongest reductions of up to 70 percent for the kinetic energy. Comparing again the scores for the changes in the confidence interval width, we find the values to vary between 0.5 and 1.4. For the potential energy the values vary between 0.7 and 1.4 and are concentrated around 1. This means the

ratio between the error in the goal and the confidence interval width is comparable on both resolutions. For the volume flux, the values concentrate around 0.8 and for the kinetic energy values lie between 0.5 and 0.7. In total, this means that although the width of the confidence intervals decreases quite substantially, the error in the goal decreases even faster.

For the Kármán vortex street setup, we show a kinetic energy goal in figure 3.14. The

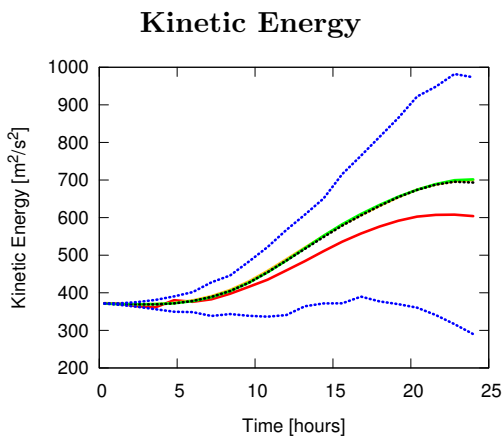


Figure 3.14: Temporal evolution of the kinetic energy for the Kármán vortex street setup at resolution $\Delta 4$. The goal calculated from the model solution $J_\Delta(q_\Delta)$ (solid red line) is shown in comparison to $J_{high}(q_{high})$ (solid green line). The confidence intervals (dashed blue lines) confine $J_{high}(q_{high})$ as well as $J_\Delta(q_\Delta)$ corrected by the scalar product $\langle q^*_\Delta, N_\Delta(q_{high})\rangle$ (dashed black line).

area which the kinetic energy is averaged over is located at a region around longitude $22E$, latitude $2S$, directly downstream of the island on the southern hemisphere. In this region a vortex is shed from the island at the starting point of the error estimation, which leads to a strong increase in the kinetic energy goal over the period of the error estimation. We find that $J_{high}(q_{high})$ is bounded by the estimated error bars at all times, our error estimator thus fulfills our main criterion. The second criterion is also fulfilled as we find that the effectivity index eff is constantly around a value of 3.

3.5.5 Discussion

In most cases, our error estimator is able to bound the 'true' goal, and we can often even correct for the error in the goal. This property is found at the beginning of the error estimation time windows and can be persistent for several hundreds of timesteps. This effect can be entirely explained by the mean-field part of the stochastic process, which is constant in time by construction. Whenever the influence of the constant mean

field on the goal dominates over the influence of the purely stochastic part on the goal, we can perform a correction of the goal's value. After the initial phase where the goal can be corrected, the ratio of the confidence interval width 3σ to the error in the goal ϵ does not stay below our chosen threshold of 10 for the volume flux and the potential energy goals in the Munk gyre experiment. This occurs only for the Munk gyre experiment because it is not in transition anymore, which means that the fluctuations in local truncation errors are only weakly correlated in time. As a result, the upper bound for the correlation coefficient given by the Cauchy-Schwarz inequality in equation (3.15) is an overestimate, which results in wide confidence intervals on the error in the goal compared to the actual error. In the context of ocean-type experiments, however, the error estimator would not be applied to an experiment that is in a steady state but to transient flows where transition and the change of viscosity parameters with resolution occur both at the same time. Under these conditions, the temporal evolution of local truncation errors will be that of a transient flow but with increased values. Based on our results, this is a situation that the error estimator should be able to cope with.

We have found strong indications that our error estimator converges, as the confidence interval widths were consistently reduced when increasing resolution. This is an important property that any error estimator should satisfy. The different rates at which this convergence takes place for the two experiments is most likely due to the resolution-dependent eddy parametrization. It prevents a fast convergence of the model solution for the Munk gyre experiment and thus also of the error estimator.

Certain properties of the random variables that we estimated from the temporal fluctuations in local truncation errors are found to be consistent for all model variables and dynamical flow regimes. All random variables can be assumed to be centered around zero. For the heights, this follows directly from the volume conservation of the model and the chosen separation of the different dynamical regimes. For the velocities, which reside on the midpoints of triangle edges and point in normal direction, this holds due to a property of the ICON-grid. For every edge with given velocity, there is an edge in its vicinity with similar velocity but opposite orientation and thus opposite sign. The resulting local truncation errors will then also be of the same order of magnitude but opposite sign. Another property we find in all random variables is that their probability distribution can be assumed to be Laplacian, which seems to be more appropriate for flows with boundaries than a Gaussian distribution.

We have shown that the confidence intervals from our error estimator and the error bounds derived from the stochastic-physics ensemble are similar to the point of being indistinguishable. An exception is the kinetic energy in the boundary region of the Munk gyre experiment for which the upper bound derived from the stochastic-physics ensemble has a higher growth rate. We can exclude the kinetic energy itself as the reason, as this does not occur for the kinetic energy in the flow around an island setup. The only explanation left is that here the perturbations due to the stochastic process

cannot be considered to be 'small' anymore. The adjoint solution is thus not fully capable of accounting for the accumulated uncertainty in the goal.

The limitation we encounter for our error estimator is due to the part of the total error in the goal that comes from the discretization of the goal itself. In the dual-weighted error estimation framework underlying our approach this error is not represented, and we thus assumed it to be negligible which is in accordance with most of our numerical results. If the error is, however, not negligible in practice, our error estimator cannot bound the 'true' goal directly. This is for instance the case when flow features that have a strong influence on the goal are underresolved compared to the 'true' solution, which is also in accordance to our finding that increasing the model resolution resolves the issue; our method thus yields asymptotic bounds. We however also find that if the discretization error in the goal is approximately constant in time, applying a bias-correction to our confidence intervals yields a valid error estimator that again bounds the 'true' goal. In summary, this effect is a known limitation of the chosen dual-weighted error estimation framework, and it requires expert knowledge about the flow to decide whether or not this limitation poses a problem for the error estimator.

3.6 Conclusions

From our numerical results we conclude that our error estimator is capable of estimating bounds on the error in a goal for shallow-water flows with lateral boundaries and an eddy viscosity parametrization. The method is shown to be a valid approximation to a full stochastic-physics ensemble by theoretical arguments as well as numerical results. Our results thus further strengthen the statements made in [41] that deterministic local truncation errors can be interpreted stochastically, and the already suspected connection to the class of ensemble methods is clarified. We find that our method requires expert knowledge in choosing a suitable separation of regions of different flow dynamics, and it is required to assess whether the error due to the discretization of the goal is negligible, and if not whether it is correctable.

The stochastic framework of our extended method was motivated by our finding that the spatial distribution of temporal fluctuations in local truncation errors can be described by a Laplace-distributed random variable. The stochastic process is then defined as the sum of these random variables that we assume to be fully correlated in time. In the context of the Mori-Zwanzig formalism, the stochastic process is a noise term. The correlation of noise components in time, and the non-Markovian and non-Gaussian character of our stochastic process are in accordance with the properties expected for a noise term by Chorin et al. [8].

In our error estimator it is only the stochastic process that is capable of providing such a memory effect introduced by the Mori-Zwanzig formalism. Furthermore, the only features of the stochastic process where a memory term could possibly be residing

in, is its constant mean field and the growth rate of its variance. This implies that for our approach, the memory term consists of a contribution from the mean field that is constant in time and a remaining part that is purely stochastic and has thus been completely converted into uncertainty.

Our algorithm employs a separation of the flow into regions of different dynamical regimes in order to separate regions of different local truncation error production rates. For this separation of dynamical regimes, we expect a typical trade-off situation between the robustness of the method and the performance in terms of ensemble spread. While a greater number of dynamical regimes should result in tighter confidence intervals, the estimated parameters of the stochastic process might lose robustness, resulting in a reduced time window for error estimation. Conversely, reducing the number of dynamical regimes might be more robust but might potentially yield a vast overestimate of the uncertainty in the respective goal.

Similarly to [41], our algorithm uses near-initial, high-resolution information, which is extrapolated to a much larger time-window. Here, we have extended this extrapolation approach from the local truncation errors themselves to their temporal fluctuations. This means that spatial knowledge about the temporal fluctuations in local truncation errors can be converted into temporal knowledge about local truncation errors at future timesteps, and we have shown that this knowledge transfer can be carried out by means of a random variable.

The strong underlying theoretical framework we have developed for our method, our understanding of its limitations, and the robustness of the obtained numerical results for the 2D shallow-water cases make us confident that we can go one step further and apply our error estimator to idealized experiments in a full 3D ocean model.

Chapter 4

The Stochastics of the Local Truncation Error in an Eddying Regime

Towards a stochastic goal-oriented error estimator for three-dimensional ocean models, we investigate the local truncation error and its stochastic representation for an eddying flow driven by baroclinic instabilities. In this regime, we argue that it is extremely difficult if not impossible to obtain the local truncation error for all state variables. We instead define a temperature-based representation of the local truncation error that is shown to be valid for density-driven flows.

For a stochastic representation of the local truncation error, the structure of the temperature-based local truncation error is divided into three components: its horizontal structure, its vertical structure and its temporal evolution. The horizontal and the vertical structure are represented by correlation matrices that we successfully merge into a stochastic representation of the three-dimensional spatial structure of the local truncation error. For the stochastic representation of the temporal evolution, the inclusion of the history of the stochastic process in the form of time-lagged correlation coefficients is argued to be crucial. Based on this analysis, a revised stochastic timestepping scheme of the stochastic process is proposed.

4.1 Introduction

The quantification of numerical errors in ocean models is a long-standing and still unresolved problem. In ocean modeling, key physical quantities of interest (goals) that characterize the physical system are of interest. These goals are linear or non-linear functionals of the model solution. A conceptual framework for the error estimation in goals is provided by dual-weighted error estimation techniques [40, 3, 15, 17, 11, 19, 51]. In the dual-weighted error estimation technique, the error in a goal is translated into a weighted sum of local truncation errors, the error in the state variables done by the model per timestep. For an oceanographic context, we have developed a stochastic extension of the dual-weighted error estimation approach for viscous, bounded shallow-water flows in chapter 3. Its key component is a stochastic process that parametrizes

the local truncation error. The purpose of this chapter is to carry the idea of the stochastic dual-weighted error estimator towards a full 3D-Ocean model, as used in a comprehensive Earth System Model, by investigating the local truncation error of an eddying three-dimensional flow and its corresponding stochastic process. The stochastic process is estimated using the learning algorithm 3.4 developed in the previous chapter.

Due to the complexity of state-of-the-art ocean models and their choice of discretization schemes, a continuous evaluation of the local truncation error or equivalent residual information is not possible. An alternative approach that replaces the local truncation error by a suitable stochastic representation was first proposed and tested for a shallow-water model and wave-type flows in [41] and later extended by us to ocean-type experiments in a shallow-water model in chapter 3. The replacement of deterministic local truncation errors by a stochastic component was motivated by the Mori-Zwanzig formalism [33, 32, 54, 18] from statistical mechanics.

Here, we aim to extend the stochastic component of our stochastic error estimator to a full 3D ocean model, namely the ICON-Ocean model [27]. The additional dimension and the overall increase in model complexity introduces several challenges for our stochastic component. The stochastic component now has to incorporate the effects of a downward energy cascade that is present in a three-dimensional flow. The model now includes an active tracer in the form of a temperature variable that influences the model dynamics and introduces the important concept of stratification to the model, and the model physics is extended by additional sub-grid scale parametrizations. The interplay of all these additional processes can result in complex spatial and temporal flow structures which will be reflected in the structure of its local truncation error.

The experiment we study in this model framework is a channel flow driven by baroclinic instabilities [38]. The occurring baroclinic instabilities are a source for the development of mesoscale eddies, ocean eddies that are nearly in geostrophic balance. As such, the experiment is a draft for dynamics that frequently occur in many regions of the ocean such as in the Antarctic Circumpolar Current or near western boundary currents like the gulf stream [30]. The mesoscale eddy field induces a heat transport in the lateral direction [24] as well as a net upward heat transport [22, 52].

For the resulting local truncation error and its stochastic representation, several challenges need to be overcome. First, we encounter that the concept of the local truncation error as we have used it in previous chapters is not directly applicable (see definition (3.8) from chapter 3) due to non-physical heating of the channel when correcting the surface elevation variable. We investigate alternative local truncation error representations that are applicable to an eddying flow regime in a 3D ocean model. Second, the occurrence of mesoscale eddies at high resolution entails a spatial structure of the local truncation error at low resolution that needs to be respected by its stochastic representation. We show how the spatial structure can be described stochastically

as a compound of sub-structures that each can be described by their own correlation matrix. And third, the persistence of mesoscale eddies for a certain lifetime suggests that a stochastic representation of the temporal structure of the local truncation error is strongly dependent on its past states. An approach to model this memory of the stochastic process is proposed that is based on a correlation matrix of time-lagged correlation coefficients.

The chapter is structured as follows: In section 2, we propose how the basic dual-weighted error estimation approach could be carried into a stochastic framework for a 3D model environment. Section 3, introduces the hydrostatic primitive equations and the ICON-Ocean model. In section 4 we introduce our channel experiment and analyze the impact of the mesoscale eddy field on the flow. In section 5, we analyze the local truncation error and its resulting stochastic representation for the channel experiment and discuss the results, and in section 6 we conclude.

4.2 Stochastic Dual-Weighted Error Estimation

For our error estimation framework, we define a general form of a continuous model and its discrete counterpart. The system

$$N\left(q\left(x,t\right)\right)=0, \tag{4.1}$$

is a continuous model, with initial conditions $q(x,0)=q^0$ and boundary conditions $q(x,t)=q^b(x,t)$ on $\partial\Omega$. The function $q\left(x,t\right)$ is defined on $\Omega\times[0,T]$ and satisfies system (4.1).
The discretization of the continuous system (4.1) can then be written as

$$N_\Delta\left(q_\Delta\right)=0, \tag{4.2}$$

with initial conditions $q^0_\Delta=P_\Delta(q^0)$ and boundary conditions $q_\Delta=P_\Delta(q^b)$ on $\partial\Omega_\Delta$. P_Δ represents a projection of the continuous solution q from the continuous function space $V(\Omega\times[0,T])$ on $\Omega\times[0,T]$ to the discretized function space $V_\Delta(\Omega_\Delta\times T_\Delta)$ on $\Omega_\Delta\times T_\Delta$.. The solution of system (4.2) is denoted by $q_\Delta=\begin{pmatrix} q^0_\Delta \\ \vdots \\ q^n_\Delta \end{pmatrix}$ and is a vector of state vectors $q^i_\Delta\in\mathbb{R}^m,\ i=0,\ldots,n$ in discrete time $T_\Delta:=\{t_0,\ldots,t_n\}$ and discrete space.

4.2.1 Errors in a Physical Quantity of Interest

A physical quantity of interest (goal) is represented by a functional J that is applied to the continuous model solution q. We denote its discrete version, which is applied to the

discrete model solution q_Δ, by J_Δ. The calculated goal $J_\Delta(q_\Delta)$ is inherently erroneous because of the error due to discretization in q_Δ. The total error ϵ in a goal J is defined as the difference between the true value of the goal $J(q)$ and its approximation $J_\Delta(q_\Delta)$

$$\epsilon := J_\Delta(q_\Delta) - J(q). \tag{4.3}$$

This error in a goal can be estimated via the dual-weighted error estimation approach given in [15]

$$\epsilon \approx \epsilon_{app} := \langle q^*_\Delta, N_\Delta(q)\rangle_{\Omega_\Delta \times T_\Delta}, \tag{4.4}$$

where $\langle\cdot\rangle_{\Omega_\Delta \times T_\Delta}$ denotes the standard euclidean scalar product in discrete space and time, q^*_Δ is the discrete adjoint solution to q_Δ, and $N_\Delta(q)$ are the local truncation errors.

In our stochastic extension of (4.4) described in chapter 3, we replace the local truncation error $N_\Delta(q)$ by a suitable discrete-time stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ of random vectors $X_{t_i}(\mathbf{p}_i)$ with parameters $\mathbf{p}_i$. This stochastic process is our model for the temporal evolution of the local truncation error, and its insertion into (4.4) yields

$$\epsilon_{app} := \langle q^*_\Delta, \begin{pmatrix} X_{t_1}(\mathbf{p}_1) \\ \vdots \\ X_{t_n}(\mathbf{p}_n) \end{pmatrix} \rangle_{\Omega_\Delta \times T_\Delta}, \tag{4.5}$$

where each random vector $X_{t_i}(\mathbf{p}_i) \in \mathbb{R}^m$ is multiplied by the transpose of vector ${q^*_\Delta}^i \in \mathbb{R}^m$at timestep i. Error estimates ϵ_{app} from (4.5) are then a measure for the uncertainty in the goal J which can be used to derive bounds $\epsilon_1, \epsilon_2 \in \mathbb{R}$ that satisfy

$$J_\Delta(q_\Delta) - \epsilon_1 < J(P_\Delta(q)) < J_\Delta(q_\Delta) + \epsilon_2. \tag{4.6}$$

Alternatively to the calculation of the scalar product (4.5), one can obtain equivalent information about the uncertainty in a goal from an ensemble of perturbed model runs of the discrete model (4.2). The perturbations applied to the model state q^i_Δ at timestep i are given by the random vector $X_{t_i}(\mathbf{p}_i)$ of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$. The relationship between the ensemble approach and the dual-weighted approach (4.5) is discussed extensively in chapter 3.

4.2.2 Parametrizing the Evolution of Local Truncation Errors by a Stochastic Process

Regardless of whether one chooses the ensemble approach or the dual-weighted approach (4.5) for goal-oriented error estimation, a suitable discrete-time stochastic pro-

cess $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ that models the temporal evolution of the local truncation error is required. The temporal character of the euclidean scalar product in equation (4.4) is emphasized when it is written as a weighted sum of local truncation errors at the different timesteps i

$$\epsilon_\Delta = \sum_{i=1}^{n} \left\langle q_\Delta^{*\,i}, N_\Delta^i(q) \right\rangle_{\Omega_\Delta}, \tag{4.7}$$

where $\langle\cdot\rangle_{\Omega_\Delta}$ is the standard euclidean scalar product in discrete space Ω_Δ, $q_\Delta^{*\,i}$ is the vector of the discrete adjoint solution at timestep i, and $N_\Delta^i(q)$ is the vector of local truncation errors at timestep i.

In contrast to the previous chapter 3, we do not aim to build the stochastic process around the separation of the local truncation error into a time-invariant bias plus the temporal fluctuations in the local truncation, but instead we build the stochastic process around the local truncation error itself. The reason is that in an eddying regime, we cannot expect to find significant time-invariant biases in the local truncation error. As a result, the complete local truncation error information will reside in the temporal fluctuations, which would then be large, which in turn would make the spread of the stochastic process large. In total, the concept that was found to be appropriate for flows dominated by strong boundary currents is not deemed to be well applicable here.

Additionally, we expect a generally more complex flow structure than the one encountered in chapter 3, which the stochastic component needs to adhere to. We now have to incorporate the vertical structure of the three-dimensional ocean experiment which will impose certain restrictions on the vertical structure of the stochastic component. For the horizontal structure in chapter 3, important horizontal correlations were dealt with by the bias of the stochastic process. Here, we also expect to encounter important horizontal structures in the local truncation error of the low resolution solution due to unresolved eddies, and we have to find an alternative way to prescribe these patterns stochastically that is not based on biases. Given that the spatial structures can be described stochastically, we additionally require a component that governs the temporal evolution of these spatial structures, a stochastic timestepping scheme which represents the temporal structure of the local truncation error. A simple approach to assume an upper bound on the temporal correlations as used in chapter 3 is not deemed applicable to a complex, eddying flow.

Our approach to model these structures of the local truncation error in a stochastic process is by means of correlation matrices. Correlation matrices provide a stochastic measure for the strength of the connection between different local truncation error values. Optimally, we would of course like to have one correlation matrix that describes the whole structure of the local truncation error. However, the amount of variables of the local truncation error is far too large to estimate a correlation matrix of this size. In our approach we thus choose to separate the full structure of the local truncation

error into three sub-structures – vertical, horizontal and time – which we each describe by correlation matrices. In this we have to make assumptions about the negligibility of certain correlations between variables of different sub-structures.

Given we have obtained a correlation matrix for each sub-structure, our aim is the calculation of stochastic fields that satisfy the complete structure of the local truncation error. Towards this, we perform a Cholesky decomposition of each obtained correlation matrices. The resulting matrices can be applied to uncorrelated fields to obtain fields that satisfy the structures of the corresponding correlation matrices. The difficulty lies now in merging the correlated fields that all contain the patterns of their respective sub-structure into one large spatio-temporal structure that can represent the local truncation error. How to link the correlation matrices of the sub-structures in order to directly obtain such a structure is the topic of this section.

We start with the definition of the correlation matrix of a given Gaussian random vector $Y \in \mathbb{R}^m$ with covariance matrix Σ

$$\Sigma := cov\left(Y\right) = E\left[\left(Y - E\left[Y\right]\right)\left(Y - E\left[Y\right]\right)^T\right], \tag{4.8}$$

dimension $\mathbb{R}^{m\times m}$, and the variances of the entries of Y lie on the diagonal of Σ. The correlation matrix of Y is then defined as the covariance of the standardized entries of Y

$$corr\left(Y\right) = \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}}\Sigma\left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}}, \tag{4.9}$$

where the matrix $diag\left(\Sigma\right)$ consists of the diagonal elements of Σ.

To be able to produce a correlated random vector $\tilde{Y}$ with desired distribution $\mathcal{N}\left(\mu, \Sigma\right)$ from an uncorrelated one $Y_{uncorrelated}$, a Cholesky decomposition of the correlation matrix can be performed. The correlation matrix associated with the covariance matrix Σ is denoted by $\mathbf{C}$. This correlation matrix can then be decomposed by Cholesky decomposition into a product of a lower and an upper triangular matrix

$$\mathbf{L}\mathbf{L}^T = \mathbf{C}. \tag{4.10}$$

This decomposition is possible because correlation matrices are always positive-semidefinite and symmetric. Given the lower triangular matrix $\mathbf{L}$, a vector $Y_{uncorrelated} \sim \mathcal{N}\left(0, \mathbb{I}\right)$ of uncorrelated normally-distributed random variables can be correlated by applying the matrix $\mathbf{L}$

$$Y_{correlated} := \mathbf{L}Y_{uncorrelated}, \tag{4.11}$$

such that the correlation matrix of $Y_{correlated}$ is

$$corr\left(Y_{correlated}\right) = \mathbf{C}. \tag{4.12}$$

Thus, the random vector

$$\tilde{Y} := \mu + (diag\,(\Sigma))^{\frac{1}{2}}\, Y_{correlated} \sim \mathcal{N}\,(\mu, \Sigma) \tag{4.13}$$

has the desired distribution and correlation structure.

The Spatial Structure of the Local Truncation Error

We separate the modeling of the spatial structure of the local truncation error into two parts, the local truncation error's horizontal structure and its vertical structure. Both parts are represented by a correlation matrix. The employed 3D grid is constructed as several layers of a chosen horizontal grid.

In our approach, the horizontal structure is governed by a correlation matrix that represents the correlations between the local truncation error in all model variables at one specific vertical layer. Which vertical layer is generally best suited is however not known *a priori*. Here, we simply choose the surface layer as vertical layer. We note that the approach described here is however not limited to the surface layer and other horizontal layers could be chosen as well, which can make sense considering that the ocean mixed layer in the upper ocean often behaves very differently from the deeper ocean. Regardless of the choice, the correlation matrix is a quadratic matrix of a dimension $\mathbb{R}^{m_{surf} \times m_{surf}}$, where m_{surf} is the number of degrees of freedom in the surface layer.

The vertical structure of the local truncation error is important as the flow will for the most part be in a stratified state, and in presence of a mesoscale eddy field, single eddies will not just occur in the surface layer but will also have a vertical extend. The local truncation error will also inherit these properties and vertical correlations are thus an integral part of its stochastic representation.

For each model variable θ, the vertical structure corresponding to each surface grid point i is represented by a correlation matrix $\mathbf{C}_{v,\,\theta,\,i}$ that contains the correlations within the entire vertical column of said model variable. The dimension of the correlation matrices is equivalent to the number of vertical levels and the entries of $\mathbf{C}_{v,\,\theta,\,i}$ are ordered, starting from the surface layer 1 and to the bottom level $\#layers$.

Both structures are connected by first creating a horizontally-correlated surface field and use it to seed the calculation of the vertical structure. More precisely, we start by creating a realization

$$y_{surf,\,uncorrelated} \in \mathbb{R}^{m_{surf}}$$

of an uncorrelated normally-distributed random vector $Y_{surf,\,uncorrelated} \sim \mathcal{N}\,(0, \mathbb{I})$. With the lower triangular matrix $\mathbf{L}_{surf}$ that we obtained from the decomposition of

$\mathbf{C}_{surf}$ into $\mathbf{L}_{surf}\mathbf{L}_{surf}^T$, we can use equation (4.11) to obtain $y_{surf,\, correlated}$

$$y_{surf,\, correlated} := \mathbf{L}_{surf} \;\; y_{surf,\, uncorrelated}, \tag{4.14}$$

a realization of the random vector $Y_{surf,\, correlated} \sim \mathcal{N}\left(0, \mathbf{C}_{surf}\right)$ that follows the correlation pattern of the surface layer.

The correlated surface field $y_{surf,\, correlated}$ is now used in the calculation of the vertical structure. First, for each model variable θ and each surface grid point i, a realization

$$y_{v,\, \theta,\, i,\, uncorrelated} \in \mathbb{R}^{\#layers}$$

of an uncorrelated random vector $Y_{v,\, \theta,\, i,\, uncorrelated} \sim \mathcal{N}\left(0, \mathbb{I}\right)$ is drawn. Then, the first entry of $y_{v,\, \theta,\, i,\, uncorrelated}$ is replaced by its corresponding value from the surface field vector $y_{surf,\, correlated}$. If we now apply the lower triangular matrix $\mathbf{L}_{v,\, \theta,\, i}$ obtained from the decomposition of $\mathbf{C}_{v,\, \theta,\, i}$, we obtain $y_{v,\, \theta,\, i,\, correlated}$

$$y_{v,\, \theta,\, i,\, correlated} := \mathbf{L}_{v,\, \theta,\, i} \;\; y_{v,\, \theta,\, i,\, uncorrelated}. \tag{4.15}$$

Due to the special construction of the matrix $\mathbf{L}_{v,\, \theta,\, i}$, the first entries of the vectors $y_{v,\, \theta,\, i,\, correlated}$ are identical to the corresponding entries in the surface layer vector $y_{surf,\, correlated}$. This is because $\mathbf{L}_{v,\, \theta,\, i}$ is a lower triangular matrix and the first entry in its first row is 1. As a consequence, the matrix $\mathbf{L}_{v,\, \theta,\, i}$ then carries the information from the horizontal surface structure associated with $y_{surf,\, correlated}$ into the vertical columns of the 3D grid. The described procedure thus results in a stochastic representation of the local truncation error's 3D structure at a specific timestep.

Of course, this choice of splitting up the spatial structure into a horizontal and a vertical part comes at a cost. Because both correlation matrices are connected by one horizontal layer only, we do not account for all correlations of the 3D flow structure. Instead, we have already implicitly made a choice that certain correlations are not taken into consideration. Direct correlations between different grid points that reside on different vertical levels are neglected. Also, apart from the correlations that are already determined by the surface layer and propagated to deeper layers, possible further correlations between different model variables θ in the vertical column have to be neglected. If these correlations are however found to be important for an experiment, the steps described above need to be adjusted. For instance, one of possible variation would be to replace the surface layer by another horizontal layer for linking the fields.

If the assumptions about the correlations hold, the approach described above yields the spatial structure of the random vectors $X_{t_i}(\mathbf{p}_i)$ of the stochastic process for the local truncation error $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$. Given the vectors $y_{v,\, \theta,\, i,\, correlated}$, all information is available to reassemble them into a base realization y_{t_i} from which a realization of the random vector $X_{t_i}(\mathbf{p}_i)$ can be obtained. De-normalization by the standard deviation

and the addition of the mean vector as given in (4.13) yields the desired realization of the random vector $X_{t_i}(\mathbf{p}_i)$. However, although the spatial structure can be obtained, the framework so far does not provide information on the temporal evolution of the random vectors $X_{t_i}(\mathbf{p}_i)$.

The Temporal Structure of the Local Truncation Error

We make certain assumptions for the stochastic timestepping of a realization of the random vector $X_{t_i}(\mathbf{p}_i)$ at timestep i to a realization of $X_{t_{i+1}}(\mathbf{p}_{i+1})$ at timestep $i+1$. We first assume that we already obtained a realization of a random vector $X_{t_i}(\mathbf{p}_i)$ that is from here on always denoted by $x_{t_i}(\mathbf{p}_i)$. We here assume that the correlation matrices $corr\left(X_{t_i}(\mathbf{p}_i)\right)$ are invariant in time, which is in accordance for what we find in the results setion. For simplicity, we also assume that the mean vectors μ_{t_i} and the covariance matrices Σ_{t_i} of the random vectors $X_{t_i}(\mathbf{p}_i)$ are invariant in time and are thus only referred to as μ and Σ from now on. If for an experiment μ and Σ were found to change in time, both can be made time-dependent without affecting the algorithm described here.

The proposed stochastic timestepping is based on directly estimating the correlation coefficient δ between the local truncation error at two consecutive timesteps. Given the correlation coefficient δ, the timestepping can be defined straightforwardly. In the first step, we construct a base realization $y_{t_{i+1}}$ of the random vector $X_{t_{i+1}}(\mathbf{p}_{i+1})$ that by construction has the correct spatial structure as described in the previous section. The vector $y_{t_{i+1}}$ is still unconnected to the realization $x_{t_i}(\mathbf{p}_i)$ of the random vector $X_{t_i}(\mathbf{p}_i)$ at the previous timestep. The second step is then to construct a new realization of the stochastic process at timestep $i+1$ from $x_{t_i}(\mathbf{p}_i)$ and $y_{t_{i+1}}$ by using the formula for correlating two random variables

$$x_{t_{i+1}}(\mathbf{p}_{i+1}) = \mu + \left(diag\left(\Sigma\right)\right)^{\frac{1}{2}} \left[\delta \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \left(x_{t_i}(\mathbf{p}_i) - \mu\right) + \sqrt{1-\delta^2} y_{t_{i+1}}\right]. \quad (4.16)$$

In this, each entry of the vector $x_{t_i}(\mathbf{p}_i)$ is propagated in time separately and the vector $y_{t_{i+1}}$ provides the rate of change in $x_{t_i}(\mathbf{p}_i)$ when performing the timestep. As was already the case for the spatial structure, we here also have to make assumptions about certain correlations being negligible. The effect that several entries of the vector $x_{t_i}(\mathbf{p}_i)$ might influence one specific entry of $x_{t_{i+1}}(\mathbf{p}_{i+1})$ is not represented in this approach.

We need to prove that a random vector $X_{t_{i+1}}(\mathbf{p}_{i+1})$ modeled in this way, indeed follows the desired probability distribution $\mathcal{N}\left(\mu, \Sigma\right)$. Equation (4.16) is basically a sum of two independent random vectors, $X_1 := \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \left(X_{t_i}(\mathbf{p}_i) - \mu\right)$ and $X_2 := Y_{t_{i+1}}$, that both have the same probability distribution $X_1,\ X_2 \sim \mathcal{N}\left(0, \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \Sigma \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}}\right)$. The covariance matrix of $X_3 := \delta X_1 + \sqrt{1-\delta^2} X_2$ follows from the summation rules

for covariance matrices

$$cov\left(X_3\right) = cov\left(\delta X_1 + \sqrt{1-\delta^2}X_2\right) \tag{4.17}$$
$$= \delta^2 cov\left(X_1\right) + \left(1-\delta^2\right) cov\left(X_2\right) \tag{4.18}$$
$$= cov\left(X_1\right) \tag{4.19}$$
$$= \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \Sigma \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \tag{4.20}$$

Thus, X_3 follows the desired distribution $X_3 \sim \mathcal{N}\left(0, \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}} \Sigma \left(diag\left(\Sigma\right)\right)^{-\frac{1}{2}}\right)$. Finally, by using (4.13) to reapply the variance and the mean, it is shown that $X_{t_{i+1}}(\mathbf{p}_{i+1}) \sim \mathcal{N}\left(\mu, \Sigma\right)$.

With this, we can now fully define the stochastic process $X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ and discuss its properties. To fully define the stochastic process, we need to assign the first random vector $X_{t_1}(\mathbf{p}_1)$. We could simply start from any realization with the correct spatial structure. In our case, it is however reasonable to assume that the local truncation error $N^1_\Delta(q)$ at timestep 1 is known, and we choose it as a realization of the random vector $X_{t_1}(\mathbf{p}_1)$ as it already satisfies the requirements in terms of the spatial structure. A fixed choice of $X_{t_1}(\mathbf{p}_1)$ is also advantageous in terms of spread of the desired goal error estimates. The reason is that the variance of the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ is then zero for the first timestep and afterwards only slowly growing in time due to the correlation coefficient δ. At least, the reduced growth rate of the error estimator's variance holds true until stochasticity takes over and the random vectors $X_{t_i}(\mathbf{p}_i)$ from a specific timestep i onwards can be considered to be uncorrelated from the initial information $N^1_\Delta(q)$.

4.3 The ICON-Ocean Model

The ICON-Ocean model is based on the hydrostatic primitive equations [21, 37]

$$\frac{\partial v}{\partial t} + \left(f+\omega\right)\vec{z}\times v + \frac{\nabla_h \left|v\right|^2}{2} + w\frac{\partial v}{\partial z} + \frac{1}{\rho_0}\nabla_h p - D_h v - \frac{\partial}{\partial z}\mathbb{K}^v_v \frac{\partial}{\partial z} v = 0 \tag{4.21}$$
$$\frac{\partial p}{\partial z} = -\rho g \tag{4.22}$$
$$\frac{\partial \eta}{\partial t} + div \int_{-B}^{\eta} v dz = 0 \tag{4.23}$$
$$div_h v + \frac{\partial w}{\partial z} = 0 \tag{4.24}$$
$$\frac{\partial C}{\partial t} + div\left(Cv\right) - div_h\left(\mathbb{K}^C_h \nabla C\right) - \frac{\partial}{\partial z}\mathbb{K}^C_v \frac{\partial}{\partial z} C = 0 \tag{4.25}$$
$$\rho = F_{eos}\left(p, \hat{t}, \hat{s}\right). \tag{4.26}$$

Staying in the notation of the continuous system (2.1) from chapter 2, the solution q is represented by the surface elevation η, the temperature $\hat{t}$, the salinity $\hat{s}$ and the horizontal velocity v. The variables and parameters of the system of equations (4.21) - (4.26) are: the Coriolis parameter f, the vorticity ω, z denotes the vertical dimension, the vertical velocity w, the density of sea water ρ and its reference value ρ_0, the hydrostatic pressure p, g is the gravitational constant, B describes the bathymetry, and C denotes a tracer, i.e. the temperature $\hat{t}$ or the salinity $\hat{s}$. Additionally, the operator D_h denotes a diffusion operator for the horizontal velocity with viscosity coefficient $\mathbb{K}_h^v$. In the same way $\mathbb{K}_v^v$ denotes the viscosity coefficient for the vertical velocity diffusion. The diffusion coefficients $\mathbb{K}_v^C$ and $\mathbb{K}_h^C$ are the vertical diffusion coefficient and the horizontal diffusion coefficient for the tracer C. The function F_{eos} is an equation of state for the density of sea water.

Our numerical model to approximate q is the ICON-Ocean model [27]. The chosen discretization is based on a mimetic scheme, which means that the discrete operators are build in a way that they mimic their continuous counterparts in key quantities such as conservation properties and a consistent behavior in terms of energy, wave propagation, vorticity dynamics and tracer transport. The timestepping is performed with a semi-implicit Adam-Bashford method. The model is implemented on an unstructured, triangular grid on the sphere, using a C-type staggering for the prognostic variables height, temperature, salinity and velocity. The discrete surface elevation η_Δ, discrete temperatures $\hat{t}_\Delta$ and the discrete salinity $\hat{s}_\Delta$ are located at the center of grid cells, while the velocities are residing on the midpoints of the triangle edges, oriented in normal direction to the triangle edge. The velocities are denoted by vn_Δ. The vertical dimension is realized as a layering of several horizontal grids, where their vertical spacing defines the thicknesses d_k of the layers $k \in \{1, \dots, \#layers\}$. In this, the layers are numbered consecutively from the surface layer $k = 1$ to the bottom level $k = \#layers$. The discretized solution at timestep i is thus a vector $q_\Delta^i = (\eta_\Delta^i, \hat{t}_\Delta^i, \hat{s}_\Delta^i, vn_\Delta^i)^T \in \mathbb{R}^m$, where a single horizontal layer consists of $\frac{m}{\#layers}$ variables.

As a sub-grid scale parametrization besides the already mentioned horizontal eddy viscosity $\mathbb{K}_h^v$ used in the diffusion operator D_h, we employ the so called PP-scheme [39] by Pacanowski and Philander for the vertical ocean mixing. With this scheme, the intensity of vertical mixing is chosen to be dependent on the Richardson number Ri, a measure for the importance of buoyancy compared to vertical shear. In ICON-Ocean model, the Richardson number is calculated as

$$Ri := max\left\{\frac{D_z\rho}{|D_z P_G vn_\Delta|^2}, 0\right\}, \tag{4.27}$$

where D_z denotes the vertical derivative and P_G is a reconstruction from the triangle edges to centers. The dependence on the Richardson number is realized by an

Table 4.1: Averages over the square roots of the triangle areas for employed grids

Resolution	Grid Spacing in km
$\Delta 4$	160
$\Delta 5$	80
$\Delta 6$	40
$\Delta 7$	20
$\Delta 8$	10

adjustment of the vertical diffusion coefficient for velocity $\mathbb{K}_v^v$

$$\mathbb{K}_v^v = \frac{\mathbb{K}_0^v}{(1 + C_{Ri} Ri)^2} + \mathbb{K}_b^v \tag{4.28}$$

and for the vertical tracer diffusion coefficient

$$\mathbb{K}_v^C = \frac{\mathbb{K}_0^C}{(1 + C_{Ri} Ri)^3} + \mathbb{K}_b^C. \tag{4.29}$$

The coefficients $\mathbb{K}_b^v$ and $\mathbb{K}_b^C$ are the respective background diffusivities, $\mathbb{K}_0^v$ and $\mathbb{K}_0^C$ are scaling factors, and C_{Ri} is an adjustable parameter.

To perform one refinement for a grid of chosen resolution, the triangle edges are bisected. As in chapter 3, the resolution of the grid is characterized by the number of bisections $N \in \mathbb{N}$ of a base grid, where we denote the N-times refined grid by ΔN. The average over the square roots of the triangle areas is given as a measure for the grid spacing in table 4.1.

For the resulting discretized system of the ICON-Ocean model, an analysis of its numerical properties and its physical behavior, we refer to the manuscript [27]. Based on the specifics of the ICON-Ocean model, we give the usual definition of the local truncation error at timestep i

$$N_\Delta^i(q) = P_\Delta\left(q(x, t_i)\right) - S_\Delta\left(P_\Delta\left(q(x, t_{i-1})\right), P_\Delta\left(q(x, t_{i-2})\right)\right). \tag{4.30}$$

Here, S_Δ is one application of the discrete time-stepping operator of the ICON-Ocean model, and P_Δ again denotes a mapping onto the discretized space Ω_Δ. In the calculation of the local truncation error we have to consider the two previous timesteps t_{i-1} and t_{i-2} in the calculation of t_i due to the chosen time discretization that is directly dependent on the two previous model states. In the case that $q(x, t_i)$ is replaced by a high-resolution solution q_{high}, we need to define P_Δ. A reasonable choice that was also used in chapter 3 is to use point-wise projection to map the height, salinity and

temperature values onto the coarse resolution and a nearest neighbor approach to map the velocities onto the coarse grid.

4.4 Experiment: Spherical Channel with Baroclinic Instabilities

We want to test our stochastic parametrization for the local truncation error that we derived in section 4.2.2 in the 3D ICON-Ocean model. As numerical experiment we want an idealized, 3D ocean flow that however still incorporates fundamentally important ocean dynamics. This is why we choose to study an experiment in which tilted density surfaces trigger baroclinic instabilities, one important source of mesoscale ocean eddies. In the process, available potential energy is converted into kinetic energy. Among other effects, the occuring mesoscale eddies have a strong influence on the heat transports in the ocean as they induce a heat transport in meridional direction [24] as well as a net upward heat transport [22, 52]. In the experiment we describe here, we focus on the differences in the net vertical heat transport due to differences in the model resolution that stem from not fully resolving the mesoscale eddy field.

Our numerical experiment is based on the experiments described in [10, 27]. The computational domain is a zonally-reentrent channel on an f-plane at 65° South. The channel has a length of $1900km$ and a width of $1650km$ and a depth of 1600m with flat bathymetry. The vertical dimension is discretized into 31 vertical layers whose thicknesses increase from $14m$ at the surface to $150m$ for the bottom layer. For this specific channel setup, the triangular grid can be constructed in a way that the lateral boundaries are entirely smooth which keeps disturbances to the channel flow due to lateral boundaries to a minimum. This enables us to further reduce the horizontal viscosity parameter $\mathbb{K}_h^v$ and thus have a stronger mesoscale eddy field.

The channel flow is initialized with a temperature profile for $\hat{t}_\Delta^0$ that increases linearly from $10.5°C$ at the southern channel boundary towards $18.5°C$ at its northern boundary. In the vertical dimension the initial temperature is linearly decreasing with depth with a rate of $-6.2 \cdot 10^{-3} K/m$. Additionally, a temperature perturbation is added with an intensity of 0.5 percent of the original field and of wavenumber 4. The salinity tracer is initialized with a constant value of $\hat{s}_\Delta^0 = 35psu$. Concerning the tracers, the density ρ is thus only dependent on changes in the temperature field. At initialization, the fluid is at rest $vn_\Delta^0 = 0$ and the ocean surface is flat $\eta_\Delta^0 = 0$.

The channel flow is driven by wind forcing and a temperature restoring. The wind forcing is analytical, follows a cosine shape, acts in zonal direction with a maximal intensity of $0.05\frac{Nm}{s}$ at the center of the channel, and diminishes to zero at its northern and southern boundaries. The initial temperature field is restored by a full field relaxation to the initial temperatures in small longitudinal bands of width $165km$ at

Table 4.2: Horizontal velocity viscosity coefficients $\mathbb{K}_h^v$ for the Laplacian and biharmonic diffusion operator respectively and corresponding timestep lengths at different resolutions. $D_h = \Delta$ denotes Laplacian diffusion and in case $D_h = \Delta^2$ a biharmonic diffusion operator is used.

Resolution	Diffusion Operator D_h	$\mathbb{K}_h^v$ in $\frac{m^2}{s}$	Δt in s
$160km$	Δ	5.5E+04 $\frac{m^2}{s}$	3600
$80km$	Δ	4.5E+03 $\frac{m^2}{s}$	1800
$40km$	Δ^2	9.6E+11 $\frac{m^4}{s}$	1200
$20km$	Δ^2	1.2E+11 $\frac{m^4}{s}$	600
$10km$	Δ^2	1.5E+10 $\frac{m^4}{s}$	300

the northern and southern channel boundaries. The restoring timescale is 3 days. The task of the temperature restoring is to replenish the available potential energy that is converted into eddy kinetic energy, the kinetic energy contained in the mesoscale eddy field.

We implement the experiment for a series of different model resolutions from 160km to 10km. The horizontal velocity diffusion coefficients $\mathbb{K}_h^v$ at the different resolutions are given in table 4.2. The other model parameters are chosen to be independent of the horizontal resolution in order to increase the comparability of the model results at different resolutions, and to be able to easier explain differences in the model solutions by the change in grid spacing. The horizontal temperature diffusion coefficient $\mathbb{K}_h^{\hat{t}}$ is chosen as $300\frac{m^2}{s}$ for all resolutions. The vertical background viscosity and diffusion coefficients are chosen to be $\mathbb{K}_b^v = 2.0 \times 10^{-5}\frac{m^2}{s}$ and $\mathbb{K}_b^{\hat{t}} = 1.0 \times 10^{-5}\frac{m^2}{s}$, and the parameter C_{Ri} is chosen as $C_{Ri} = 0.5^{-2}$. The scaling factors in the PP-scheme are $\mathbb{K}_0^v = 3.2 \cdot 10^{-3}\frac{m^2}{s}$ for vertical viscosity and the corresponding scaling factor for the vertical diffusion of temperature is $\mathbb{K}_0^{\hat{t}} = 3.0 \cdot 10^{-3}\frac{m^2}{s}$. Additionally, the forcings are balanced by bottom boundary friction in the form of a bottom drag coefficient of $3.2 \cdot 10^{-3}$.

Figure 4.1 shows snapshots of the final state for the surface elevation, the temperature at 105 m depth and the vertical velocities at 110m depth at the 10km model resolution. All flow fields clearly indicate the presence of mesoscale eddies and the geostrophic turbulence in the flow. This first impression is further strengthened by the corresponding snapshot of the vertical structure shown for a cross-section of the potential temperature field in figure 4.2. In the absence of salinity changes, this is equivalent to the pattern of density. The perturbations to the density field due to the passing of mesoscale eddies extend to the entire vertical column in our experimental setup and are visible in figure 4.2 at $28S$, $5S$, $7N$, $20N$ and $32N$ as undulations in the

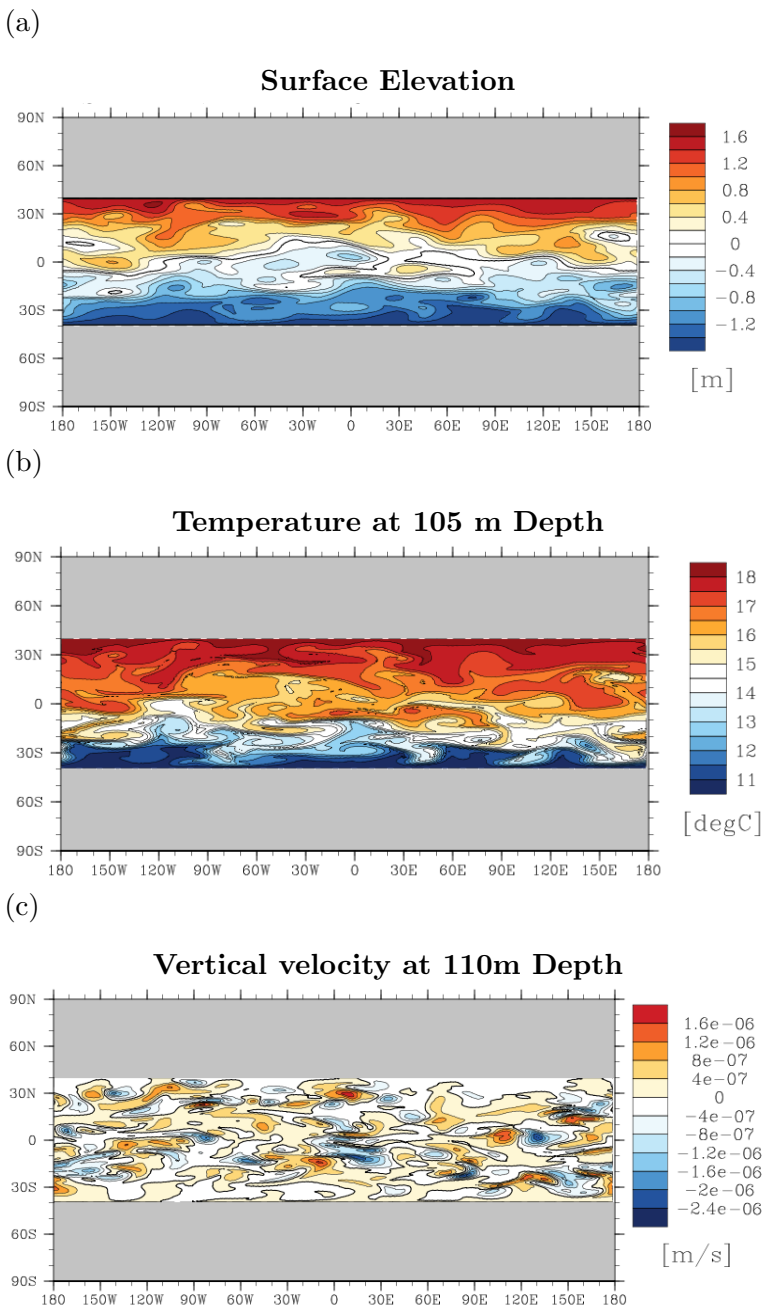


Figure 4.1: State of the channel flow after an integration time of 2 years at a resolution of 10km. The plots show snapshots of surface elevation (a), potential temperature at 105m depth (b), and vertical velocities at 110 meters depth (c). The channel is mapped to $[180W, 180E] \times [40S, 40N]$.

Vertical profile of potential temperature in meridional cross-section

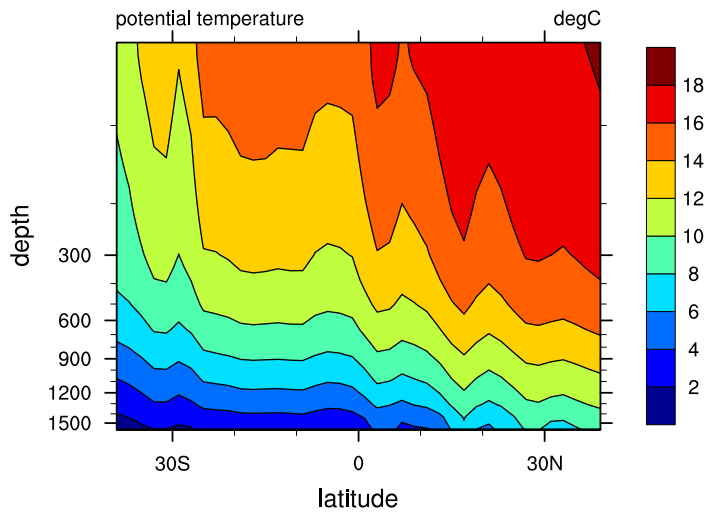


Figure 4.2: Snapshot of the vertical structure of the potential temperature for a cross-section from the southern to the northern boundary after an integration time of 2 years on a resolution of 10km. The channel is mapped to $[40S, 40N]$.

temperature field.

We now want to quantify this rather qualitative description of the channel flow by evaluating the strength of the mesoscale eddy field and its resulting effects on the distribution of heat. A good indicator for the strength of the mesoscale eddy field is the kinetic energy contained in the eddies, referred to as Eddy Kinetic Energy (EKE). We here define the volume-averaged eddy kinetic energy where we choose the volume to be the entire volume of the channel. For the volume-averaged eddy kinetic energy, the velocities vn^i_Δ, residing on the edges, are first interpolated into the cell centers. This yields one zonal velocity component u^i_Δ and one meridional velocity component v^i_Δ per grid cell. With these velocities we can define a volume-averaged eddy kinetic energy in a region $\Omega_{\Delta R}$ with volume $V(\Omega_{\Delta R})$ by:

$$J_{EKE}(q^i_\Delta) := \frac{1}{V(\Omega_{\Delta R})} \sum_{C^i \in \Omega_{\Delta R}} V(C^i)((u^i_\Delta - \bar{u^i_\Delta})^2 + (v^i_\Delta - \bar{v^i_\Delta})^2). \tag{4.31}$$

The C^is denote grid cells of thickness $d(C^i)$ and volume $V(C^i)$. Usually, $\bar{u^i_\Delta}$ and $\bar{v^i_\Delta}$ denote temporal averages of u^i_Δ and v^i_Δ, respectively. However, due to the zonally symmetric structure of the channel, we are able to replace the temporal averages with

spatial averages of u^i_Δ and v^i_Δ over latitudinal bands. The reasoning behind this choice is that, in a zonally symmetric channel, the grid cells at one specific latitudinal band all share the same statistical properties.

The resulting behavior of EKE for the channel flow at different model resolutions is shown in figure 4.3 for a series of model resolutions, ranging from $10km$ to $160km$. At the beginning of the integration time, the EKE is small for all model resolutions

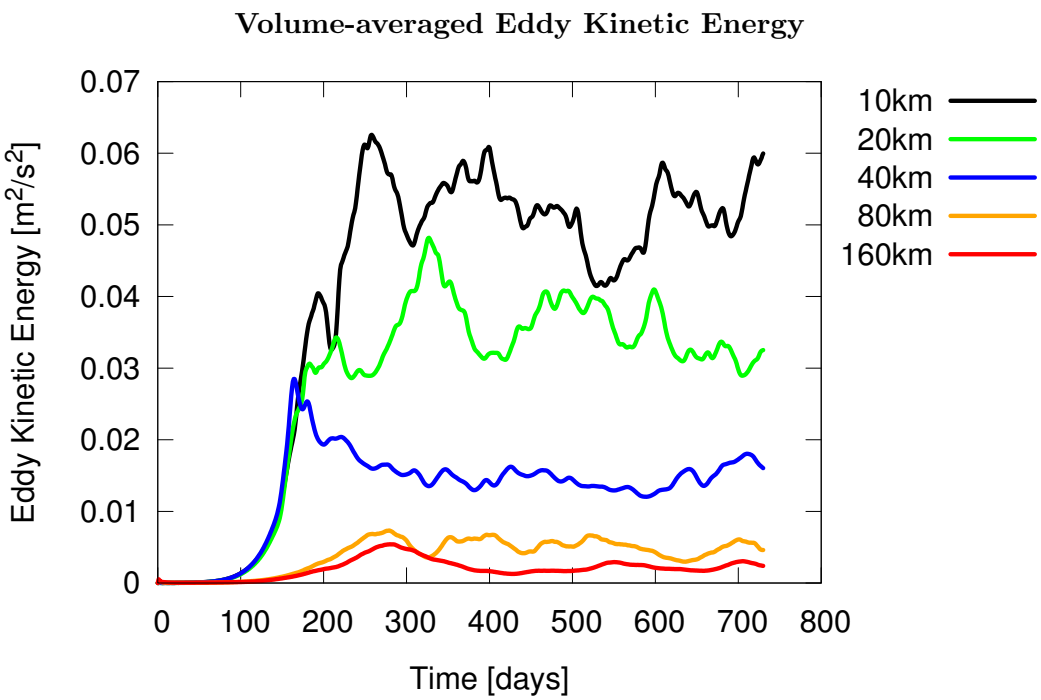


Figure 4.3: The Eddy Kinetic Energy averaged over the entire channel volume is shown for the entire integration time of 2 years for different model resolutions: 160km (red), 80km (orange), 40km (blue), 20km (green), 10km (black)

until 100 days of integration time. After that, the flows at model resolution 40km and higher start to develop a significantly different behavior than the 80km and 160km experiments. For the higher model resolutions, mesoscale eddies emerge in the flow in order to even out the sloped temperature profile. The highest rate of increase in EKE is found at about 150 days of integration time. The temperature profile is however constantly replenished by the temperature forcing in form of the temperature relaxation to the initial temperature field. As a result, the mesoscale eddy field is further spun up until the system eventually runs into an equilibrium state where the rate at which available potential energy is restored is on par with the rate at which it is converted into EKE. This state is obtained after around one year of integration time for the 40km, 20km and 10km resolution. The strengths of the resulting mesoscale

eddy fields are however quite different. While the strength at the 40km resolution is about 0.015 m^2/s^2, the mesoscale eddy field at the next higher model resolution 20km is already more than 100 percent stronger, and the 10km resolution is again about 60 percent stronger in terms of EKE compared to the 20km resolution. Compared to these resolutions, the EKE for the 160km resolution (0.0025 m^2/s^2) and the 80km resolution (0.005 m^2/s^2) is significantly weaker. However the EKE is not just lower at the coarser resolutions, its behavior in time is also completely different. The channel flow at the two low model resolutions does not undergo the same transition in terms of the spin up of the mesoscale eddy field that starts at a 100 days of integration time. There is an indication of an onset of a similar transition at later times at around 150 to 200 days of integration time, but this transition is much weaker and progresses slower as it reaches its peak after 270 days of integration time. These differences between the higher and lower resolutions might be in part explained by the different forms of the horizontal velocity diffusion operator D_h, but ultimately they are a consequence of the horizontal model resolution.

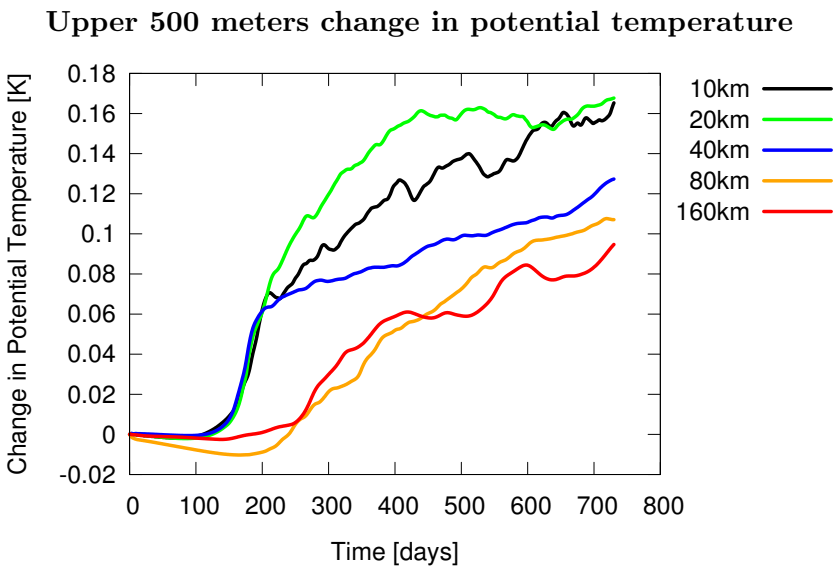
The differences in the EKE have a strong influence on the heat distribution in the channel. For the analysis of the heat distribution, we compare timeseries of the change in regionally-averaged potential temperatures compared to the initial temperature profile. We define this quantity by

$$J_{PT}(q^i_\Delta) := \frac{1}{V(\Omega_\Delta)} \sum_{C^i \in \Omega_{\Delta R}} V(C^i) \left(\hat{t}^i_\Delta - \hat{t}^0_\Delta\right). \tag{4.32}$$

In figure 4.4 a) , we show the evolution of $J_{PT}(q^i_\Delta)$ for the upper 500 meters of the entire channel at different model resolutions. The general trend is an average warming of the upper 500 meters for all model resolutions. The rates at which this region of the channel warms differ depending on model resolution. After the 2 year of integration time period, we obtain the lowest warming of 0.95° K for the 160km resolution, which then steadily increases by about $0.01^\circ - 0.03^\circ$ K with every refinement step of the horizontal grid. An exception is the 20km resolution for which the warming is on par with the 10km resolution. For the lack of difference between the 10km and the 20km resolution we can however not draw final conclusions because after 2 years of integration time, none of the model simulations is in equilibrium yet. The reason for the differences in the warming at the five different model resolutions can be clearly tracked down to the strength of the mesoscale eddy field. This can be best seen for the model resolutions higher than 40km for which the strong increase in warming after 150 days of integration time is lead by the strong increase in EKE that was discussed previously in this section.

The effects of the mesoscale eddy field on the vertical heat distribution are also found for the lower 1000 meters of the channel, see figure 4.4 b). While the averaged potential

(a)



(b)

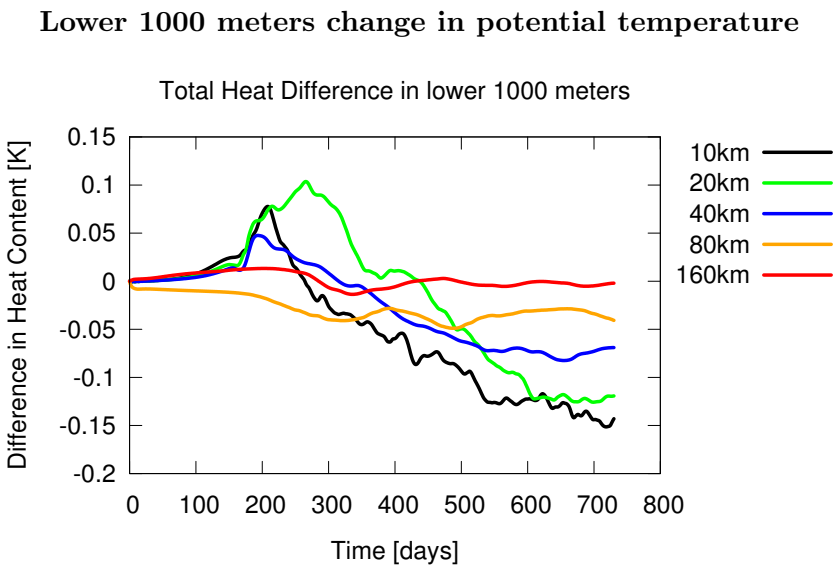


Figure 4.4: The change in the averaged potential temperatures compared to the initial temperature profile is shown for the upper 500 meters (a), and for the lower 1000 meters (b) for the entire integration time of 2 years and different model resolutions: 160km (red), 80km (orange), 40km (blue), 20km (green), 10km (black)

temperature in the lower 1000 meters is constant for the 160km resolution, the values

fluctuate for the higher model resolutions with an eventual cooling at the end of the integration time. This cooling becomes stronger with increased model resolution. The cooling is suspected to result from a lack of warming of the deeper layers when the warmer waters at the northern boundaries are constantly advected by mesoscale eddy activity instead of being diffused to warm the surrounding deep waters.

4.5 Estimation of the Problem-specific Stochastic Process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$

For the channel experiment described in the previous section, we want to perform the estimation of a problem-specific stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ that could be used in the goal error estimation procedures described in 4.2.1. As we do not yet have an adjoint model of the ICON-Ocean model that could provide an adjoint solution for the required timescales of weeks to months, we aim at deriving a stochastic process for the stochastic-physics ensemble error estimator that corresponds to our stochastic dual-weighted error estimator (4.5). Because we do not work with the linear, adjoint-based error estimator but instead create an ensemble of perturbed model simulations at the chosen low model resolution, the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ needs to be even stronger constrained to behave physically sound, as unphysical behavior now has a direct negative impact on the underlying flow.

The chosen structure of the stochastic process is described in 4.2.2, but the task to calculate the local truncation error and estimate the required parameters $\mathbf{p}_i$ remains and is described in this section.

First, the calculation of the local truncation error is discussed that we need for the estimation of the parameters $\mathbf{p}_i$. We discuss the local truncation error and how we can represent it in our model. We show that the usual approach that we described in (4.30) to simply insert the complete true state vector into the timestepping operator S_Δ is not applicable and we discuss alternatives.

Afterwards, we analyze the stochastic process and its derivation in the following way. First, we thoroughly describe the parameter learning process and the choices made, in order to guarantee that the parameters are robust estimates. Then we analyze the resulting structures of the random vectors $X_{t_i}(\mathbf{p}_i)$. In a first step, we compare the spatial structure of the local truncation error to realizations of the spatial structure we obtain for the random vectors $X_{t_i}(\mathbf{p}_i)$. For this, we compare snapshots of both spatial structures and discuss their patterns, value ranges and the general visual impression. In a second step, we discuss how these spatial structures are connected in time by the stochastic timestepping scheme of the stochastic process. This is done by analyzing a realization of the stochastic process in the context of the ensemble-type error estimator described in 4.2.1. We describe the impact of the stochastic forcing on

the channel circulation, and discuss the differences between the temporal evolution of the local truncation error and the stochastic process by numerical as well as analytical arguments.

To estimate the required parameters, we use the parameter learning algorithm 3.4 we developed in chapter 3 for the shallow-water flows. The algorithmic steps to obtain the parameters of the stochastic process have been proven to work successfully and we find that they only need to be slightly adjusted to work in our 3D environment. This adjustment is not about conceptual changes but we now estimate different stochastic quantities and thus the fourth step of the processing step of the described learning algorithm needs to be extended. The changes are that the parameter set $\mathbf{p}_i$ is now a associated with the full local truncation error instead of its temporal fluctuations, and additionally to our standard stochastic quantities, variance and mean, we now need estimates of the correlation matrices for the spatial structure of the local truncation error and an estimate of the correlation coefficient for the time-correlations of the local truncation error.

In the estimation of all of these stochastic quantities we make use of the zonally-symmetric structure of the channel to increase the number of available observations of the local truncation error at each timestep. The idea is that the local truncation error values in a specific latitudinal band are treated as independent observations of the same random variable, which as a result largely increases the available observations per random variable in a given timestep. This is especially important for the estimation of the correlation matrices such as the matrix for the surface structure $\mathbf{C}_{surf}$ that requires a large amount of observations because a necessary condition for the Cholesky-decomposition is that the number of observations is larger than its dimension. The dimension, however, is dependent on the chosen model resolution for the channel experiment.

We need to choose two model resolutions, one coarse resolution model for which we eventually want to perform goal-oriented error estimation, and its high-resolution reference model resolution. As coarse model resolution, we choose the 40km resolution model, as this is targeted to be the new standard resolution for future ocean models in climate simulations. Also, this is the coarsest resolution at which we obtained a significant mesoscale eddy field. As a corresponding high-resolution reference model resolution that will be known for the near-initial-learning period to estimate the parameter set $\mathbf{p}_i$ from, we choose the 10km model resolution, a model resolution which is considered to be eddy-resolving. The local truncation error will contain the differences between these two model resolution by insertion of the 10km model solution, denoted by $q_{\Delta,\,10km}$, into the 40km model operator.

The parameter estimation process is started after about 340 days of integration time, which corresponds to 24500 timesteps at the 40km resolution. This time-window is of interest because the flow already has a fully developed mesoscale eddy field (see figure

4.3), but at the same time the temperature field is in transition into a mean warming of the upper layers of the channel and a cooling for the deeper layers (as shown in figures 4.4 a) and b)). A suitable stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ would have to be able to represent the connection between the differences in warming respective cooling and the strength of the mesoscale eddy field at different model resolutions.

4.5.1 The Representation of the Local Truncation Error

For our 3D ocean model, we find the straightfoward formulation of the local truncation error by (4.30) that uses all model variables to be not applicable. We will illustrate the problem we encountered with this formulation and propose an alternative formulation of the local truncation error that is then shown to be valid for an eddying flow regime driven by density differences.

We implement the local truncation error (4.30) calculation as a cycle that is operated while running the model. We assume that we have already obtained a timeseries of corrected model states until timestep $i-1$ and we are at the i-th timestep. The cycle then starts with correcting the current state variables of the 40km resolution model at the i-th timestep, i.e. add the difference between the true 10km model state at timestep i and the model state at timestep i as a forcing term. Second, with the corrected 40km model states at timestep i and $i-1$ we can apply S_Δ once to obtain $S_\Delta\left(P_\Delta\left(q(x,t_i)\right), P_\Delta\left(q(x,t_{i-1})\right)\right)$ from which we can calculate the local truncation error. Third, we correct the 40km model state at timestep $i+1$ by the true 10km model state at timestep $i+1$ and move to the next timestep $i+2$.

The Local Truncation Error using the Full State Vector

We use the described cycle to create the local truncation error by correcting all state variables at the same time. However, as a result we obtain unphysical values for the local truncation error which can be traced back to the surface elevation variable η. To isolate the effect, we only correct the surface elevation variable η in the local truncation error calculation and show snapshots of the resulting temperature fields in which the unphysical behavior occurs (figure 4.5). The effects are however not limited to a specific timestep but occur frequently at every timestep. The highest temperature values are found at the surface at $40N$ in figure 4.5 a), and we argue that these temperature values indicate artifacts in the solution. This is because we obtain values larger than $19.0°C$, with the maximum value being $19.38°C$, where the maximum temperature in the channel in the absence of an additional temperature forcing can only be $18.5°C$ by construction. These artificial temperature are found to occur throughout the channel and not just at the northern boundary. They can just not be easily identified as such because they can be easily mistaken as mesoscale eddies. For instance, in the vertical cross-section in figure 4.5 b), the vertical extend of the surface heating at $10S, 70E$

is shown. It strongly resembles the temperature undulations due to mesoscale eddies. However, these temperatures in the surface layer can grow larger than $18.5°C$, the heating is strongly confined to the surface, and it is several degrees warmer than its surroundings. What we have not shown here is that these high temperatures are not transported to this specific location, but they just occur seemingly random in the flow and must thus be a consequence of our applied correction of the surface elevation. The

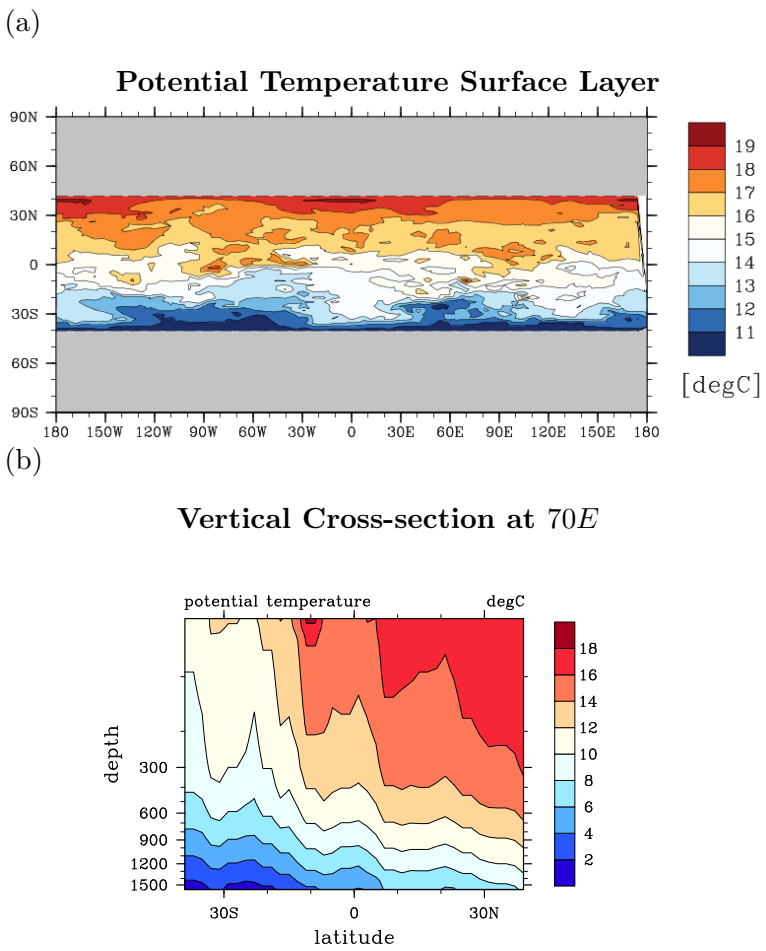


Figure 4.5: Potential Temperature of height-corrected simulation at 40km resolution for (a) the surface layer and (b) a vertical cross-section at $70E$ at timestep 24487.

presence of these artifacts in the temperature variable would of course then also be present in the local truncation error which means that many of the indicated errors in the model solution could just be artifacts.

A Temperature-based Local Truncation Error Representation

We thus need an alternative formulation of the local truncation error that bypasses the correction of the surface elevation variable η. As an alternative, we choose to omit the corrections for the surface elevations η and the velocities vn entirely from the local truncation error calculation. The omitted variables have to follow the changes in the temperature field, and we show this to be a suitable assumption for the channel flow that is for the most part density-driven. The local truncation error and thus the stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ then only consist of the the temperature variable $\hat{t}$.

We show that, for the channel experiment, our choice to only take into account the temperature variable in the calculation of the local truncation error is justified. Towards that aim, we compare the resulting surface elevation fields η from the 10km high-resolution reference solution to the η fields of a 40km model resolution solution whose temperature variable is corrected to the corresponding true value from the 10km reference solution at the beginning of each new timestep. If the resulting 40km model solution is close to the 10km reference solution in the model variable η, we consider the assumption to omit the surface elevation field η and the velocity field vn to be justified. For the comparison, we show snapshots of the surface elevations η from the same timestep for the 10km resolution and the temperature-corrected 40km resolution in figure 4.6. To get a better picture for the changes in the surface elevation, we additionally show the original field at a 40km model resolution and show difference plots between the two different 40km solutions and the 10km solution in figure 4.7. From figure 4.6 we can see that the temperature-corrected 40km resolution field and the 10km reference resolution field generally show a very similar pattern for the entire surface layer. Although similar, the temperature-corrected 40km field is generally weaker and differs in the details, i.e. it differs in the exact position and shape of the surface elevations. However, in comparison to the uncorrected 40km resolution surface elevation, the full impact of the temperature correction in the 40km resolution solution becomes visible. Although both are being run using the same 40km resolution grid, we can state that there is not much of a resemblance between the uncorrected and the temperature-corrected fields. Where the uncorrected 40km resolution field does not indicate much of a mesoscale eddy field, the temperature-corrected 40km resolution flow in contrast clearly indicates a geostrophically-turbulent flow. This interpretation is strengthened by the difference plots in figure 4.7. The difference between the uncorrected 40km solution and the 10km solution (figure 4.7 a) can be up to $1m$, which is exactly double compared to the maximum difference between the temperature-corrected 40km solution and the 10km solution (figure 4.7 b) of $0.5m$. The difference is not just significantly smaller, the difference in surface elevation also looks much smoother, indicating that much of the high-resolution dynamics are represented, the effects are just represented consistently weaker. We find these results to be a general description, regardless of

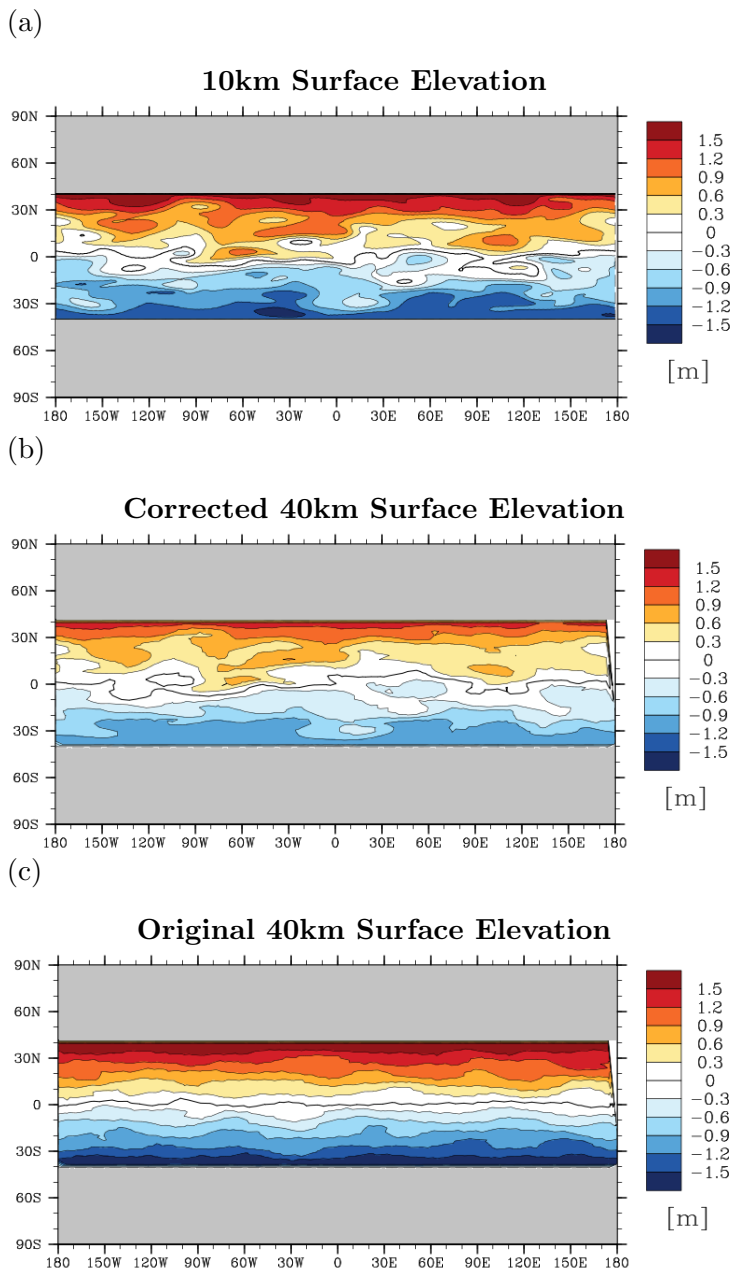


Figure 4.6: Snapshot of the surface elevation η_Δ after 24487 timesteps of integration time at (a) the 10km high-resolution reference resolution, (b) the 40km model resolution, and (c) the 40km model resolution whose temperature values are corrected by the 10km reference resolution values at the beginning of each timestep. The channel is mapped to $[180W, 180E] \times [40S, 40N]$.

the chosen timestep. The temperature-corrected 40km resolution solution is thus undoubtedly much closer to the 10km resolution solution than it is to the uncorrected 40km resolution solution. We conclude that the temperature variable conveys enough information about the 10km resolution flow to reliably correct the 40km resolution flow. Omitting the surface elevation η and the velocities vn from the calculation of the local truncation error is thus justified. This holds especially true if one is eventually interested in error estimates for temperature-based goals.

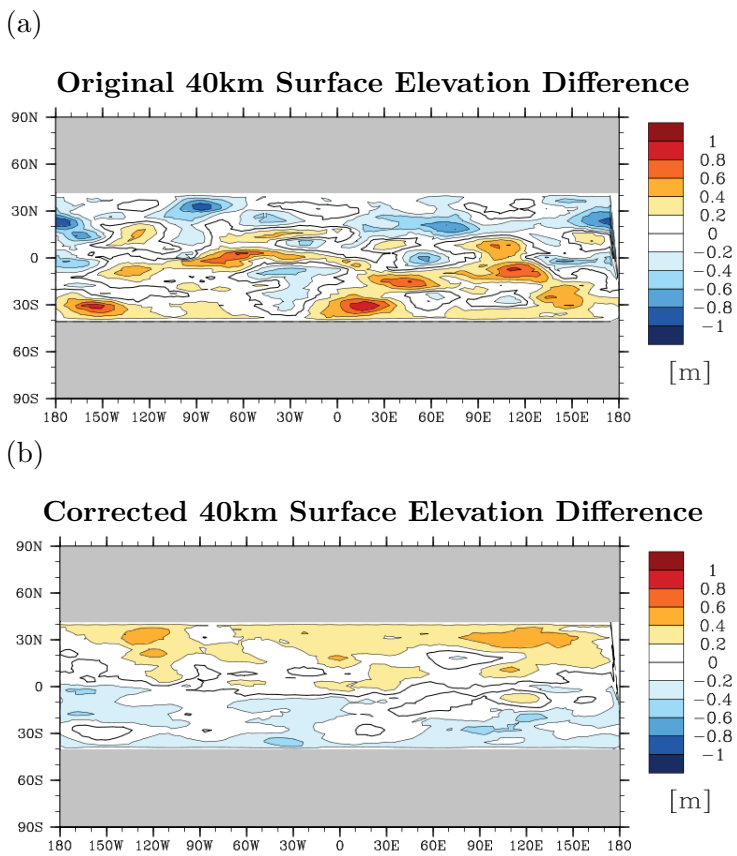


Figure 4.7: Snapshots of differences in the surface elevation η_Δ after 24487 timesteps of integration time between (a) the 10km high-resolution reference resolution and the original 40km solution, (b) the 10km high-resolution reference resolution and the temperature-corrected 40km solution, whose temperature values are corrected by the 10km reference resolution values at the beginning of each timestep. The channel is mapped to $[180W, 180E] \times [40S, 40N]$.

4.5.2 Learning the Parameters of the Stochastic Process

We now follow the four algorithmic steps that are part of the processing cycle of algorithm 3.4.

Step 1: Separate the Flow into Different Dynamical Flow Regimes

The first algorithmic step of the processing step is the separation of the flow into different dynamical flow regimes. We choose to separate the channel flow into latitudinal bands and to separate their vertical layers. We choose each latitudinal band to have the width of a grid cell, which means each band extends only over one band of temperature variables that all reside on the same latitude. We denote the number of resulting latitudinal bands by $\#Lats$ and number them consecutively from the first latitudinal band at the southern boundary of the channel to the $\#Lats$'s band at its northern boundary. The total number of chosen dynamical flow regimes is then $\#Lats \cdot \#layers$. For the channel at 40km resolution the number of latitudinal bands is $\#Lats = 31$ and the channel consists of 32 vertical layers, which makes 992 dynamical flow regimes, which we denote by the subscript $k := \{1, \ldots, 992\}$.

Step 2: Choose the Length of the Near-initial Learning Phase

The second step of the algorithm is the choice of the near-initial learning phase, which we here set to be $n_l := 1520$ timesteps, about 21 days of integration time, after the initial 14500 timesteps. For these 21 days the algorithm knows the 10km high-resolution reference solution $q_{\Delta,\,10km}$ and thus the local truncation error.

We require the knowledge about the local truncation error for a near-initial learning phase of 21 days, which is long compared to the previous chapter 3. The reasons are the required correlation matrix $\mathbf{C}_{surf}$, the mean vector of the random vectors $X_{t_i}(\mathbf{p}_i)$, and the correlation coefficient for the time-correlations of the local truncation error. First, the dimension of $\mathbf{C}_{surf}$ is $\mathbb{R}^{1922\times1922}$, where 1922 is the number of temperature variables in the surface layer of the channel at 40km resolution. Thus, we require at least 1922 observation. Second, the mean vectors cannot be ignored as has been done in chapter 3 and instead need to be estimated robustly, because the structure of the stochastic process is such that its variances do not grow strong-monotonously in time, which increases the importance of the means in relation to the variances. Unreliable estimates of the means would introduce unwanted drifts of the stochastic process, possibly rendering the entire stochastic process $\{X_{t_i}(\mathbf{p}_i)\}_{i\in\mathbb{N}}$ useless. Third, the correlation coefficient in time requires a large timeseries of observation, i.e. a timeseries of local truncation errors. These requirements are met by the longer near-initial learning phase that alone contributes 1520 observations, and by using the mentioned symmetry of the channel flow. Using the symmetry increases the number of available observations per

dynamical flow regime k of 1520 by a factor of 62. The factor 62 is the total number of temperature variables that belong to one latitudinal band 1922/31, i.e. $\#Lons := 62$ longitude values belong to each dynamical flow regime k. Similarly to the ordering of the latitudinal bands, these longitude values are numbered consecutively from west to east by $j \in \{0, \dots 61\}$.

Estimators for the Stochastic Quantities in the Channel Experiment

With these specification, we are able to define the specific estimators for the stochastic quantities that we need to estimate for our algorithm. The stochastic quantities are the mean vector μ, the variances $diag\,(\Sigma)$, the horizontal correlation matrix $\hat{\mathbf{C}}_{surf}$, the vertical correlation matrices $\hat{\mathbf{C}}_{v,\,\hat{t},\,Lat}$, and the temporal correlation coefficient δ. For their estimation, we use that each position of a temperature variable in the channel is now uniquely defined by a combination of one dynamical flow regime number $k \in \{1, \dots 992\}$ and one longitude number $j \in \{0, \dots 61\}$.

The mean μ of the random vector $X_{t_i}(\mathbf{p}_i)$ is estimated separately for each dynamical flow regime, i.e. we estimate one mean value for each latitudinal band and vertical layer. The estimator $\hat{\mu}_k$ for the mean value in dynamical flow regime k is defined as

$$\hat{\mu}_k := \frac{1}{62 \cdot n_l} \sum_{i=1}^{n_l} \sum_{j=0}^{61} N_{\Delta}^{i} \left(P_{\Delta} \left(q_{\Delta,\,10km} \right) \right)_{k,\,j}, \tag{4.33}$$

where n_l is the number of near-initial learning timesteps of 1520. $N_{\Delta}^{i} \left(P_{\Delta} \left(q_{\Delta,\,10km} \right) \right)$ is the local truncation error vector at learning timestep i. The attached subscript $_{k,\,j}$ denotes the value of the local truncation error vector that belongs to the k-th dynamical flow regime and resides at the j-th longitude position. An estimate for the complete mean vector μ can be constructed from the set of estimates $\hat{\mu}_k$, $k = \{1, \dots, 992\}$.

The estimator for the variance respectively the standard deviation of the dynamical flow regime k is defined accordingly

$$\hat{\sigma}_k^2 := \frac{1}{62 \cdot n_l - 1} \sum_{i=1}^{n_l} \sum_{j=0}^{61} \left(N_{\Delta}^{i} \left(P_{\Delta} \left(q_{\Delta,\,10km} \right) \right)_{k,\,j} - \hat{\mu}_k \right)^2. \tag{4.34}$$

The estimator uses the previously defined estimator for the means $\hat{\mu}_k$. From the estimates of $\hat{\sigma}_k^2$, we can construct an estimate for the matrix $diag\,(\Sigma)$, which has the variances of the random variables of the random vector $X_{t_i}(\mathbf{p}_i)$ on the diagonal.

Next, we give the definition of the estimators for the correlations in time and in space. We define these estimators by the matrix entries of their respective sample correlation matrices.

The entry at the matrix position with row l and column m of the sample correlation

matrix $\hat{\mathbf{C}}_{surf}$ is defined by

$$\hat{\mathbf{C}}_{surf}(l,m) := \frac{1}{62\cdot n_l - 1}\sum_{i=1}^{n_l}\sum_{j=0}^{61}\left(N_\Delta^i\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k(l),\,(j(l)+j)mod\,62} - \hat{\mu}_{k(l)}\right) \\ \left(N_\Delta^i\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k(m),\,(j(m)+j)mod\,62} - \hat{\mu}_{k(m)}\right)/\sqrt{\hat{\sigma}_{k(l)}\hat{\sigma}_{k(m)}}. \quad (4.35)$$

The attached subscript $k(m)$ denotes the subset of the local truncation error vector that is associated with the unique dynamical flow regime number $k = \{1,\ldots,992\}$ whose element is the surface temperature variable m. $k(l)$ is defined accordingly. Similarly, $j(l)$ and $j(m)$ denote the longitude at which the temperature variables l and m reside. Additional to the n_l observations, the whole latitudinal bands of the dynamical flow regimes $k(l)$ and $k(m)$ are used. This is represented by the modulo functions $(j(l)+j)\,mod\,62$ and $(j(m)+j)\,mod\,62$ that ensure the correct choice of temperature variables throughout the summation. As a result of this increase in available observations, some matrix entries appear multiple times in the correlation matrix $\hat{\mathbf{C}}_{surf}$.

The set of correlation matrices for the vertical correlations $\mathbf{C}_{v,\,\theta,\,i}$ for each surface variable grid point i is reduced to the correlation matrices $\mathbf{C}_{v,\,\hat{t},\,Lat}$ for each latitudinal band $Lat \in \{1,\ldots,32\}$ for our choice of dynamical flow regimes k and a local truncation error that only uses the temperature variable $\hat{t}$. For the sample correlation matrix $\hat{\mathbf{C}}_{v,\,\hat{t},\,Lat}$, the matrix entry at the position with row d_1 and column d_2 represents the spatial correlation of the temperature variable between layer d_1 and layer d_2 for the latitudinal band $Lat \in \{1,\ldots,32\}$ and is defined by

$$\hat{\mathbf{C}}_{v,\,\hat{t},\,Lat}(d_1,d_2) := \frac{1}{62\cdot n_l - 1}\sum_{i=1}^{n_l}\sum_{j=0}^{61}\left(N_\Delta^i\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k(d_1,\,Lat),\,j} - \hat{\mu}_{k(d_1,\,Lat)}\right) \\ \left(N_\Delta^i\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k(d_2,\,Lat),\,j} - \hat{\mu}_{k(d_2,\,Lat)}\right)/\sqrt{\hat{\sigma}_{k(d_1,\,Lat)}\hat{\sigma}_{k(d_2,\,Lat)}}. \quad (4.36)$$

The attached subscript $k(d_1,\ Lat)$ denotes the subset of the local truncation error vector that is associated with the unique dynamical flow regime number $k = \{1,\ldots,992\}$ defined by layer d_1 and latitude lat. $k(d_2,\ Lat)$ is defined accordingly. Similarly, j denotes the longitude at which the temperature variable reside. In total, this makes 32 different correlation matrices $\hat{\mathbf{C}}_{v,\,\hat{t},\,Lat}$ at the 40km resolution.

The estimator for the temporal correlation coefficient δ is defined by

$$\hat{\delta} := \frac{1}{62 \cdot 992 \cdot (n_l - 1) - 1} \sum_{i=1}^{n_l-1} \sum_{k=1}^{992} \sum_{j=0}^{61} \left(N_\Delta^i \left(P_\Delta \left(q_{\Delta,\, 10km} \right) \right)_{k,\, j} - \hat{\mu}_k \right) \left(N_\Delta^{i+1} \left(P_\Delta \left(q_{\Delta,\, 10km} \right) \right)_{k,\, j} - \hat{\mu}_k \right) / \sqrt{\hat{\sigma}_k \hat{\sigma}_k}, \tag{4.37}$$

where k again denotes a dynamical flow regime number and j denotes the j-th longitude position in the dynamical flow regime with number k.

Step 3: Choose the Underlying Distribution of $\{X_{t_i}(\mathbf{p}_i)\}_{i \in \mathbb{N}}$

In the third step of our algorithm, we discuss the underlying probability distribution of the random vectors $X_{t_i}(\mathbf{p}_i)$ and thus of the stochastic process. In chapter 3, we chose a Laplace distribution that was better at capturing the probability in the tails of the probability distribution. For the channel experiment, we base the choice of the probability distribution on histograms of the real distribution of the local truncation error that we compare to a corresponding Gaussian-distributed random variable and a Laplace-distributed random variable in figure 4.8. These random variables are estimated using (4.33) and (4.34). These figures are taken from latitude number 15 at the center of the channel and different depths. The same plots for other dynamical flow regimes of the channel are comparable and thus not shown. It is to note here that the tails of the distribution of the local truncation error in the temperature value are very long. The figures are capped to ± 6 times the standard deviation, but we observe outliers that can even go up to ± 10 times the standard deviation. We decide to neglect these extreme values and focus on representing the bulk of the probability. Neglecting the outliers, both the Gaussian and the Laplace distribution fit reasonably well to the local truncation error data and can both be used for the stochastic process. The Laplace distribution does a slightly better job at representing the tails of the distribution, especially in the upper layers. However, we here decide on the Gaussian distribution for one reason, which is the repeated applications of the lower triangular matrices L_{surf} and $L_{v,\, \hat{t},\, Lat}$ to realizations of the chosen probability distribution. If we were to choose a Laplace distribution instead of a Gaussian distribution, these linear combinations of random variables would converge to a Gaussian distribution, due to the central limit theorem. As the matrices are lower triangular, depending on the position of the random variable in the matrices, the convergence towards a Gaussian distribution would occur to varying degrees, which would introduce inconsistencies in the probability distribution of the stochastic process.

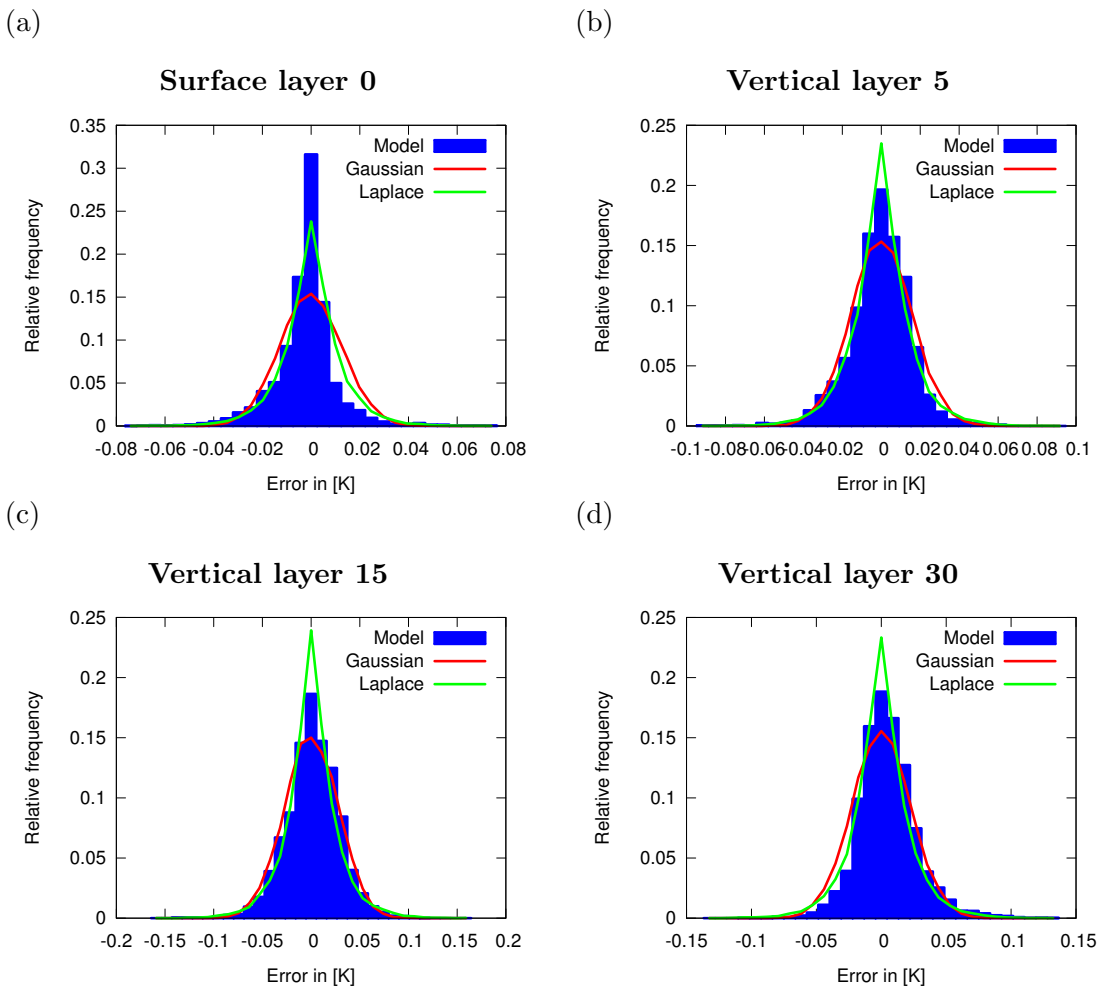


Figure 4.8: Distribution of the local truncation error $N_{\Delta}^{1520}\left(P_{\Delta}\left(q_{\Delta,\,10km}\right)\right)$ (blue bars) at the end of the near-initial learning phase for latitude $Lat = 15$ in the center of the channel, the corresponding Gaussian-distributed random variable (solid red line), and the corresponding Laplace-distributed random variable (solid green line). The comparison is shown for (a) the surface layer, (b) vertical layer 5, (c) vertical layer 15, (d) vertical layer 30.

Step 4: Estimate the Parameters of the Stochastic Process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$

We find that all parameters $\mathbf{p}_i$ can be assumed to be constant in time, and we thus omit the subscript i.

The mean values of $X_{t_i}(\mathbf{p})$ are estimated from (4.33) and are shown in figure 4.9 for all dynamical flow regimes. The means increase in intensity from the northern to the southern channel boundary. In total, the upper layers are generally cooled, while the lower layers are generally warmed. Noticeable, there is a vertical band structure in the

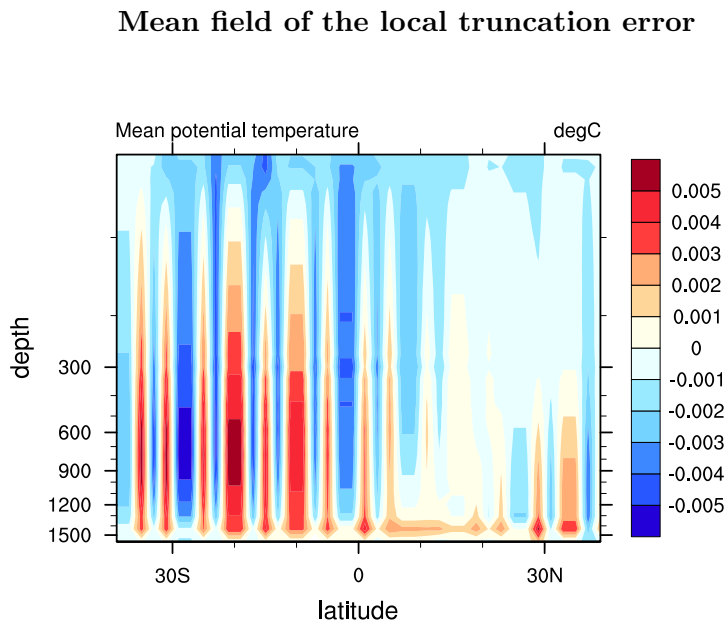


Figure 4.9: Estimates for the mean value of the local truncation error in the temperature variable for all dynamical flow regimes shown as a vertical cross-section. The channel is mapped to $[40S, 40N]$.

mean field. The persistence in the vertical supports the reasoning that this is not an artifact of a non-converged estimate affected by outliers but a robust result in the local truncation error.

The standard deviations $(diag\,(\Sigma))^{\frac{1}{2}}$ that are estimated from (4.34) are shown in figure 4.10. The values for the standard deviation differ by one order of magnitude, where the maximum values of each vertical column is located at a depth of 600 to 900 meters. The drastic reduction in standard deviation between latitudes $40S$ to $32S$ and $32N$ to $40N$ and their vicinities clearly indicates the temperature forcing by the 3D temperature relaxation to the initial conditions. This type of forcing acts the same at all resolutions and directly restricts the solution and thus the local truncation error. Above and below the band of high standard deviation between 600 and 900 meters, the standard deviation decreases again. The structure of the decrease in standard deviation above 600 meters depth roughly follows the mean density structure in the channel. Thus in this region, the local truncation error increases with density and follows the mean stratification contours. In the channel below 900 meters, the standard deviation in the local truncation error decreases again. It can be suspected that this decrease is again connected to constraints on the local truncation error in form of a forcing, as the flow in the lower layers is strongly influenced by the bottom boundary friction. Thus, in

Standard deviations of the local truncation error

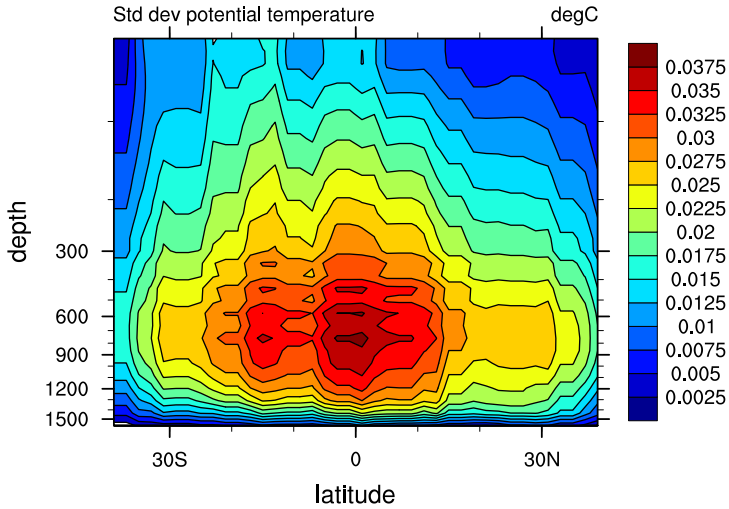


Figure 4.10: Estimates for the standard deviation of the local truncation error in the temperature variable for all dynamical flow regimes shown as a vertical cross-section. The channel is mapped to $[40S, 40N]$.

general, for our channel setup, the further the distance to the forcings, the larger the standard deviation of the local truncation error. Thus, the applied forcings seem to act as constraints on the variance of the local truncation error.

The Spatial Structure of the Local Truncation Error

The estimates of the correlation matrices for the local truncation error at the surface and in the vertical dimension are best discussed in the context of the resulting stochastic forcing that is created from the application of both. This is why we compare the spatial structure of an arbitrarily chosen local truncation error at one timestep to a realization of $X_{t_i}(\mathbf{p})$ for the surface layer and a vertical cross-section. Figure 4.11 shows the resulting fields. As discussed for the underlying probability distributions, the long tails of the distribution are neglected. Consequently, the shown values for the surface layer had to be capped to $\pm 0.048°C$, while the values actually range to $\pm 0.07°C$. For the vertical cross-section of the local truncation error, the actual values reach up to $\pm 0.08°C$. Nevertheless, the resemblance of the real and the modeled structures is obvious for both, the surface layer as well as in the vertical cross-section. For the surface layer, the spatial extend of positive and negative temperature anomalies and their general structure is very similar. For the local truncation error at the surface and

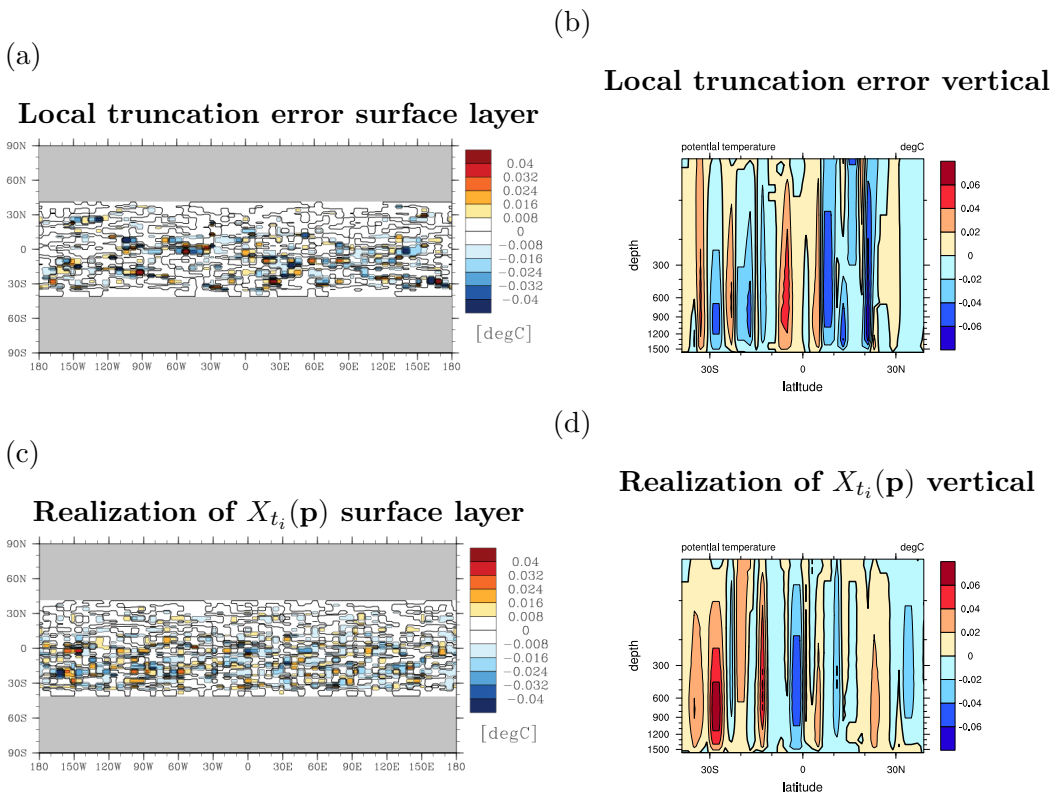


Figure 4.11: The local truncation error $N^i_\Delta\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)$ at an arbitrary timestep is shown at (a) the surface layer and (b) a vertical cross-section and compared to a selected realization of $X_{t_i}(\mathbf{p})$ for (c) the surface layer and (d) a vertical cross-section

its stochastic representation, the values increase towards the southern boundary of the channel. For the vertical cross-sections, the band-like structures in the local truncation error are well represented in shape as well as in strength. We again observe the general increase in the local truncation error in the depth range of 600 to 900 meters that is also found in its stochastic representation. Overall, the structure obtained by using the estimated correlation matrices $\hat{\mathbf{C}}_{surf}$ and $\hat{\mathbf{C}}_{v,\,\hat{t},\,Lat}$ is a good estimate for the structure of the local truncation error in the horizontal as well as in the vertical.

The Temporal Structure of the Local Truncation Error

To fully describe the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$, the modeled spatial structures of the local truncation error need to be connected in time. Our approach requires an estimate of the correlation coefficient δ that describes the connectivity of the local truncation error in time. The estimate for the coefficient δ from the near-initial learning phase is 0.99989. The value is robust for all dynamical flow regimes. The local

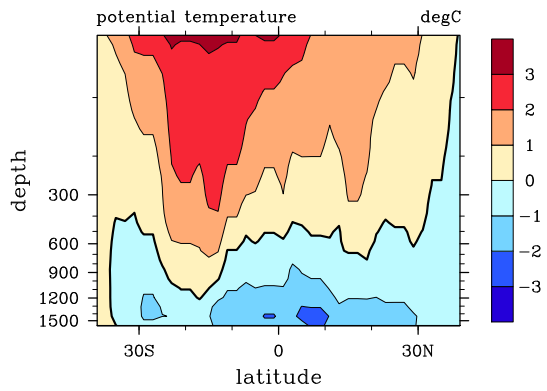
truncation errors at consecutive timesteps are highly correlated and are thus to a large part identical.

Although we have now estimated all required stochastic quantities, we find that we cannot create a model ensemble of perturbed goals to estimate the error in a goal with the stochastic forcing provided by the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$. The temperature values in the perturbed model simulations continuously grow beyond physically reasonable values after several hundreds of timesteps. We illustrate the problem by investigating the time-mean of an applied realization of the stochastic process and compare it to the change in vertical heat distribution due to the application of this temperature forcing, see figure 4.12. For the change in heat distribution in the channel, figure 4.12 a), we find that the upper layers of the channel warm strongly, while the lower levels cool by an equivalent amount. This heating respective cooling is directly connected to the temperature forcing by the stochastic process figure 4.12 b). The latitude positions with the strongest positive temperature forcing correspond to the latitude positions with the strongest warming of the channel, and we find the equivalent to be true for the latitudes with the strongest cooling. We note here that the time-averaged temperature forcing of the first 15 days of integration time and the forcing at the specific timestep 15 are indistinguishable, the same vertical columns are thus heated respectively cooled constantly at every timestep. The forcing is strongest at 600m to 900m depth, which is not surprising as it corresponds to the locations of maximum variance, see figure 4.10. The mechanism by which the excessive heat due to the temperature forcing is then transported from the lower levels to the surface is undoubtedly convection activated by the unstable stratification in the respective columns. Equivalently the same mechanism governs the cooling of the the bottom layer, when the negative temperature forcing cools the column with a minimum at 600m to 900m depth, this triggers convection with the deeper layers, cooling them in the process. In this way, a different circulation is created in the channel that dominates the former channel flow with mesoscale eddies.

On top of these general upper-level warming respective lower-level cooling in the channel, heating and cooling spikes are created in the flow that are a local phenomenon and are rarely occurring in space, figure 4.13. For the snapshot in figure 4.13 a), the most prominent of these locations is $65W, 5S$ where the temperature increases by more than $12°C$ within 15 timesteps. This can be tracked down directly to the time-averaged temperature forcing by the stochastic process in the same location, see figure 4.13 b). For these specific locations we always find a strong forcing with the same sign in its vicinity that is consistently at a high value over hundreds of timesteps. Equivalently, the cooling spikes in the bottom layer can reach up to $-5°C$. These singularities in heating respectively cooling are the direct causes for the crashes of the perturbed model simulations.

(a)

Zonally-averaged Potential Temperature



(b)

Time-averaged Realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$

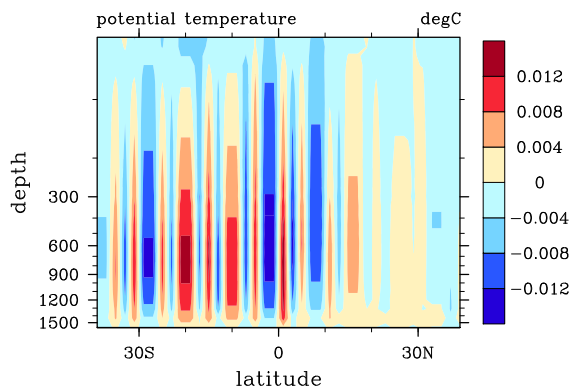


Figure 4.12: Vertical cross-sections of zonally-averaged potential temperature of the flow forced by a realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ after 15 days of integration time (a), and zonally-averaged time-mean of a realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ averaged over 15 days of integration time (b)

To better understand why an ensemble of perturbed model solutions with $\delta = 0.99989$ cannot be run successfully, we study the effect of the time correlation coefficient δ by running a series of experiments with lower correlation coefficients. The result is that the creation of a model ensemble of perturbed goals with the choices $\delta = 0.995$, $\delta = 0.990$, and even with $\delta = 0.950$ fails in the same way as has been described for $\delta = 0.99989$.

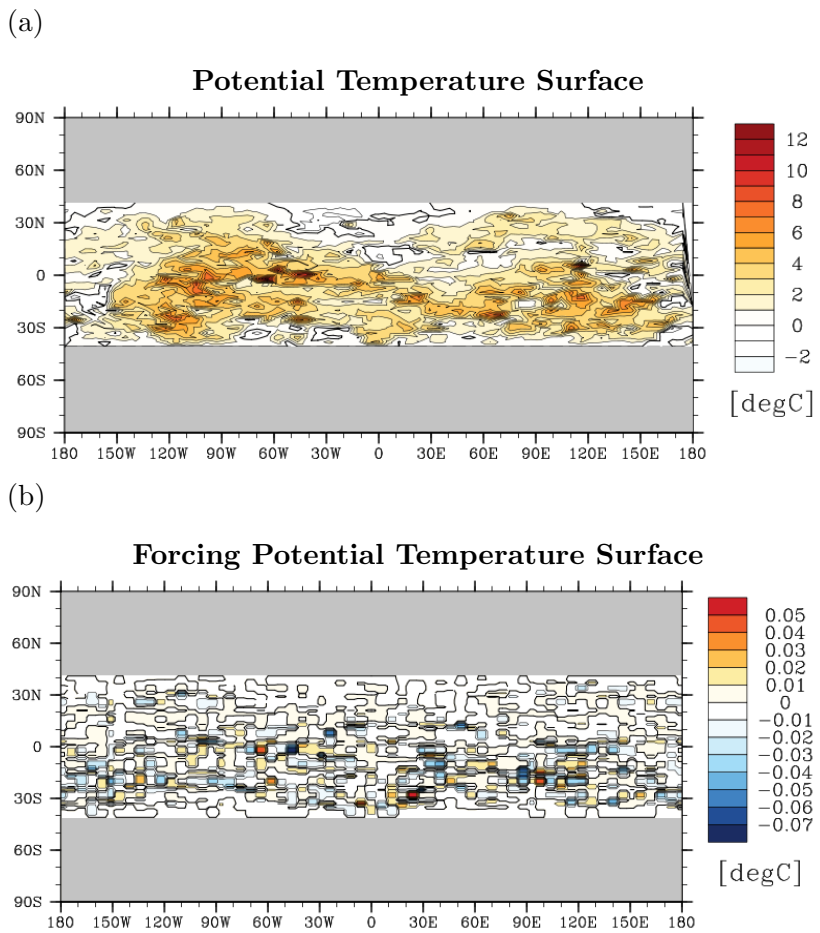


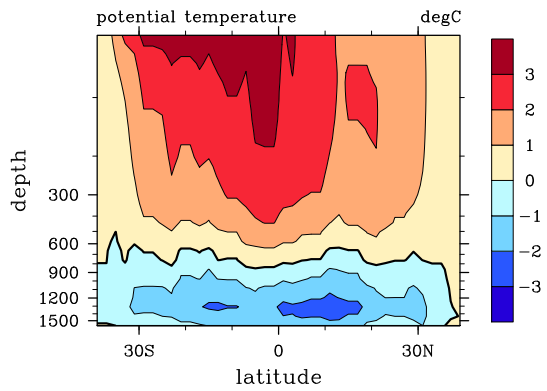
Figure 4.13: Snapshot for the change in surface temperature from timestep 1 to timestep 15 of the stochastically-forced flow (a), and the corresponding surface temperature forcing by the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ at the surface layer (b)

We find that we have to go as low as $\delta = 0.70$ to not produce too strong heating respectively cooling spikes in the ensemble members.

But although it is possible to stably run the perturbed ensemble with $\delta = 0.70$, the resulting ensemble still cannot be used for goal-error estimation purposes, because under this stochastic forcing we still obtain a channel flow with a significant heating in the upper layers and cooling in the lower layers, see figure 4.14. For the change in the vertical heat distribution with $\delta = 0.70$ (figure 4.14 a), we observe a similar distribution pattern as for $\delta = 0.99989$. The pattern is however more stably separated into the heating upper layers and the cooling lower levels at a constant depth of $600m$ to $900m$. Also the change in heat distribution occurs 4 times slower for $\delta = 0.70$ than for $\delta = 0.99989$. In contrast to the results found for $\delta = 0.99989$, this heating pattern is not

(a)

Zonally-averaged Potential Temperature for $\delta = 0.7$



(b)

Time-averaged Realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ for $\delta = 0.7$

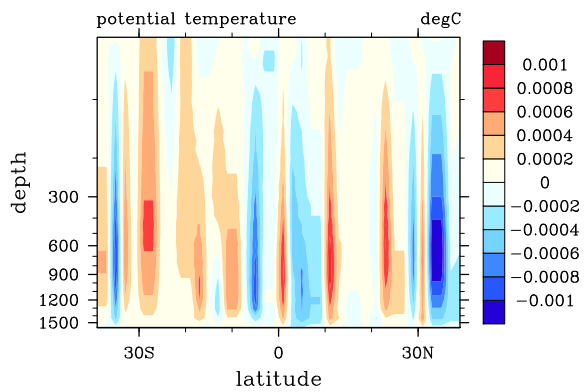


Figure 4.14: Vertical cross-sections of zonally-averaged potential temperature of the flow forced by a realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ with $\delta = 0.7$ after 60 days of integration time (a), and zonally-averaged time-mean of a realization of $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ with $\delta = 0.7$ averaged over the first 60 days of integration time (b)

directly reflected in the time-averaged temperature forcing by the stochastic process (figure 4.14 b) anymore. Also, the time-averaged forcing patterns are found to be one order of magnitude weaker. This reduction in the mean strength is due to stronger averaging effects for $\delta = 0.70$ as the temperature forcings at specific timesteps are not as connected in time anymore as for $\delta = 0.99989$. Thus, although the temperature forcing

is now changing more frequently in time, we sill obtain regions where convection occurs that transports heat from the layers at a depth of $600m$ to $900m$ to the surface or cools the levels deeper than $600m$ to $900m$. As the temperature forcing changes faster in time for $\delta = 0.70$ than for $\delta = 0.99989$, the resulting heating respectively cooling due to convection is now more evenly distributed over all latitudes. The reason why the ensemble with $\delta = 0.70$ can be run at all is that the stochastic forcing changes fast enough to not produce significant heating or cooling spikes.

In summary, although a useful perturbed ensemble cannot be run for $\delta = 0.70$ as well, we have seen that reducing the connectivity of the temperature forcings in time makes the resulting flow behave more reasonable as the undesired heating respectively cooling due to convection as well as the heating and cooling spikes are reduced. At the same time, however, the parameter δ has to faithfully represent physical phenomena that are missing in the 40km resolution model solution and it should therefore not be changed arbitrarily.

Our results thus strongly indicate that the problem lies in the time-evolution of the local truncation error compared to the time-evolution of our stochastic process. To study the differences in the temporal evolutions, we analyze the evolution of the time-lagged correlation coefficients of the local truncation error for different lag times. If the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ is a suitable model for the local truncation error, the time-evolutions of their time-lagged correlation coefficients should coincide. The time-lagged correlation coefficients of the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ for a parameter δ can be obtained analytically, as it is the correlation coefficient of $X_{t_1}(\mathbf{p})$ that is obtained from multiple applications of the stochastic timestepping scheme (4.16). It can then be seen that the random vector $X_{t_i}(\mathbf{p})$ at timestep i and the initial random vector $X_{t_1}(\mathbf{p})$ are correlated by the correlation coefficient $\delta_{lag}(i) := \delta^i$.

The values of the function $\delta_{lag}(i)$ for different choices of δ are compared to the time-lagged correlation coefficient of the local truncation error in figure 4.15 with a maximum time-lag of 14 days of integration time. The time-lagged correlation coefficient for the local truncation error drops off immediately and reaches zero after a lag time of 270 timesteps which is about 4 days of integration time. Afterwards, the time-lagged correlation coefficient is negative, with a minimum of -0.074 at a lag time of 390 timesteps and a value of -0.029 for a lag time of 14 days. If compared to the lag-time correlation coefficient of our stochastic process, it becomes clear that the choice $\delta = 0.99989$ is not suitable as it completely overestimates the correlations in time for longer time lags. We find that $\delta = 0.995$ or $\delta = 0.990$ are a better fit. Both, however, drop off too quickly for small lag times of up to 150 to 200 timesteps but then decrease too slowly for large time lags. The time-lagged correlation coefficient for $\delta = 0.7$ immediately drops off to 0.01 after only 13 timesteps. Additionally, the time-lagged correlation coefficient never become negative as our stochastic timestepping scheme (4.16) does not provide negative correlations by construction.

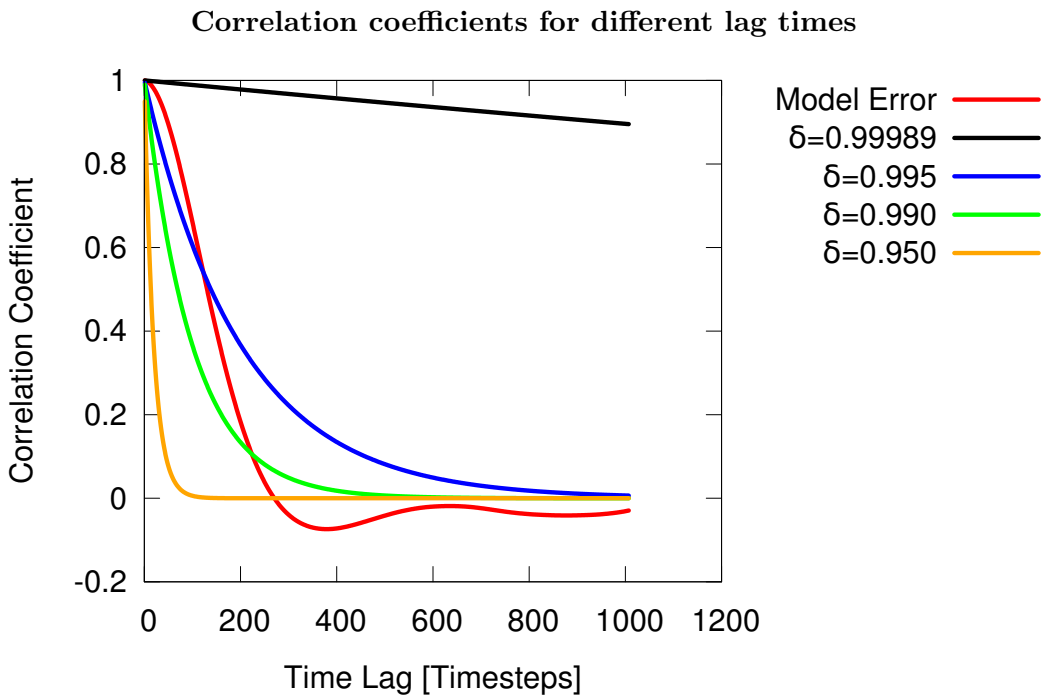


Figure 4.15: Time-lagged correlation coefficients are shown for the actual local truncation error (red) and the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ with different choices of δ, $\delta = 0.99989$ (black), $\delta = 0.995$ (blue), $\delta = 0.990$ (green), $\delta = 0.950$ (orange)

We need to get an idea about the connection between the time-lagged correlation coefficients of the local truncation error and the occurring convection in a vertical column of the channel when we apply our stochastic temperature forcing. Towards this aim, we provide a rough analytical estimate for the total heat applied to the vertical column by a constant temperature forcing within a given timespan. More specifically, we are interested in the needed timespan for the deep layers to have warmed to the temperature of the surface layer. This would give us a rough timescale of when convection can be expected to occur. As an example, we orient along a vertical column at latitude $10S$ as seen in figure 4.2. The typical temperature values at the surface are about $15°C$, the potential temperature at a depth of 900 meters is about $7°C$, and a high but realistic difference in temperature forcing between the surface and 900 meters depth is about $0.08°C$ per timestep. From these numbers we can get the estimate for the timescales at which the deep levels at 900 meters would have warmed to the value of the surface temperature of about 100 timesteps. Of course, here we have neglected the effects of heat advection and diffusion, also the temperature forcing will not be constantly at this high level. But these effects will just make this timescale longer, and thus the important fact is that the timescale for convection in the channel due to the

temperature forcing is of the order of hundreds of timesteps.

The predicted order of magnitude of a hundred timesteps fits well to the time-lagged correlation coefficients of the local truncation error becoming negative after 270 timesteps. The negative time-lagged correlation coefficients thus seem to indicate the timescale when a specific correction of the heat distribution in a vertical column by the local truncation error typically ends or even needs to be slightly negated. Further heating or cooling of the vertical column would eventually result in the occurrence of convection, which occurs on a similar timescale. It is thus important that a stochastic process for the local truncation error can precisely model the time-evolution of the local truncation error in order to avoid the timescale on which convection would eventually occur. We argue that this could be possible by incorporating the negative time-lagged correlation coefficients into the stochastic process. With this, the stochastic process would act similarly to the shown stochastic process with reduced $\delta = 0.7$, but the effects of negative correlations would be stronger and thus more directed towards actually hindering further forcing of the same kind and as a result should prevent the occurance of convection. Our stochastic process however lacks this feature of negative time-correlations and needs to be extended.

Stochastic Timestepping Revisited

Based on our results, we propose a revised time-handling for the stochastic process that is capable of incorporating the history of the stochastic process in the form of, possibly negative, time-lagged correlation coefficient.

We assume that we have knowledge about a timeseries of realizations $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ of the correct stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$. Our aim is to define the entries of the vector $x_{t_i}(\mathbf{p})$ at timestep i by the corresponding entries of the vectors $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ and a correlation matrix $\mathbf{C}_{time}$. We again assume that we already obtained a base realization $y_{t_i}(\mathbf{p})$ of $x_{t_i}(\mathbf{p})$ that satisfies the spatial structure of the local truncation error. For simplicity, we assume that the vectors $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ have already been normalized by subtracting the means $\hat{\mu}_k$ and division by $\hat{\sigma}_k$.

The sample correlation matrix $\hat{\mathbf{C}}_{time} \in \mathbb{R}^{q\times q}$ is defined by its entries at the matrix positions with rows t_1 and columns t_2

$$\begin{aligned}\hat{\mathbf{C}}_{time}(t_1,t_2) :=& \frac{1}{992\cdot 62\cdot (n_l-q)-1}\sum_{i=1}^{n_l-q}\sum_{k=1}^{992}\sum_{j=0}^{61}\left(N_\Delta^{i+t_1}\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k,\,j}-\hat{\mu}_k\right)\\ &\left(N_\Delta^{i+t_2}\left(P_\Delta\left(q_{\Delta,\,10km}\right)\right)_{k,\,j}-\hat{\mu}_k\right)/\sqrt{\hat{\sigma}_k\hat{\sigma}_k},\end{aligned} \tag{4.38}$$

where t_1 and t_2 denote the time lags $t_1,\ t_2 \in \{1,\ldots,q\}$. Because of the special construction, the time-lagged correlations between future timestep i and all past timesteps $i-j,\ j=1,\ldots,q$ are found in the last row of $\hat{\mathbf{C}}_{time}$. With this, we have included the

possibility that some of the time-lagged correlation coefficients might become negative.

For the stochastic timestepping, we first perform a Cholesky decomposition of $\hat{\mathbf{C}}_{time}$ into

$$\mathbf{L}_{time}\mathbf{L}_{time}^T = \hat{\mathbf{C}}_{time} \tag{4.39}$$

The special property of the last row of $\mathbf{L}_{time}$ is now that it provides us with a relation of the entries of the vector $x_{t_i}(\mathbf{p})$ to the corresponding entries of the previous q states $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$.

However, the entries of $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ are already correlated in time and thus cannot be used directly for the estimation of the entries of $x_{t_i}(\mathbf{p})$. This can be illustrated if the whole system is written in matrix notation for the j-th entry of the vectors $x_{t_i}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$

$$\mathbf{L}_{time}\begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \\ y_{t_i}(\mathbf{p})_j \end{pmatrix}_{uncorrelated} = \begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \\ x_{t_i}(\mathbf{p})_j \end{pmatrix}, \tag{4.40}$$

where the vector on the left-hand side with subscript *uncorrelated* is the decorrelated vector of the vector on the right hand side, and it is obvious that it is required for the calculation of $x_{t_{i-1}}(\mathbf{p})_j$.

The entries of the vector can be decorrelated by application of their so-called whitening matrix. The required information for the whitening is already present in the lower triangular matrix $\mathbf{L}_{time}$. More specifically, the quadratic sub-matrix given by the first $q-1$ rows and the first $q-1$ columns, which we denote by $\mathbf{L}_{white}$ with

$$\mathbf{L}_{white}\begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \end{pmatrix}_{uncorrelated} = \begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \end{pmatrix}, \tag{4.41}$$

needs to be inverted. Insertion of $\mathbf{L}_{white}^{-1}$ results in

$$\mathbf{L}_{time}\begin{pmatrix} \mathbf{L}_{white}^{-1}\begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \end{pmatrix} \\ y_{t_i}(\mathbf{p})_j \end{pmatrix} = \begin{pmatrix} x_{t_{i-q}}(\mathbf{p})_j \\ \vdots \\ x_{t_{i-1}}(\mathbf{p})_j \\ x_{t_i}(\mathbf{p})_j \end{pmatrix}, \tag{4.42}$$

which yields the vector $x_{t_i}(\mathbf{p})$ that is correlated to the timeseries $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ according to the time-lagged correlation coefficients until timestep $i-q$.

The new realization $x_{t_i}(\mathbf{p})$ satisfies the spatial structure of the local truncation error,

because it is a linear combination of whitened vectors of $x_{t_{i-1}}(\mathbf{p}), \ldots, x_{t_{i-q}}(\mathbf{p})$ that each satisfy the spatial structure, which can proven by multiple applications of (4.17). The matrix $\mathbf{L}_{white}^{-1}$ does not need to be calculated explicitly, because $\mathbf{L}_{white}$ is a lower triangular matrix. However it should be kept in mind that due to the high lagged-time correlation coefficients the values on the main diagonal of $\mathbf{L}_{white}$ might become small, which would mean that the condition number becomes large.

We can connect the revised stochastic timestepping to the original timestepping approach (4.16). We find that the approach (4.16) is the special case of choosing $q = 1$. For this choice, the whitening matrix results in $\mathbf{L}_{white}^{-1} = 1$ and the lower triangular matrix $\mathbf{L}_{time}$ is

$$\mathbf{L}_{time} = \begin{pmatrix} 1 & 0 \\ \delta & \sqrt{1-\delta^2} \end{pmatrix}, \tag{4.43}$$

which yields (4.16). Our revised stochastic timestepping approach thus includes our original approach (4.16) and is its generalization.

4.5.3 Discussion

Our approach to deriving a stochastic process for the local truncation error in an eddying regime is based on stochastically describing the full local truncation error. In comparison, in chapter 3 the local truncation error is represented by a time-invariant bias plus a temporal fluctuation. Considering the temporal evolution of the time-lagged correlation coefficients we find this change of the stochastic process to be a reasonable choice. This is because in the channel flow, the local truncation errors at different timesteps are only correlated for a limited amount of time before they are uncorrelated and eventually even negatively correlated. In this regime, estimated bias information would be mostly lost after just 300 timesteps because the time-lagged correlations have reached zero. At the same time, the temporal fluctuations would have to be very large to capture the large changes in the local truncation error in time. The approach from chapter 3 is thus not appropriate.

We find that for our channel experiment we can calculate reasonable local truncation error information from just the temperature variable. The other fields follow the prescribed temperature structure as shown for the case of the surface elevation η whose horizontal gradient is connected to the resulting velocity field vn. In the context of parametrizing these local truncation errors by a stochastic process, this is however not the only criterion by which to judge the local truncation error. If we have a closer look at the resulting order of magnitude for the local truncation error, we find that the values can be up to $0.1°C$, which represents a huge amount of change in a temperature value within a single timestep. Given that a stochastic representation is never going to be

perfect, we cannot expect the associated stochastic forcing to always be fully applied in the correct regions and at the right times even in a best case scenario. This can cause or enhance problems like the observed heating respectively cooling in the model solution we have described for our estimated stochastic process. It might thus be advisable to again include other model variables in the calculation of the local truncation error to separate the single large temperature forcing into several smaller forcings. This would distribute the risk as each forcing alone would be less likely to produce effects like the strong heating and cooling, even if it were applied slightly incorrectly.

The distribution of the temperature-based local truncation error is chosen to be represented by Gaussian distributed random variables. A Laplace-distributed random variable seems to fit slightly better because of the longer tails of the distribution, but the choice of the probability distribution is not as clear as for the probability distributions encountered in chapter 3. This could be related to the differences in experiments. While the channel experiment is dominated by the sloped density field and its resulting mesoscale-eddy field, the experiments in chapter 3 are dominated by lateral boundaries and their associated boundary currents. The presence of strong boundary currents seems to enhance a deviation from Gaussianity of the local truncation error towards longer and significant tails in the probability distribution. We here mention significance because in this argument we have so far neglected the very long tails in the local truncation error for the channel experiment. These tails however contain almost no probability and could be a property of the temperature variable itself and might not even be directly connected to the flow.

We find that the spatial structure of the local truncation error can be successfully modeled by our approach to separate the horizontal structure from the vertical structure and only let both structures communicate at the surface layer. The required correlation matrices are robustly estimated for the channel experiment by making use of the symmetric structure of the channel that we exploit here. The answer to the question why it is sufficient to only link the vertical and the horizontal structure at the surface layer can be found in the extend of the vertical structures. In accordance to the local truncation error, the emerging vertical bands extend almost over the entire vertical column. This indicates that the local truncation error represents missing barotropic dynamics, which we find is in accordance with the observed vertical density structure shown in figure 4.2 where densitiy anomalies extend to the entire column depth. Consequently, all correlation coefficients throughout the vertical column must be high, which in our approach is then equivalent to a strong information transfer from the surface layer to the deep channel layers. Only linking the structures at the surface layer is thus sufficient because the surface layer already explains the entire vertical column for the most part. In general the surface layer might however be a bad choice as the ocean mixed layer has significantly different flow properties compared to the deeper ocean. The question of how to generally choose the horizontal layer at which to link the vertical and the

horizontal structure for a given experiment is thus the task of finding the horizontal layer that maximizes the information throughput into the vertical column.

The combination of the spatial and the temporal structure of the local truncation error cannot be performed by a single temporal correlation coefficient. The time evolution of the stochastic process needs to be extended to respect the entire temporal evolution of the time-lagged correlations. These correlations extend back several hundreds of timesteps but are essential in learning the restrictions the physical system imposes on the time-evolution of the local truncation error. In chapter 3, we restricted the temporal evolution of the stochastic process by an upper bound on the correlation coefficient. Now, the required stochastic process needs to be restricted in time in a very particular way. Possibly, the correlations in time could be modeled by using the proposed correlation matrix for the temporal correlations $\mathbf{C}_{time}$. Regardless of the extension to the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$, it would of course need to be build in a way that our formulated requirements, such as the spatial structure at every timestep i, are still satisfied.

Assuming that such a stochastic process could be constructed for the channel experiment and its parameters could be estimated from a near-initial learning phase, its realizations would be very strongly confined in time by the shear amount of involved correlation coefficients in space and time. The problem we see with these constraints is that it further reduces the amount of stochasticity of the stochastic process, a property that has made our stochastic approach to error estimation that appealing. The idea has always been to replace the difficult to impossible problem of correcting an under-resolved physical system of high-complexity by an easy to get, low-complexity stochastic representation that is computationally cheap. Every further restriction that we have to impose on the stochastic process shifts the stochastic representation towards a higher level of complexity. Higher complexity of the stochastic process in turn further decreases its robustness to changes in its parameters as potentially every small change in one of the key parameter estimates that constrains the stochastic process might make the stochastic process unusable. This is already visible for the time-lagged correlation coefficients as the occurring negative correlation coefficient of -0.07 might be considered small, but if one were not able to robustly estimate it or had to neglect it, the stochastic process seems to not be applicable in terms of our error estimation method. The same could possibly hold true for a large number of parameters of the stochastic process that would not fit anymore to a however slightly changed flow. This could result in a non-robust stochastic representation of the local truncation error.

4.6 Conclusions

We find that for an eddying flow in an ocean model, the concept of the local truncation error needs to be broadened, which we have successfully performed by a temperature-

based local truncation error. A stochastic representation of the local truncation error is obtained for the local truncation error's spatial structure, but a stochastic representation of its temporal evolution proves challenging. To represent the temporal evolution in a stochastic model for an eddying flow regime, we argue that the history of time-lagged correlation coefficients needs to be included. Only then can the stochastic process be used in the ensemble-based stochastic error estimator. A loophole could be provided by the adjoint-based error estimator (4.5) because it does not require a stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$ that is physically-correct in all aspects. However, we would need the discrete adjoint solution q^*_Δ to do so.

We have shown that the zonally-symmetric structure of the channel can vastly increase the number of observations for the parameter estimation process. Additionally, the symmetry provides an easy and natural way to choose the separation into different dynamical flow regimes. Vice versa, for flows where the experimental setup does not provide these symmetries, a far lower amount of observations would be available which would pose difficulties for the estimation of the horizontal correlation matrix $\mathbf{C}_{surf}$. In the absence of these exploitable symmetries, the obvious choice is then to further increase the length of the near-initial learning phase, which however would further increases the computational cost, making the application of a stochastic error estimator less profitable.

From a comparison to the stochastic process used in chapter 3, we conclude that the choice of the stochastic process is flow-dependent. For a given flow, one can imagine that even different regions of the same flow might require different types of stochastic processes, which would introduce a new, additional algorithmic step to to the Goal Error Ensemble Method algorithm 3.4. The choice which specific stochastic process is selected for a specific region will then have to be performed in a meta-level dynamical flow regime separation.

The additional algorithmic step and the overall increase in complexity of the stochastic process increasingly requires decisions based on educated guesses by the user. For 3D ocean experiments, such as our channel experiment, the user needs an in-depth understanding of the studied experiment and needs to develop a good idea of what the present ocean dynamics mean in terms of the local truncation error and how this would be translated into a stochastic representation. We observe that the initial idea of a learning algorithm for the local truncation error that does the learning task to a large part independently, as intended by [41], turns more and more into a serious stochastic modeling undertaking.

Independent of how the temporal evolution will finally be implemented in the stochastic process $\{X_{t_i}(\mathbf{p})\}_{i\in\mathbb{N}}$, the memory term of the Mori-Zwanzig formalism will be much more delicate than the fully time-correlated one introduced in chapter 3. Compared to the stochastic process from chapter 3, the new stochastic process will typically result in a smaller spread for comparable tasks. However, it will be dependent on a large

amount of correlation coefficients in time as well as in space that keep the stochastic process in this tightly constrained state. The tradeoff is thus the discussed increase in complexity which threatens the robustness of the resulting stochastic process and eventually the robustness of the targeted stochastic error estimator.

The increased complexity of the stochastic process stems from the complexity of the structure of the local truncation error in time and space. The local truncation error is just a measure for the difference between two deterministic model solutions with differently resolved dynamics. The local truncation error is thus inherently deterministic and follows its own dynamics, the combined dynamics of both models at different resolutions. The main features of these deterministic local truncation error dynamics then have to be matched by the stochastic process. For a dynamically-rich but still idealized experiment such as our channel flow, we already expect to have to constrain the stochastic process by several hundred of its past states. Given the large temporal scales of the dynamics in the real ocean, we expect a further increase in complexity for the stochastic representation of the local truncation error .

Chapter 5

Conclusions and Outlook

5.1 The Quintessence

We have successfully brought the concept of goal-oriented error estimation into the field of ocean modeling.

For flows with lateral boundaries in a shallow-water environment, we have shown that our developed stochastic error estimator works reliably for a selection of ocean-type experiments that were designed to verify the error estimator on essentials of two-dimensional large-scale ocean dynamics.

For the baroclinic instability in the 3D ocean model environment, the results do not yet allow statements on goal-oriented error estimation specifically. What we can conclude is that we have successfully understood and modeled many pieces of the puzzle that is the stochastic component of our error estimator. The current results clearly indicate a missing piece to be in the correct propagation of the stochastic component in time, i.e. the memory of the stochastic process.

For all ocean experiments we have found the need to instil the concept of memory into the stochastic component of our error estimator. The memory's curve of complexity, starting from no memory in [41], over a memory induced by an upper bound on temporal correlations in chapter 3, to a memory represented by the interplay of thousands of correlation coefficients in time and in space in chapter 4, has still not reached its limit and could potentially grow further. From this observation we conclude that the correct handling of memory effects is going to be the linchpin for all future work on this type of stochastic dual-weighted error estimation methods.

In the context of our stochastic error estimator, it is important to keep in mind that the local truncation error in a grid element at a specific timestep is a deterministic measure between two determinist model solutions. However, we continuously find that, within a dynamical flow regime, its characterization as a stochastic quantity is justified. This has held true for all investigated variables, for non-eddying as well as eddying flows, and even under the influence of additional sub-grid scale parametrizations.

5.2 The Answers to the Research Questions

We here provide the answers to the questions we posed in the introductory part of this thesis in section 2.4.

- What probability distribution can be used to model the local truncation error?

  We have found that a Laplace or a Gaussian distribution fit well to the distribution of the local truncation error for all flows we encountered. If possible, we would favor a Laplace distributed random variable, because the long tails of the local truncation error distribution are better represented by a it. In general, we can however confirm that the choice from [41] who used a Gaussian distribution for their experiments can be justified for a wide range of ocean-type experiments.

- What is the relation between the stochastic error estimator (2.22) and a classical stochastic-physics ensemble?

  The stochastic dual-weighted error estimator is proven to be a linearized stochastic-physics ensemble, and both are shown to be interchangeable. As of yet, we have not found the limits of said connection, but it is unquestionable that increasing the non-linearity of the studied flows will eventually lead to the connection's severance.

- How big is the effect of the discretization of the goal itself compared to the error in the goal ϵ?

  For our results with the shallow-water model, we have encountered cases where the error due to the discretization of the goal itself is not negligible. We expect to come across similar cases for other experiments and goals, because the issue is connected to under-resolution of the model solutions, a problem that is constantly encountered in ocean modeling.

- What part of the stochastic process represents the memory term that is predicted by the Mori-Zwanzig formalism?

  In the shallow-water framework as well as in the full 3D Ocean model, the memory is initially represented by the first local truncation error, and then governed by temporal correlations. The strength of the temporal correlations and the length by which they reach into the past defines the memory. The effect of memory can be illustrated for the initial states in the different model frameworks. While in the shallow-water experiments the employed fully-time correlation ensures that the initial state is never erased from the memory, in an eddying flow it is necessary that the initial state is rapidly forgotten to be in accordance with the temporal evolution of the local truncation error.

- How can spatial and temporal correlations be represented in the stochastic process? And, can these correlations be estimated reliably?

  Spatial and temporal correlations in the form of correlation coefficients are at the core of the stochastic component of our stochastic error estimator. In chapter 3, it was sufficient to make use of correlation coefficients in the form of an upper bound on the temporal correlations. For eddying, three-dimensional flows such as in chapter 4, we find that not even a collection of multiple correlation matrices in space and a correlation coefficient in time is sufficient to model the temporal evolution of the local truncation error. Based on our analysis, our proposed extension revolves around an additional correlation matrix in time. These matrices need to be constructed and intertwined in a special way to make the evolving structures coherent.

- Does this approach lead us to valid error estimates for physically relevant goals?

  Yes, for ocean-type experiments in a shallow-water framework, our approach leads to reliable estimates for bounds on the error in a goal in the form of confidence intervals. For the full 3D ocean model, we have not yet finalized the stochastic component of the local truncation error, and are thus not yet at the stage where we can obtain goal error estimates.

- Does the concept of the local truncation error even make sense in an environment where the local truncation error primarily results from only partly resolving the mesoscale eddy field?

  We find that the straight-forward definition of the local truncation error is not applicable directly. We formulate a temperature-based local truncation error formulation as an alternative and provide numerical evidence that this is a reasonable choice for an eddying flow that is driven by density differences.

- How can the vertical structure, especially stratification, of the ocean experiment be preserved under the influence of the stochastic process?

  If there is a vertical structure present in the local truncation error which the stochastic component needs to adhere to, it will be reflected in the vertical correlation coefficients, a quantity that is incorporated in the stochastic component and shown to yield the desired vertical patterns. We have found that the estimated vertical patterns need to be correctly propagated in time to not introduce unwanted effects, such as constantly triggering convection in the vertical column.

- Is it possible to use a stochastic process to model the local truncation error that is due to not fully resolving the mesoscale eddy field?

  Until now, we can state that we have understood much of its spatio-temporal

structure, and based on our analysis, we have been able to model its spatial structure. For the temporal structure we have identified a key mechanism for which we have then proposed a stochastic model. Thus, we are confident that a stochastic process can be developed for this flow regime.

5.3 Concluding Remarks & Outlook

The described increase in the curve of complexity that we find for our results raises the question whether our stochastic dual-weighted error estimator is reasonably applicable in experiments of even higher complexity. sIn this, we think about realistic ocean simulations where phenomena have timescales of years, decades and even higher, whereas in comparison the lifetime of the mesoscale eddies studied in this thesis is in the range of days to weeks. A long timeseries of time-lagged correlation coefficients might be required, i.e. a long history past states of the stochastic component. This would mean a hugely prolonged near-initial learning phase and a strongly constrained stochastic process, which is not desirable due to the computational cost and due to a possibly less robust stochastic component because of the large amount of dependencies.

On this note, we feel that we, in the role of the algorithm user, have to perform very much of the model's work in the complex three-dimensional environment. We basically have to teach the stochastic component the most important aspects of the underlying model physics, i.e. prescribe how to learn the patterns present in the local truncation error. For an eddying flow, the physics we have to teach the stochastic process is how the patterns of the unresolved part of the mesoscale eddy field look like and how they develop over time. Because this step happens outside of the physics-governed environment of the ocean model, the major features of this unresolved part have to be identified by us and prescribed in a physically sound way. Sound representations of physical systems have however never been the strong side of stochastic modeling.

Thus, we argue that we need to find ways to outsource some of this stochastic modeling work back into the domain of the physical model, i.e. the ocean model. Our motivation that this step is promising stems from our channel experiment described in chapter 4, which shows that the employed model resolutions (40km resolution and reference 10km resolution) are actually quite similar in terms of the dynamics. The coarse resolution model solution is just the high-resolution solution's filtered version in terms of the mesoscale eddy field, which in turn is then reflected in the slightly reduced impact on the vertical heat distribution. In summary, it can be said that the coarse model solution performs quite well and shows the correct dynamics. Yet, the associated local truncation error indicates large point-wise errors per timestep, presumably because of errors in the actual shape of the eddies or their correct positioning. Thus, instead of enhancing the already available flow and work with the available model dynamics, the local truncation error to a large part rewrites the variable fields at every timestep to

make them look exactly like the high-resolution solution's fields. We however argue that this correction towards point-wise identity of the fields is not of much relevance here, because goals in an ocean modeling context are typically averaged quantities over large areas or timescales. This means that the goals can be interpreted as a function of the probability distribution of the underlying fields. In terms of these goals, a successful correction of a flow fields would already be achieved if their probability distributions were corrected instead of their point-wise values. We thus propose the investigation of stochastic components for our error estimator that do not take the detour over the point-wise local truncation error which we then describe stochastically anyway, but that instead try to quantify the error in the probability distribution in a dynamical flow regime per timestep and directly correct the probability distributions of the fields.

Another research path that might be worth pursuing is the investigation of our stochastic dual-weighted approach for other dual-weighted methods than the one described in [15]. This step might not simplify the stochastic modeling part of our error estimator but could make the estimation of the parameters of the stochastic process easier. In this, we think about the dual-weighted approach described in [11] that formulates the error estimator from the point of view of the high-resolution grid instead of the coarse resolution grid. Instead of a mapping P_Δ from the high-resolution grid to the coarse grid, we would have to define a reconstruction R_{high}, which maps fields from the coarse grid to the high-resolution grid. In this approach, the Taylor expansions (2.7) and (2.10) are both formulated in terms of the high-resolution goal J_{high} and the high-resolution discrete model $N_{high}\left(q_{high}\right)$. The residual information is then obtained by inserting the coarse resolution solution q_Δ into the high-resolution operator N_{high}

$$N_{high}\left(R_{high}(q_\Delta)\right).$$

However, the disadvantage of this approach is the reconstruction R_{high} of a typically under-resolved coarse model solution q_Δ, which will probably come with a high interpolation error. How big the effect of interpolation is for ocean model experiment, should be tested and quantified.

Bibliography

[1] Babuška, I., and W. C. Rheinboldt, 1978: A-posteriori error estimates for the finite element method. *International Journal for Numerical Methods in Engineering*, **12 (10)**, 1597–1615.

[2] Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, **139 (674)**, 1132–1161, doi:10.1002/qj.2063, URL http://dx.doi.org/10.1002/qj.2063.

[3] Becker, R., and R. Rannacher, 2001: An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica 2001*, **10**, 1–102.

[4] Berner, J., G. Shutts, M. Leutbecher, and T. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *Journal of the Atmospheric Sciences*, **66 (3)**, 603–626.

[5] Berner, J., and Coauthors, 2015: Stochastic parameterization: Towards a new view of weather and climate models. *arXiv preprint arXiv:1510.08682*.

[6] Bonaventura, L., and T. Ringler, 2005: Analysis of discrete shallow-water models on geodesic Delaunay grids with C-type staggering. *Monthly Weather Review*, **133 (8)**, 2351–2373.

[7] Buizza, R., M. Milleer, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **125 (560)**, 2887–2908.

[8] Chorin, A. J., and O. H. Hald, 2013: Estimating the uncertainty in underresolved nonlinear dynamics. *Mathematics and Mechanics of Solids*, 1081286513505465.

[9] Corliss, G., C. Faure, A. Griewank, L. Hascoet, and U. Naumann, 2013: *Automatic Differentiation of Algorithms: From Simulation to Optimization*. Springer Science & Business Media.

[10] Danilov, S., 2012: Two finite-volume unstructured mesh models for large-scale ocean modeling. *Ocean Modelling*, **47**, 14–25.

[11] Fidkowski, K. J., and D. L. Darmofal, 2011: Review of output-based error estimation and mesh adaptation in computational fluid dynamics. *AIAA journal*, **49 (4)**, 673–694.

[12] Flato, G., and Coauthors, 2013: Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, book section 9, 741–866, doi: 10.1017/CBO9781107415324.020, URL www.climatechange2013.org.

[13] Fraysse, F., J. De Vicente, and E. Valero, 2012: The estimation of truncation error by τ-estimation revisited. *Journal of Computational Physics*, **231 (9)**, 3457–3482.

[14] Gent, P. R., and J. C. Mcwilliams, 1990: Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, **20 (1)**, 150–155.

[15] Giles, M., 1998: On adjoint equations for error analysis and optimal grid adaptation in CFD. *Frontiers of Computational Fluid Dynamics 1998*, World Scientific Pub Co Inc, 155–169.

[16] Giles, M. B., and N. A. Pierce, 2003: Adjoint error correction for integral outputs. *Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics*, T. Barth, and H. Deconinck, Eds., Lecture Notes in Computational Science and Engineering, Vol. 25, Springer Berlin Heidelberg, 47–95, doi:10.1007/978-3-662-05189-4_2, URL http://dx.doi.org/10.1007/978-3-662-05189-4_2.

[17] Giles, M. B., and E. Süli, 2002: Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, **11**, 145–236.

[18] Givon, D., R. Kupferman, and A. Stuart, 2004: Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, **17 (6)**, R55.

[19] Grätsch, T., and K.-J. Bathe, 2005: Review: A posteriori error estimation techniques in practical finite element analysis. *Comput. Struct.*, **83 (4-5)**, 235–265, doi:10.1016/j.compstruc.2004.08.011, URL http://dx.doi.org/10.1016/j.compstruc.2004.08.011.

[20] Griewank, A., and A. Walther, 2008: *Evaluating derivatives: principles and techniques of algorithmic differentiation*. Siam.

[21] Griffies, S. M., and A. J. Adcroft, 2008: Formulating the equations of ocean models. *Ocean Modeling in an Eddying Regime*, 281–317.

[22] Griffies, S. M., and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, **28 (3)**, 952–977.

[23] Hecht, M. W., M. R. Petersen, B. A. Wingate, E. Hunke, and M. Maltrud, 2013: Lateral mixing in the eddying regime and a new broad-ranging formulation. *Ocean Modeling in an Eddying Regime*, American Geophysical Union, 339–352, doi:10.1029/177GM20, URL http://dx.doi.org/10.1029/177GM20.

[24] Jayne, S. R., and J. Marotzke, 2002: The oceanic eddy heat transport*. *Journal of Physical Oceanography*, **32 (12)**, 3328–3345.

[25] Kalnay, E., 2003: *Atmospheric modeling, data assimilation, and predictability*. Cambridge University Press.

[26] Köhl, A., 2015: Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic. *Quarterly Journal of the Royal Meteorological Society*, **141 (686)**, 166–181, doi:10.1002/qj.2347, URL http://dx.doi.org/10.1002/qj.2347.

[27] Korn, P., 2016: A class of mimetic ocean Primitive Equation Models: The hydrodynamical kernel. *in preparation*.

[28] Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, **32 (4)**, 363–403.

[29] McWilliams, J. C., 2006: *Fundamentals of geophysical fluid dynamics*. Cambridge University Press.

[30] McWilliams, J. C., 2008: The nature and consequences of oceanic eddies. *Ocean Modeling in an Eddying Regime*, 5–15.

[31] Meidner, D., and T. Richter, 2015: A posteriori error estimation for the fractional step theta discretization of the incompressible navier–stokes equations. *Computer Methods in Applied Mechanics and Engineering*, **288**, 45–59.

[32] Mori, H., 1965: Transport, collective motion, and Brownian motion. *Progress of Theoretical Physics*, **33 (3)**, 423–455.

[33] Mori, H., H. Fujisaka, and H. Shigematsu, 1974: A new expansion of the master equation. *Progress of Theoretical Physics*, **51 (1)**, 109–122.

[34] Munk, W. H., 1950: On the wind-driven ocean circulation. *Journal of Meteorology*, **7 (2)**, 80–93.

[35] Oberkampf, W. L., and T. G. Trucano, 2002: Verification and validation in computational fluid dynamics. *Progress in Aerospace Sciences*, **38 (3)**, 209–272.

[36] Oden, J. T., and S. Prudhomme, 2002: Estimation of modeling error in computational mechanics. *Journal of Computational Physics*, **182 (2)**, 496–515.

[37] Olbers, D., J. Willebrand, and C. Eden, 2012: *Ocean dynamics*. Springer Science & Business Media.

[38] Orlanski, I., and M. D. Cox, 1972: Baroclinic instability in ocean currents. *Geophysical & Astrophysical Fluid Dynamics*, **4 (1)**, 297–332.

[39] Pacanowski, R., and S. Philander, 1981: Parameterization of vertical mixing in numerical models of tropical oceans. *Journal of Physical Oceanography*, **11 (11)**, 1443–1451.

[40] Prudhomme, S., and J. T. Oden, 1999: On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, **176 (1)**, 313–331.

[41] Rauser, F., J. Marotzke, and P. Korn, 2015: Ensemble-type numerical uncertainty information from single model integrations. *Journal of Computational Physics*, **292**, 30–42, doi:http://dx.doi.org/10.1016/j.jcp.2015.02.043, URL http://www.sciencedirect.com/science/article/pii/S0021999115001138.

[42] Rauser, F., J. Riehme, K. Leppkes, P. Korn, and U. Naumann, 2010: On the use of discrete adjoints in goal error estimation for shallow water equations. *Procedia Computer Science*, **1 (1)**, 107–115.

[43] Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *Journal of Physical Oceanography*, **12 (10)**, 1154–1158.

[44] Ringler, T., M. Petersen, R. L. Higdon, D. Jacobsen, P. W. Jones, and M. Maltrud, 2013: A multi-resolution approach to global ocean modeling. *Ocean Modelling*, **69**, 211 – 232, doi:http://dx.doi.org/10.1016/j.ocemod.2013.04.010, URL http://www.sciencedirect.com/science/article/pii/S1463500313000760.

[45] Rípodas, P., and Coauthors, 2009: Icosahedral Shallow Water Model (ICOSWM): results of shallow water test cases and sensitivity to model parameters. *Geoscientific Model Development*, **2 (2)**, 231–251.

[46] Roache, P. J., 1997: Quantification of uncertainty in computational fluid dynamics. *Annual Review of Fluid Mechanics*, **29 (1)**, 123–160.

[47] Roy, C. J., 2010: Review of discretization error estimators in scientific computing. *in: AIAA Paper 2010-0126*.

[48] Sirkes, Z., and E. Tziperman, 1997: Finite difference of adjoint or adjoint of finite difference? *Monthly Weather Review*, **125 (12)**, 3373–3378.

[49] Stammer, D., M. Balmaseda, P. Heimbach, A. Köhl, and A. Weaver, 2016: Ocean data assimilation in support of climate applications: Status and perspectives. *Annual review of marine science*, **8**, 491–518.

[50] Vallis, G. K., 2006: *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press.

[51] Venditti, D. A., and D. L. Darmofal, 2000: Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. *Journal of Computational Physics*, **164 (1)**, 204–227.

[52] Wolfe, C., P. Cessi, J. McClean, and M. Maltrud, 2008: Vertical heat transport in eddying ocean models. *Geophysical Research Letters*, **35 (23)**.

[53] Wunsch, C., and P. Heimbach, 2007: Practical global oceanic state estimation. *Physica D: Nonlinear Phenomena*, **230 (1 - 2)**, 197–208, doi:http://dx.doi.org/10.1016/j.physd.2006.09.040, URL http://www.sciencedirect.com/science/article/pii/S0167278906003496.

[54] Zwanzig, R., 1973: Nonlinear generalized Langevin equations. *Journal of Statistical Physics*, **9 (3)**, 215–220.

Acknowledgements

First and foremost I would like to thank my PhD-supervisors Peter Korn and Jochem Marotzke for giving me the opportunity to work on this challenging and interesting topic, your continuous support throughout my entire PhD time, and introducing me to the world of scientific research. I have thoroughly enjoyed working with you for these past four years and looking forward to keep in contact with you also in the future.

I thank Prof. Michael Hinze for taking over the task of thesis evaluator on such short notice and for his constructive feedback on its contents.

I would like to thank Jan Riehme, Ralf Müller, and Leonidas Linardakis for your support in all the countless technical questions, adjoint computing, and the occasional ICON-related challenges.

Antje, Connie and Wiebke thank you for all your help and your time, the important and fruitful discussions, helping me to make the right decisions, for always being there, and generally for making the IMPRS the way it is.

I would like to especially thank my colleagues and friends Josiane Salameh, Felix Bunzel, Mathias Heinze, Thomas Keitzl, Max Popp, Fabio Cresto Aleina, Andreas Veira, Jessica Engels, Matthias Bittner, Dirk Olonscheck, Zhuhua Li, Janja Nahrstaedt, and Mirjana Sakradzija for all your support, the nice coffee breaks, countless lunches, fun hours watching soccer, having Wednesday beers, the times spent at AGU and EGU, and of course all the nice discussions in office 318 :-).

And of course I cannot thank Freja enough for being one of my pillars in life these past years. You have strongly influenced me in a positive way and continue doing so, and have significantly contributed to help me get to the point of finishing this thesis.

Last but definitely not least, I want to thank my parents and my sister for always believing in me and for always being there to keep me on track whenever times got turbulent.