Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

A maximum entropy approach to the parameterization of subgrid-scale processes in two-dimensional flows

P. Kalverla

KNMI Internal Report IR-2015-01

A maximum entropy approach to the parameterization of subgrid-scale processes in two-dimensional flows

Peter Kalverla

May 26, 2015

A maximum entropy approach to the parameterization of subgrid-scale processes in two-dimensional flows

Summary

One of the problems that inevitably needs to be dealt with in numerical modelling of the atmosphere, ocean or any other geophysical flow system, is the representation of processes that act on scales that are smaller than those resolved by the model. This problem is called parameterization. Over the last decades, several methods have been proposed to address this issue. Often, these methods strongly depend on the system under consideration and usually some form of tuning is necessary to optimize the parameterization. The latter is, in principle, not necessary for parameterizations that are based on the principle of maximum entropy, i.e., on the assumption that the probability distribution of the unresolved scales (unresolved degrees of freedom of the system) should be in a state of maximum entropy. This has been shown to work well for a simple schematic model proposed by Lorenz (Verkley, 2011).

In this study, we applied the same procedure to a somewhat more realistic model: a two-dimensional, doubly-periodic flow system. It is assumed that the flow is incompressible and therefore divergence free, so that the system can be described completely by the vorticity equation. The system is forced by a fixed vorticity pattern and damped by linear damping and viscosity. The numerical description is based on a spectral form of the vorticity equation with sine and cosine functions in both directions and uses a fourth order Runge Kutta time integration scheme. The truncation that is chosen, i.e. the maximum values of the wave numbers in the sine and cosine functions, determines which scales are resolved and which scales are unresolved.

We first ran the model with a high truncation, so that the part of the spectrum that we are interested in has converged to a stable solution. The time evolution that is described by this model is referred to as the reference

run and is considered as reality, which we then tried to approximate with a model with a lower truncation. The main research question was: *"To what extent can the model with the lower truncation represent the corresponding scales in the model with the higher truncation and does the inclusion of a parameterization of the unresolved scales (everything between the lower and higher truncation), based on the maximum entropy assumption, lead to an improvement of this representation?"*

The results show that without any parameterization, the lower truncation leads to an overestimation of the energy and enstrophy in the smallest resolved scales of the model. This is reflected in a grainy, unphysical texture of the flow field and quick decorrelation with the reference run. The new parameterization leads to a damping that acts specifically on these smallest resolved scales. The simulated flow field is more realistic and correlation with the reference run is maintained longer. We also show that the climate of the model, expressed as probability density functions of the energy and enstrophy of the system, is improved by the new parameterization.

Contents

Foreword

This report is the final product of my internship at KNMI, the Royal Dutch Meteorological Institute. I worked here from January to May 2015 under the supervison of Wim Verkley and Camiel Severijns. The internship is part of the MSc program 'Earth and Environment' at Wageningen University (WUR). Within this study programme I specialized in meteorology. On behalf of WUR, Leo Kroon was the internship supervisor.

I worked with a two-dimensional flow model, written in Fortran77 code. Part of this report can be considered as a user manual for this model. All the relevant equations are derived, the construction of the model is outlined and also the settings as we used them are listed and explained.

Another important ingredient of this report is the theory of maximum information entropy. It states that the probability density function that describes the (unknown) state of a system should be as broad as possible, given any (known) constraints on that system. Wim used this concept to derive an expression for the unresolved scales of our model. In this report, I develop this derivation in a way that is understandable for me and I hope also for students with a similar background.

My most important contribution to this study was the implementation of the maximum entropy parameterization in the Fortran77 model. The rest of my internship was devoted to testing the new parameterization and documenting the results. In anticipation of what follows, I can say that they are at least very promising.

I had an enjoyable time at KNMI. I had the opportunity to see a lot of what is going on here: the weather room, the 3D-visualisation, the radar tower, the ozone balloon, the wind tunnel and the seismology department. The initiative to bring interns together in a mailing list resulted in very sociable lunches, pitch presentations and spontaneous Friday afternoon drinks. I want to thank Wim for all his time and patience and both Wim, Camiel and Leo for their excellent feedback and teaching.

Enjoy reading :-)

1 Introduction

In understanding geophysical fluid flows, numerical models that simulate these flows have become indispensable instruments. A challenging part of many of these models is the question how to deal with the unresolved scales in the model. This problem is called parameterization. In this report, we present a new type of parameterization. It is based on the assumption that the information entropy of the probability density function representing the unresolved scales should be maximal, given the state of the resolved scales. We have implemented this parameterization in a two-dimensional flow model on a bi-periodic domain. This chapter deals with the basic concepts of atmospheric modelling and the need for parameterization. As such, it provides the background and motivation for this study.

1.1 Two-dimensional flow

In this project we are dealing with a flat surface, a square of dimensions $2\pi L * 2\pi L$, where L is an arbitrary length scale and periodic in both directions. Moreover, this surface does not move or rotate. In that sense, it is different from the atmosphere or ocean, which are both three-dimensional and experience an ambient rotation. The turning of the globe introduces interesting phenomena referred to as the Coriolis effect and centrifugal acceleration, but the treatment of these is reserved for later studies. Nevertheless, this square surface has important similarities to the more complex

three-dimensional atmosphere and therefore it is still useful for exploratory studies like this.

One of the reasons that the atmosphere can be relatively well described by this simple model is that the atmosphere is, in fact, very thin. The day-to-day weather variations occur in a layer called the boundary layer, with a typical depth of the order of 1 km. Larger-scale weather phenomena such as anvil clouds and cyclones may reach up to the troposphere, which is located at approximately 12 km. In recent years the stratospheric circulation has received increasing attention in the context of climate variability. Then, we are dealing with a height of around 50 km. Still, compared to the radius of the earth (6371 km) and the horizontal extent of the atmosphere ($2\pi * 6371$ km) this is indeed very thin. Another reason why atmospheric flows behave as if they were two-dimensional is that air pressure decreases with height. A common simplification in meteorology is that the atmosphere is *barotropic*, which means that density (and temperature) is only dependent on pressure. Then, the atmosphere can be viewed as a stack of surfaces, each with its own density. If a parcel of air is displaced from one surface to another, it will be 'pushed' back because the ambient pressure is higher or lower than the pressure of the parcel. This is called the buoyancy effect. The result is that moving air tends to follow these surfaces of equal density. In essence, this is also what is done in meteorological models: stacked surfaces with interaction terms for the cases in which the barotropic assumption is not justified.

Indeed, a model for two-dimensional flow gives results that are similar to what we observe in the atmosphere (Wayne, 2011). An example is given in Figure 1.1. Both the atmosphere and the two-dimensional model form rotating structures called vortices, which move with the mean flow and interact when they come close to each other, often forming long filaments on the run. Extensive and illustrative descriptions of this behaviour can be found in e.g. Kraichnan and Montgomery (1980); McWilliams (1984) and Bouchet and Venaille (2012).

There are numerous textbooks about numerical modelling of geophysical flows (e.g. Coiffier (2011) or Cushman-Roisin and Beckers (2011)) and models based on the assumption that the motion is quasi two-dimensional are widely applied. For example, the dynamic core of the European Centre for Medium-Range Weather Forecast (ECMWF) model, albeit much more extensive, is in essence similar to the model that is used for this particular study. Models are quite suitable for describing the large-scale motions of the atmosphere, but it is obvious that they will never be able to describe features up to the individual motion of each air molecule on the planet. Small-scale motions, however, do have an effect on the large scale flow. To be able to represent these effects, atmospheric modellers have introduced the concept of parameterization.

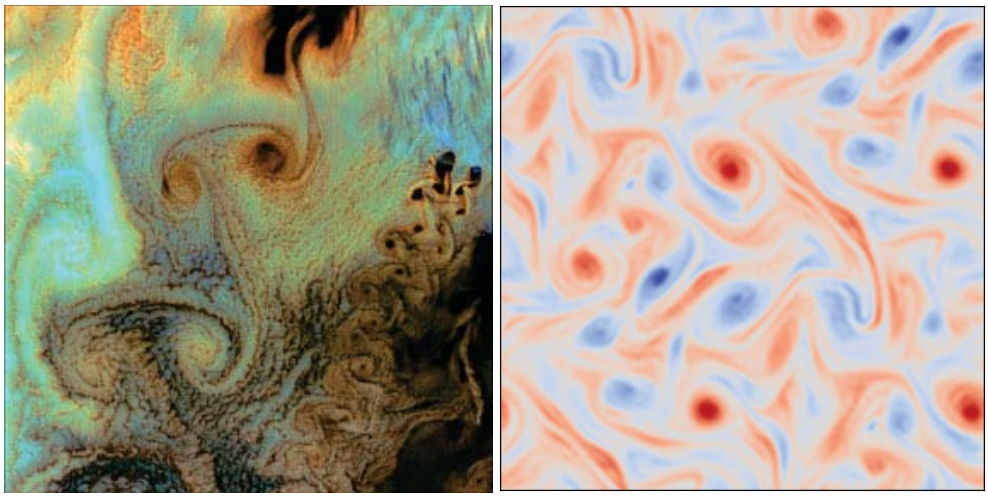


Figure 1.1: Left: Landsat image of real atmospheric vortices (Wayne, 2011). Right: Modelled two-dimensional vorticity field.

1.2 Parameterization

The goal of parameterization is to derive expressions that enable us to account for the effects of small-scale processes without explicitly resolving them. The effect of these unresolved processes is usually a form of damping and the most straightforward type of parameterization, still widely used, is to add an appropriate damping term to the equations (Cushman-Roisin and Beckers, 2011). A disadvantage of this type of parameterization is that it needs to be tuned for each resolution separately. Several studies further elaborate on this kind of parameterization (see, for example, Smagorinsky (1963); Sadourny and Basdevant (1985); Frederiksen and Davies (1997); Frederiksen and Kepert (2006) and Bihlo *et al.* (2014)).

From a more fundamental perspective, we can think of the unresolved scales as variables that are in an unknown state that can be statistically described by a probability density function (PDF). This PDF may be conditioned on the large-scale variables. In a deterministic approach we can then average the equations over this PDF to find the mean effect of the unresolved scales. In a stochastic approach, random samples are taken from the PDF to represent the effect of small-scale processes (e.g. Crommelin and Vanden-Eijnden 2008; Christensen *et al.* 2014). In both approaches, the question arises how to choose this PDF.

1.3 The maximum entropy approach

In this study, we derive an expression for the PDF using the principle of maximum entropy. The information entropy is a measure for the uncertainty of the state of a system first introduced by Shannon (1948). Jaynes (1957) showed that maximization of the entropy can be used to infer best estimates for the state of unknown variables in a system. Verkley and Lynch (2009) adopted this principle to derive best estimates for energy and enstrophy spectra of geostrophic flows and to find long-term global means for relative vorticity. Two years later, Verkley (2011) proposed to use the maximum information entropy to address the problem of parameterization. For this study, he used the Lorenz '96 model (Lorenz, 1996), which was referred to by Crommelin and Vanden-Eijnden (2008) as "a test bed for parameterization algorithms".

1.4 This study

The aim of this study is to extend the results of Verkley (2011) to a more realistic model, in particular to two-dimensional flow. We might formulate the central research question as: *"Can the simulation of two-dimensional flow be improved by using a parameterization that is based on the principle of maximum entropy?"* To answer this question, we first run the model using a high resolution and consider this simulation as 'reality'. Subsequently, we run the model with exactly the same settings, but on a lower resolution. This is the 'unparameterized' model. Then, we increase the value of the viscosity parameter. This is a commonly used parameterization, which we refer to as 'conventional'. Finally, we implement the new parameterization and compare the results, both on a short term, relevant to weather forecasting applications, and on a long term, relevant to climate simulations.

1.5 About this report

The outline of the following chapters is as follows. In chapter 2 we derive the vorticity transport equation, which is the basis for our model. Chapter 3 is concerned with the numerical implementation of this equation and chapter 4 outlines the reasoning behind the maximum entropy approach. Chapter 5 explains our methodology, the results are presented in chapter 6 and finally chapter 7 contains a discussion and conclusions.

2

The vorticity transport equation

In this chapter, we will present a derivation of the vorticity transport equation (VTE). The vorticity transport equation is a simplification of the full vorticity equation, valid for two-dimensional, incompressible fluid flow in an inertial frame of reference. The VTE is derived from the Navier-Stokes equations for fluid motion and is formulated in terms of (relative) vorticity and a stream function. As already mentioned in the introduction, the VTE can be extended with extra terms that account for the Coriolis effect and centrifugal acceleration, but this is not within the scope of the current study.

2.1 Equations of fluid motion

The equations of motions are derived from Newton's second law, which states that the time rate of change of momentum of an object is equal to the sum of all forces acting on this object. Since momentum equals mass times velocity, we can obtain a prognostic equation for the velocity by expressing the forces per unit of mass (assuming that the mass of the object itself is independent of time). In mathematical form, this is:

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \sum_i \frac{\mathbf{F}_i}{m}, \tag{2.1}$$

where $\mathbf{v}$ represents the three-dimensional velocity vector, $\frac{\mathrm{d}}{\mathrm{d}t}$ is the total derivative with respect to time, $\mathbf{F}_i$ are force vectors and m is the mass of

the particle or parcel. The relevant forces are the pressure gradient force, viscous force and gravitational force, which can be represented respectively as (Holton and Hakim, 2013):

$$\frac{\mathbf{F}_p}{m} = -\frac{1}{\rho}\nabla p, \tag{2.2}$$

$$\frac{\mathbf{F}_v}{m} = \nu\nabla^2\mathbf{U}, \tag{2.3}$$

$$\frac{\mathbf{F}_g}{m} = \mathbf{g}, \tag{2.4}$$

so that the prognostic equation for velocity becomes:

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{v} + \mathbf{g}, \tag{2.5}$$

where ρ is the density of the fluid, p is the pressure, ν is the viscosity of the fluid and $\mathbf{g}$ is the acceleration due to gravity. To simplify what follows, we now choose to continue the analysis in two-dimensional Cartesian coordinates. With the total derivative written out the equation above takes the form

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), \tag{2.6a}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -\frac{1}{\rho}\frac{\partial p}{\partial y} + \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right). \tag{2.6b}$$

Gravity is absent in these equations because it only acts in the vertical direction. We assume that density and viscosity are constant. Constant density implies that the divergence of the velocity field is zero. This will be used in the following section (Equation (2.11)).

2.2 Vorticity

Vorticity is a measure of fluid rotation that is defined for each individual point in a fluid. It is defined as the curl of the velocity field. In two dimensions, this is simply a scalar field:

$$\zeta \equiv \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \tag{2.7}$$

To obtain a prognostic equation for the vorticity, we subtract the y-derivative of Equation (2.6a) from the x-derivative of Equation (2.6b):

$$\frac{\partial}{\partial y}\left[\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right]=\frac{\partial}{\partial y}\left[-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)\right], \tag{2.8a}$$

$$\frac{\partial}{\partial x}\left[\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}\right]=\frac{\partial}{\partial x}\left[-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right)\right]. \tag{2.8b}$$

Writing out term by term and changing the order of the equations for easier subtraction results in:

$$\frac{\partial}{\partial x}\frac{\partial v}{\partial t}+\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+u\frac{\partial^2 v}{\partial x^2}+\frac{\partial v}{\partial x}\frac{\partial v}{\partial y}+v\frac{\partial}{\partial x}\frac{\partial v}{\partial y}= -\frac{1}{\rho}\frac{\partial}{\partial x}\frac{\partial p}{\partial y}+\nu\left(\frac{\partial}{\partial x}\frac{\partial^2 v}{\partial x^2}+\frac{\partial}{\partial x}\frac{\partial^2 v}{\partial y^2}\right), \tag{2.9a}$$

$$\frac{\partial}{\partial y}\frac{\partial u}{\partial t}+\frac{\partial u}{\partial y}\frac{\partial u}{\partial x}+u\frac{\partial}{\partial y}\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\frac{\partial u}{\partial y}+v\frac{\partial^2 u}{\partial y^2}= -\frac{1}{\rho}\frac{\partial}{\partial y}\frac{\partial p}{\partial x}+\nu\left(\frac{\partial}{\partial y}\frac{\partial^2 u}{\partial x^2}+\frac{\partial}{\partial y}\frac{\partial^2 u}{\partial y^2}\right). \tag{2.9b}$$

It is immediately clear that the pressure gradient terms will cancel upon subtraction. This is an important advantage of using the vorticity-formulation of the Navier-Stokes equations. Subtraction and rearrangement results in:

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)+u\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)+v\frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)+ \left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)=\nu\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right). \tag{2.10}$$

The last term on the left-hand side of the above equation is zero because of the continuity equation:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0. \tag{2.11}$$

In all other terms we substitute Equation (2.7) and we recognise the Laplacian on the right hand side of the equation which leaves us with:

$$\frac{\partial \zeta}{\partial t}+u\frac{\partial \zeta}{\partial x}+v\frac{\partial \zeta}{\partial y}=\nu\nabla^2\zeta. \tag{2.12}$$

2.3 The stream function

For two-dimensional, incompressible flows it is possible to define a stream function ψ, such that:

$$u = -\frac{\partial \psi}{\partial y} \quad \text{and} \quad v = \frac{\partial \psi}{\partial x}. \tag{2.13}$$

The use of a stream function has several advantages. First of all, it is possible to replace both components of velocity with only one scalar field. Secondly, use of the stream function automatically ensures that the continuity equation (2.11) is satisfied. Indeed, using the stream function to write u and v gives:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = \frac{\partial}{\partial x}\left(-\frac{\partial \psi}{\partial y}\right) + \frac{\partial}{\partial y}\left(\frac{\partial \psi}{\partial x}\right) = 0. \tag{2.14}$$

Finally, vorticity can also be expressed in terms of the stream function:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \nabla^2 \psi. \tag{2.15}$$

The last equation is called a Poisson equation. This equation will allow us to compute the velocity field from the vorticity. Substituting the stream function in the advection terms in Equation (2.12) yields:

$$\frac{\partial \zeta}{\partial t} - \frac{\partial \psi}{\partial y}\frac{\partial \zeta}{\partial x} + \frac{\partial \psi}{\partial x}\frac{\partial \zeta}{\partial y} = \nu \nabla^2 \zeta. \tag{2.16}$$

Finally we use the definition of the Jacobian:

$$\mathcal{J}(\psi, \zeta) = \frac{\partial \psi}{\partial x}\frac{\partial \zeta}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial \zeta}{\partial x}, \tag{2.17}$$

to obtain:

$$\frac{\partial \zeta}{\partial t} + \mathcal{J}(\psi, \zeta) = \nu \nabla^2 \zeta. \tag{2.18}$$

This is the vorticity transport equation, which states that the rate of change of relative vorticity following the flow is only due to viscous forces. In the absence of other source terms, this implies that vorticity will slowly dissipate until all gradients are smoothed.

We want the model to remain in a fully turbulent state, so we will use a slightly different version of the vorticity transport equation, namely:

$$\frac{\partial \zeta}{\partial t} + \mathcal{J}(\psi, \zeta) = \nu \nabla^2 \zeta + \mu(F - \zeta). \tag{2.19}$$

Two terms are added on the right-hand side: F is an arbitrary forcing that drives the model to a given state and the last term is a linear damping. μ is a linear coefficient that determines the magnitude of the forcing and damping. After sufficiently long integration, the cumulative contributions of forcing, linear and viscous damping drive the model to a state of statistical equilibrium. The next chapter will be devoted to the numerical implementation of this equation.

3

Construction of the model

In this chapter we explain how our model is constructed. Equation (2.19) is evaluated on a biperiodic domain, i.e. between $-\pi$ to $+\pi$ in both the x- and y-direction. The model is formulated in terms of non-dimensional variables, which means that the real grid boundaries are obtained by multiplying with a length scale L. The advantage of the non-dimensionalized model is that we can choose any length scale, for example $L \sim$2000 km such that the domain stretches over North-West Europe[1]. Section 3.1 elaborates on the steps that must be taken to make the model dimensionless. In Section 3.2, expressions for the energy and enstrophy are derived, which we will need later on. The model is a spectral model, which means that all information is stored in so-called spectral coefficients. These coefficients are independent of space and are related to a grid using certain basis functions. This is explained in more detail in Section 3.3. Here, we also discuss the concept of truncation. Finally, we elaborate on the numerical implementation of the model in Section 3.4.

3.1 Non-dimensional quantities

In order to make the model dimensionless we need scaling parameters. We choose a time scale related to the rotation rate of the earth and a length

[1] In this example, the curvature of the earth's surface is neglected.

scale equal to the earth's radius:

$$T = \Omega^{-1} = \frac{\text{sidereal day}}{2\pi} = \frac{1}{7.292 \times 10^{-5}} \text{ s} \tag{3.1}$$

$$L = a = 6371 \times 10^3 \text{ m} \tag{3.2}$$

The choice for these parameters is arbitrary, but this particular choice will make the extension to more sophisticated, global and non-inertial models easier. The dimensionless units of time t' and lengths x' and y' are obtained by dividing by these scaling parameters:

$$t' = \frac{t}{T} \qquad x' = \frac{x}{L} \qquad y' = \frac{y}{L} \tag{3.3}$$

Recalling the units of the model parameters, we can use these expressions to obtain the following non-dimensional model parameters:

$$u' = u\frac{T}{L} \qquad v' = v\frac{T}{L} \tag{3.4a-b}$$

$$\zeta' = \zeta T \qquad \psi' = \psi\frac{T}{L^2} \tag{3.4c-d}$$

$$\nu' = \nu\frac{T}{L^2} \qquad \mu' = \mu T \tag{3.4e-f}$$

$$F' = FT \tag{3.4g}$$

Also the definitions of the Jacobian and Laplace operators must be adapted, as illustrated by the following example:

$$\mathcal{J}(\psi,\zeta) = \frac{\partial\psi}{\partial x}\frac{\partial\zeta}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial\zeta}{\partial x} \tag{3.5a}$$

$$= \frac{\partial\psi'}{\partial x'}\frac{L^2}{TL}\frac{\partial\zeta'}{\partial y'}\frac{1}{TL} - \frac{\partial\psi'}{\partial y'}\frac{L^2}{TL}\frac{\partial\zeta'}{\partial x'}\frac{1}{TL} \tag{3.5b}$$

$$= \frac{1}{T^2}\frac{\partial\psi'}{\partial x'}\frac{\partial\zeta'}{\partial y'} - \frac{1}{T^2}\frac{\partial\psi'}{\partial y'}\frac{\partial\zeta'}{\partial x'} = \frac{1}{T^2}\mathcal{J}'(\psi',\zeta'). \tag{3.5c}$$

Likewise,

$$\nu\nabla^2\zeta = \frac{1}{T^2}\nu'\nabla'^2\zeta', \quad \text{where} \tag{3.6}$$

$$\nabla'^2 = \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2}. \tag{3.7}$$

It can be checked that upon substitution of the expressions (3.4) to (3.6) in the original equation (2.19), all scaling parameters cancel and we are left with exactly the same equation, with now all variables dimensionless. If we want to convert to dimensional values of e.g. vorticity, we just have to use the expressions (3.4).

3.2 Energy and enstrophy

Before we continue, we will also derive expressions for the energy and enstrophy and their evolution. Both are important quantities in the characterisation of two-dimensional flow and they will later be used to derive the maximum entropy parameterization.

Kinetic energy is a measure for the intensity of motion in the system. The total kinetic energy in the flow field is given by

$$E = \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\frac{1}{2}\rho\mathbf{v}^2, \tag{3.8}$$

where $\mathcal{D}$ is used to indicate integration over the model domain, ρ is the mass per unit area and $\mathbf{v}$ is the velocity vector in two dimensions. In meteorology, it is common practice to express energy per unit area. Bearing this in mind we introduce the energy scale

$$E^* = (2\pi L)^2\rho\left(\frac{L}{T}\right)^2 = \left(\frac{2\pi L^2}{T}\right)^2\rho \tag{3.9}$$

and express the dimensionless energy as

$$E' = \frac{E}{E^*} = \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}'} \mathrm{d}x'\,\mathrm{d}y'\,\frac{1}{2}\mathbf{v}'^2 \tag{3.10}$$

It will later be convenient to express the energy equation in terms of vorticity and the stream function. To this end, we note that $\mathbf{v}'$ equals

$$\mathbf{v}' = \mathbf{k}\times\nabla\psi'. \tag{3.11}$$

Further, because

$$\mathbf{v}'\cdot\mathbf{v}' = \nabla\psi'\cdot\nabla\psi', \tag{3.12}$$

we have

$$E' = \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}'} \mathrm{d}x'\,\mathrm{d}y'\,\frac{1}{2}\nabla\psi'\cdot\nabla\psi' \tag{3.13}$$

Using the vector identities

$$\nabla\cdot(\psi'\nabla\psi') = \psi'\nabla\cdot\nabla\psi' + \nabla\psi'\cdot\nabla\psi' \tag{3.14a}$$

$$\nabla\cdot\nabla\psi' = \nabla^2\psi' \tag{3.14b}$$

and recalling that the vorticity is the Laplacian of the stream function, we can write

$$\mathbf{v}' \cdot \mathbf{v}' = \nabla \cdot (\psi' \nabla \psi') - \psi' \nabla \cdot \nabla \psi' = \nabla \cdot (\psi' \nabla \psi') - \psi' \zeta', \tag{3.15}$$

to express Equation (3.13) as:

$$E' = \frac{1}{2} \left(\frac{1}{2\pi} \right)^2 \left[\iint\limits_{\mathcal{D}'} \mathrm{d}x' \, \mathrm{d}y' \, \nabla \cdot (\psi' \nabla \psi') - \iint\limits_{\mathcal{D}'} \mathrm{d}x' \, \mathrm{d}y' \, \psi' \zeta' \right], \tag{3.16}$$

Gauss' theorem states that for a closed area $\mathcal{D}'$ with boundary $\mathcal{C}'$,

$$\iint\limits_{\mathcal{D}'} \mathrm{d}x' \mathrm{d}y' \, \nabla \cdot \mathbf{A} = \oint\limits_{\mathcal{C}'} \mathrm{d}s' \, \mathbf{n} \cdot \mathbf{A}, \tag{3.17}$$

where $\mathbf{A}$ denotes a vector field. We can apply this theorem to the first integral in Equation (3.16). Since we will use periodic boundary conditions there is no net flow out of our system and we can say that

$$\iint\limits_{\mathcal{D}'} \mathrm{d}x' \, \mathrm{d}y' \, \nabla \cdot (\psi' \nabla \psi') = \oint\limits_{\mathcal{C}'} \mathrm{d}s' \, \mathbf{n} \cdot (\psi' \nabla \psi') = 0, \tag{3.18}$$

so that we can finally write

$$E' = - \left(\frac{1}{2\pi} \right)^2 \iint\limits_{\mathcal{D}'} \mathrm{d}x' \, \mathrm{d}y' \, \frac{1}{2} \psi' \zeta' \tag{3.19}$$

Enstrophy is a measure of the amount of structure or detail in the flow. High values of enstrophy correspond to a flow field with many small vortices. Enstrophy is associated with dissipation or the decay rate of kinetic energy (e.g. Zhu and Antonia, 1996). At the end of this section we will demonstrate this. Enstrophy turns out to be a useful quantity in the description of two-dimensional flow (Bouchet and Venaille, 2012). In two dimensions, it is defined as the square of the vorticity in two dimensions:

$$Z = \iint\limits_{\mathcal{D}} \mathrm{d}x \, \mathrm{d}y \, \frac{1}{2} \rho \zeta^2 \tag{3.20}$$

which can be rewritten in terms of our non-dimensional model parameters as

$$Z' = \frac{Z}{Z^*} = \left(\frac{1}{2\pi} \right)^2 \iint\limits_{\mathcal{D}'} \mathrm{d}x' \, \mathrm{d}y' \, \frac{1}{2} \zeta'^2 \tag{3.21}$$

with

$$Z^* = (2\pi L)^2 \rho \left(\frac{1}{T}\right)^2 = \left(\frac{2\pi L}{T}\right)^2 \rho \tag{3.22}$$

At this point have derived non-dimensional expressions for all important model parameters that we will work with. In the remainder of this report we will omit the primes and work only with the non-dimensional parameters, unless explicitly stated otherwise.

Prognostic equations for energy and enstrophy provide valuable insight in the system of equations. Moreover, the time derivative of energy will be used later in this report to derive an important expression in the maximum entropy parameterization. Let us first examine the energy. Following equation (3.10), we get:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\frac{1}{2}\mathbf{v}^2, \tag{3.23a}$$

$$= \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\mathbf{v}\cdot\frac{\partial}{\partial t}\mathbf{v}, \tag{3.23b}$$

$$= \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\nabla\psi\cdot\frac{\partial}{\partial t}\nabla\psi. \tag{3.23c}$$

Again using the vector identities (3.14) and Gauss' theorem (3.17), this can be rewritten as

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\psi\frac{\partial\zeta}{\partial t}. \tag{3.24}$$

Substituting (2.19) we find

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\psi\left[-\mathcal{J}(\psi,\zeta) + \nu\nabla^2\zeta + \mu(F-\zeta)\right]. \tag{3.25}$$

We can eliminate the Jacobian term in this equation by realising that it is another form of writing vorticity advection for a divergence-free fluid:

$$\mathcal{J}(\psi,\zeta) = \mathbf{v}\cdot\nabla\zeta. \tag{3.26}$$

Using the product rule (3.14a) twice we find subsequently

$$\psi\left[\mathbf{v}\cdot\nabla\zeta\right] = \psi\left[\nabla\cdot(\mathbf{v}\zeta) - \nabla\cdot\mathbf{v}\zeta\right], \tag{3.27a}$$

$$\psi\left[\nabla\cdot(\mathbf{v}\zeta)\right] = \nabla\cdot\left[\psi(\mathbf{v}\zeta)\right] - \nabla\psi\cdot\mathbf{v}\zeta, \tag{3.27b}$$

where the last term in Equation (3.27a) vanishes because of incompressibility and the last term in Equation (3.27b) vanishes because $\mathbf{v}$ and $\nabla\psi$ are perpendicular. Finally, we apply Gauss' theorem once again to completely eliminate the Jacobian term in the integral, which leaves us with

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\left(\frac{1}{2\pi}\right)^2 \iint\limits_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\psi\left[\nu\nabla^2\zeta + \mu(F-\zeta).\right] \tag{3.28}$$

A similar procedure leads to a prognostic equation for enstrophy. Using

$$\zeta[\mathbf{v}\cdot\nabla\zeta] = \mathbf{v}\cdot\nabla\frac{1}{2}\zeta^2, \tag{3.29}$$

$$= \nabla\cdot\left[\mathbf{v}\frac{1}{2}\zeta^2\right] - \frac{1}{2}\zeta^2\nabla\cdot\mathbf{v}, \tag{3.30}$$

$$= \nabla\cdot\left[\mathbf{v}\frac{1}{2}\zeta^2\right] \tag{3.31}$$

and again applying Gauss' theorem, it follows that

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = -\left(\frac{1}{2\pi}\right)^2 \iint\limits_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\,\zeta\left[\nu\nabla^2\zeta + \mu(F-\zeta)\right]. \tag{3.32}$$

The physical interpretation of the above expressions is that the rate of change of energy and enstrophy is determined only by the forcing and friction terms. If both μ and ν equal zero, energy and enstrophy are conserved in two-dimensional flow. That enstrophy is conserved in unforced-undamped two-dimensional turbulence is an essential property that distinguishes it from three-dimensional turbulence (Kraichnan and Montgomery, 1980).

Let us quickly examine a conceptual system without forcing or linear damping. Then, Equation (3.28) can be written as

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -2\nu Z. \tag{3.33}$$

This equation expresses the relation between enstrophy and the dissipation of energy, which we mentioned earlier in this section.

3.3 Spectral representation

In this section, we explain the concept of a spectral model and how it is derived for the vorticity transport equation.

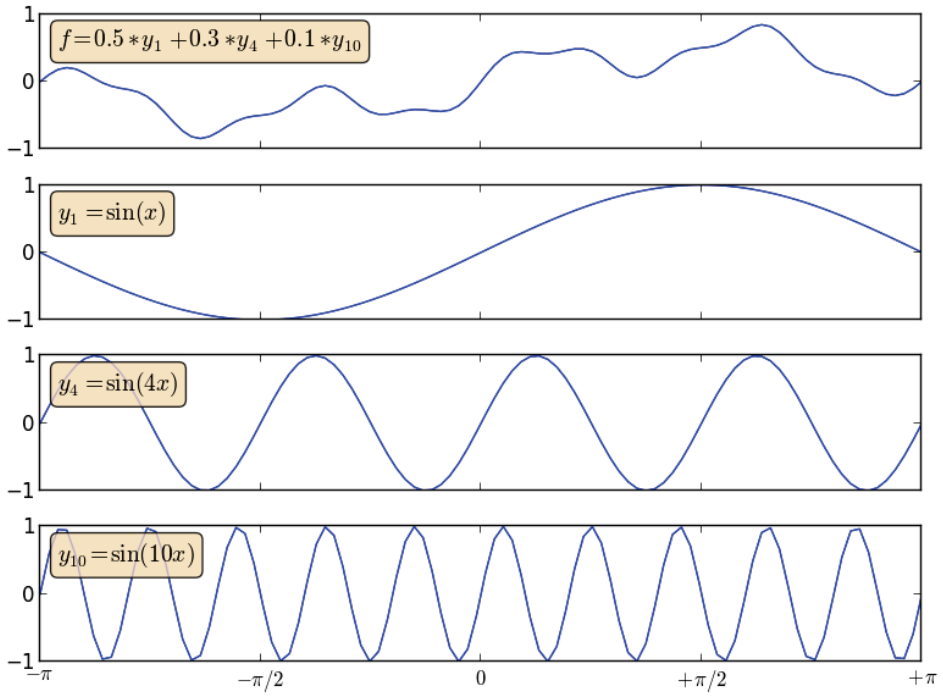


Figure 3.1: A seemingly complicated function can be decomposed into simple waves with different wave number and amplitude.

3.3.1 Basis functions

The concept of a spectral model is based on the idea that any (seemingly complicated) scalar field or function can be represented by a summation of orthonormal basis functions, for example a set of sine and cosine waves. This is illustrated in Figure 3.1 for a one-dimensional field, but the concept can easily be extended to two (or more) dimensions. We choose our basis functions to be

$$Y_{m,n}(x,y) = X_m(x)X_n(y), \tag{3.34}$$

with

$$X_m(x) = \begin{cases} \sqrt{2}\cos(|m|x) & \text{if } m > 0 \\ 1 & \text{if } m = 0 \\ \sqrt{2}\sin(|m|x) & \text{if } m < 0 \end{cases} \tag{3.35}$$

and likewise for $X_n(y)$. Here, m and n are components of the wave vector $\mathbf{k}$. It is then possible to express the vorticity as

$$\zeta(x,y,t) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \zeta_{m,n}(t) Y_{m,n}(x,y), \tag{3.36}$$

where $\zeta_{m,n}$ are called the spectral coefficients. They can be interpreted as amplitudes or weight factors associated with each wave. Note that the above expression for $X_m(x)$ is closely related to the complex exponential form

$$X_m(x) = \mathrm{e}^{imx}. \tag{3.37}$$

Both representations are useful and it is possible to switch between the two using a relation between the real coefficients associated with the real basis function and the complex coefficients associated with the exponential basis functions. We show this derivation in Appendix A. The advantage of the exponential form is that we can use a Fast-Fourier Transformation (FFT) to switch between the grid-point representation and the spectral representation of ζ. The basis functions (3.34) have the convenient properties that

$$\frac{\partial X_m}{\partial x} = -mX_{-m}(x) \quad , \quad \frac{\partial^2 X_m}{\partial x^2} = -m^2 X_m(x) \quad \text{and} \tag{3.38a-b}$$

$$\nabla^2 Y_{mn}(x,y) = -(m^2+n^2) Y_{m,n}(x,y) \tag{3.38c}$$

and since $\zeta_{m,n}(t)$ is not dependent on x or y and we used absolute values of m and n in the definitions of the real basis functions, we can write

$$\frac{\partial \zeta(x,y,t)}{\partial x} = \frac{\partial}{\partial x} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \zeta_{m,n}(t) Y_{m,n}(x,y) \tag{3.39a}$$

$$= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \zeta_{m,n}(t) \frac{\partial}{\partial x} Y_{m,n}(x,y) \tag{3.39b}$$

$$= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \zeta_{m,n}(t) [-mY_{-m,n}(x,y)] \tag{3.39c}$$

$$= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} m\zeta_{-m,n}(t) Y_{m,n}(x,y). \tag{3.39d}$$

Likewise it can be checked that

$$\frac{\partial \zeta(x,y,t)}{\partial y} = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} n\zeta_{m,-n}(t) Y_{m,n}(x,y), \tag{3.40a}$$

$$\frac{\partial^2 \zeta(x,y,t)}{\partial x^2} = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} -m^2\zeta_{m,n}(t) Y_{m,n}(x,y), \tag{3.40b}$$

$$\frac{\partial^2 \zeta(x,y,t)}{\partial y^2} = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} -n^2\zeta_{m,n}(t) Y_{m,n}(x,y), \tag{3.40c}$$

$$\nabla^2 \zeta(x,y,t) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} -(m^2+n^2)\zeta_{m,n}(t) Y_{m,n}(x,y), \tag{3.40d}$$

If we recall that the vorticity is the Laplacian of the stream function, we can easily find the spectral coefficients of the stream function from the spectral coefficients of the vorticity by

$$\psi_{m,n}(t) = \frac{-1}{(m^2+n^2)} \zeta_{m,n}(t). \tag{3.41}$$

3.3.2 Finding the spectral coefficients

In this section, we derive an expression for the relation between the spectral coefficients $\zeta_{m,n}(t)$ and the corresponding field $\zeta(x,y,t)$. We define the inner product of two arbitrary vector fields A and B as:

$$\langle A, B \rangle = \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\, \mathrm{d}y\, A(x,y) B(x,y), \tag{3.42}$$

where we divide by the area of $(2\pi)^2$ of the dimensionless domain in line with the energy equation. As mentioned above, the basis functions are said to be orthonormal, which means that by virtue of the above expression for the inner product,

$$\langle Y_{m,n}, Y_{m',n'} \rangle = \delta_{m,m'} \delta_{n,n'} \tag{3.43}$$

with

$$\delta_{m,m'} = \begin{cases} 1 & \text{if } m = m' \\ 0 & \text{if } m \neq m' \end{cases} \tag{3.44}$$

and likewise for $\delta_{n,n'}$. We will not prove this here, but an excellent explanation can be found, for example, on Paul Dawkins' online math tutorials[2]. We

[2] http://tutorial.math.lamar.edu/terms.aspx (accessed 16 February, 2015)

will now use this property to find an expression for $\zeta_{m,n}(t)$. Let us consider the inner product

$$\langle Y_{m,n}(x,y), \zeta(x,y,t)\rangle = \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\, Y_{m,n}(x,y)\zeta(x,y,t). \tag{3.45}$$

Substituting the spectral expression (3.36), we can rewrite the left hand side as

$$\langle Y_{m,n}(x,y), \zeta(x,y,t)\rangle = \langle Y_{m,n}(x,y), \sum_{m'=-\infty}^{+\infty}\sum_{n'=-\infty}^{+\infty} \zeta_{m',n'}(t)Y_{m',n'}(x,y)\rangle \tag{3.46}$$

$$= \sum_{m'=-\infty}^{+\infty}\sum_{n'=-\infty}^{+\infty} \zeta_{m',n'}(t)\langle Y_{m,n}, Y_{m',n'}\rangle. \tag{3.47}$$

We recognise that by virtue of (3.43) this inner product is only non-zero for $m = m'$ and $n = n'$, thus the summation only involves m and n so that we can write

$$\zeta_{m,n}(t) = \langle Y_{m,n}(x,y), \zeta(x,y,t)\rangle = \left(\frac{1}{2\pi}\right)^2 \iint_{\mathcal{D}} \mathrm{d}x\,\mathrm{d}y\, Y_{m,n}(x,y)\zeta(x,y,t). \tag{3.48}$$

3.3.3 Truncation and the need for parameterization

The theory above assumes that we use an infinite number of waves to represent our function. In practice this is not possible and we limit our model to $2N+1$ waves in both the x- and y-direction:

$$\zeta(x,y,t) = \sum_{m=-N}^{+N}\sum_{n=-N}^{+N} \zeta_{m,n}(t)Y_{m,n}(x,y) \tag{3.49}$$

As a result of truncation the effects introduced by interactions of smaller waves (small scale/sub-grid scale processes) are not represented by the model. This is visualized in Figure 3.2. Truncation can have considerable influence on the model results. Using a higher truncation improves the model performance, but comes at the cost of extra computation time. Therefore, we wish to somehow implement the effect of the unresolved scales without explicitly computing them. This is referred to as *parameterization*. Small-scale processes tend to perturb the mean flow. Therefore, the effect of the unresolved scales is usually a damping of the larger scales. In that perspective, it is

similar to friction and the simplest (and commonly used) parameterization is to manually adjust the viscosity parameter. However, this method requires tuning of the viscosity parameter. Moreover, it is not universal, in the sense that the tuning has to be done for each individual model and must be validated for different regimes in which the system can find itself. Our aim is to find and test a parameterization for which this is not necessary. We will explain this in Chapter 4.

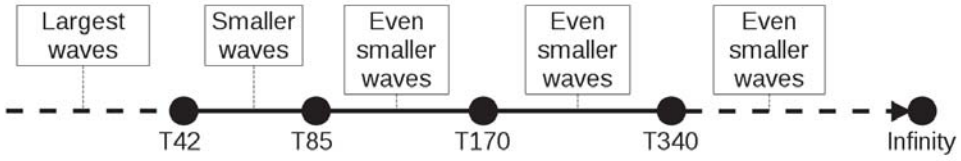


Figure 3.2: Illustration of the wavenumber continuum. Black dots represent a number of commonly used truncation levels, e.g. T42 means that waves up to wavenumber 42 are accounted for. Unresolved scales are all waves between the truncation level and infinity.

3.4 Numerical implementation

3.4.1 Time stepping routine

A flowchart of the Fortran77 model that we use for our numerical simulations is given in Figure 3.3. The program first reads an initial field, which may be zero or a set of spectral coefficients taken from a previous run that has reached statistical equilibrium. From the initial vorticity field, the backward Laplace routine computes the spectral coefficients of the stream function. Also, a forcing field is defined. In our model, the forcing field is simply a single wave, which can be denoted with the wave vector $\mathbf{k} = \binom{5}{5}$. The two coefficients $(m = 5, n = 5)$ correspond to the wave numbers in x- and y-direction. The amplitude of the forcing is $\frac{1}{2}\sqrt{2}$. From Equation (2.19) it can be seen that there are four terms contributing to the tendency of ζ: the redistribution term (Jacobian), friction, forcing and linear damping. The model computes these terms separately and then adds them to find the overall tendency of the vorticity field. This tendency calculation is the core of the model. It is used in a fourth order Runge-Kutta time integration scheme to predict future values of vorticity. This scheme runs through the following steps:

1. Load initial vorticity field ζ_0 (or set it to 0)

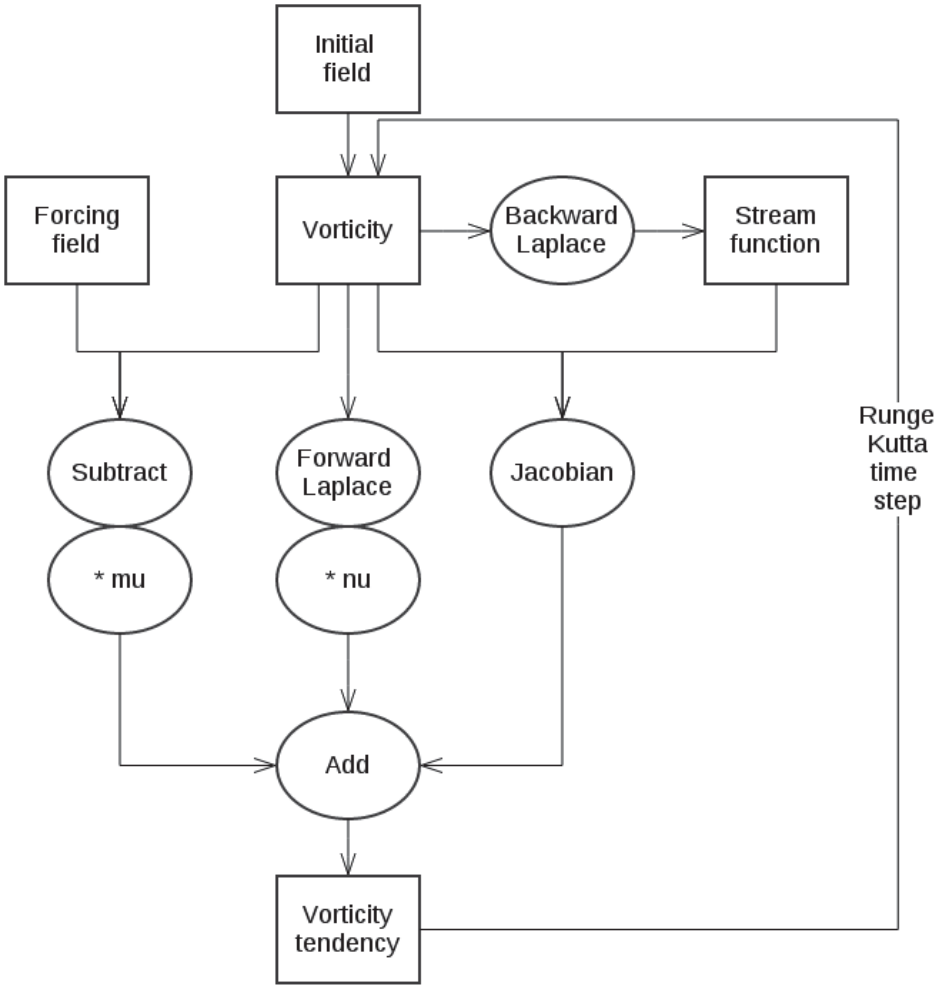


Figure 3.3: Flowchart of the original numerical model. Squares indicate fields of spectral coefficients, ellipses represent functions (fortran subroutines).

2. Compute the tendency of ζ_0, call it k_1
3. Use k_1 to calculate $\zeta_a = \zeta_0 + \frac{1}{2}\Delta t k_1$
4. Compute the tendency of ζ_a, call it k_2
5. Use k_2 to calculate $\zeta_b = \zeta_0 + \frac{1}{2}\Delta t k_2$
6. Compute the tendency of ζ_b, call it k_3

7. Use k_3 to calculate $\zeta_c = \zeta_0 + \Delta t k_2$
8. Compute the tendency of ζ_c, call it k_4
9. Use k_1, k_2, k_3, k_4 to calculate $\zeta_1 = \zeta_0 + \frac{\Delta t}{6}(k_1 + 2k_2 + 2k_3 + k_4)$
10. Repeat step 2 - 9 with the output of step 9 as new input to step 2, until the desired integration time is reached.

During the computation of one time step Δt, the tendency calculation routine is called four times. These intermediate tendencies are weighted to compute one overall tendency that is used for the time step.

3.4.2 Tendency routines

Most routines that are visualized in Figure 3.3 are straightforward, linear operations. For example, the backward Laplacian is given by Equation (3.41), and the forward Laplacian by the inverse. Only the Jacobian is a non-linear term and requires some extra explanation. Since it is a computationally expensive task to determine the non-linear term in spectral space, the derivatives of the fields represented by the spectral coefficients $\zeta_{m,n}$ and $\psi_{m,n}$ are transformed to a grid by a Fast Fourier Transform. On the grid, they are multiplied and the resulting product is transformed back to spectral space. This is illustrated in Figure 3.4. For this approximation of the non-linear terms to be exact, the number of grid points in each direction must be equal to or larger than $N \geq 3M + 1$ where M is the number of spectral coefficients corresponding to that direction (Coiffier, 2011). This is also known as the Orszag two-thirds rule (see Boyd, 2001, chapter 11). The FFT as implemented in this model works most efficient when the number of grid points is a power of 2, so we work with a grid of $N = 128 = 2^7 \geq 3 * 42 + 1$ for T42 and for T85 with $N = 256 = 2^8 \geq 3 * 85 + 1$.

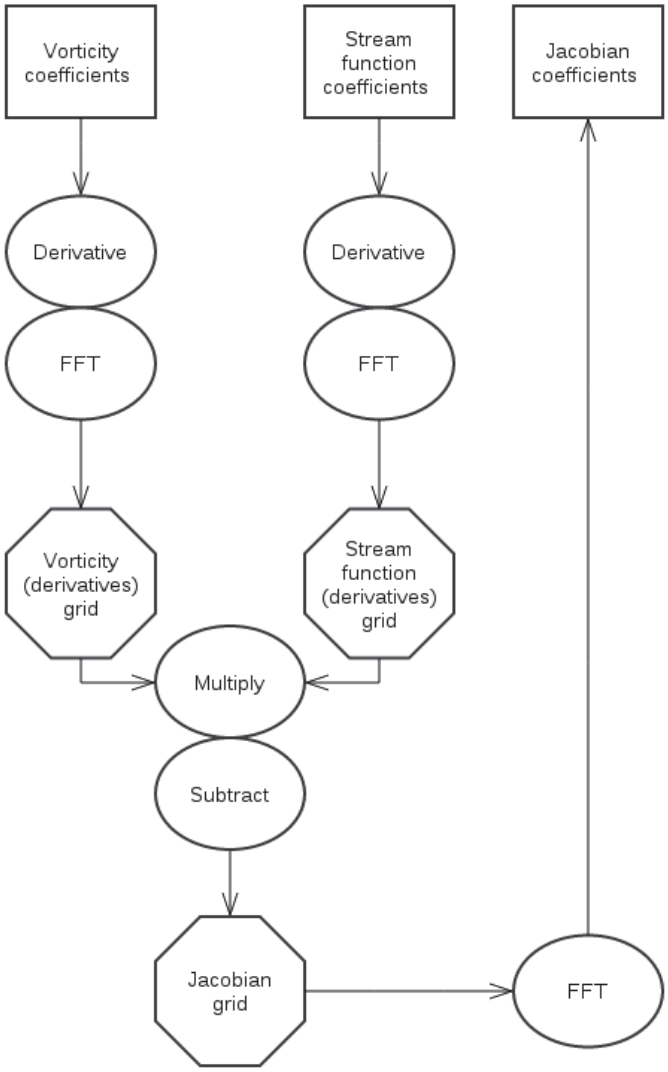


Figure 3.4: Flowchart of the Jacobian subroutine. Like figure 3.3 with the octagons representing fields in grid space.

4

The maximum entropy approach

In this chapter, we derive a new expression for the parameterization of the effect of subgrid-scale processes in our two-dimensional flow model. First, we split Equation (2.19) in a resolved ($\mathcal{R}$) and an unresolved ($\mathcal{U}$) part. We then approximate $\psi^{\mathcal{U}}$ and $\zeta^{\mathcal{U}}$ with averages over a probability density function. To justify this approach we assume that the unresolved processes act on time scales that are sufficiently short compared to the larger scale (resolved) processes. On the time scale of the resolved processes the unresolved scales are assumed to be in a statistically stationary state. Then, the input and output of energy in the unresolved scales will balance on average.

4.1 Resolved and unresolved scales

Before we can derive expressions for unresolved variables, we need to know how they influence the model. To this end, we define the resolved part of the spectrum as the set $\mathcal{R}$ and the unresolved part as the set $\mathcal{U}$, as is visualized in Figure 4.1. Then, we can write

$$\zeta = \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}, \tag{4.1}$$

$$\zeta^{\mathcal{R}} = \sum_{m,n \in \mathcal{R}} \zeta_{m,n} Y_{m,n}, \tag{4.2}$$

$$\zeta^{\mathcal{U}} = \sum_{m,n \in \mathcal{U}} \zeta_{m,n} Y_{m,n} \tag{4.3}$$

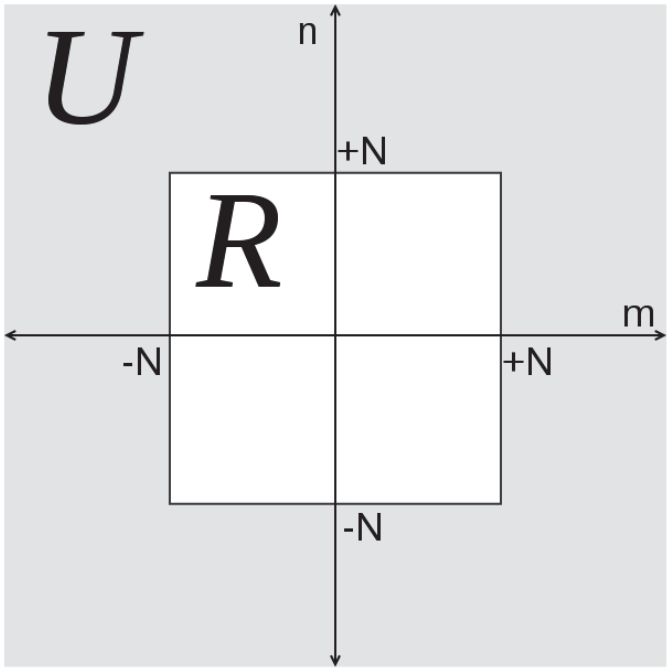


Figure 4.1: The spectral grid with truncation level N. $\mathcal{R}$ represents the spectral domain, i.e. the resolved part of the grid: $\mathcal{R} = \{\{m,n\}\,|m,n \in \mathbb{Z} \wedge -M \leq m,n \leq +M\}$. The model that we consider as 'reality' is truncated at $N = 85$. The other models are truncated at $N = 42$. The unresolved variables are represented by the set $\mathcal{U}$ which contains all wavenumbers that are part of the T85 spectrum, but fall outside of the T42 spectrum.

and likewise for ψ and F. Substituting these forms in Equation (2.19), we get:

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + \frac{\partial \zeta^{\mathcal{U}}}{\partial t} + \mathcal{J}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) = \nu\nabla^2(\zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) + \mu(F^{\mathcal{R}} + F^{\mathcal{U}} - \zeta^{\mathcal{R}} - \zeta^{\mathcal{U}}). \quad (4.4)$$

The contribution of the Jacobian can be split in a resolved and unresolved part as well:

$$\mathcal{J}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}}) = \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}}) + \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}}). \quad (4.5)$$

The superscripts $\mathcal{R}$ and $\mathcal{U}$ are added to the Jacobian operators to indicate that only the projection of this operator in the resolved resp. unresolved part of the spectrum is considered. The projection in the unresolved part

of the spectrum arises from the transformation back to the spectral space, as illustrated in figure 3.4. The other terms do not introduce interactions between the resolved and the unresolved part of the spectrum, so that we can split Equation (4.4) in two:

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) = \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}), \tag{4.6a}$$

$$\frac{\partial \zeta^{\mathcal{U}}}{\partial t} + \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) = \nu\nabla^2\zeta^{\mathcal{U}} + \mu(F^{\mathcal{U}} - \zeta^{\mathcal{U}}). \tag{4.6b}$$

The first equation is explicitly solved by the numerical model, while the second equation represents subgrid-scale processes. It can be seen that interactions between the resolved and unresolved variables occur only through the non-linear advection term. If the effect of the unresolved scales is neglected ($\psi^{\mathcal{U}} = \zeta^{\mathcal{U}} = 0$), the first expression (4.6a) is equivalent to the truncated spectral model without any parameterization. As we will see in the results section, this leads to an overestimation of energy and enstrophy in the smallest scales. We propose to represent the effect of unresolved processes by averaging Equation (4.6a) over a probability density function (PDF) of the unresolved scales. For the Jacobian, we write:

$$\begin{aligned}\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) =& \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}}) \\ &+ \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}) + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}).\end{aligned} \tag{4.7}$$

Averaging over the probability density function of the unresolved scales then gives

$$\begin{aligned}\overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})} =& \overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})} + \overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}})} \\ &+ \overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}})} + \overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})} && \text{(4.8a)} \\ =& \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \overline{\zeta^{\mathcal{U}}}) \\ &+ \mathcal{J}^{\mathcal{R}}(\overline{\psi^{\mathcal{U}}}, \zeta^{\mathcal{R}}) + \mathcal{J}^{\mathcal{R}}(\overline{\psi^{\mathcal{U}}}, \overline{\zeta^{\mathcal{U}}}) && \text{(4.8b)} \\ =& \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \overline{\psi^{\mathcal{U}}}, \zeta^{\mathcal{R}} + \overline{\zeta^{\mathcal{U}}}). && \text{(4.8c)}\end{aligned}$$

Although it is common in statistics and probability theory to denote averages and expectation values with angle brackets, we choose to work with overbars to avoid confusion between the average and the inner product later on (normally, overbars are used for the arithmetic mean). In going from (4.8a) to (4.8b) we made the assumption that $\overline{\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})} = \mathcal{J}^{\mathcal{R}}(\overline{\psi^{\mathcal{U}}}, \overline{\zeta^{\mathcal{U}}})$. We will come back to this assumption later. The model for the resolved part then becomes

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \overline{\psi^{\mathcal{U}}}, \zeta^{\mathcal{R}} + \overline{\zeta^{\mathcal{U}}}) = \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}). \tag{4.9}$$

Thus, our aim is to replace $\zeta^{\mathcal{U}}$ and $\psi^{\mathcal{U}}$ in Equation (4.6a) with averages over a probability density function $\overline{\zeta^{\mathcal{U}}}$ and $\overline{\psi^{\mathcal{U}}}$. In the following sections, we explain how this PDF is constructed.

4.2 The probability density function (PDF)

The probability density function $\mathcal{P}(\zeta^{\mathcal{U}})$ gives the probability for the unresolved scales to assume a given value for $\zeta^{\mathcal{U}}$ (see Figure 4.2). The PDF has the properties that the integral over all possible states equals unity and can therefore be used to construct a weighted average:

$$\int_{-\infty}^{+\infty} \mathrm{d}\zeta^{\mathcal{U}}\, \mathcal{P}(\zeta^{\mathcal{U}}) = 1, \tag{4.10a}$$

$$\overline{G} = \int_{-\infty}^{+\infty} \mathrm{d}\zeta^{\mathcal{U}}\, G(\zeta^{\mathcal{U}})\mathcal{P}(\zeta^{\mathcal{U}}), \tag{4.10b}$$

where G can be any function of $\zeta^{\mathcal{U}}$. Note that in these expressions $\zeta^{\mathcal{U}}$ is represented by its spectral coefficients $\zeta_{m,n}$ with $m, n \in \mathcal{U}$ and that the integral is a multiple integral over all these coefficients, i.e.:

$$\int_{-\infty}^{+\infty} \mathrm{d}\zeta^{U} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \mathrm{d}\zeta_{-N,-N}\, \mathrm{d}\zeta_{-N,-N+1} ... \mathrm{d}\zeta_{N,N} \tag{4.11}$$

The information entropy of the unresolved scales is defined as

$$S_I^{\mathcal{U}} = \int_{-\infty}^{+\infty} \mathrm{d}\zeta^{\mathcal{U}}\, \mathcal{P}(\zeta^{\mathcal{U}}) \ln\left(\frac{\mathcal{P}(\zeta^{\mathcal{U}})}{\mathcal{M}(\zeta^{\mathcal{U}})}\right). \tag{4.12}$$

Here, $\mathcal{M}$ is an *a priori* PDF for which a constant value is used (assuming no a priori information). In the derivation that follows, this value will turn out to be irrelevant. According to Jaynes (1957), the information entropy should be maximal given suitable constraints. As constraints, we use the normalization condition (4.10a) and the the assumption that the time derivative of energy in the unresolved scales equals zero on average. Using Equation (4.10b) this leads to the following expression for this average:

$$\overline{\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t}} = \int_{-\infty}^{+\infty} \mathrm{d}\zeta^{\mathcal{U}}\, \frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t}(\zeta^{\mathcal{U}})\mathcal{P}(\zeta^{\mathcal{U}}) = 0 \tag{4.13}$$

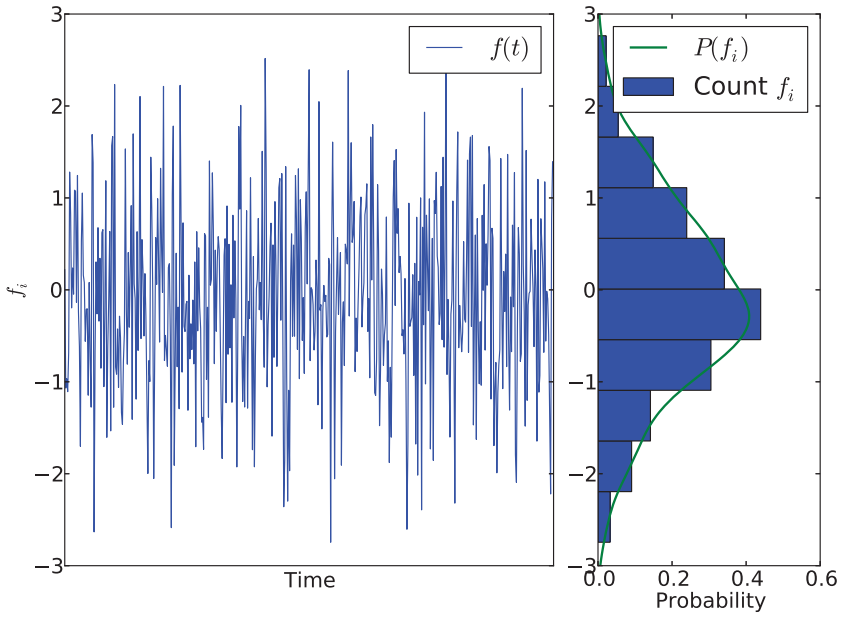


Figure 4.2: The system $f(t)$ moves through the phase space of $f(t)$ in time. The probability that a system $f(t)$ is in a certain state f_i, i.e., between $f_i - \Delta/2$ and $f_i + \Delta/2$ is given by the probability density function $P(f_i)\Delta$.

In the following section, we will use these constraints to derive the maximum entropy parameterization.

4.3 Finding the maximum entropy PDF

We will first show that the time derivative of $E^{\mathcal{U}}$ can be written in the form $A = ax^2 + bx$ where $x = \zeta^{\mathcal{U}}$. Then, we demonstrate that the maximization of S_I (Equation (4.12)) under the constraints (4.10a) and (4.13) leads to an expression for the PDF-averaged vorticity of the form $\overline{x} = \frac{-b}{2a}$.

Like vorticity and stream function, energy can also be split in a resolved and an unresolved part. To see this, it is convenient to write E in the form of an inner product as defined in (3.42):

$$E = -\left(\frac{1}{2\pi}\right)^2 \iint_D \mathrm{d}x\,\mathrm{d}y\,\frac{1}{2}\psi\zeta = -\frac{1}{2}\langle\psi,\zeta\rangle \tag{4.14a}$$

$$= -\frac{1}{2}\langle\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}\rangle \tag{4.14b}$$

$$= -\frac{1}{2}\langle\psi^{\mathcal{R}},\zeta^{\mathcal{R}}\rangle - \frac{1}{2}\langle\psi^{\mathcal{R}},\zeta^{\mathcal{U}}\rangle - \frac{1}{2}\langle\psi^{\mathcal{U}},\zeta^{\mathcal{R}}\rangle - \frac{1}{2}\langle\psi^{\mathcal{U}},\zeta^{\mathcal{U}}\rangle \tag{4.14c}$$

$$= -\frac{1}{2}\langle\psi^{\mathcal{R}},\zeta^{\mathcal{R}}\rangle - \frac{1}{2}\langle\psi^{\mathcal{U}},\zeta^{\mathcal{U}}\rangle \tag{4.14d}$$

$$= E^{\mathcal{R}} + E^{\mathcal{U}}. \tag{4.14e}$$

with $E^{\mathcal{R}} = -\frac{1}{2}\langle\psi^{\mathcal{R}},\zeta^{\mathcal{R}}\rangle$ and $E^{\mathcal{U}} = -\frac{1}{2}\langle\psi^{\mathcal{U}},\zeta^{\mathcal{U}}\rangle$. The second and third term in Equation (4.14c) cancel because projections on $\mathcal{R}$ and $\mathcal{U}$ are perpendicular by virtue of property (3.43) of the basis functions (3.34). Considering the time evolution of E, we have (see the derivation leading to Equation (3.24)):

$$\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left[-\frac{1}{2}\langle\psi^{\mathcal{R}},\zeta^{\mathcal{R}}\rangle\right]$$

$$= -\langle\psi^{\mathcal{R}}, \frac{\partial\zeta^{\mathcal{R}}}{\partial t}\rangle \tag{4.15a}$$

$$\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left[-\frac{1}{2}\langle\psi^{\mathcal{U}},\zeta^{\mathcal{U}}\rangle\right]$$

$$= -\langle\psi^{\mathcal{U}}, \frac{\partial\zeta^{\mathcal{U}}}{\partial t}\rangle. \tag{4.15b}$$

Substituting (4.6) gives

$$\begin{aligned}\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} &= -\langle\psi^{\mathcal{R}}, -\mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}}) + \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}}-\zeta^{\mathcal{R}})\rangle \\ &= \langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle - \nu\langle\psi^{\mathcal{R}}, \nabla^2\zeta^{\mathcal{R}}\rangle - \mu\langle\psi^{\mathcal{R}}, (F^{\mathcal{R}}-\zeta^{\mathcal{R}})\rangle\end{aligned} \tag{4.16a}$$

$$\begin{aligned}\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t} &= -\langle\psi^{\mathcal{U}}, -\mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}}) + \nu\nabla^2\zeta^{\mathcal{U}} + \mu(F^{\mathcal{U}}-\zeta^{\mathcal{U}})\rangle \\ &= \langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle - \nu\langle\psi^{\mathcal{U}}, \nabla^2\zeta^{\mathcal{U}}\rangle - \mu\langle\psi^{\mathcal{U}}, (F^{\mathcal{U}}-\zeta^{\mathcal{U}})\rangle.\end{aligned} \tag{4.16b}$$

The Jacobian terms can be expanded as

$$\begin{aligned}\langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle =& \langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}})\rangle + \\ & \langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{R}}, \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle\end{aligned} \tag{4.17a}$$

$$\begin{aligned}\langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle =& \langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}})\rangle + \\ & \langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{U}}, \mathcal{J}^{\mathcal{U}}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle.\end{aligned} \tag{4.17b}$$

The inner product is by definition zero for two orthogonal fields, so we can discard the superscript for the Jacobian. This allows us to use the following identities:

$$\langle a, \mathcal{J}(a, b)\rangle = 0, \tag{4.18a}$$

$$\langle a, \mathcal{J}(b, c)\rangle = \langle\mathcal{J}(a, b), c\rangle, \tag{4.18b}$$

$$\mathcal{J}(x, y) = -\mathcal{J}(y, x). \tag{4.18c}$$

The first two terms in Equation (4.17a) and the last two terms in Equation (4.17b) vanish because of (4.18a) and we can write the remaining terms in Equation (4.17) as

$$\langle\psi^{\mathcal{R}}, \mathcal{J}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle = -\langle\psi^{\mathcal{U}}, \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{R}}, \mathcal{J}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle, \tag{4.19}$$

$$\langle\psi^{\mathcal{U}}, \mathcal{J}(\psi^{\mathcal{R}}+\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}+\zeta^{\mathcal{U}})\rangle = \langle\psi^{\mathcal{U}}, \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle - \langle\psi^{\mathcal{R}}, \mathcal{J}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle. \tag{4.20}$$

The final results are

$$\begin{aligned}\frac{\mathrm{d}E^{\mathcal{R}}}{\mathrm{d}t} =& -\langle\psi^{\mathcal{U}}, \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle + \langle\psi^{\mathcal{R}}, \mathcal{J}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle \\ & - \nu\langle\psi^{\mathcal{R}}, \nabla^2\zeta^{\mathcal{R}}\rangle - \mu\langle\psi^{\mathcal{R}}, (F^{\mathcal{R}}-\zeta^{\mathcal{R}})\rangle,\end{aligned} \tag{4.21a}$$

$$\begin{aligned}\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t} =& \langle\psi^{\mathcal{U}}, \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})\rangle - \langle\psi^{\mathcal{R}}, \mathcal{J}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle \\ & - \nu\langle\psi^{\mathcal{U}}, \nabla^2\zeta^{\mathcal{U}}\rangle - \mu\langle\psi^{\mathcal{U}}, (F^{\mathcal{U}}-\zeta^{\mathcal{U}})\rangle.\end{aligned} \tag{4.21b}$$

The first two terms in both expressions have opposite signs and cancel if both expressions are added. These terms are responsible for the transport of energy between the resolved and the unresolved scales. Using $c_{m,n} = m^2 + n^2$ and $\psi_{m,n} = -\zeta_{m,n}/c_{m,n}$, Equation (4.21b) can be written in terms of spectral coefficients as

$$\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t} = \sum_{m,n\in\mathcal{U}} \left[\frac{\zeta_{m,n}}{-c_{m,n}} \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} - \nu \frac{\zeta_{m,n}}{-c_{m,n}} (-c_{m,n}) \zeta_{m,n} \right.$$
$$\left. -\mu \frac{\zeta_{m,n}}{-c_{m,n}} (F_{m,n} - \zeta_{m,n}) \right] - \sum_{m,n\in\mathcal{R}} \frac{\zeta_{m,n}}{-c_{m,n}} \mathcal{J}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})_{m,n}. \quad (4.22)$$

The last term leads to correlations in the unresolved scales. This is not necessarily zero on average, but assuming that it is small compared to the other terms greatly simplifies the analysis (it was used earlier in going from Equation (4.8a) to Equation (4.8b)). We can then write Equation (4.22) as

$$-\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t} = \sum_{m,n\in\mathcal{U}} \frac{\zeta_{m,n}}{c_{m,n}} \mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} + \nu \frac{\zeta_{m,n}}{c_{m,n}} c_{m,n} \zeta_{m,n} - \mu \frac{\zeta_{m,n}}{c_{m,n}} (F_{m,n} - \zeta_{m,n}) \quad (4.23)$$

$$= \sum_{m,n\in\mathcal{U}} \left(\nu + \frac{\mu}{c_{m,n}} \right) \zeta_{m,n}^2 + \left(\frac{\mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} - \mu F_{m,n}}{c_{m,n}} \right) \zeta_{m,n}. \quad (4.24)$$

Note that Equation (4.24) is a summation of quadratic functions in $\zeta_{m,n}$. Let us examine a simple case, where only one variable is considered. We denote this variable with x and the time rate of change of the energy will be called $\mathcal{A}$. Then, Equation (4.24) can be written as

$$\mathcal{A} = ax^2 + bx \text{ , with} \quad (4.25a)$$

$$a = \left(\nu + \frac{\mu}{c_{m,n}} \right) \text{ and} \quad (4.25b)$$

$$b = \left(\frac{J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} - \mu f_{m,n}}{c_{m,n}} \right) \quad (4.25c)$$

The information entropy in the case of a single variable x is given by

$$S_I = -\int_{-\infty}^{\infty} \mathrm{d}x \, \mathcal{P}(x) \ln \left(\frac{\mathcal{P}(x)}{\mathcal{M}(x)} \right). \quad (4.26)$$

Recall that this quantity should be maximal. Furthermore, (4.13) reduces to

$$\int_{-\infty}^{\infty} \mathrm{d}x \, \mathcal{P}(x) \mathcal{A}(x) = 0 \quad (4.27)$$

and the normalization constraint becomes

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) = 1. \tag{4.28}$$

The maximum entropy approach now combines these constraints to show that the probability density function $\mathcal{P}(x)$ has the form of a Gaussian distribution with mean

$$\overline{x} = -\frac{b}{2a}. \tag{4.29}$$

Also, because all variables $\zeta_{m,n}$ are statistically independent, the result is equally valid for the multivariate case (Equation (4.24)). The complete derivation is given in Appendix B of this report and can also be found in Verkley and Lynch (2009) and Verkley (2011).

Our final model is thus:

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + \mathcal{J}^{\mathcal{R}}(\psi^{\mathcal{R}} + \overline{\psi^{\mathcal{U}}}, \zeta^{\mathcal{R}} + \overline{\zeta^{\mathcal{U}}}) = \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}) \tag{4.30a}$$

$$\overline{\zeta^{\mathcal{U}}} = \sum_{m,n\in\mathcal{U}} \overline{\zeta_{m,n}} Y_{m,n} \tag{4.30b}$$

$$\overline{\psi^{\mathcal{U}}} = \sum_{m,n\in\mathcal{U}} \frac{-\overline{\zeta_{m,n}}}{c_{m,n}} Y_{m,n} \tag{4.30c}$$

$$\overline{\zeta_{m,n}} = -\left(\frac{\mathcal{J}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} - \mu F_{m,n}}{2(\nu c_{m,n} + \mu)}\right) \tag{4.30d}$$

4.4 Numerical implementation

We will now discuss the implementation of the new parameterization in the Fortran77 model. The expression for the unresolved scales unfortunately still includes a Jacobian contribution in the unresolved part of the spectrum. Therefore, the simplest way to implement it is to use a full T85 model with a mask function that sets the part of the spectrum outside the T42 truncation to zero. The new tendency routine is visualised in figure 4.3. We are aware that this implementation is likelier to increase computation time than to make it more efficient, but our primary objective is to check whether it works. Later on we may try and find a shortcut in the calculation of the Jacobian.

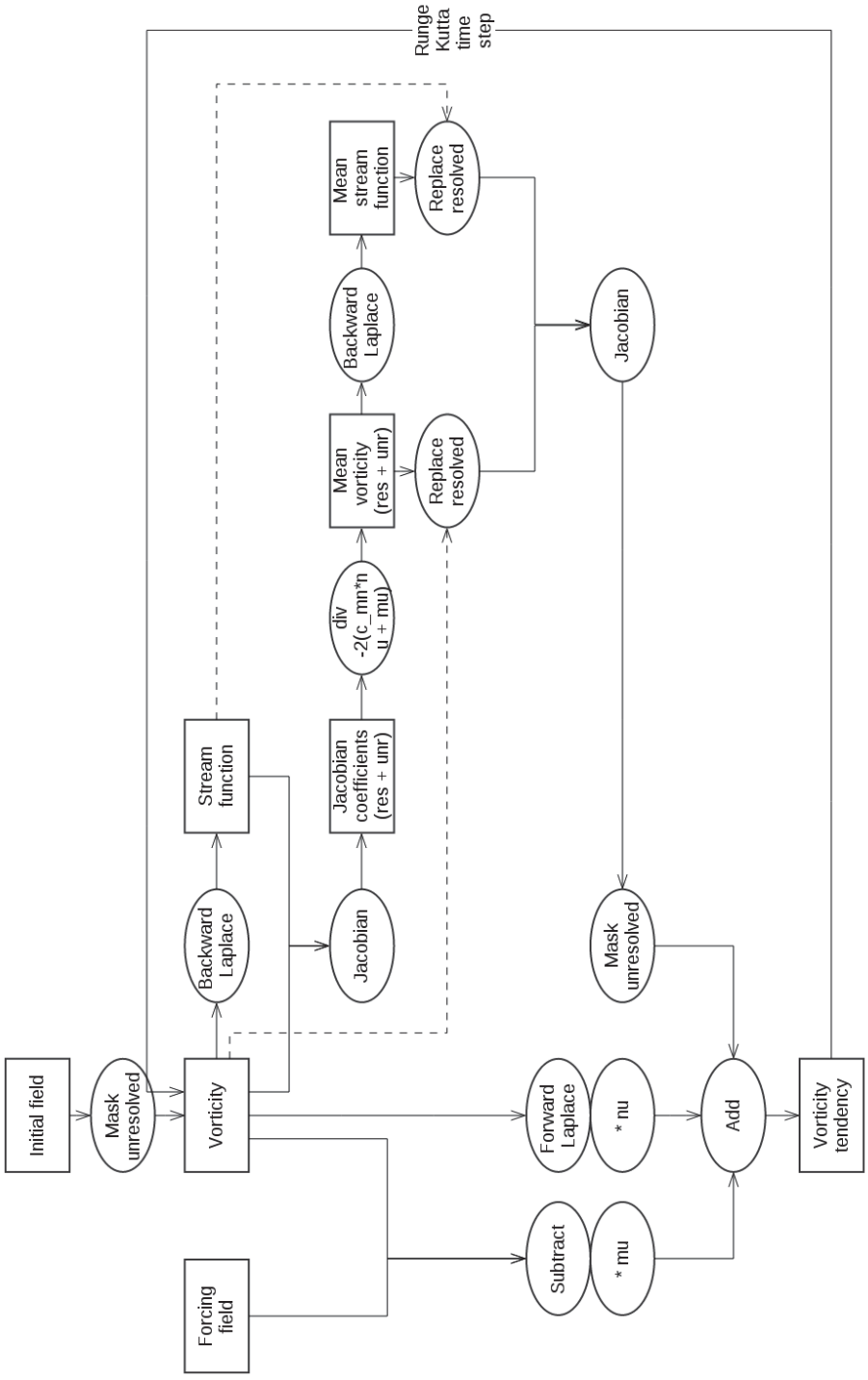


Figure 4.3: Flow chart of the model including maximum entropy parameterization.

5 Methodology

In this chapter, the model settings as used in our experiments are outlined. We do two different kinds of experiments. The first set of experiments is intended to analyse the forecast skill on the short term, i.e. approximately 50 model days[1]. The central question in here is *'How long does it take before the model run deviates substantially from reality?'*[2]

The second set of experiments is used to describe the 'climate' of the model. This climate can be described as a mean state in which the model resides and an associated standard deviation (compare Figure 4.2). The central question in here is *'Does the new parameterization improve the simulation of the state in which the model resides?'* If this is the case, it is likely that also the short range forecast is improved.

5.1 Model parameters

In Table 5.1 we list a number of model parameters as encountered in the Fortran77 model. We use a time step equivalent to 15 minutes, which cor-

[1]Note that there is no relation between model days and real days. We just need a measure to express the time it takes for the model to decorrelate from the reference.

[2]We recall here that reality is given by a model run at high resolution and that this model run is compared with a model run at a lower resolution, with or without a given parameterization.

Table 5.1: Important Fortran77 model parameters with explanation.

LT	Total integration time (e.g. $25 * 96 = 25$ days)
MN	Total spectral domain (85 coefficients in all directions)
MR	Fixed grid resolution (256 grid points in both directions)
MS	Base two exponent for fast fourier transform (default = 7)
MNP	Inner domain: 85 or 42 coefficients (outer domain is masked)
ME	Resolution for energy spectra (default 60 for T42 spectrum)
AF,AE	Forcing and damping (AE=AF=μ)
AY	Viscosity (=ν)
IA	Starting choice: start from zero or read initial field EST
LPAR	Use the maximum entropy parameterization (0 = off, 1 = on)
KE	Interval for writing energy/enstrophy to output files
KC	Interval for writing vorticity field to output files
AN	Resolution-dep. viscosity ($85 * 85$ or $42 * 42$)
SQ	Amplitude of the forcing ($f_{5,5} = \frac{1}{2}\sqrt{2}$).

responds to 900 seconds. To make it non-dimensional[3], we scale it with $T = \Omega^{-1}$

$$\Delta t = 15 * 60 * \Omega = 0.065629036 \tag{5.1}$$

As to the value of the T85 viscosity, we choose a value that damps oscillations on the smallest model scale with an e-folding time[4] of five days. This gives an expression for the non-dimensional viscosity:

$$\nu_{T85} = \frac{1}{24 * 60 * 60} * \frac{1}{\Omega} * \frac{1}{5} * \frac{1}{85^2} \tag{5.2}$$

where we recognize both a conversion from seconds to days for convenience, the scaling parameter $T = \Omega^{-1}$ and the inverse proportionality to the smallest waves. Likewise, the time scale of the linear forcing and damping is 90 days:

$$\mu = \frac{1}{24 * 60 * 60} * \frac{1}{\Omega} * \frac{1}{90} \tag{5.3}$$

The forcing is a single wave with wave number 5 in both directions and amplitude $F_{5,5} = \frac{1}{2}\sqrt{2}$.

[3]Actually, in our model simulations the time step was scaled with a solar day instead of a sidereal day. The time step then becomes 0.065449847, which does not make a substantial difference, but it is in fact a bit improper.

[4]The e-folding time is the time it takes for the oscillations to decrease by a factor of e.

5.2 Experiments

We use a T85 simulation as reference run. We will treat this run as 'reality' and try to approximate it using a model with T42 truncation. First, we run the T42 model with the same viscosity as used in the T85 model. We call this the unparameterized run. However, the original model used a resolution-dependent viscosity. More precisely, the viscosity was inversely proportional to the maximum wave number squared. This adaptive viscosity is a common method to account for the effect of unresolved scales (and numerical instabilities). Sometimes, this increased viscosity is referred to as eddy viscosity (e.g. Cushman-Roisin and Beckers, 2011). We will also use this adaptive viscosity so that we can compare results of the new parameterization with a conventional one. Thus, we perform the following experiments

1. Reference run on T85 resolution
2. Model 1: Unparameterized T42 model
3. Model 2: T42 with resolution-dependent viscosity
4. Model 3: T42 with maximum entropy parameterization

As mentioned before, we do two different kind of analyses. For short-range forecast analysis, the vorticity field is output on a high frequency of two times a day to inspect it in detail. To substantiate the results of this experiment, we perform all the experiments five times with different initial fields (see Section 5.3). For climate analysis, we use energy and enstrophy to characterize the state in which the model resides. A preliminary experiment pointed out that in order to obtain a reliable mean and standard deviation of the system, an integration over at least 5000 model days is desirable.

5.3 Initialization

When the vorticity transport model is initiated from a zero field, the forcing will first lead to a vorticity field increasing in strength. The total energy of the model increases until the flow becomes unstable. This happens after approximately 100 days. At that point, the vorticity field becomes turbulent and the total energy of the field decreases until the forcing and damping reach an equilibrium. This happens after about 250 days. Then, the energy and enstrophy fluctuate about an equilibrium value as is visualized in Figure 5.1. We performed this simulation on a T85 resolution and output the vorticity fields on day 200, 400, 600, 800 and 1000. These fields are used as input fields

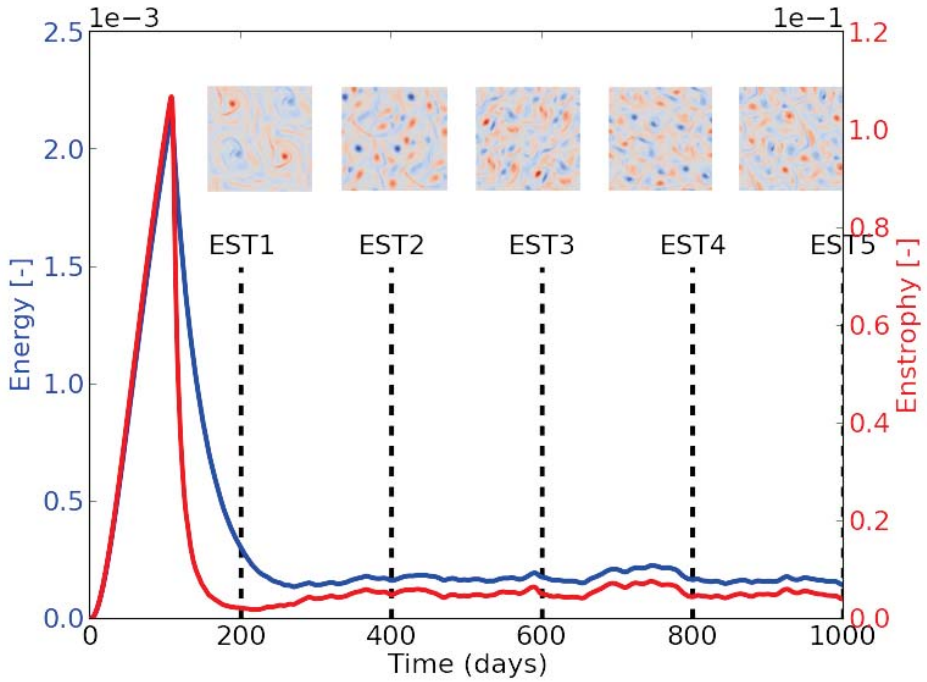


Figure 5.1: Time evolution of total kinetic energy and enstrophy. The energy first increases, then the flow becomes unstable and changes into a turbulent state which evolves to statistical equilibrium. The EST# fields are used as initial fields for the other simulations. For an impression, the EST fields have been added as insets to the figure.

for the five forecast simulations. The field at day 600 is used to initialize the climate simulation. Note that the field at day 200 has not yet reached statistical equilibrium. Still, it is interesting to investigate the effect of our new parameterization for this run as well.

6 Results

This chapter gives an overview of the results of the experiments. We start with a detailed analysis of the simulations with forecast settings. We continue with an analysis of the model climate as expressed by means and variances of the model's energy and enstrophy. Before presenting the results, it must be noted that we encountered several crashes while using the model with the maximum entropy parameterization. To avoid this numerical instability, we decreased the time step for these runs by a factor of 3.

6.1 Forecast simulations

6.1.1 Qualitative analysis of vorticity fields

As mentioned in Chapter 5, we ran four different models with five different initial fields (Set 1-5), totalling 20 simulations in forecast mode. Following initial field number 3, the calculated vorticity fields after 10, 20, 30, 40, 50 and 60 days are shown in Figures 6.1 and 6.2.

A feature that stands out is the grainy texture of the flow field in the unparameterized run. There are large gradients between adjacent grid cells at the smallest scales of the model (near the truncation limit). This behaviour seems unphysical and a possible explanation might be that the smallest waves of the spectrum are exaggerated. Both parameterizations reduce this noise, especially the conventional parameterization is very effective and rigorously

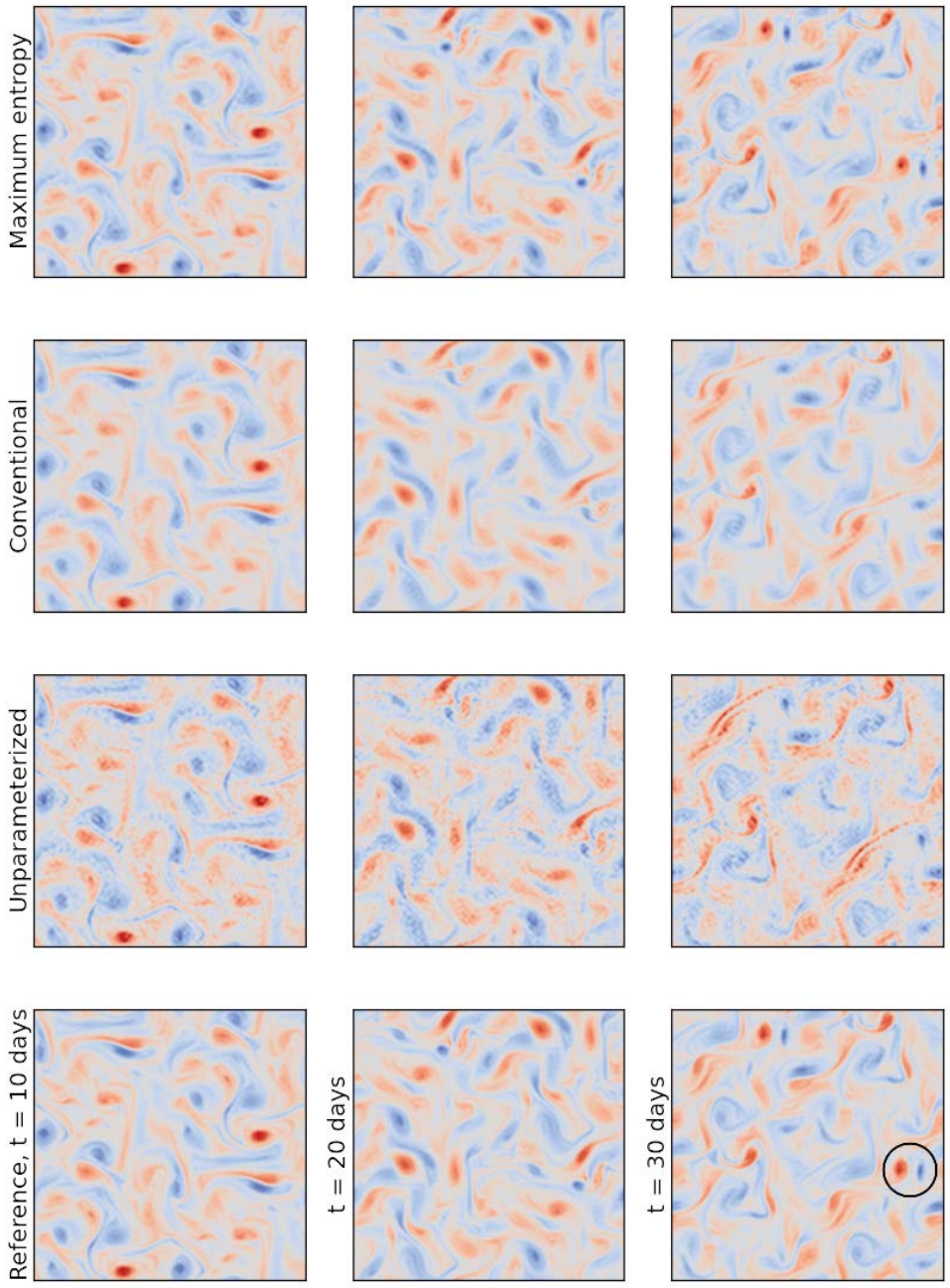


Figure 6.1: Vorticity fields from model simulations at time intervals of 10 days up to 60 days (continued on page 48). A T42 truncation has been used to plot the vorticity fields, also for the reference.

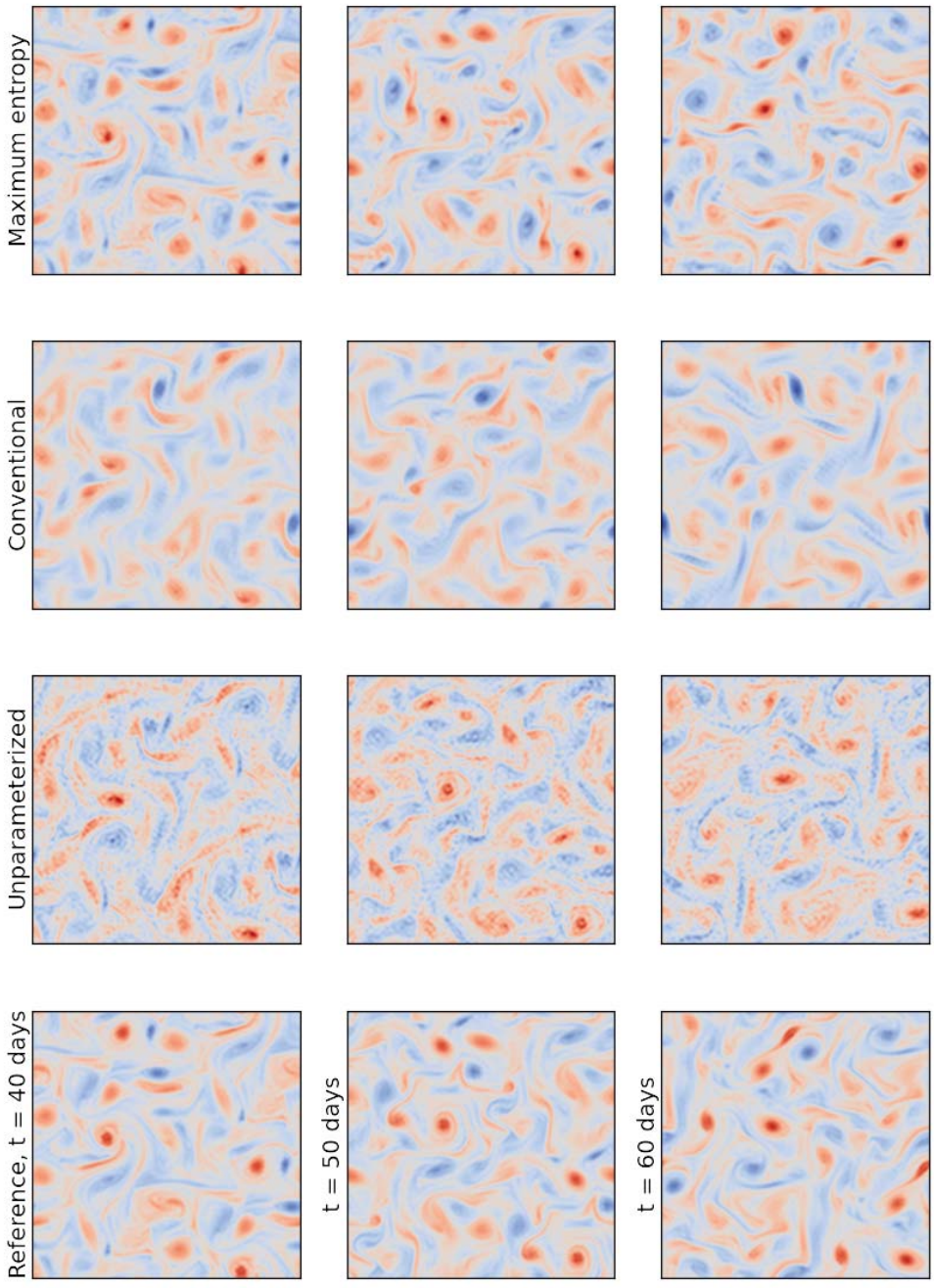


Figure 6.2: Vorticity fields from model simulations at time intervals of 10 days up to 60 days (continued).

damps all the small scale effects. The maximum entropy parameterization is not completely smooth, but it is still a considerable improvement over the unparameterized run. We also note that the conventional parameterization tends to damp out the extremes in the vorticity fields. In contrast, the maximum entropy parameterization realistically retains the vorticity minima and maxima.

If we analyse the vorticity fields in the course of time and focus on individual vortices, we see that at day 10 all fields are still more or less the same. After 20 days, the first discrepancies can be identified, but they are more clear after 30 days. We might focus, for example, on the two small, strong vortices of opposite sign near the bottom boundary (circled in the plot of the reference run). They are nearly identical in the reference and maximum entropy simulations, whereas the other two runs fail to reproduce this feature. At day 50 the maximum entropy simulation still closely resembles the reference run, whereas the conventional parameterization has damped out most small scale features. At day 60, most of the correlation with the reference run is lost for all parameterizations.

6.1.2 Quantitative analysis of vorticity fields

There are several measures to quantify model performance. Two indices that are frequently applied are correlation and root mean square difference (RMSD). Correlation is a measure for the similarity between two fields, whereas RMSD is a measure for the difference between two fields. We obtained time series of these quantities by directly comparing the vorticity values of each grid point of the model to the same grid point in the reference run. Representing the model output values with x_i and the corresponding cells of the reference run with y_i, the correlation and RMSD are given, respectively, by

$$corr = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \overline{y})^2}} \tag{6.1}$$

$$RMSD = \sqrt{\sum_{i=1}^{N}\frac{1}{N}(x_i - y_i)^2} \tag{6.2}$$

where N is the total number of grid cells and the overbars denote the mean value of the field. The results for Set 3 are shown in Figure 6.3. All models

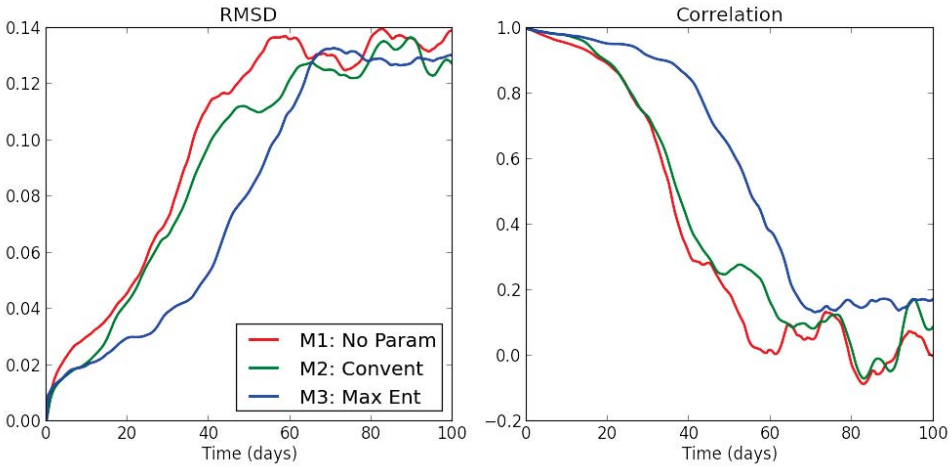


Figure 6.3: Time evolution of RMSD (left) and correlation (right) with respect to reference run for three models, started from initial field number 3.

show a steep rise in RMSD and a decrease in correlation within the first 70 days. In line with our previous qualitative analysis, we see that both the rise in RMSD and the drop in correlation come about 20 days later in the maximum entropy simulation than in the other runs. To substantiate these results, we repeated the experiment for all initial states mentioned in Figure 5.1 and averaged over the small ensemble. The results are shown in Figure 6.4. Again, the ensemble average of the maximum entropy simulation shows a later increase in RMSD and a later decrease in correlation, albeit slightly less pronounced than in Set 3 alone. There is one set that deviates slightly from the other sets. This is the simulation that follows from initial field number 1. Here, the initialisation run had not reached statistical equilibrium yet. However, also for this set, the maximum entropy parameterization improves the model performance. Also note that in terms of correlation, the conventional parameterization does not seem to improve the simulation at all.

6.1.3 Model performance in a Taylor diagram

There is a trigonometric relation between RMSD and correlation that allows us to plot them in a single graph, called a Taylor diagram, after its inventor (Taylor, 2001). Such a graph is shown in Figure 6.5 for both Set 3 and the ensemble mean. In these plots, the reference is indicated by a star. In

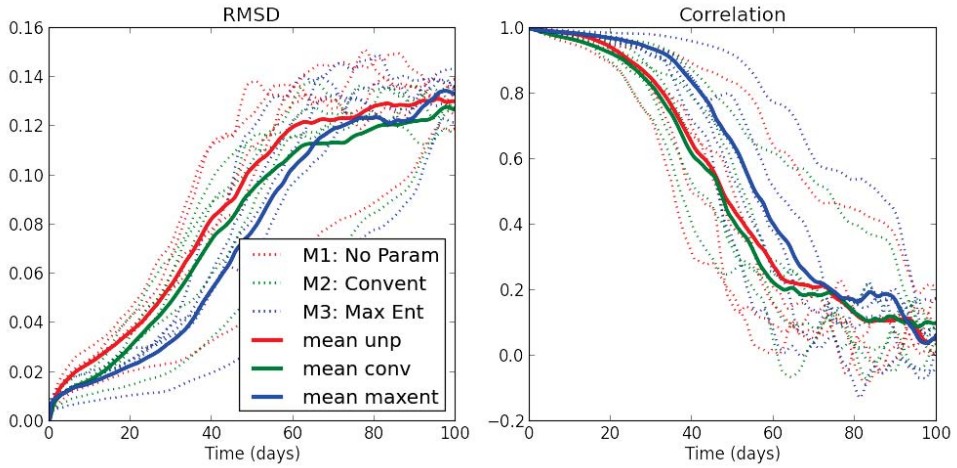


Figure 6.4: Time evolution of (mean) RMSD (left) and correlation (right) with respect to reference run for three models, started from five initial fields.

the course of time, the models start to deviate from the reference and the absolute distance to the reference star becomes larger. If we compare points from different models at the same time, the model that is closest to the reference point has the best score. For example, after 30 days of integration in the diagram for Set 3, the maximum entropy simulation is found near the 0.04 RMSD radius, whereas the other two models are closer to the 0.08 RMSD radius. This indicates that the maximum entropy parameterization improves the accuracy of the simulation. Focusing on the ensemble mean, it can be seen that the maximum entropy simulation had approximately the same accuracy at day 40 as the other runs had after 30 days of integration. This again indicates a substantial improvement of the model performance.

6.1.4 Energy and enstrophy

Instead of looking at the vorticity fields, we can also look at the total energy and enstrophy of the flow. These are shown in Figure 6.6. The most striking aspect of these graphs is the poor quality of the conventional parameterization. It is immediately clear that the extra damping introduced by the higher viscosity is far too large, resulting in too low energy and enstrophy values of the field. In contrast, the maximum entropy parameterization closely resembles the reference run up to 60 days of integration.

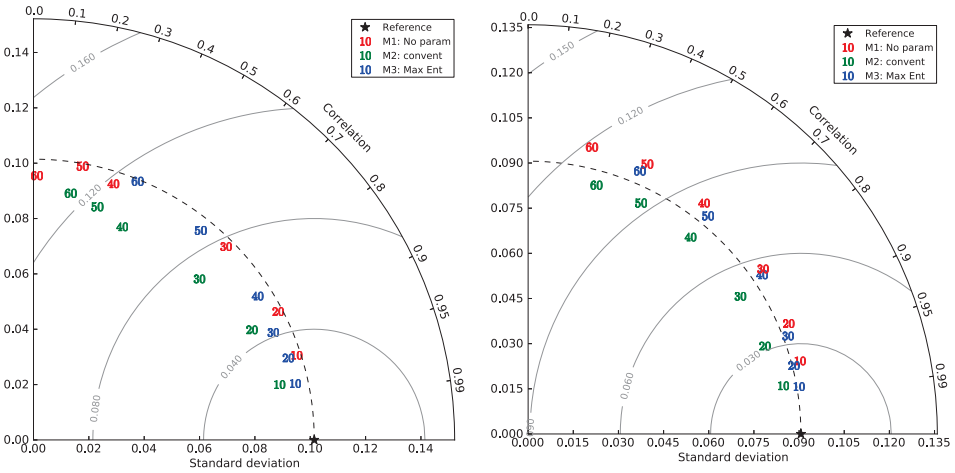


Figure 6.5: Taylor diagrams of the runs initiated from initial condition 3 (set 3, left) and of the ensemble mean of the runs from initial conditions 1-5, right. Numbers represent days after the start of the simulation. Radial grid lines are RMSD values.

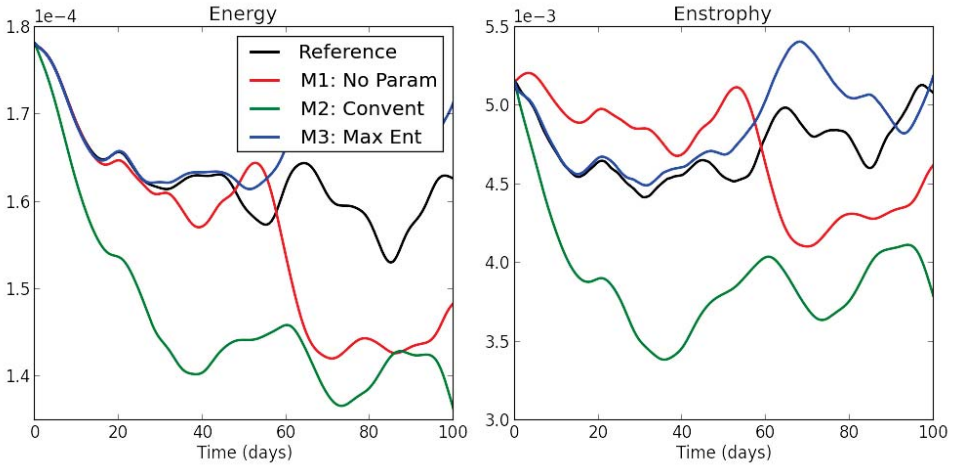


Figure 6.6: Time evolution of kinetic energy (left) and enstrophy (right) for simulations of Set 3.

6.2 Climate

6.2.1 Energy and enstrophy distributions

We could continue by showing an ensemble mean of the energy and enstrophy evolution. However, we believe that it is more informative to present the statistics of these variables that result from a very long integration. After initialisation, the energy and enstrophy will start to fluctuate around an equilibrium value as in Figure 4.2. This equilibrium can be characterized by a mean value and a variance, which can in turn be used to construct a normal distribution curve representing the probability density function for the total energy and enstrophy[1]. These distribution curves are shown in Figure 6.7. It appears that the conventional parameterization results in a model climate with too little energy and enstrophy, as was already indicated by the results of the previous section. The mean energy in the unparameterized run is quite similar to the reference run, but the distribution is somewhat broader. The enstrophy in the unparameterized run is overestimated. The near coincidence of the maximum entropy and reference distributions of both energy and enstrophy signifies that the maximum entropy parameterization substantially improves the modelled climate.

6.2.2 Energy and enstrophy spectra

Finally, we consider the energy and enstrophy spectra. The energy and enstrophy are calculated per wave number and averaged over the climate simulation. Figure 6.9 illustrates how the spectra are constructed. The rectangle represents the spectral space. Each combination of m and n forms a wave vector $\mathbf{k}$ with length (magnitude) $\sqrt{m^2+n^2}$: the wave number. The energy/enstrophy spectrum is constructed by binning waves with approximately the same wave number and computing the energy/enstrophy of all the individual bins. When the wave number of the bins exceeds N (the first dotted line in Figure 6.8), there are less coefficients, simply because they fall outside the truncation that the model used for the integration. Outside $\sqrt{2}N$ there is no data at all (the second dotted line in Figure 6.8).

The results are displayed in Figure 6.8 and show that the energy and enstrophy in the larger scales are nearly identical in all simulations, but on the smaller scales, substantial differences occur. The unparameterized run has too much energy in the smallest scales. This finding supports our hypothesis

[1]We checked that after a long integration as used here (5000 days), the histograms of energy and enstrophy indeed converged towards the bell-shape characteristic for a normal distribution.

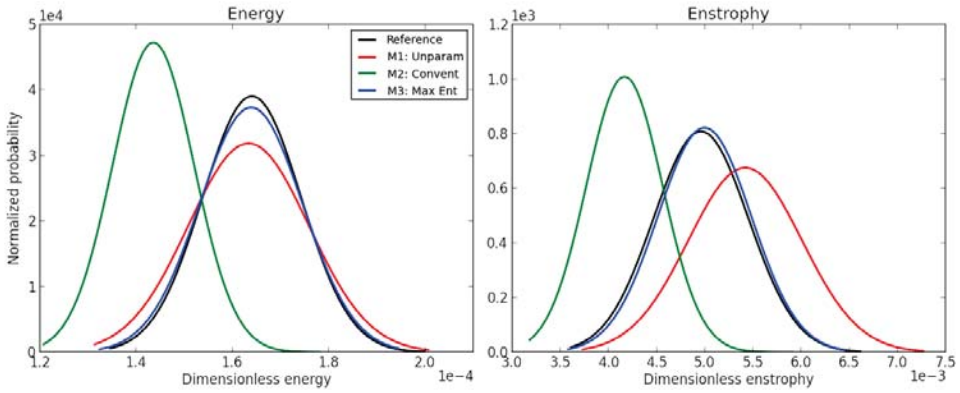


Figure 6.7: Gaussian distributions based on mean and variance of kinetic energy and enstrophy from climate simulations with four different models.

that the grainy texture of the flow field is a result of excessive representation of the smallest waves. The conventional parameterization imposes a damping that is too large. Again, the maximum entropy parameterization improves the simulation: the representation of the spectra is fairly accurate, even in the smallest scales. This substantiates the results presented in Figure 6.7.

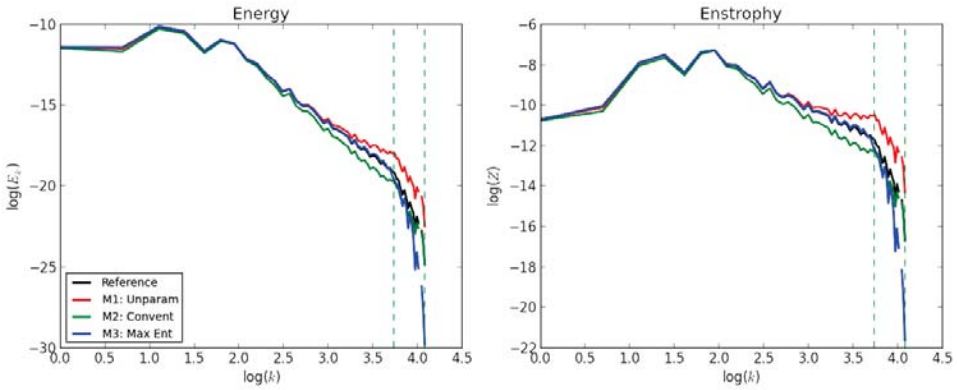


Figure 6.8: Energy and enstrophy spectra of the flow field as simulated with four different models. The dotted vertical lines indicate two truncation levels, as illustrated in Figure 6.9.

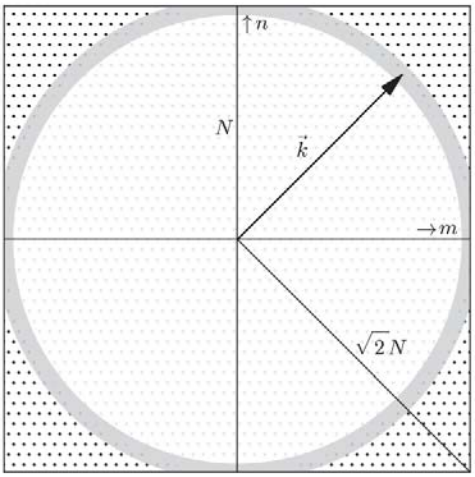


Figure 6.9: Illustration of the calculation of the spectra. See text for explanation.

7

Discussion and conclusions

We have implemented a new parameterization for the effects of unresolved scales on the resolved scales in a simple model of two-dimensional fluid flow. The parameterization is based on the maximum entropy principle, which states that the information entropy of the probability density function that describes the unresolved scales should be maximal under suitable constraints. We used the constraints of normalization and the assumption that the energy in the unresolved scales should have a time derivative that is equal to zero. The results so far seem quite promising. The parameterization leads to a physically realistic flow field without affecting the intensity of the strongest vortices (Figures 6.1-6.2). Compared to a model that uses a conventional parameterization, the forecasting skill of the maximum entropy model (analyzed in terms of the root mean square difference and correlation with the reference run) is substantially higher. This is illustrated in Figures 6.3-6.5. We also performed a climate simulation experiment. The maximum entropy parameterization improves the representation of energy and enstrophy of the flow (Figures 6.7-6.8). Especially near the truncation limit, the energy and enstrophy spectra are more realistic. In this chapter, we will make some side notes to the new parameterization and we look ahead to what the next steps might be.

To investigate the effect of the new parameterization when it operates in isolation, we performed a simulation without any forcing or damping, except for the parameterization of the unresolved scales. The energy of this run remains nearly constant while the enstrophy decreases to a new, much

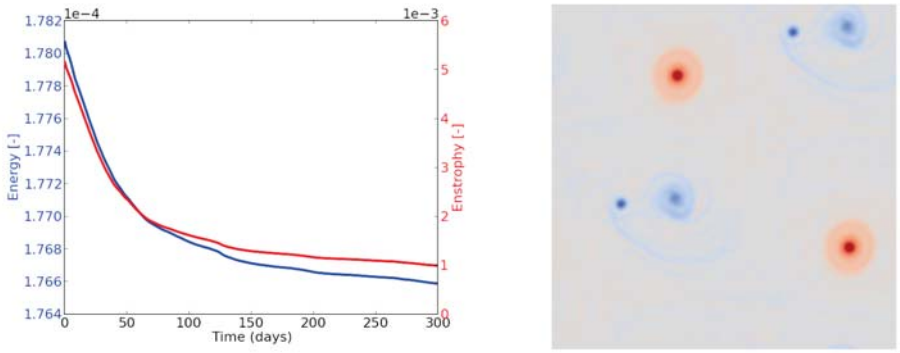


Figure 7.1: Energy and enstrophy of the flow field for a simulation of 300 days. Started from EST3, ran without forcing and linear damping and without viscosity (left). The only remaining source/sink is the maximum entropy parameterization. Note the small range of the energy-axis. Right: final vorticity field of this simulation.

lower equilibrium, as shown in Figure 7.1. The vorticity field converges to a state which is dominated by only a few large vortices. These results are comparable to the decaying two-dimensional flow solution of McWilliams (1984) and confirm that the new parameterization indeed acts as a damping on the smallest scales, such that all energy will concentrate in the larger scales. This is an essential property of two-dimensional turbulence. In three-dimensional turbulence, energy will flow continuously towards the smaller scales (Bouchet and Venaille, 2012).

We note that the formulation of our parameterization is very similar to the anticipated potential vorticity method (AVPM) introduced by Sadourny and Basdevant (1985). Their Equation (8) expresses the damping resulting from subgrid-scale processes as

$$D = \theta r^{-1/2} L(r^{-1/2}\mathbf{V} \cdot \nabla\eta), \tag{7.1}$$

which for our model would reduce to

$$D = \theta L \mathcal{J}(\psi, \zeta) \tag{7.2}$$

Here, θ is a time scale and L is a 'nondimensional nonnegative definite linear operator'. If we assume the contribution of $\overline{\psi^U}$ is negligible, our Equations (4.30) can be rewritten in exactly the same form. However, whereas Sadourny and Basdevant have to make an educated guess for the value of their θ and L, the maximum entropy approach completely determines these

parameters:

$$\theta = \frac{1}{2\mu} \quad \text{and} \quad L = P^U \left(1 - \frac{\nu}{\mu}\nabla^2\right)^{-1}, \tag{7.3}$$

where P^U is a projection operator used to indicate that only spectral coefficients in the unresolved domain are considered. Frederiksen *et al.* (1996) tested the APVM parameterization for a barotropic vorticity model (their Section 6.2). Our results show similar properties to theirs, in the sense that enstrophy is dissipated while the mean energy is nearly unaffected.

With the current implementation, the maximum entropy parameterization does not improve the forecast speed. Instead, the simulations take (much[1]) longer. The reason that the Jacobian term is so expensive in the spectral model is because it requires a transformation to and from the grid space. There are also models that compute everything in grid space, in which case this drawback is smaller. For the spectral model, there are several ways in which the speed can be boosted. As shown in Figure 4.3, the model now goes through the Jacobian routine twice. The first time, only the resulting Jacobian coefficients in the unresolved domain are used and those of the resolved domain are discarded. It is not too difficult to rewrite the Fortran code in such a way that only the unresolved variables are returned when calculating the first Jacobian. This would already lead to a gain in computation time.

A much more effective method, however, is to let only a few waves, that are *just* outside the truncation limit, interact with the resolved variables. To explain this more clearly, we return to Figure 3.2. The non-linear Jacobian operator, when applied to T42 variables, also has a contribution in the T85 part of the spectrum. In the unparameterized T42 model, we simply neglect these contributions. In the maximum entropy model, we average them with Equation (4.30d) and account for them when going through the Jacobian for the second time (the T85 waves that result from the second call of the Jacobian routine, are discarded anyway). However, we could also discard everything outside, say, a T50 truncation and only apply Equation (4.30d) to the coefficients between T42 and T50. We ran some test simulations and it appears that using only 3 or 5 extra rings of coefficients (T42+3, T42+5) already leads to a comparable gain in accuracy. Combining these two modifications leads us to believe that we will be able to run the model on T42 resolution and speed while maintaining a T85 accuracy.

As mentioned in the beginning of Chapter 6, we encountered some numerical instabilities with the new parameterization, forcing us to reduce the time

[1] approximately twice as long.

step with a factor 3. These instabilities look like typical Courant-Friedrichs-Lewy (CFL) violations, but further investigation of this issue is necessary.

The reader may have noticed a certain symmetry in the vorticity fields shown in this report. It probably results from the simple, regular forcing pattern that we prescribed for our simulations. Indeed, it would be better to use a more sophisticated forcing (e.g. a stochastic forcing acting on multiple scales) but we leave this to future studies on the subject.

One of such studies would be to extend our results to a shallow-water or reduced-gravity model. This would allow the study of the effects of divergence, e.g. vortex stretching, and also to implement the latitudinal dependence on the earth's rotation through the Coriolis parameter.

Appendices

A

Fourier transformations

We will prove here that expressing a field $\zeta(x, y, t)$ as in Equation (3.37) can be done by using complex exponentials. If we consider only one dimension, we split the summation in (3.36) in three and substitute the real expressions for $X_m(x)$:

$$\zeta(x) = \sum_{m=-\infty}^{+\infty} \zeta_m X_m(x) \tag{A.1}$$

$$= \sum_{m=-\infty}^{-1} \zeta_m X_m(x) + \sum_{m=0} \zeta_m X_m(x) + \sum_{m=1}^{+\infty} \zeta_m X_m(x) \tag{A.2}$$

$$= \sum_{m=1}^{+\infty} \zeta_{-m}\sqrt{2}\sin(mx) + \zeta_0 + \sum_{m=1}^{+\infty} \sqrt{2}\zeta_{+m}\cos(mx) \tag{A.3}$$

Now, we use

$$e^{imx} = \cos(mx) + i\sin(mx) \tag{A.4}$$

$$\cos(-x) = \cos(x) \tag{A.5}$$

$$\sin(-x) = -\sin(x) \tag{A.6}$$

to show that

$$\begin{aligned} e^{imx} + e^{-imx} &= \cos(mx) + i\sin(mx) + \cos(-mx) + i\sin(-mx) \\ &= \cos(mx) + i\sin(mx) + \cos(mx) - i\sin(mx) \qquad \text{(A.7)} \\ &= 2\cos(mx) \\ e^{imx} - e^{-imx} &= \cos(mx) + i\sin(mx) - \cos(-mx) - i\sin(-mx) \\ &= \cos(mx) + i\sin(mx) - \cos(mx) + i\sin(mx) \qquad \text{(A.8)} \\ &= 2i\sin(mx) \end{aligned}$$

Next, we replace the sine and cosine in equation (A.1) with the expressions above, to find:

$$\zeta(x) = \zeta_0 + \sum_{m=1}^{\infty} \zeta_{+m}(e^{imx} + e^{-imx})/\sqrt{2} \qquad \text{(A.9)}$$

$$+ \sum_{m=1}^{+\infty} \zeta_{-m}(e^{imx} - e^{-imx})/(\sqrt{2}i) \qquad \text{(A.10)}$$

Taking the two summations together yields

$$\zeta(x) = \zeta_0 + \sum_{m=1}^{\infty} \frac{1}{\sqrt{2}}(\zeta_{+m} - \zeta_{-m}i)e^{imx} + \frac{1}{\sqrt{2}}(\zeta_{+m} + \zeta_{-m}i)e^{-imx} \qquad \text{(A.11)}$$

where we used that $(\frac{1}{i}) = (-i)$. Now, we define the complex coefficient $\widehat{\zeta}_m$ as

$$\widehat{\zeta}_{+m} = (\zeta_{+m} - \zeta_{-m}i)/\sqrt{2} \qquad \text{(A.12)}$$

$$\widehat{\zeta}_0 = \zeta_0 \qquad \text{(A.13)}$$

$$\widehat{\zeta}_{-m} = (\zeta_{+m} + \zeta_{-m}i)/\sqrt{2} \qquad \text{(A.14)}$$

so we can replace the loop over m to ∞ with a loop over $-\infty$ to $+\infty$ and drop out the terms with negative exponentials. Then we are left with the one-dimensional, complex exponential form of equation (3.36).

The two-dimensional equivalent of this is:

$$
\begin{aligned}
\zeta(x,y) = \sum_{m,n} \\
\frac{1}{2}\mathrm{e}^{imx}\mathrm{e}^{iny}(-\zeta_{--} - i\zeta_{-+} - i\zeta_{+-} + \zeta_{++}) + \\
\frac{1}{2}\mathrm{e}^{imx}\mathrm{e}^{-iny}(\zeta_{--} - i\zeta_{-+} + i\zeta_{+-} + \zeta_{++}) + \\
\frac{1}{2}\mathrm{e}^{-imx}\mathrm{e}^{iny}(\zeta_{--} + i\zeta_{-+} - i\zeta_{+-} + \zeta_{++}) + \\
\frac{1}{2}\mathrm{e}^{-imx}\mathrm{e}^{-iny}(-\zeta_{--} + i\zeta_{-+} + i\zeta_{+-} + \zeta_{++}) + \\
\frac{1}{\sqrt{2}}\mathrm{e}^{imx}\ (-i\zeta_{-0} + \zeta_{+0}) + \\
\frac{1}{\sqrt{2}}\mathrm{e}^{-imx}(i\zeta_{-0} + \zeta_{+0}) + \\
\frac{1}{\sqrt{2}}\mathrm{e}^{iny}(-i\zeta_{0-} + \zeta_{0+}) + \\
\frac{1}{\sqrt{2}}\mathrm{e}^{-iny}(i\zeta_{0-} + \zeta_{0+}) + \\
\zeta_{00}
\end{aligned}
\tag{A.15}
$$

where the subscript ζ_{--} represents the coefficients $\zeta_{-m,-n}$, etc., so that the complex version of the spectral coefficient is defined as:

$$\widehat{\zeta}_{++} = (-\zeta_{--} - i\zeta_{-+} - i\zeta_{+-} + \zeta_{++})/2 \tag{A.16}$$
$$\widehat{\zeta}_{+-} = (\zeta_{--} - i\zeta_{-+} + i\zeta_{+-} + \zeta_{++})/2 \tag{A.17}$$
$$\widehat{\zeta}_{-+} = (\zeta_{--} + i\zeta_{-+} - i\zeta_{+-} + \zeta_{++})/2 \tag{A.18}$$
$$\widehat{\zeta}_{--} = (-\zeta_{--} + i\zeta_{-+} + i\zeta_{+-} + \zeta_{++})/2 \tag{A.19}$$
$$\widehat{\zeta}_{+0} = (-i\zeta_{-0} + \zeta_{+0})/\sqrt{2} \tag{A.20}$$
$$\widehat{\zeta}_{-0} = (i\zeta_{-0} + \zeta_{+0})/\sqrt{2} \tag{A.21}$$
$$\widehat{\zeta}_{0+} = (-i\zeta_{0-} + \zeta_{0+})/\sqrt{2} \tag{A.22}$$
$$\widehat{\zeta}_{0-} = (i\zeta_{0-} + \zeta_{0+})/\sqrt{2} \tag{A.23}$$
$$\widehat{\zeta}_{00} = \zeta_{00} \tag{A.24}$$

This will be used in for the Fast-Fourier transform in the Fortran program.

B

Maximum entropy theory

We may illustrate how the principle of maximum entropy is applied by considering a simple case in which there is only one unresolved variable; this variable will be called x. Denoting $-\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t}$ by $\mathcal{A}$, the expression (4.24) then becomes:

$$\mathcal{A} = ax^2 + bx \text{ , with} \tag{B.1a}$$

$$a = \left(\nu + \frac{\mu}{c_{m,n}}\right) \text{ and} \tag{B.1b}$$

$$b = \left(\frac{J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})_{m,n} - \mu f_{m,n}}{c_{m,n}}\right) \tag{B.1c}$$

The information entropy in the case of a single variable x is given by

$$S_I = -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln\left(\frac{\mathcal{P}(x)}{\mathcal{M}(x)}\right), \tag{B.2}$$

where $\mathcal{P}(x)$ is the probability density function. For the so-called a-priori probability density function $\mathcal{M}(x)$, we take $\mathcal{M}(x) = \frac{1}{c}$, where c is a constant with the same dimension as x. This means, effectively, that we assume no

a-priori information. We then get:

$$S_I = -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln\left[\mathcal{P}(x)c\right] \tag{B.3a}$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \left[\ln(\mathcal{P}(x)) + \ln(c)\right] \tag{B.3b}$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln(\mathcal{P}(x)) - \int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln(c) \tag{B.3c}$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln(\mathcal{P}(x)) - \ln(c) \int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \tag{B.3d}$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \ln(\mathcal{P}(x)) - \ln(c) \tag{B.3e}$$

The last result follows from the fact that we always require

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) = 1. \tag{B.4}$$

We see that the value of c only leads to an offset of S_I. In maximizing S_I we may just as well take $c = 1$.

The principle of maximum entropy states that S_I should be maximum under suitable constraints. The first constraint is the normalization condition (B.4). The second constraint is

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)\mathcal{A}(x) = 0, \tag{B.5}$$

i.e., that the average value of $\mathcal{A}$ (averaged over $\mathcal{P}$) is equal to zero. Let us

define:

$$\mathcal{B}(\mathcal{P}) \equiv \int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x) \tag{B.6a}$$

$$\mathcal{C}(\mathcal{P}) \equiv \int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)\mathcal{A}(x) \tag{B.6b}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)(ax^2 + bx). \tag{B.6c}$$

The condition that S_I is maximal (or, more generally, extremal) can be expressed as:

$$\delta S_I = 0, \tag{B.7}$$

where δS_I is a perturbation of S_I as a result of any allowable perturbation $\delta\mathcal{P}$. Now, we require that the perturbation of S_I is zero under the constraints that the normalization remains 1 (also the perturbed probability density function has to be normalized) and that the average value of $\mathcal{A}$ remains 0. This implies that the perturbations that are allowed should satisfy

$$\delta\mathcal{B} = 0, \tag{B.8a}$$

$$\delta\mathcal{C} = 0. \tag{B.8b}$$

The trick is now to find values of λ_1 and λ_2 such that for *any* perturbation (i.e. without constraint) we have

$$\delta S_I + \lambda_1\delta\mathcal{B} + \lambda_2\delta\mathcal{C} = 0. \tag{B.9}$$

If values of λ_1 and λ_2 (nonzero) can be found then it follows that if we limit ourselves to perturbations for which $\delta\mathcal{B} = 0$ and $\delta\mathcal{C} = 0$, then also $\delta S_I = 0$. This is the of Lagrange multipliers. In our case we have two Lagrange multipliers: λ_1 and λ_2. Now, we have:

$$\delta S_I = -\int_{-\infty}^{\infty} \mathrm{d}x\, \delta\mathcal{P}(x)\ln(\mathcal{P}(x)) - \int_{-\infty}^{\infty} \mathrm{d}x\, \delta\mathcal{P}(x), \tag{B.10a}$$

$$\delta\mathcal{B} = -\int_{-\infty}^{\infty} \mathrm{d}x\, \delta\mathcal{P}(x), \tag{B.10b}$$

$$\delta\mathcal{C} = -\int_{-\infty}^{\infty} \mathrm{d}x\, \delta\mathcal{P}(x)(ax^2 + b), \tag{B.10c}$$

so that (for convenience we changed the signs of λ_1 and λ_2)

$$\delta S_I - \lambda_1 \delta \mathcal{B} - \lambda_2 \delta \mathcal{C} = - \int_{-\infty}^{\infty} \mathrm{d}x \, \delta \mathcal{P}(x) \left[\ln(\mathcal{P}(x)) + 1 + \lambda_1 + \lambda_2 (a x^2 + b x) \right] . \quad \text{(B.11)}$$

If this has to be zero for any perturbation $\delta\mathcal{P}$, then the expression between square brackets has to be zero. This implies:

$$\ln(\mathcal{P}(x)) + 1 + \lambda_1 + \lambda_2 (a x^2 + b x) = 0 \Rightarrow \quad \text{(B.12a)}$$

$$\ln(\mathcal{P}(x)) = -1 - \lambda_1 - \lambda_2 (a x^2 + b x) \Rightarrow \quad \text{(B.12b)}$$

$$\mathcal{P}(x) = \exp(-1 - \lambda_1 - \lambda_2 (a x^2 + b x)) \quad \text{(B.12c)}$$

$$= \exp(-1 - \lambda_1) \exp(-\lambda_2 (a x^2 + b x)) \quad \text{(B.12d)}$$

$$= \exp(-1 - \lambda_1) \exp\left(-\lambda_2 a \left(x^2 + \frac{b}{a} x \right) \right) \quad \text{(B.12e)}$$

$$= \exp(-1 - \lambda_1) \exp\left(-\lambda_2 a \left[\left(x + \frac{b}{2a} \right)^2 - \left(\frac{b}{2a} \right)^2 \right] \right) \quad \text{(B.12f)}$$

$$= \exp(-1 - \lambda_1) \exp\left(\lambda_2 a \left(\frac{b}{2a} \right)^2 \right) \exp\left(-\lambda_2 a \left(x + \frac{b}{2a} \right)^2 \right) . \quad \text{(B.12g)}$$

Now, compare this expression with the expression of the normal (Gaussian) distribution:

$$\mathcal{N}(\mu, \sigma, x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2} \right) , \quad \text{(B.13)}$$

with mean μ and standard deviation σ. Then we see that the maximum entropy probability density function is equal to a Gaussian distribution with mean

$$\mu = -\frac{b}{2a} \quad \text{(B.14)}$$

and variance

$$\sigma^2 = \frac{1}{2\lambda_2 a} \quad \text{(B.15)}$$

We see that the mean of the distribution is already given in terms of the system parameters a and b. To obtain the standard deviation σ, we note

that for a normal distribution we have

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)x = \mu, \tag{B.16a}$$

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)x^2 = \mu^2 + \sigma^2. \tag{B.16b}$$

We now use the constraint:

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)(ax^2 + b) = 0 \Rightarrow \tag{B.17a}$$

$$a\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)x^2 + b\int_{-\infty}^{\infty} \mathrm{d}x\, \mathcal{P}(x)x = 0 \Rightarrow \tag{B.17b}$$

$$a(\mu^2 + \sigma^2) + b\mu = 0. \tag{B.17c}$$

This gives, after substituting the known value of μ,

$$a\left(\left(\frac{b}{2a}\right)^2 + \sigma^2\right) - b\left(\frac{b}{2a}\right) = 0 \Rightarrow \tag{B.18a}$$

$$\left(\frac{b}{2a}\right)^2 + \sigma^2 - \frac{b^2}{2a^2} = 0 \Rightarrow \tag{B.18b}$$

$$-\left(\frac{b}{2a}\right)^2 + \sigma^2 = 0 \Rightarrow \tag{B.18c}$$

$$\sigma^2 = \left(\frac{b}{2a}\right)^2 \tag{B.18d}$$

Now σ is known, we know λ_2. The value of λ_1 can be obtained by using that

$$\exp(-1-\lambda_1)\exp\left(\lambda_2 a\left(\frac{b}{2a}\right)^2\right) = \frac{1}{\sigma\sqrt{2\pi}} \tag{B.19}$$

This completely determines the maximum entropy probability density function. The calculation above can be easily extended to the multivariate (more variables) case as long as all the variables are statistically independent. We are, actually, in this situation because the expression for$-\frac{\mathrm{d}E^{\mathcal{U}}}{\mathrm{d}t}$ (see (4.24)) does not contain any cross terms. This implies that if we would apply our

maximum entropy analysis to this case, the resulting probability density function would be a product of independent Gaussians with means and variances that are determined, individually, by $A_{m,n}$ and $B_{m,n}$.

Intuitively, we can also see that constraint (B.5) leads to a nonzero mean of $\mathcal{P}(x)$. Since the integral over $\mathcal{P}(x) = 1$, the expression for $\mathcal{A}$ should 'compensate' to get zero. When the PDF has a Gaussian form, and $\mathcal{A}$ is a quadratic function, this is possible when the mean of $\mathcal{P}$ coincides with the location where the quadratic expression $\mathcal{A} = 0$ has its minimum (see figure B.1). For the minimum of a quadratic function, we have

$$\frac{\partial}{\partial x}\left[ax^2 + bx\right] = 0 \tag{B.20}$$

$$2ax = -b \tag{B.21}$$

$$x = -\frac{b}{2a} \tag{B.22}$$

which is equivalent to Equation (B.13)

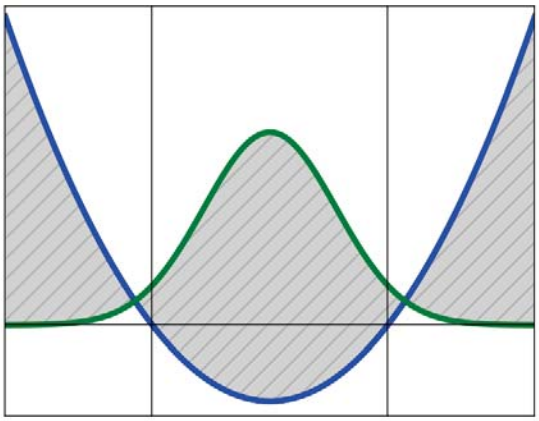


Figure B.1: Shape of the probability density function (green) and the quadratic function (blue). The product of the two functions is negative where the quadratic function is negative, and positive elsewhere. With some imagination, one can see that under certain conditions, the positive and negative products cancel and condition (4.13) is met. This is only possible if the locations of the extremes coincide.

Bibliography

Bihlo A, Cardoso-Bihlo EDS, Popovych RO. 2014. Invariant parameterization and turbulence modeling on the beta-plane. *Physica D: Nonlinear Phenomena* **269**: 48–62.

Bouchet F, Venaille A. 2012. Statistical mechanics of two-dimensional and geophysical flows. *Physics reports* **515**(5): 227–295.

Boyd JP. 2001. *Chebyshev and fourier spectral methods*. Courier Corporation.

Christensen H, Moroz I, Palmer T. 2014. Simulating weather regimes: impact of stochastic and perturbed parameter schemes in a simple atmospheric model. *Climate Dynamics* : 1–20.

Coiffier J. 2011. *Fundamentals of numerical weather prediction*. Cambridge University Press.

Crommelin D, Vanden-Eijnden E. 2008. Subgrid-scale parameterization with conditional markov chains. *Journal of the Atmospheric Sciences* **65**(8): 2661–2675.

Cushman-Roisin B, Beckers JM. 2011. *Introduction to geophysical fluid dynamics: physical and numerical aspects*, vol. 101. Academic Press.

Frederiksen JS, Davies AG. 1997. Eddy viscosity and stochastic backscatter parameterizations on the sphere for atmospheric circulation models. *Journal of the atmospheric sciences* **54**(20): 2475–2492.

Frederiksen JS, Dix MR, Kepert SM. 1996. Systematic energy errors and the tendency toward canonical equilibrium in atmospheric circulation models. *Journal of the atmospheric sciences* **53**(6): 887–904.

Frederiksen JS, Kepert SM. 2006. Dynamical subgrid-scale parameterizations from direct numerical simulations. *Journal of the atmospheric sciences* **63**(11): 3006–3019.

Holton JR, Hakim GJ. 2013. *An introduction to dynamic meteorology*. Academic press.

Jaynes ET. 1957. Information theory and statistical mechanics. *Physical review* **106**(4): 620.

Kraichnan Robert H, Montgomery D. 1980. Two-dimensional turbulence. *Rep. Prog. Phys* **43**.

Lorenz EN. 1996. Predictability: A problem partly solved. In: *Proc. Seminar on predictability*, vol. 1.

McWilliams JC. 1984. The emergence of isolated coherent vortices in turbulent flow. *Journal of Fluid Mechanics* **146**: 21–43.

Sadourny R, Basdevant C. 1985. Parameterization of subgrid scale barotropic and baroclinic eddies in quasi-geostrophic models: Anticipated potential vorticity method. *Journal of the atmospheric sciences* **42**(13): 1353–1363.

Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical Journal* **27**: 379–423.

Smagorinsky J. 1963. General circulation experiments with the primitive equations: I. the basic experiment*. *Monthly weather review* **91**(3): 99–164.

Taylor KE. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres (1984–2012)* **106**(D7): 7183–7192.

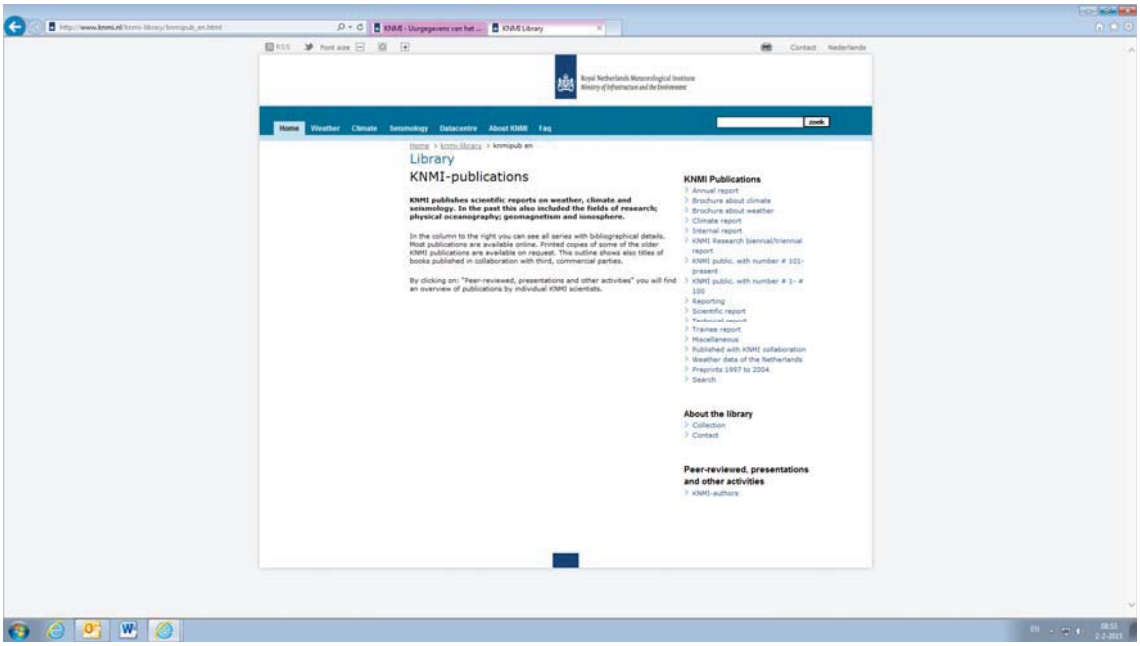
Verkley W. 2011. A maximum entropy approach to the problem of parametrization. *Quarterly Journal of the Royal Meteorological Society* **137**(660): 1872–1886.

Verkley W, Lynch P. 2009. Energy and enstrophy spectra of geostrophic turbulent flows derived from a maximum entropy principle. *Journal of the Atmospheric sciences* **66**(8): 2216–2236.

Wayne CE. 2011. Vortices and two-dimensional fluid motion. *Notices of the AMS* **58**(1): 10–19.

Zhu Y, Antonia R. 1996. On the correlation between enstrophy and energy dissipation rate in a turbulent wake. *Applied scientific research* **57**(3-4): 337–347.

A complete list of all KNMI-publications (1854 – present) can be found on our website www.knmi.nl/knmi-library/knmipub_en.html



The most recent reports are available as a PDF on this site.