RMetS

Royal Meteorological Society

A maximum entropy approach to the parametrization of subgrid processes in two-dimensional flow

W. T. M. Verkley,* P. C. Kalverla and C. A. Severijns

Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

*Correspondence to: W. T. M. Verkley, KNMI, PO Box 201, 3730 AE De Bilt, The Netherlands. E-mail: verkley@knmi.nl

In numerical models of geophysical fluid systems, parametrization schemes are needed to account for the effect of unresolved processes on processes that are resolved explicitly. Usually, these parametrization schemes require tuning of their parameters to achieve optimal performance. We propose a new type of parametrization that requires no tuning, as it is based on an assumption that is not specific to any particular model. The assumption is that the unresolved processes can be represented by a probability density function that has maximum information entropy under the constraints of zero average time derivatives of key integral quantities of the unresolved processes. In the context of a model of a simple fluid dynamical system, it is shown that this approach leads to definite expressions for the mean effect of unresolved processes on processes that are resolved. The merits of the parametrization, regarding both short-range forecasting and long-term statistics, are demonstrated by numerical experiments in which a low-resolution version of the model is used to simulate the results of a high-resolution version of the model. For the fluid dynamical system that is studied, the proposed parametrization turns out to be related to the anticipated potential vorticity method with definite values of its parameters.

1. Introduction

In the numerical modelling of geophysical fluid systems, one has to deal with processes that cannot be resolved explicitly. These processes are to be represented in some approximate sense, a problem that is called parametrization. The problem is difficult because of the large variety of processes that need to be parametrized and because parametrization schemes have to cope with different flow regimes and model parameters, sometimes necessitating the use of different schemes under different conditions. As discussed in section 5.6 of Stensrud (2009), even in the limited field of turbulence and horizontal diffusion many issues still remain. In the relatively simple context of two-dimensional turbulence, the situation is discussed by Thuburn *et al.* (2014).

An issue of particular importance is that parametrization schemes usually have one or more parameters that need to be tuned. For practical reasons, only limited datasets can be used for this purpose, so that their performance will not be optimal in all cases. It would thus be useful if tuning could be avoided or limited as much as possible. We will argue that such tuning might not be necessary if enough is known of the unresolved processes to formulate global budgets of key integral quantities that characterize the state of the unresolved processes. When applied in a statistically average sense, these budgets can be used to constrain the maximization of the information entropy of the probability density function of the unresolved processes. The

probability density function can then be used to calculate the average interaction of the unresolved processes with the resolved processes. The procedure has proved to work well for a simple schematic model proposed by Lorenz (Verkley, 2011).

In this study, we apply the same procedure to a model that is somewhat more realistic than the Lorenz model: a two-dimensional, doubly periodic flow system in flat geometry, the same system as was studied by Thuburn *et al.* (2014). The flow is divergence-free and described by the vorticity equation. The system is forced by a fixed vorticity pattern and loses energy and enstrophy by viscosity and an additional linear damping term. The numerical model is a spectral form of the vorticity equation, based on trigonometric functions. Following the general outline of Frederiksen and Kepert (2006), we use a high-resolution version of this model to generate an artificial 'reality'. A low-resolution version of the model is then considered to describe the low-resolution part of this reality, identified as the resolved processes (resolved scales). The remaining scales of the high-resolution model are identified with the unresolved processes (unresolved scales). In the context of this idealized set-up, the aim of parametrization is to reformulate the low-resolution model in such a way that it approximates as closely as possible the resolved scales generated by the high-resolution model.

The numerical model is described in section 2. In section 3, the distinction between resolved and unresolved scales is discussed, paying particular attention to the energy and enstrophy equations related to these scales. In section 4, we outline the general

formulation of a deterministic parametrization in terms of a probability density function for the unresolved scales. In section 5, the required probability density function is obtained by maximizing its information entropy using the averaged energy and enstrophy equations of the unresolved scales as constraints. In section 6, we show that the resulting parametrization turns out to be related to the anticipated potential vorticity method (APVM), introduced by Sadourny and Basdevant (1985). The numerical details of the high-resolution model that is used to test the method and demonstrate its merits are specified in section 7. The results of several numerical experiments are presented in two parts: the short-term model forecasting skill is assessed in section 8 and the long-term model climate statistics in section 9. We give our conclusions in section 10.

2. The model

The system that we study is the same as the system considered by Thuburn *et al.* (2014): a forced-dissipative version of the vorticity equation for two-dimensional incompressible flow.* The dissipation is by viscosity and by a linear damping term. The forcing F is combined with the linear damping to form a linear relaxation to a given vorticity field:

$$\frac{\partial \zeta}{\partial t} + \mathbf{v} \cdot \nabla \zeta = \nu \nabla^2 \zeta + \mu (F - \zeta). \tag{1}$$

Here, ζ is the vertical component of the vorticity, $\zeta = \mathbf{k} \cdot \nabla \times \mathbf{v}$. The second term on the left-hand side of Eq. (1) represents advection by the horizontal velocity $\mathbf{v}$, which, as it is assumed to satisfy $\nabla \cdot \mathbf{v} = 0$, can be written as $\mathbf{v} = \mathbf{k} \times \nabla \psi$ with ψ denoting the stream function. We may thus write $\zeta = \nabla^2 \psi$, where ∇^2 is the horizontal Laplace operator, and $\mathbf{v} \cdot \nabla \zeta = J(\psi, \zeta)$, where J is the Jacobi operator. The expressions for these operators are

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2}, \tag{2}$$

$$J(\psi, \zeta) = \frac{\partial \psi}{\partial x} \frac{\partial \zeta}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial \zeta}{\partial x}. \tag{3}$$

The magnitude of the viscosity and of the relaxation to the forcing F are given by ν and μ, respectively.

The system's flow domain is periodic with period $2\pi L$ in both x- and y-directions. Here, L is a length-scale which, together with a time-scale T, is used to non-dimensionalize the system. The original, dimensional, variables can then be obtained from the non-dimensional variables by $\zeta_d = \zeta_n T^{-1}$, $\mathbf{v}_d = \mathbf{v}_n L T^{-1}$, $\psi_d = \psi_n L^2 T^{-1}$, etc., where the subscripts 'd' and 'n' refer to dimensional and non-dimensional variables, respectively. With the conversions $\nu_d = \nu_n L^2 T^{-1}$ and $\mu_d = \mu_n T^{-1}$ for the parameters ν and μ, the vorticity equation assumes the same form in both dimensional and non-dimensional variables. From now on, we will refer to Eq. (1) in its non-dimensional form.

The field ζ is represented numerically by a sum of the type

$$\zeta(x, y, t) = \sum_{m=-N}^{+N} \sum_{n=-N}^{+N} \zeta_{mn}(t) Y_{mn}(x, y), \tag{4}$$

where the coefficients $\zeta_{mn}(t)$ depend on time and N is the truncation limit that determines the accuracy (resolution) of the numerical representation. The values ζ_{mn} can be considered as values on a rectangular grid, called $\mathcal{T}$, as illustrated by Figure 1.

The functions $Y_{mn}(x, y)$ depend on the two spatial coordinates and are defined by
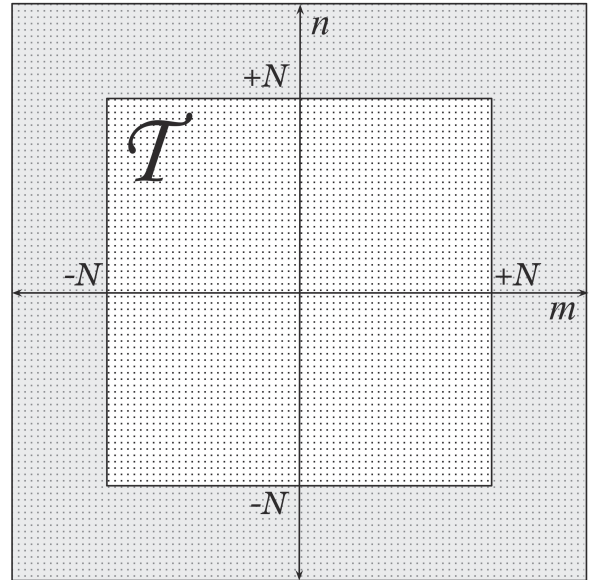
$$Y_{mn}(x, y) = X_m(x) X_n(y), \tag{5}$$

Figure 1. The spectral grid, where N is the truncation level and $\mathcal{T}$ represents the spectral domain.

with

$$X_m(x) = \begin{cases} \sqrt{2} \cos(|m|x) & \text{if } m > 0, \\ 1 & \text{if } m = 0, \\ \sqrt{2} \sin(|m|x) & \text{if } m < 0, \end{cases} \tag{6}$$

and likewise for $X_n(y)$. The functions Y_{mn} are orthonormal with respect to the following inner product:

$$(\xi, \chi) = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \xi(x, y) \chi(x, y) \, dx \, dy, \tag{7}$$

i.e. they satisfy

$$(Y_{mn}, Y_{m'n'}) = \delta_{mm'} \delta_{nn'}, \tag{8}$$

where $\delta_{mm'}$ and $\delta_{nn'}$ are Kronecker delta functions. We use the parenthesis notation for inner products to avoid confusion with the probabilistic average that will be used later. From the orthonormality of the functions Y_{mn}, it follows that the coefficients ζ_{mn} can be obtained from the field ζ by the following expression:

$$\zeta_{mn} = (Y_{mn}, \zeta). \tag{9}$$

Writing out the inner product, this yields

$$\zeta_{mn}(t) = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} Y_{mn}(x, y) \zeta(x, y, t) \, dx \, dy, \tag{10}$$

which constitutes the inverse of Eq. (4).

A major advantage of a spectral model is that derivatives of the model fields can be calculated exactly. In particular, it can be checked that we have

$$\left(\nabla^2 \psi\right)_{mn} = -(m^2 + n^2) \psi_{mn}, \tag{11}$$

in which $(\ldots)_{mn}$ denotes the spectral coefficient of the expression between the brackets. The advection or Jacobi term can also be calculated exactly. We have

$$(J(\psi, \zeta))_{mn} = (Y_{mn}, J(\psi, \zeta)) = \left(\frac{1}{2\pi}\right)^2$$
$$\times \int_0^{2\pi} \int_0^{2\pi} Y_{mn}(x, y) J(\psi(x, y, t), \zeta(x, y, t)) \, dx \, dy. \tag{12}$$

Now, if ψ and ζ are written as in Eq. (4) and if we limit ourselves to the coefficients of the Jacobian for which m and n lie between $-N$ and $+N$, then the integrand can be written in the same form as Eq. (4) but with N replaced by $3N$. In other words, the integrand is a trigonometric series in x and y with a maximum wave number $3N$. This implies that the integration over x and y can be carried out exactly by replacing it by a summation over K equidistant values of x and y, with weights $(2\pi)/K$, if $K \geq 3N + 1$ (Coiffier, 2011, see his eq. (4.72)). As an illustration: if K is taken to be a power of 2, in which case the fast Fourier transform routines that are used in the different sums are most efficient, we have for $K = 2^8 = 256$ that N should not be larger than 85.

Writing the advection term by means of a Jacobian and evaluating its spectral representation in the manner described above, a spectral model of Eq. (1) would thus assume the form

$$\frac{\partial \zeta}{\partial t} + PJ(\psi, \zeta) = \nu \nabla^2 \zeta + \mu(F - \zeta). \tag{13}$$

Here, P is an orthogonal projection operator that projects on the space of functions $\mathcal{T}$ in terms of which ζ is represented. The operator P in front of the Jacobian formalizes the fact that we keep only the spectral coefficients within $\mathcal{T}$, so that ζ does not leave this space of functions. For the full infinite-dimensional system in which $N \to \infty$, the projection operator P can be replaced by the identity I or left out altogether.

Energy and enstrophy are important global quantities of two-dimensional flow. In the absence of forcing and dissipation, i.e. if the flow is inviscid, both energy and enstrophy are conserved. Inviscid conservation of enstrophy is a characterizing feature of two-dimensional flow that distinguishes it from three-dimensional flow (Bouchet and Venaille, 2012). For two-dimensional doubly periodic flow, the energy and enstrophy (more precisely, energy and enstrophy per unit area) are given by

$$E = -\frac{1}{2}(\psi, \zeta), \tag{14}$$

$$Z = \frac{1}{2}(\zeta, \zeta). \tag{15}$$

The energy E and enstrophy Z change in time according to

$$\frac{dE}{dt} = -\nu(\zeta, \zeta) - \mu(\psi, (F - \zeta)), \tag{16}$$

$$\frac{dZ}{dt} = \nu(\nabla^2 \zeta, \zeta) + \mu(\zeta, (F - \zeta)). \tag{17}$$

These results are derived in the Appendix. Equations (16) and (17) show that the spectral model conserves energy and enstrophy if ν and μ are zero.

3. Resolved and unresolved scales

As remarked above, the coefficients ζ_{mn} and ψ_{mn}, by means of which the vorticity and the stream function are represented, can be visualized as values on a rectangular grid of pairs (m, n), where m and n are integers running from $-N$ to $+N$, as illustrated by Figure 1. For the purpose of this study, we will consider the behaviour of the corresponding spectral model, of which the evolution equation is given by Eq. (13), as 'reality'. The coefficients associated with a smaller rectangular grid, in which the integers m and n vary between $-M$ and $+M$ with $M < N$, are to be associated with the 'model' and are supposed to describe the resolved scales. All other coefficients within the large rectangular grid are to be associated with the unresolved scales. The set of coefficients that is to be identified with the resolved scales is denoted by $\mathcal{R}$, while the set of coefficients that is to be identified with the unresolved scales is denoted by $\mathcal{U}$. As the set of all coefficients within the large rectangle was called $\mathcal{T}$, we might say that $\mathcal{T} = \mathcal{R} \cup \mathcal{U}$. The situation is visualized in Figure 2.
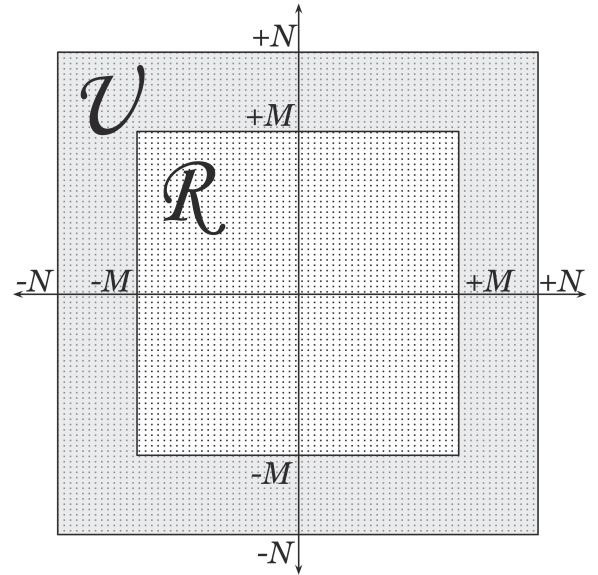

Figure 2. The spectral grid $\mathcal{T}$, corresponding to a truncation level N and visualized in Figure 1, is decomposed into a resolved part $\mathcal{R}$, corresponding to a truncation level M with $M < N$, and an unresolved part $\mathcal{U}$ consisting of the complement of $\mathcal{R}$ in $\mathcal{T}$, i.e. consisting of all other coefficients within $\mathcal{T}$.

For convenience, we will write Eq. (4) more concisely as

$$\zeta = \sum_{(m,n) \in \mathcal{T}} \zeta_{mn} Y_{mn}. \tag{18}$$

We next decompose ζ into a resolved part $\zeta^{\mathcal{R}}$ and an unresolved part $\zeta^{\mathcal{U}}$:

$$\zeta = \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}, \tag{19}$$

in which we define

$$\zeta^{\mathcal{R}} = \sum_{(m,n) \in \mathcal{R}} \zeta_{mn} Y_{mn}, \tag{20}$$

$$\zeta^{\mathcal{U}} = \sum_{(m,n) \in \mathcal{U}} \zeta_{mn} Y_{mn}. \tag{21}$$

A similar decomposition is applied to ψ and F. With the sets of coefficients $\mathcal{R}$ and $\mathcal{U}$, we associate orthogonal projection operators $P^{\mathcal{R}}$ and $P^{\mathcal{U}}$ such that $P = P^{\mathcal{R}} + P^{\mathcal{U}}$. These projection operators allow us to decompose the Jacobian:

$$PJ = (P^{\mathcal{R}} + P^{\mathcal{U}})J = P^{\mathcal{R}}J + P^{\mathcal{U}}J. \tag{22}$$

As the Laplacian is diagonal in the basis functions (see Eq. (11)), the model equation, Eq. (13), can be split into two separate equations:

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + P^{\mathcal{R}}J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})$$
$$= \nu \nabla^2 \zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}), \tag{23}$$

$$\frac{\partial \zeta^{\mathcal{U}}}{\partial t} + P^{\mathcal{U}}J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})$$
$$= \nu \nabla^2 \zeta^{\mathcal{U}} + \mu(F^{\mathcal{U}} - \zeta^{\mathcal{U}}). \tag{24}$$

These two equations, taken together, are mathematically equivalent with Eq. (13). We see that the interaction between the two sets of coefficients, $\mathcal{R}$ and $\mathcal{U}$, takes place in the nonlinear advection terms.

Substituting the decomposition Eq. (19) into Eqs (14) and (15) and making use of the orthogonality of the resolved and unresolved scales, we may decompose the energy and enstrophy into resolved and unresolved contributions:

$$E = -\frac{1}{2}(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) = E^{\mathcal{R}} + E^{\mathcal{U}}, \quad (25)$$

$$Z = \frac{1}{2}(\zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) = Z^{\mathcal{R}} + Z^{\mathcal{U}}, \quad (26)$$

where $E^{\mathcal{R}} = -\frac{1}{2}(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})$ and $E^{\mathcal{U}} = -\frac{1}{2}(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})$ will be referred to as the resolved and unresolved energy and $Z^{\mathcal{R}} = \frac{1}{2}(\zeta^{\mathcal{R}}, \zeta^{\mathcal{R}})$ and $Z^{\mathcal{U}} = \frac{1}{2}(\zeta^{\mathcal{U}}, \zeta^{\mathcal{U}})$ as the resolved and unresolved enstrophy. For the time derivatives of the resolved energy and enstrophy, we have

$$\frac{dE^{\mathcal{R}}}{dt} = -(\psi^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) + (\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}))$$
$$- \nu(\zeta^{\mathcal{R}}, \zeta^{\mathcal{R}}) - \mu(\psi^{\mathcal{R}}, (F^{\mathcal{R}} - \zeta^{\mathcal{R}})), \quad (27)$$

$$\frac{dZ^{\mathcal{R}}}{dt} = (\zeta^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) - (\zeta^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}))$$
$$+ \nu(\nabla^2\zeta^{\mathcal{R}}, \zeta^{\mathcal{R}}) + \mu(\zeta^{\mathcal{R}}, (F^{\mathcal{R}} - \zeta^{\mathcal{R}})). \quad (28)$$

These relations are derived in the Appendix. For the time derivatives of the unresolved energy and enstrophy, we have, analogously,

$$\frac{dE^{\mathcal{U}}}{dt} = (\psi^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) - (\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}))$$
$$- \nu(\zeta^{\mathcal{U}}, \zeta^{\mathcal{U}}) - \mu(\psi^{\mathcal{U}}, (F^{\mathcal{U}} - \zeta^{\mathcal{U}})), \quad (29)$$

$$\frac{dZ^{\mathcal{U}}}{dt} = -(\zeta^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) + (\zeta^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}))$$
$$+ \nu(\nabla^2\zeta^{\mathcal{U}}, \zeta^{\mathcal{U}}) + \mu(\zeta^{\mathcal{U}}, (F^{\mathcal{U}} - \zeta^{\mathcal{U}})). \quad (30)$$

If we add Eq. (27) to Eq. (29), we end up with Eq. (16). Similarly, if we add Eq. (28) to Eq. (30) we end up with Eq. (17). The terms that cancel are the terms that involve Jacobians; these are the terms that exchange energy and enstrophy between resolved and unresolved scales.

4. Parametrization of unresolved scales

In the following, we will focus on the behaviour of the vorticity field $\zeta^{\mathcal{R}}$, as described by Eq. (23). Both detailed time evolution ('weather') and statistical properties ('climate') will be considered. Of course, if limitations in terms of computer resources did not exist, the optimal approach to Eq. (23) is to solve Eq. (24) alongside Eq. (23) and to use the resulting unresolved scales in the Jacobian of Eq. (23). This is what is actually done if we solve the full problem Eq. (13). The aim of parametrization is to avoid solving Eq. (24) and instead simulate its effect on Eq. (23) by properly modifying the latter equation.

The simplest approach is just to neglect the unresolved scales in Eq. (23) and to set $\psi^{\mathcal{U}} = \zeta^{\mathcal{U}} = 0$. However, as we will see in section 9, this leads to an overestimation of the energy and enstrophy in the smallest scales of the resolved model. A more refined method is the use of a scale-dependent (hyper)viscosity (McWilliams, 1984), sometimes called an eddy viscosity (Cushman-Roisin and Beckers, 2011). This is quite effective, although artificial, and it requires tuning of the viscosity parameter. We propose instead to reduce the dynamics of the unresolved scales to their average energy and enstrophy budgets and to describe the unresolved scales by a probability density function, conditioned on the resolved scales, that takes these budgets into account. On their own, these budgets do not specify the probability density function fully and we therefore require, as our closure assumption, that the probability density function shall have maximum information entropy.

Replacing the spectral coefficients of $\psi^{\mathcal{U}}$ and $\zeta^{\mathcal{U}}$ by a probability density function over these coefficients and then averaging the equation for the resolved coefficients, Eq. (23), we obtain

$$\langle \frac{\partial \zeta^{\mathcal{R}}}{\partial t} \rangle + \langle P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) \rangle$$
$$= \langle \nu\nabla^2\zeta^{\mathcal{R}} \rangle + \langle \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}) \rangle, \quad (31)$$

where the brackets $\langle ... \rangle$ denote the average. As the probability density function only involves the unresolved scales, we have

$$\langle \frac{\partial \zeta^{\mathcal{R}}}{\partial t} \rangle = \frac{\partial \zeta^{\mathcal{R}}}{\partial t}, \quad (32)$$

$$\langle \nu\nabla^2\zeta^{\mathcal{R}} \rangle = \nu\nabla^2\zeta^{\mathcal{R}}, \quad (33)$$

$$\langle \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}) \rangle = \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}). \quad (34)$$

With regard to the Jacobian term, we first write

$$P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})$$
$$= P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) + P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}})$$
$$+ P^{\mathcal{R}} J(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}) + P^{\mathcal{R}} J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}). \quad (35)$$

We then notice that, because the procedures of averaging and projecting may be interchanged and the Jacobian is a bilinear operator, we have

$$\langle P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) \rangle = P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}), \quad (36)$$

$$\langle P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}}) \rangle = P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \langle \zeta^{\mathcal{U}} \rangle), \quad (37)$$

$$\langle P^{\mathcal{R}} J(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}}) \rangle = P^{\mathcal{R}} J(\langle \psi^{\mathcal{U}} \rangle, \zeta^{\mathcal{R}}), \quad (38)$$

$$\langle P^{\mathcal{R}} J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}) \rangle = P^{\mathcal{R}} \langle J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}) \rangle, \quad (39)$$

so that we can write

$$\langle P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}}) \rangle$$
$$= P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \langle \psi^{\mathcal{U}} \rangle, \zeta^{\mathcal{R}} + \langle \zeta^{\mathcal{U}} \rangle)$$
$$+ P^{\mathcal{R}} \langle J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}) \rangle - P^{\mathcal{R}} J(\langle \psi^{\mathcal{U}} \rangle, \langle \zeta^{\mathcal{U}} \rangle). \quad (40)$$

The latter result, i.e. Eq. (40), leads to the following equation for the resolved vorticity $\zeta^{\mathcal{R}}$:

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \langle \psi^{\mathcal{U}} \rangle, \zeta^{\mathcal{R}} + \langle \zeta^{\mathcal{U}} \rangle)$$
$$+ P^{\mathcal{R}} \langle J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}}) \rangle - P^{\mathcal{R}} J(\langle \psi^{\mathcal{U}} \rangle, \langle \zeta^{\mathcal{U}} \rangle)$$
$$= \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}). \quad (41)$$

In the next section, we will use the principle of maximum entropy to obtain a probability density function for the unresolved scales and calculate the mean fields involved in the expressions above. We will see that these depend on the system parameters and on the state of the resolved scales, so that we end up with a closed system in terms of the resolved scales.

5. The maximum entropy approach

The basis of the probability density function with which we wish to represent the state of the unresolved variables is the energy and enstrophy equations, Eq. (29) and Eq. (30). To write these equations in terms of spectral coefficients, we use the following expression of the inner product for unresolved scales:

$$(\xi^{\mathcal{U}}, \chi^{\mathcal{U}}) = \sum_{(m,n)\in\mathcal{U}} \xi_{mn}\chi_{mn}, \quad (42)$$

which follows from the orthonormality of the basis functions Y_{mn}. Introducing the notation $c_{mn} = (m^2 + n^2)$, we obtain as a result

$$\frac{dE^{\mathcal{U}}}{dt} = \sum_{(m,n)\in\mathcal{U}} \psi_{mn}(Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) \quad (43)$$
$$- \sum_{(m,n)\in\mathcal{U}} \sum_{(m',n')\in\mathcal{U}} \psi_{mn}\zeta_{m'n'}(\psi^{\mathcal{R}}, J(Y_{mn}, Y_{m'n'}))$$
$$- \nu \sum_{(m,n)\in\mathcal{U}} \zeta_{mn}^2 - \mu \sum_{(m,n)\in\mathcal{U}} \psi_{mn}(F_{mn} - \zeta_{mn}),$$

$$\frac{dZ^{\mathcal{U}}}{dt} = - \sum_{(m,n)\in\mathcal{U}} \zeta_{mn}(Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) \qquad (44)$$

$$+ \sum_{(m,n)\in\mathcal{U}} \sum_{(m',n')\in\mathcal{U}} \psi_{mn}\zeta_{m'n'}(\zeta^{\mathcal{R}}, J(Y_{mn}, Y_{m'n'}))$$

$$- \nu \sum_{(m,n)\in\mathcal{U}} c_{mn}\zeta_{mn}^2 + \mu \sum_{(m,n)\in\mathcal{U}} \zeta_{mn}(F_{mn} - \zeta_{mn}).$$

After some rearranging and replacing ψ_{mn} by $-\zeta_{mn}/c_{mn}$, we may write

$$-\frac{dE^{\mathcal{U}}}{dt} = \sum_{(m,n)\in\mathcal{U}} \left[\left(\frac{\mu + \nu c_{mn}}{c_{mn}} \right) \zeta_{mn}^2 + \frac{\chi_{mn}}{c_{mn}} \zeta_{mn} \right]$$

$$- \sum_{(m,n)\in\mathcal{U}} \sum_{(m',n')\in\mathcal{U}} \frac{\xi_{mnm'n'}}{c_{mn}} \zeta_{mn}\zeta_{m'n'}, \qquad (45)$$

$$-\frac{dZ^{\mathcal{U}}}{dt} = \sum_{(m,n)\in\mathcal{U}} \left[(\mu + \nu c_{mn}) \zeta_{mn}^2 + \chi_{mn} \zeta_{mn} \right]$$

$$+ \sum_{(m,n)\in\mathcal{U}} \sum_{(m',n')\in\mathcal{U}} \frac{\eta_{mnm'n'}}{c_{mn}} \zeta_{mn}\zeta_{m'n'}, \qquad (46)$$

where we have defined

$$\chi_{mn} = (Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) - \mu F_{mn}, \qquad (47)$$

$$\xi_{mnm'n'} = (\psi^{\mathcal{R}}, J(Y_{mn}, Y_{m'n'})), \qquad (48)$$

$$\eta_{mnm'n'} = (\zeta^{\mathcal{R}}, J(Y_{mn}, Y_{m'n'})). \qquad (49)$$

We see that both $dE^{\mathcal{U}}/dt$ and $dZ^{\mathcal{U}}/dt$ are quadratic expressions in the unresolved coefficients ζ_{mn}.

The principle of maximum entropy (Jaynes, 2003, chapters 11 and 12, in particular Eq. 12.8) requires that the probability density function for the unresolved scales, $\mathcal{P}(\zeta^{\mathcal{U}})$, should have maximal information entropy S_I, where S_I is defined by

$$S_I = - \int \mathcal{P}(\zeta^{\mathcal{U}}) \log \frac{\mathcal{P}(\zeta^{\mathcal{U}})}{\mathcal{M}(\zeta^{\mathcal{U}})} d\zeta^{\mathcal{U}}. \qquad (50)$$

Here, we use a rather abstract notation in which $\zeta^{\mathcal{U}}$ denotes the set of coefficients ζ_{mn} with $(m, n) \in \mathcal{U}$ and $d\zeta^{\mathcal{U}}$ denotes an integration element of a multiple integral over all unresolved coefficients. The measure $\mathcal{M}$, necessary for dimensional consistency, embodies any *a priori* information on the values of the coefficients and, since no such information is assumed, is taken to be a product of constants. The first constraint is the normalization condition:

$$\langle 1 \rangle = \int \mathcal{P}(\zeta^{\mathcal{U}}) d\zeta^{\mathcal{U}} = 1. \qquad (51)$$

The other two constraints are $\langle dE^{\mathcal{U}}/dt \rangle = 0$ and $\langle dZ^{\mathcal{U}}/dt \rangle = 0$. This means that we assume that, on the time-scale of the resolved scales, the unresolved scales are in a statistically stationary state that is characterized by a mean balance between sources and sinks of energy and enstrophy:

$$\langle \frac{dE^{\mathcal{U}}}{dt} \rangle = \int \mathcal{P}(\zeta^{\mathcal{U}}) \frac{dE^{\mathcal{U}}}{dt}(\zeta^{\mathcal{U}}) d\zeta^{\mathcal{U}} = 0, \qquad (52)$$

$$\langle \frac{dZ^{\mathcal{U}}}{dt} \rangle = \int \mathcal{P}(\zeta^{\mathcal{U}}) \frac{dZ^{\mathcal{U}}}{dt}(\zeta^{\mathcal{U}}) d\zeta^{\mathcal{U}} = 0. \qquad (53)$$

The maximization problem can be solved by the use of Lagrange multipliers. The variations of the entropy, normalization condition and the two conditions on the time rate of change of energy and enstrophy are

$$\delta S_I = - \int \delta \mathcal{P}(\zeta^{\mathcal{U}}) \left[\log \frac{\mathcal{P}(\zeta^{\mathcal{U}})}{\mathcal{M}(\zeta^{\mathcal{U}})} + 1 \right] d\zeta^{\mathcal{U}}, \qquad (54)$$

$$\delta \langle 1 \rangle = \int \delta \mathcal{P}(\zeta^{\mathcal{U}}) d\zeta^{\mathcal{U}}, \qquad (55)$$

$$\delta \langle \frac{dE^{\mathcal{U}}}{dt} \rangle = \int \delta \mathcal{P}(\zeta^{\mathcal{U}}) \frac{dE^{\mathcal{U}}}{dt}(\zeta^U) d\zeta^{\mathcal{U}}, \qquad (56)$$

$$\delta \langle \frac{dZ^{\mathcal{U}}}{dt} \rangle = \int \delta \mathcal{P}(\zeta^{\mathcal{U}}) \frac{dZ^{\mathcal{U}}}{dt}(\zeta^U) d\zeta^{\mathcal{U}}. \qquad (57)$$

Using a Lagrange multiplier $-\rho$ for the normalization and Lagrange multipliers α and β for the other two constraints, we arrive at the following condition:

$$\delta S_I - \rho\delta\langle 1 \rangle + \alpha\delta\langle \frac{dE^{\mathcal{U}}}{dt} \rangle + \beta\delta\langle \frac{dZ^{\mathcal{U}}}{dt} \rangle = 0. \qquad (58)$$

When the expressions above are substituted, this gives (leaving out the arguments $\zeta^{\mathcal{U}}$ for readability)

$$\int \delta\mathcal{P} \left[\log\frac{\mathcal{P}}{\mathcal{M}} + 1 + \rho - \alpha\frac{dE^{\mathcal{U}}}{dt} - \beta\frac{dZ^{\mathcal{U}}}{dt} \right] d\zeta^{\mathcal{U}} = 0. \qquad (59)$$

As this should be valid for any variation $\delta\mathcal{P}$, we should have

$$\log\frac{\mathcal{P}}{\mathcal{M}} + 1 + \rho - \alpha\frac{dE^{\mathcal{U}}}{dt} - \beta\frac{dZ^{\mathcal{U}}}{dt} = 0. \qquad (60)$$

or

$$\mathcal{P} = \mathcal{M}\exp\left[-1 - \rho + \alpha\frac{dE^{\mathcal{U}}}{dt} + \beta\frac{dZ^{\mathcal{U}}}{dt} \right]. \qquad (61)$$

Now, the factor $\exp(-1-\rho)$ is usually written as $\mathcal{Z}^{-1}$, so that (restoring the arguments $\zeta^{\mathcal{U}}$)

$$\mathcal{P}(\zeta^{\mathcal{U}}) = \frac{1}{\mathcal{Z}}\mathcal{M}\exp\left[\alpha\frac{dE^{\mathcal{U}}}{dt}(\zeta^{\mathcal{U}}) + \beta\frac{dZ^{\mathcal{U}}}{dt}(\zeta^{\mathcal{U}}) \right]. \qquad (62)$$

The function $\mathcal{Z}$ is called the partition function and is determined by the normalization condition Eq. (51). The Lagrange multipliers α and β are determined by the conditions on the time derivatives of the unresolved energy and enstrophy, i.e. by Eqs (52) and (53). We will take $\mathcal{M} = 1$.

In principle, the probability density function is now given and from the form above, in combination with Eqs (45) and (46), we see that it is a multivariate normal probability density function in the variables ζ_{mn} with $(m, n) \in U$. Now, in order to simplify the calculations, we make the approximation that in Eqs (45) and (46) the last terms on the right-hand sides can be disregarded. This is a rather serious approximation that is difficult to justify *a priori* and has to be judged on the basis of its consequences. With this approximation, the probability density function becomes

$$\mathcal{P}(\zeta^{\mathcal{U}}) = \frac{1}{\mathcal{Z}}$$
$$\times \exp\{-\alpha \sum_{(m,n)\in\mathcal{U}} \left[\left(\frac{\mu + \nu c_{mn}}{c_{mn}} \right) \zeta_{mn}^2 + \frac{\chi_{mn}}{c_{mn}} \zeta_{mn} \right]$$
$$- \beta \sum_{(m,n)\in\mathcal{U}} \left[(\mu + \nu c_{mn}) \zeta_{mn}^2 + \chi_{mn} \zeta_{mn} \right]\}. \qquad (63)$$

Collecting terms that are quadratic and linear in ζ_{mn}, we obtain

$$\mathcal{P}(\zeta^{\mathcal{U}}) = \frac{1}{\mathcal{Z}}$$
$$\times \exp\{ \sum_{(m,n)\in\mathcal{U}} -\frac{1}{c_{mn}}(\mu + \nu c_{mn})(\alpha + \beta c_{mn})$$
$$\times (\zeta_{mn}^2 + \frac{\chi_{mn}}{\mu + \nu c_{mn}}\zeta_{mn})\}. \qquad (64)$$

Completing the square in the factors involving ζ_{mn} then gives

$$
\mathcal{P}(\zeta^{\mathcal{U}}) = \frac{1}{\mathcal{Z}}
$$
$$
\times \exp\Big\{ \sum_{(m,n)\in\mathcal{U}} -\frac{1}{c_{mn}}(\mu + \nu c_{mn})(\alpha + \beta c_{mn})
$$
$$
\times ([\zeta_{mn} + \frac{1}{2}\frac{\chi_{mn}}{\mu + \nu c_{mn}}]^2 - \frac{1}{4}[\frac{\chi_{mn}}{\mu + \nu c_{mn}}]^2)\Big\}. \quad (65)
$$

Now, defining

$$
\lambda_{mn} = -\frac{1}{2}\chi_{mn}\left(\frac{1}{\mu + \nu c_{mn}}\right), \quad (66)
$$

$$
\sigma_{mn}^2 = \frac{1}{2}c_{mn}\left(\frac{1}{\mu + \nu c_{mn}}\right)\left(\frac{1}{\alpha + \beta c_{mn}}\right), \quad (67)
$$

the properly normalized expression of $\mathcal{P}(\zeta^{\mathcal{U}})$ reads

$$
\mathcal{P}(\zeta^{\mathcal{U}}) = \prod_{(m,n)\in U} \mathcal{N}(\lambda_{mn}, \sigma_{mn}, \zeta_{mn}), \quad (68)
$$

where the functions $\mathcal{N}$ are the normal distributions:

$$
\mathcal{N}(\lambda, \sigma, x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\lambda)^2}{2\sigma^2}\right). \quad (69)
$$

This immediately gives us the corresponding partition function $\mathcal{Z}$:

$$
\mathcal{Z} = \prod_{(m,n)\in U}(2\pi\sigma_{mn}^2)^{1/2}\exp\left(\frac{\lambda_{mn}^2}{2\sigma_{mn}^2}\right). \quad (70)
$$

Except for the Lagrange multipliers α and β, the probability density function of the unresolved scales is now known: it is a product of independent normal distributions with means (66) and variances (67).

Quite remarkably, the Lagrange multipliers α and β do not appear in the expression for λ_{mn} and, as we will see later, it is not necessary to determine them, as only the mean λ_{mn} is actually needed. However, for completeness and with an eye on possible applications to stochastic parametrizations, we explain how α and β can be obtained from the constraints $\langle dE^{\mathcal{U}}/dt\rangle = 0$ and $\langle dZ^{\mathcal{U}}/dt\rangle = 0$. Using the fact that for a probability density function of the form (68), i.e. a product of independent normal distributions, we have

$$
\langle \zeta_{mn}\rangle = \lambda_{mn}, \quad (71)
$$

$$
\langle \zeta_{mn}\zeta_{m'n'}\rangle = \lambda_{mn}\lambda_{m'n'} + \sigma_{mn}^2\delta_{mm'}\delta_{nn'}, \quad (72)
$$

we find

$$
-\langle dE^{\mathcal{U}}/dt\rangle
$$
$$
= \sum_{(m,n)\in\mathcal{U}}\left(\frac{\mu + \nu c_{mn}}{c_{mn}}\right)(\lambda_{mn}^2 + \sigma_{mn}^2) + \frac{\chi_{mn}}{c_{mn}}\lambda_{mn}, \quad (73)
$$

$$
-\langle dZ^{\mathcal{U}}/dt\rangle
$$
$$
= \sum_{(m,n)\in\mathcal{U}}(\mu + \nu c_{mn})(\lambda_{mn}^2 + \sigma_{mn}^2) + \chi_{mn}\lambda_{mn}. \quad (74)
$$

If we then substitute Eqs (66) and (67) for λ_{mn} and σ_{mn}^2, the constraints $\langle dE^{\mathcal{U}}/dt\rangle = 0$ and $\langle dZ^{\mathcal{U}}/dt\rangle = 0$ reduce to

$$
\sum_{(m,n)\in\mathcal{U}}\left[\frac{1}{\alpha + \beta c_{mn}} - \frac{1}{2}\chi_{mn}^2\frac{1}{c_{mn}}\frac{1}{\mu + \nu c_{mn}}\right] = 0, \quad (75)
$$

$$
\sum_{(m,n)\in\mathcal{U}}\left[\frac{c_{mn}}{\alpha + \beta c_{mn}} - \frac{1}{2}\chi_{mn}^2\frac{1}{\mu + \nu c_{mn}}\right] = 0. \quad (76)
$$

These are two nonlinear relationships that allow us to obtain α and β and therewith the complete probability density function of the unresolved scales. Note that the probability density function depends parametrically on the resolved scales $\psi^{\mathcal{R}}$ and $\zeta^{\mathcal{R}}$ via χ_{mn} as given by Eq. (47).

We now return to the equation for the resolved scales, Eq. (41). Expressions for $\langle \psi^{\mathcal{U}}\rangle$ and $\langle \zeta^{\mathcal{U}}\rangle$ are

$$
\langle \psi^{\mathcal{U}}\rangle = \sum_{(m,n)\in\mathcal{U}}\langle\psi_{mn}\rangle Y_{mn} = \sum_{(m,n)\in\mathcal{U}} -\frac{\lambda_{mn}}{c_{mn}}Y_{mn}, \quad (77)
$$

$$
\langle \zeta^{\mathcal{U}}\rangle = \sum_{(m,n)\in\mathcal{U}}\langle\zeta_{mn}\rangle Y_{mn} = \sum_{(m,n)\in\mathcal{U}}\lambda_{mn}Y_{mn}. \quad (78)
$$

For $\langle J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})\rangle$, we have, expanding $\psi^{\mathcal{U}}$ and $\zeta^{\mathcal{U}}$ in terms of the basis functions Y_{mn},

$$
\langle J(\psi^{\mathcal{U}}, \zeta^{U})\rangle
$$
$$
= -\sum_{(m,n)\in\mathcal{U}}\sum_{(m',n')\in\mathcal{U}}\frac{J(Y_{mn}, Y_{m'n'})}{c_{mn}}\langle\zeta_{mn}, \zeta_{m'n'}\rangle
$$
$$
= -\sum_{(m,n)\in\mathcal{U}}\sum_{(m',n')\in\mathcal{U}}\frac{J(Y_{mn}, Y_{m'n'})}{c_{mn}}\lambda_{mn}\lambda_{m'n'}
$$
$$
= J(\langle\psi^{\mathcal{U}}\rangle, \langle\zeta^{U}\rangle). \quad (79)
$$

In the second equality, we made use of Eq. (72) and the fact that only the first contribution of this expression is retained, as the Jacobian of two identical fields is zero. As a consequence, we have

$$
P^{\mathcal{R}}\langle J(\psi^{\mathcal{U}}, \zeta^{U})\rangle = P^{\mathcal{R}}J(\langle\psi^{\mathcal{U}}\rangle, \langle\zeta^{U}\rangle), \quad (80)
$$

so that the third and fourth terms on the left-hand side of Eq. (41) cancel. The equation for the resolved scales thus simplifies to

$$
\frac{\partial\zeta^{\mathcal{R}}}{\partial t} + P^{\mathcal{R}}J(\psi^{\mathcal{R}} + \langle\psi^{\mathcal{U}}\rangle, \zeta^{\mathcal{R}} + \langle\zeta^{\mathcal{U}}\rangle)
$$
$$
= \nu\nabla^2\zeta^{\mathcal{R}} + \mu(F^{\mathcal{R}} - \zeta^{\mathcal{R}}). \quad (81)
$$

We recall that the fields $\langle\zeta^{\mathcal{U}}\rangle$ and $\langle\psi^{\mathcal{U}}\rangle$ are given by Eqs (77) and (78), respectively, where λ_{mn} is given by (combining Eq. (47) with Eq. (66))

$$
\lambda_{mn} = -\frac{1}{2}\left(\frac{1}{\mu + \nu c_{mn}}\right)
$$
$$
\times \big((Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) - \mu F_{mn}\big). \quad (82)
$$

If these expressions for $\langle\psi^{\mathcal{U}}\rangle$ and $\langle\zeta^{\mathcal{U}}\rangle$ are used in Eq. (81), we obtain a closed system in terms of the resolved scales.

6. The anticipated potential vorticity method[†]

It is of some interest to note that our parametrization is similar to the APVM introduced by Sadourny and Basdevant (1985). In the context of our model, their parametrized potential vorticity equation (i.e. their eq. (5)) would take the form

$$
\frac{\partial\zeta^{\mathcal{R}}}{\partial t} + P^{\mathcal{R}}J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) - P^{\mathcal{R}}J(\psi^{\mathcal{R}}, D) = 0, \quad (83)
$$

with

$$
D = \theta\mathcal{L}J(\psi^R, \zeta^R). \quad (84)
$$

Here, θ is a time-scale and $\mathcal{L}$ a non-dimensional non-negative linear operator. The characterizing property of this

[†]Readers eager to see the numerical results might skip this paragraph on a first reading.

parametrization is that the enstrophy $Z^{\mathcal{R}}$ of the system decreases, whereas the energy $E^{\mathcal{R}}$ stays constant, as can be checked quite readily. To compare this parametrization with ours, we make the additional approximation that the average $\langle \psi^{\mathcal{U}} \rangle$ in Eq. (81) can be neglected. Our parametrized system then becomes

$$\frac{\partial \zeta^{\mathcal{R}}}{\partial t} + P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}) + P^{\mathcal{R}} J(\psi^{\mathcal{R}}, \langle \zeta^{\mathcal{U}} \rangle)$$
$$= \nu \nabla^2 \zeta^{\mathcal{R}} + \mu (F^{\mathcal{R}} - \zeta^{\mathcal{R}}). \quad (85)$$

Assuming, furthermore, that there is no forcing in the unresolved variables, we would have, for $\langle \zeta_{mn} \rangle = \lambda_{mn}, (m, n) \in \mathcal{U}$,

$$\langle \zeta_{mn} \rangle = -\frac{1}{2} \left(\frac{1}{\mu + \nu c_{mn}} \right) (Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) \quad (86)$$

$$= -\frac{1}{2\mu} \left(\frac{1}{1 + (\nu/\mu) c_{mn}} \right) (Y_{mn}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})). \quad (87)$$

For the field $\langle \zeta^{\mathcal{U}} \rangle$, we may thus write, more abstractly,

$$\langle \zeta^{\mathcal{U}} \rangle = -\frac{1}{2\mu} \left(1 - \frac{\nu}{\mu} \nabla^2 \right)^{-1} P^{\mathcal{U}} J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}). \quad (88)$$

The relationship with the term D of the APVM is clear if we identify

$$\theta = \frac{1}{2\mu}, \quad (89)$$

$$\mathcal{L} = \left(1 - \frac{\nu}{\mu} \nabla^2 \right)^{-1} P^{\mathcal{U}}. \quad (90)$$

The important difference with the APVM is that the time-scale θ and the operator $\mathcal{L}$ – which is indeed a non-negative linear operator – do not have to be chosen and tuned but are given directly in terms of the system parameters ν and μ. We also note that the quite special form that our operator assumes is rather different from the forms considered by Sadourny and Basdevant (1985) and Thuburn *et al.* (2014).

7. Numerical experiments

For the length-scale L of the model, we choose the Earth's radius, i.e. $L = a$, where $a = 6.371 \times 10^6$ m; for the time-scale T we choose the inverse of the Earth's angular velocity, i.e. $T = \Omega^{-1}$, where $\Omega = 7.292 \times 10^{-5}$ s^{-1}. These choices are completely arbitrary, but it allows us to express time in terms of days, hours and minutes using that 1 day is $24 \times 60 \times 60$ s by definition, corresponding to $24 \times 60 \times 60 \times \Omega = 6.300288 \approx 2\pi$ non-dimensional time units. The value of the viscosity ν of the reference model is chosen such that a single vorticity wave, described by one basis function Y_{mn} with $c_{mn} = m^2 + n^2 = 85^2$, is exponentially damped with an e-folding time of 5 days. The value μ of the reference model is chosen such that any vorticity wave, described by a basis function Y_{mn}, is exponentially damped with an e-folding time of 90 days. For the non-dimensional values of ν and μ, we thus have

$$\nu = \frac{1}{24 \times 60 \times 60} \frac{1}{\Omega} \frac{1}{85^2} \frac{1}{5} = 4.394 \times 10^{-6}, \quad (91)$$

$$\mu = \frac{1}{24 \times 60 \times 60} \frac{1}{\Omega} \frac{1}{90} = 1.764 \times 10^{-3}. \quad (92)$$

The forcing F is a single basis function with $m = n = 5$, i.e. $F = F_{5,5} Y_{5,5}$ with $F_{5,5} = 2^{-1/2} = 0.707\,106$. The time stepping is by fourth-order Runge–Kutta with a time step of $0.065\,449$ non-dimensional time units, corresponding to 14.959 min.‡

---

‡The question of why 15 min was not chosen has the somewhat unsatisfactory answer that it is the consequence of some confusion concerning sidereal and solar days at the beginning of this work. It has no serious consequences for anything that follows, except that when the horizontal axis of some of the figures shows 'days', actually sidereal days are meant.
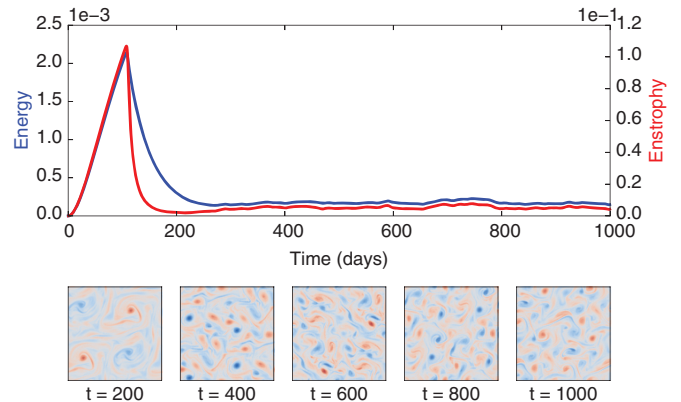
Figure 3. Time evolution of the energy and enstrophy of the reference model, run at truncation $N = 85$. The energy and enstrophy first increase, but after some time the flow becomes unstable and changes into a turbulent state that later becomes statistically stationary. The small panels below the graphs are the vorticity fields at $t = 200, 400, 600, 800$ and 1000 days. These fields, referred to as initial states 1, 2, 3, 4 and 5, are used as initial fields for the simulations with the different model versions.

The model that we consider as 'reality' will be run with $N = 85$, using a 256×256 computational grid to calculate the Jacobian without aliasing. This model, of which the fields are used at the lower truncation $M = 42$, will be called the 'reference' model. The same model, run with $M = 42$, will be referred to as the 'model' and will be considered in three variants. If the truncation $M = 42$ is all that distinguishes it from the reference model, the model will be referred to as the 'unparametrized' model. If the viscosity is increased by replacing 85 in Eq. (91) by 42 (giving a value $\nu = 1.800 \times 10^{-5}$), we obtain the second variant of the model, called the 'conventional' model. If we use Eq. (81) with $\langle \psi^{\mathcal{U}} \rangle$ and $\langle \zeta^{\mathcal{U}} \rangle$ given by Eqs (77), (78) and (82) instead of increasing the viscosity, we obtain the third variant, referred to as the 'maximum entropy' model. In all three variants of the model (unparametrized, conventional and maximum entropy), the computational grid of 256×256 points of the reference model is used for reasons of simplicity and consistency, although for the unparametrized and conventional model a coarser grid would have sufficed. The same grid is used to produce plots of vorticity fields and to calculate the statistics.

The basis of the numerical experiments to be discussed in the next two sections is an integration of the reference model for a period of 1000 days, starting from the state of rest. The result in terms of the energy and enstrophy as a function of time and five snapshots of the vorticity field are shown in Figure 3. These five snapshots are referred to as initial states 1, 2, 3, 4 and 5 and are the vorticity fields at $t = 200, 400, 600, 800$ and 1000 days. We will first investigate how the three low-resolution models perform in reproducing the low-resolution vorticity fields as generated by the reference model. We then consider how well the low-resolution models reproduce the climate of the reference run, where the climate is defined in terms of the mean and variance of the energy and enstrophy as well as their spectra. More details on the numerical experiments can be found in Kalverla (2015).

8. Forecast skill

For each of the five initial states shown in Figure 3, we start a 100 days run with the reference model, the unparametrized model, the conventional model and the maximum entropy model. For the latter three models, the initial fields are truncated to $M = 42$. The aim of these small ensembles of runs is to investigate to what extent the three models, with their different parametrizations, are able to reproduce the resolved scales as generated by the reference model.
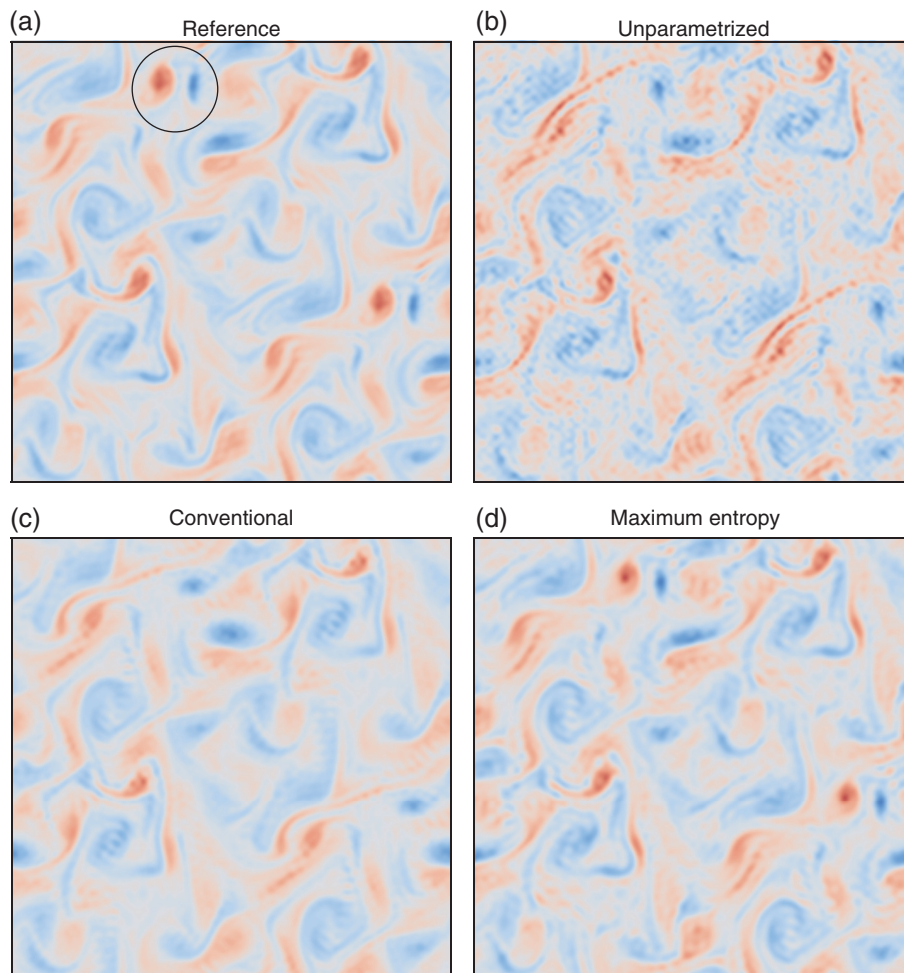
Figure 4. Vorticity fields from four model simulations at day 30 of the integration, starting from initial field 3, i.e. from the reference run after 600 days. A *T*42 truncation has been used to plot the vorticity fields, including the reference, on a 256 × 256 grid from 0−2π in both directions. (a) Reference run, (b) unparametrized run, (c) conventional run and (d) maximum entropy run. The circle in the upper left panel highlights a dipole structure that is only reproduced in the maximum entropy run.

8.1. Qualitative analysis of vorticity fields

The vorticity fields at day 30 of the model runs that started from initial state 3, i.e. from $t = 600$ days in Figure 3, are shown in Figure 4. A feature that stands out is the grainy texture of the flow field in the unparametrized run. This is a manifestation of the smallest scales of the model (near the truncation limit) and is clearly unrealistic. Both parametrizations reduce this noise; the conventional parametrization, in particular, is very effective and damps all the small-scale structures rigorously. The maximum entropy parametrization is not completely smooth, but it is still a considerable improvement over the unparametrized run. We also note that the conventional parametrization tends to reduce the extremes in the vorticity fields rather drastically. In contrast, the maximum entropy parametrization is less diffusive and keeps the vorticity extremes more closely at their values in the reference run.

If we analyze the vorticity fields in the course of time (not shown) 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 top boundary, highlighted by the circle in Figure 4(a). 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 parametrization has damped out most small-scale features. At day 60, most of the correlation with the reference run is lost for all parametrizations.

8.2. Quantitative analysis of vorticity fields

To quantify the performance of the three low-order models in reproducing the time evolution of the high-resolution model, we calculated the root-mean-square difference (RMSD) and the correlation (CORR) between the vorticity fields of the latter models and the reference model. This was done on the basis of the low-resolution representation of the fields and making use of the 256 × 256 grid for the calculations. For each model, the results are averaged over the ensemble of five runs and are shown in Figure 5. All models 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. Also note that, in terms of correlation, the conventional parametrization does not seem to improve the simulation at all.

9. Model climate

We now consider the question as to what extent the three low-resolution models reproduce the statistics of the high-resolution reference model. We do this by analyzing long 5000 day integrations with the reference model and the other models, all starting from initial state 3, i.e. from $t = 600$ days in Figure 3. We will define the climate of the models in terms of the distribution of the energy and enstrophy as well as their spectra. The results are obtained by sampling the runs twice every day for the energy and enstrophy distribution and once every ten days for the spectra.
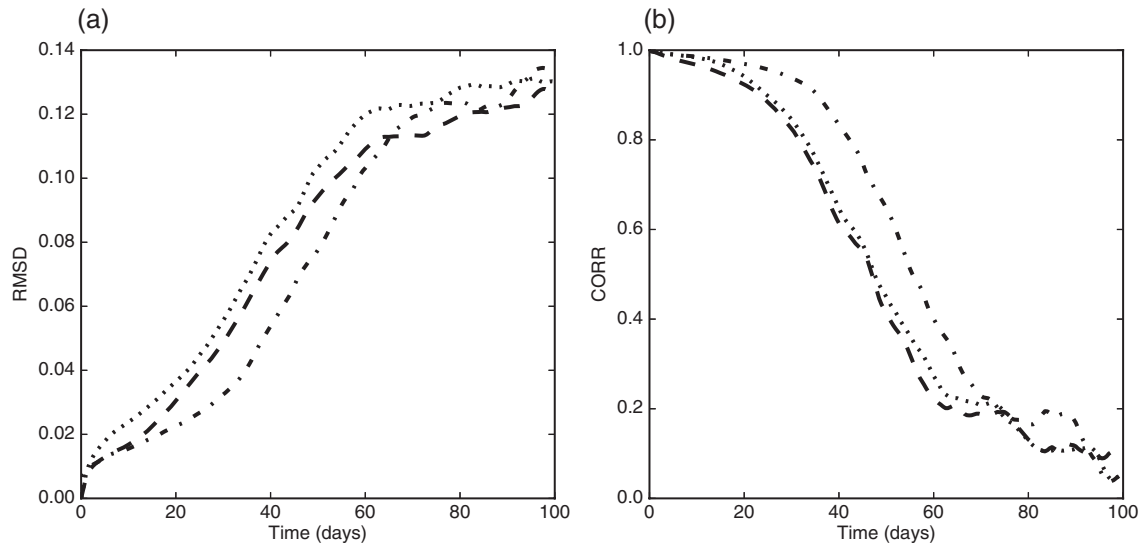
Figure 5. (a) Time evolution of the root-mean-square difference (RMSD) between the vorticity fields generated by the three models (unparametrized, conventional and maximum entropy) and the reference model, calculated on the basis of 256×256 grid points using truncation $M = 42$ and averaged over the ensemble of five runs. (b) Time evolution of the ensemble-averaged correlation (CORR) between the three models and the reference model, based on the same grid and truncation. The dotted curves refer to the unparametrized model, the dashed curves to the conventional model and the dash–dotted curves to the maximum entropy model.
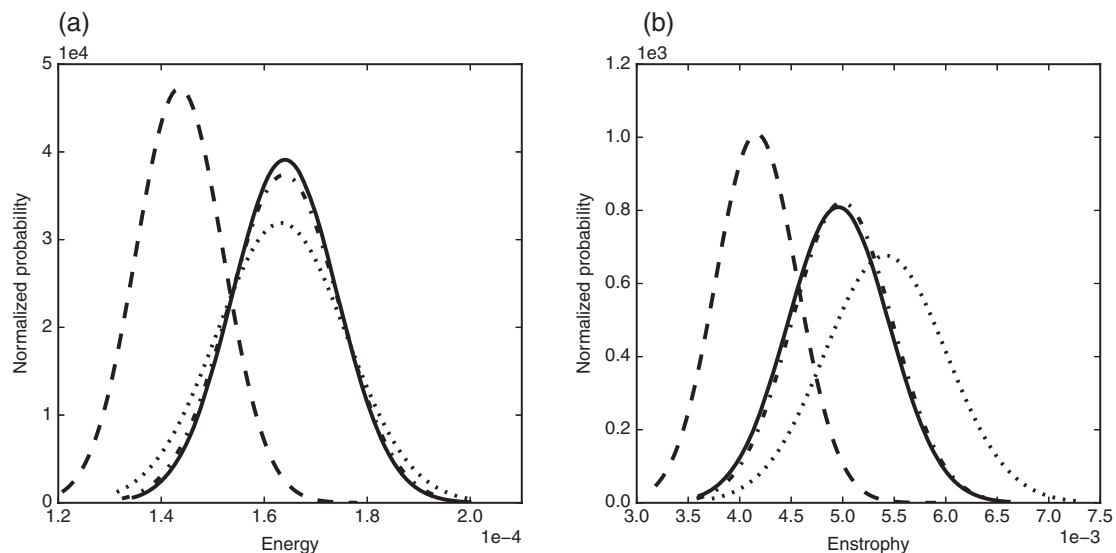

Figure 6. Normal distributions based on the mean and variance of a large series of energy and enstrophy values from climate simulations with four different models. The data are obtained from 5000 day runs starting from initial state 3, sampling the output twice every day. The distribution of the reference run is displayed by a solid curve, the distribution of the unparametrized run by a dotted curve, the distribution of the conventional run by a dashed curve and the distribution of the maximum entropy run by a dash–dotted curve. Panel (a) refers to the energy, panel (b) to the enstrophy.

9.1. Energy and enstrophy distributions

In all model runs, the energy and enstrophy fluctuate around a mean value. The fluctuations can be characterized by a variance which, together with the mean, can be used to construct normal distributions representing the probability density function for the energy and enstrophy.[§] These distributions are shown in Figure 6. The conventional parametrization 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 unparametrized run is quite similar to the reference run, but the distribution is somewhat broader. The enstrophy in the unparametrized run is overestimated. The near-coincidence of the maximum entropy and reference distributions of both energy and enstrophy was not possible to achieve by modifying the viscosity in the conventional parametrization and demonstrates

that the maximum entropy parametrization performs best in reproducing the reference climate.

9.2. Energy and enstrophy spectra

The energy and enstrophy spectra of the four model runs are shown in Figure 7. The spectra are calculated by adding the energy and enstrophy contributions of all spectral coefficients (m, n) for which $k - 1/2 \leq (m^2 + n^2)^{1/2} < k + 1/2$, for $k = 1$ to $k = 2^{1/2}M$ with $M = 42$ and then averaged over all samples. Note that, between $k = M$ and $k = 2^{1/2}M$, marked by the vertical dashed lines in Figure 7, all spectra steepen because of the rectangular shape of $\mathcal{R}$. This is an artefact of the square truncation and this part of the spectra will not be considered. The results show that the energy and enstrophy on larger scales are nearly identical in all model simulations, but on smaller scales substantial differences occur. For the model with no parametrization, too much energy and enstrophy piles up in the smallest scales, but the conventional parametrization, on the other hand, reduces the energy and enstrophy at the smallest scales too much. Again, the maximum entropy parametrization improves the simulation:
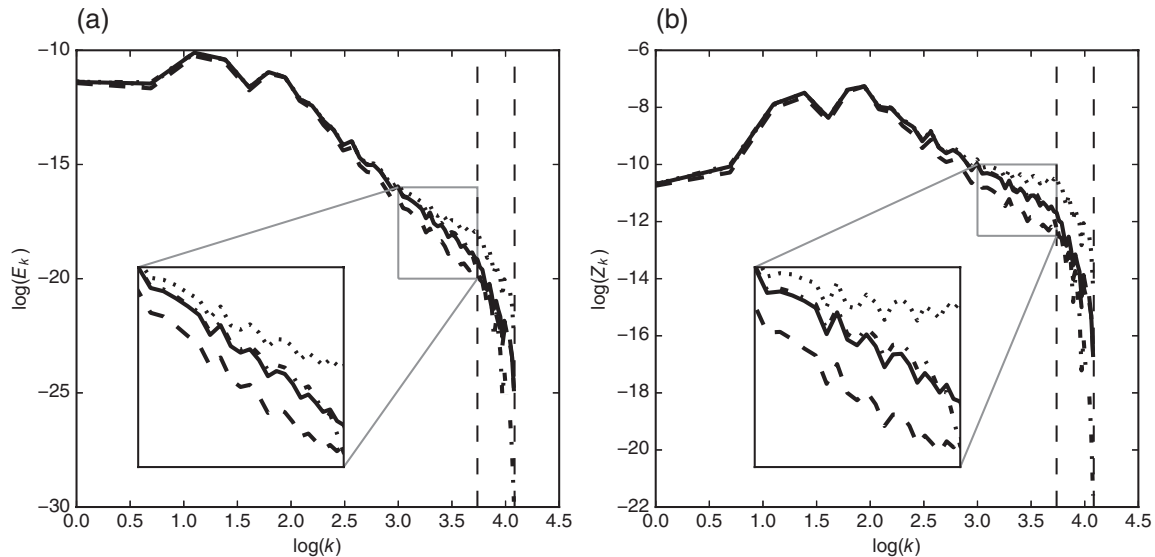

[§]We checked that, after a long integration as used here (5000 days), the values of the mean and variance stabilized and the histograms of energy and enstrophy converged towards the bell-shape characteristic of a normal distribution.

Figure 7. Energy and enstrophy spectra of the flow field as simulated with the reference model (solid lines), the unparametrized model (dotted lines), the conventional model (dashed lines) and the maximum entropy model (dash–dotted lines). All spectra are based on long 5000 day integrations, starting from initial state 3 in Figure 3, the spectra being averages over the data ten days apart. The insets show the differences between the spectra close to the truncation limit. Panel (a) refers to the energy, panel (b) to the enstrophy.

the representation of the spectra is rather accurate, even at the smallest scales.

10. Conclusions

We have implemented a new parametrization of the effects of the unresolved scales on the resolved scales in a simple model of a fluid dynamical system. The model is a spectral representation of a forced-dissipative version of the vorticity equation on a doubly periodic flow domain. The parametrization that we propose is based on a probability density function for the unresolved scales. This probability density function is derived from the principle of maximum entropy, in which zero-average time derivatives of the energy and enstrophy at unresolved scales are used as constraints. If terms in the energy and enstrophy equations that involve Jacobians of unresolved variables are ignored, the probability density function is a product of independent normal distributions, the parameters of which depend on the model and the state of the resolved scales. The resulting parametrization is a closed system in terms of the resolved scales that resembles the anticipated potential vorticity method (APVM) of Sadourny and Basdevant (1985).

We have shown that, in terms of both short-range performance and long-range statistics, the maximum entropy parametrization represents the interaction of the unresolved scales with the resolved scales realistically in the simple model that we considered. It performs better than a model with no interaction at all or a model with a conventional parametrization in terms of a higher value of viscosity. Most importantly, however, the method achieves this without any tuning of the parameters. Once the decision is made to base the unresolved-scale probability density function on maximum entropy in combination with constraints on the time derivatives of energy and enstrophy, the procedure leads to unique values of the parameters. In particular, it fixes the form of the non-negative linear operator in the corresponding APVM, which, for that matter, is rather different from the class of linear operators considered by Sadourny and Basdevant (1985) and Thuburn *et al.* (2014).

Requiring that the probability density function of the unresolved scales has maximum entropy, constrained by the average budgets of energy and enstrophy, is the fundamental closure assumption of our approach. We made an additional simplifying approximation by neglecting terms in the budgets that involve Jacobians and would lead to correlations between

the unresolved variables if retained. This approximation – which in future work might be avoided – enabled us to proceed with the calculations rather straightforwardly. The resulting absence of correlations implied that the equation for the resolved scales assumed a form in which only the average stream function and vorticity of the unresolved scales are involved. That this worked so well is actually quite surprising. It might be related to the particular choice of basis functions (sines and cosines) that naturally fit the dynamics and boundary conditions of the two-dimensional flow system. This and other aspects of the problem need further scrutinizing.

Appendix

Energy and enstrophy

In this Appendix, we recapture the derivation of a few results regarding energy and enstrophy conservation in two-dimensional incompressible flow.

The energy (actually energy density) is defined by

$$E = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \frac{1}{2}\mathbf{v}^2 \, dx \, dy$$
$$= -\left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi}\int_0^{2\pi} \frac{1}{2}\psi\zeta \, dx \, dy = -\frac{1}{2}(\psi,\zeta). \quad (A1)$$

The second equality in Eq. (A1) is the result of the following series of identities:

$$\mathbf{v}\cdot\mathbf{v} = \nabla\psi\cdot\nabla\psi = \nabla\cdot(\psi\nabla\psi) - \psi\nabla^2\psi$$
$$= \nabla\cdot(\psi\nabla\psi) - \psi\zeta, \quad (A2)$$

in combination with Gauss' theorem and the assumption of doubly periodic flow. The third equality in Eq. (A1) is the result of the definition of the inner product, i.e. Eq. (7). In an analogous

way, we have for the time derivative of the energy

$$
\begin{aligned}
\frac{dE}{dt} &= \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \mathbf{v} \cdot \frac{\partial \mathbf{v}}{\partial t} \, dx \, dy \\
&= -\left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \psi \frac{\partial \zeta}{\partial t} \, dx \, dy = -(\psi, \frac{\partial \zeta}{\partial t}). \quad \text{(A3)}
\end{aligned}
$$

Substituting Eq. (13), we obtain:

$$
\begin{aligned}
\frac{dE}{dt} &= (\psi, PJ(\psi, \zeta)) \\
&\quad - \nu(\psi, \nabla^2 \zeta) - \mu(\psi, (F - \zeta)). \quad \text{(A4)}
\end{aligned}
$$

This expression can be simplified by making use of the fact that both the orthogonal projection operator P and the Laplace operator ∇^2 are self-adjoint with respect to the inner product defined in Eq. (7), i.e.

$$
(\xi, P\chi) = (P\xi, \chi), \quad \text{(A5)}
$$

$$
(\xi, \nabla^2 \chi) = (\nabla^2 \xi, \chi). \quad \text{(A6)}
$$

Furthermore, for the Jacobian it can be shown by integration by parts and by using the periodic boundary conditions that

$$
(\psi, J(\xi, \chi)) = (J(\psi, \xi), \chi). \quad \text{(A7)}
$$

For the first term on the right-hand side of Eq. (A4), we thus have

$$
(\psi, PJ(\psi, \zeta)) = (P\psi, J(\psi, \zeta)) = (\psi, J(\psi, \zeta)) = 0. \quad \text{(A8)}
$$

In the second equality of Eq. (A8), we use that $P\psi = \psi$, as ψ already belongs to the space of functions on to which P projects. The last equality of Eq. (A8) is a consequence of Eq. (A7) and the fact that the Jacobian of two identical fields is zero. For the second term on the right-hand side of Eq. (A4), we have

$$
(\psi, \nabla^2 \zeta) = (\nabla^2 \psi, \zeta) = (\zeta, \zeta). \quad \text{(A9)}
$$

Taken together, this yields Eq. (16).

The enstrophy of the flow is defined as the integrated squared vorticity. Its density is given by

$$
Z = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \frac{1}{2}(\zeta)^2 \, dx \, dy = \frac{1}{2}(\zeta, \zeta). \quad \text{(A10)}
$$

For its time derivative, we have

$$
\frac{dZ}{dt} = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \zeta \frac{\partial \zeta}{\partial t} \, dx \, dy = (\zeta, \frac{\partial \zeta}{\partial t}). \quad \text{(A11)}
$$

When substituting Eq. (13) we obtain, in analogy to Eq. (A4),

$$
\begin{aligned}
\frac{dZ}{dt} &= -(\zeta, PJ(\psi, \zeta)) \\
&\quad + \nu(\zeta, \nabla^2 \zeta) + \mu(\zeta, (F - \zeta)). \quad \text{(A12)}
\end{aligned}
$$

As in the case of the energy, the Jacobian term vanishes:

$$
(\zeta, PJ(\psi, \zeta)) = (P\zeta, J(\psi, \zeta)) = (\zeta, J(\psi, \zeta)) = 0. \quad \text{(A13)}
$$

The last equality follows if we interchange the arguments of the Jacobian and use Eq. (A7) and the fact that the Jacobian of two identical fields is zero. The result is Eq. (17).

We next consider the energy and enstrophy equations for the resolved and unresolved scales of the model. Taking the time derivative of the resolved energy, the expression for which is given after Eqs (25) and (26), and then substituting Eq. (23), we obtain

$$
\begin{aligned}
\frac{dE^{\mathcal{R}}}{dt} &= -(\psi^{\mathcal{R}}, \frac{\partial \zeta^{\mathcal{R}}}{\partial t}) = (\psi^{\mathcal{R}}, P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})) \\
&\quad - \nu(\psi^{\mathcal{R}}, \nabla^2 \zeta^{\mathcal{R}}) - \mu(\psi^{\mathcal{R}}, (F^{\mathcal{R}} - \zeta^{\mathcal{R}})). \quad \text{(A14)}
\end{aligned}
$$

The Jacobian term can be worked out as follows:

$$
\begin{aligned}
&(\psi^{\mathcal{R}}, P^{\mathcal{R}} J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})) \\
&= (P^{\mathcal{R}} \psi^{\mathcal{R}}, J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})) \\
&= (\psi^{\mathcal{R}}, J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})). \quad \text{(A15)}
\end{aligned}
$$

For the latter expression, we have

$$
\begin{aligned}
&(\psi^{\mathcal{R}}, J(\psi^{\mathcal{R}} + \psi^{\mathcal{U}}, \zeta^{\mathcal{R}} + \zeta^{\mathcal{U}})) \\
&= (\psi^{\mathcal{R}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) + (\psi^{\mathcal{R}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{U}})) \\
&+ (\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}})) + (\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})). \quad \text{(A16)}
\end{aligned}
$$

The first two of the four terms on the right-hand side of this expression vanish as a consequence of Eq. (A7) and the fact that the Jacobian of two identical fields is zero. The third term can be rewritten as

$$
\begin{aligned}
&(\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{R}})) = -(\psi^{\mathcal{R}}, J(\zeta^{\mathcal{R}}, \psi^{\mathcal{U}})) \\
&= -(J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}}), \psi^{\mathcal{U}}) = -(\psi^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})). \quad \text{(A17)}
\end{aligned}
$$

This yields for the time derivative of the resolved energy, also using Eq. (A6) to rewrite the viscosity term,

$$
\begin{aligned}
\frac{dE^{\mathcal{R}}}{dt} &= -(\psi^{\mathcal{U}}, J(\psi^{\mathcal{R}}, \zeta^{\mathcal{R}})) + (\psi^{\mathcal{R}}, J(\psi^{\mathcal{U}}, \zeta^{\mathcal{U}})) \\
&\quad - \nu(\zeta^{\mathcal{R}}, \zeta^{\mathcal{R}}) - \mu(\psi^{\mathcal{R}}, (F^{\mathcal{R}} - \zeta^{\mathcal{R}})). \quad \text{(A18)}
\end{aligned}
$$

The equations for the time derivative of the unresolved energy and the resolved and unresolved enstrophy can be obtained in an analogous way.

References

Bouchet F, Venaille A. 2012. Statistical mechanics of two-dimensional and geophysical flows. *Phys. Rep.* **515**: 227–295.

Coiffier J. 2011. *Fundamentals of Numerical Weather Prediction*. Cambridge University Press: Cambridge, UK.

Cushman-Roisin B, Beckers JM. 2011. *Introduction to Geophysical Fluid Dynamics – Physical and Numerical Aspects*, Vol. 101. Academic Press: Cambridge, MA.

Frederiksen JS, Kepert SM. 2006. Dynamical subgrid-scale parameterizations from direct numerical simulations. *J. Atmos. Sci.* **63**: 3006–3019.

Jaynes ET. 2003. *Probability Theory – The Logic of Science*. Cambridge University Press: Cambridge, UK.

Kalverla P. 2015. 'A maximum entropy approach to the parameterization of subgrid-scale processes in two-dimensional flows'. KNMI Internal Report IR-2015-01. http://bibliotheek.knmi.nl/internrapport.html (accessed 5 May 2016).

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

Sadourny R, Basdevant C. 1985. Parameterization of subgrid scale barotropic and baroclinic eddies in quasi-geostrophic models: Anticipated potential vorticity method. *J. Atmos. Sci.* **42**: 1353–1363.

Stensrud DJ. 2009. *Parameterization Schemes – Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press: Cambridge, UK.

Thuburn J, Kent J, Wood N. 2014. Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. *Q. J. R. Meteorol. Soc.* **140**: 626–638.

Verkley WTM. 2011. A maximum entropy approach to the problem of parametrization. *Q. J. R. Meteorol. Soc.* **137**: 1872–1886.