



---

# An Energy Conserving Restoration Scheme for the Shallow Water Equations

James Kent<sup>a,b\*</sup>, Christiane Jablonowski<sup>b</sup>, John Thuburn<sup>c</sup> and Nigel Wood<sup>d</sup>

<sup>a</sup> *Computing and Mathematics, University of South Wales, Pontypridd, CF37 1DL, UK*

<sup>b</sup> *Climate and Space, University of Michigan, Ann Arbor, MI 48109-2143, USA*

<sup>c</sup> *College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QE, UK*

<sup>d</sup> *Met Office, FitzRoy Road, Exeter EX1 3PB, UK*

\*Correspondence to: Computing and Mathematics, University of South Wales, Pontypridd, CF37 1DL, UK. E-mail:

james.kent@southwales.ac.uk

---

**The numerical methods that solve the governing equations in an atmospheric dynamical core are designed to dissipate potential enstrophy and prevent the buildup of kinetic energy at the grid scale. A side effect of this is the dissipation of total energy which should be conserved. Energy fixers are used in climate models to replace the dissipated energy by modifying the temperature in the thermodynamic equation, and stochastic backscatter schemes have also been developed for use in weather prediction models. Here, we present the first steps towards designing a deterministic energy conserving restoration scheme that considers the conversion of kinetic energy to heat, replacing kinetic energy lost due to model error, and the backscatter of kinetic energy.**

**The energy conserving restoration scheme (ECRS) is presented in the context of the shallow water equations on the sphere. It is designed to be used with any existing shallow water equation scheme (called the preliminary scheme) which can adequately dissipate potential enstrophy, and in this paper we use a semi-implicit semi-Lagrangian (SISL) scheme. For each prognostic variable a spatial pattern is chosen; this is added to the preliminary scheme solution, and the amount added is calculated to ensure energy conservation. Results from short-term test cases show that ECRS and SISL have very similar error norms. For long-term simulations ECRS conserves energy to a good approximation whereas SISL dissipates energy.**

*Key Words:* Conservation; Dissipation; Dynamical Core; Backscatter; Spherical Geometry

*Received ...*

## 1. Introduction

Many quantities are conserved by the continuous adiabatic frictionless governing equations. It is desirable for atmospheric model dynamical cores to have discrete analogues of the most important of these conservation properties (Thuburn, 2008). Some of these quantities, such as mass, can straightforwardly be conserved in the discrete approximation to the continuous equations. Nonlinear quantities such as energy, however, may be less straightforward.

The numerical methods that comprise dynamical cores are designed to dissipate quantities that should be transferred to the unresolved subgrid scales, such as potential enstrophy. Often however, this also leads to the dissipation of total energy. The conservation of total energy is important for long climate simulations (Boville, 2000; Jablonowski and Williamson, 2011), although exact conservation may be unnecessary for short term weather prediction models. Energy is usually lost by a model due to the dissipation of kinetic energy. Kinetic energy is dissipated by the model to prevent a buildup of noise at the grid scale. This dissipation is interpreted as capturing the downscale transfer of kinetic energy to the subgrid scales, and then to the turbulent microscale where it is converted into heat. To conserve total energy, the dissipated kinetic energy may be restored by adding it to the temperature field in the thermodynamic equation. This can be achieved through the use of frictional heating terms (Burkhardt and Becker, 2006) and/or energy fixers.

Traditional energy fixers conserve total energy by adding heat to the temperature field, although the choice of heat function can have a large (even detrimental) effect on the circulation (Williamson *et al.*, 2009). These fixers therefore assume that all of the kinetic energy dissipation is ‘physical’, and hence should be replaced by heat. However, in many models the kinetic energy dissipation is a side-effect of dissipating the potential enstrophy, and therefore much of the kinetic energy dissipated by a model should be considered as model error. This implies that some of the replaced energy should be added back to the kinetic energy. In the atmosphere there is also an upscale cascade of kinetic energy, called backscatter. The effects of backscatter and the fixing of overly diffusive numerics have been included in short term

ensemble prediction models using stochastic methods (Berner *et al.*, 2009; Bowler *et al.*, 2009; Shutts, 2008). **As with the energy fixers, many stochastic backscatter schemes are mostly correcting the kinetic energy dissipated due to model error. While true backscatter is associated with the unknown subgrid scales, the resolved scales and the model errors on those scales are known, at least in principle. Therefore, a deterministic approach to fixing the energy based on the error of the underlying model may be better justified than a stochastic approach.** Here, in this paper, we present a deterministic method that restores the dissipated energy and conserves total energy. This is the first step towards designing an energy fixer that considers the combined effect of: the conversion of kinetic energy to heat, replacing kinetic energy lost due to model error, and the backscatter of kinetic energy. This fixer or “restoration scheme” is built upon the method for the barotropic vorticity equation of (Thuburn *et al.*, 2014), and is presented as a model for the shallow water equations.

Thuburn *et al.*, (2014) proposed a simple energy fixer for numerical models of the barotropic vorticity equation. They explored how the effectiveness of the fixer depended on the length scales at which energy is restored for different vorticity advection schemes. A major simplification in that case comes from the fact that only kinetic energy needs to be considered. Our extension of the Thuburn *et al.*, (2014) method is for the shallow water equations. The shallow water equations on the sphere are commonly used as a first step towards building an atmospheric dynamical core (Williamson *et al.*, 1992). The shallow water equations possess many of the same properties as the primitive equations, for example total energy and potential enstrophy are conserved in the continuous equations. For an energy fixer for the shallow water case, both kinetic and potential energy must be considered, in particular the partitioning of the restored energy between the two. Moreover, potential energy may be decomposed into available and unavailable components (Lorenz, 1955), and again how the restored energy is partitioned must be considered. A closely related issue is that, at least approximately, energy may be decomposed into a balanced component associated with Rossby waves and coherent vortices, and an unbalanced component associated with inertio-gravity waves. Energy restored into the two components will affect the dynamics through different

mechanisms and on different timescales. Although the energy conserving restoration scheme presented in our paper is developed for the shallow water equations, a goal of this work is the extension of the method for application to dynamical cores of atmospheric models.

The paper is structured as follows. Section 2 describes the shallow water equations on the sphere. The energy conserving restoration scheme is described in Section 3, and results from well-known shallow water equation test cases are presented in Section 4. Section 5 investigates changes in model parameters, and the conclusions and discussion are in Section 6.

## 2. The Shallow Water Equations on the Sphere

The shallow water equations are made up of the continuity equation and some form of the momentum equations. The continuity equation, for fluid depth  $h$ , is

$$\frac{\partial h}{\partial t} + \frac{1}{a \cos \varphi} \left[ \frac{\partial uh}{\partial \lambda} + \frac{\partial vh \cos \varphi}{\partial \varphi} \right] = 0, \quad (1)$$

where  $u$  and  $v$  are the zonal and meridional velocities respectively,  $\lambda$  and  $\varphi$  are the longitude and latitude respectively,  $t$  is time and  $a$  is the Earth's radius. The momentum equations can be written in a number of different forms (for example, see Williamson *et al.*, (1992)). The advective form is

$$\frac{\partial u}{\partial t} + \frac{u}{a \cos \varphi} \frac{\partial u}{\partial \lambda} + \frac{v}{a} \frac{\partial u}{\partial \varphi} - \left( f + \frac{u}{a} \tan \varphi \right) v + \frac{g}{a \cos \varphi} \frac{\partial H}{\partial \lambda} = 0, \quad (2)$$

$$\frac{\partial v}{\partial t} + \frac{u}{a \cos \varphi} \frac{\partial v}{\partial \lambda} + \frac{v}{a} \frac{\partial v}{\partial \varphi} + \left( f + \frac{u}{a} \tan \varphi \right) u + \frac{g}{a} \frac{\partial H}{\partial \varphi} = 0, \quad (3)$$

where  $H$  is the height of the free surface ( $H = h + h_0$ , where  $h_0$  denotes the height of any underlying mountains),  $g$  is gravity, and  $f = 2\Omega \sin \varphi$  is the Coriolis parameter (where  $\Omega$  is the rotation rate of the Earth). The momentum equations can also be written in terms of the vorticity (or potential vorticity, PV, see Bates *et al.*, (1995)) and divergence, where the relative vorticity,  $\zeta$ , potential vorticity,  $q$ , and divergence,  $D$ , are defined as

$$\zeta = \frac{1}{a \cos \varphi} \left[ \frac{\partial v}{\partial \lambda} - \frac{\partial u \cos \varphi}{\partial \varphi} \right], \quad (4)$$

$$q = \frac{\zeta + f}{h}, \quad (5)$$

$$D = \frac{1}{a \cos \varphi} \left[ \frac{\partial u}{\partial \lambda} + \frac{\partial v \cos \varphi}{\partial \varphi} \right]. \quad (6)$$

The vorticity and divergence can be inverted to calculate the velocities using the stream function,  $\psi$ , and velocity potential,  $\chi$ ,

$$\nabla^2 \psi = \zeta, \quad (7)$$

$$\nabla^2 \chi = D, \quad (8)$$

$$u = -\frac{1}{a} \frac{\partial \psi}{\partial \varphi} + \frac{1}{a \cos \varphi} \frac{\partial \chi}{\partial \lambda}, \quad (9)$$

$$v = \frac{1}{a \cos \varphi} \frac{\partial \psi}{\partial \lambda} + \frac{1}{a} \frac{\partial \chi}{\partial \varphi}, \quad (10)$$

where the Laplacian operator on the sphere is given as

$$\nabla^2() = \frac{1}{a^2 \cos^2 \varphi} \frac{\partial^2 ()}{\partial \lambda^2} + \frac{1}{a^2 \cos \varphi} \frac{\partial}{\partial \varphi} \left( \cos \varphi \frac{\partial ()}{\partial \varphi} \right). \quad (11)$$

As this study investigates energy and potential enstrophy transfers, we use the total energy defined by

$$E = \int_A \frac{1}{2} h (u^2 + v^2) + \frac{1}{2} g (H^2 - h_0^2) dA, \quad (12)$$

and the total potential enstrophy defined by

$$Z = \int_A \frac{1}{2} h q^2 dA, \quad (13)$$

where the integral is over the whole domain

$$\int_A dA = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} a^2 \cos \varphi d\varphi d\lambda. \quad (14)$$

The total energy and potential enstrophy are conserved by the continuous shallow water equations (1)-(3).

## 3. Energy Conserving Restoration Scheme

The energy conserving restoration scheme is designed to be used with an existing dissipative shallow water scheme. The existing shallow water scheme is called the preliminary scheme. The

preliminary scheme, for example a scheme with explicit hyper-diffusion or a scheme with diffusion implicit in the numerics, is used to dissipate potential enstrophy and model the downscale cascade from resolved to unresolved scales. In general this preliminary scheme will also dissipate energy. The application of the restoration scheme will replace the energy that is lost through the diffusive preliminary scheme.

The energy conserving restoration scheme replaces the lost energy by adding an additional term to the velocity and depth fields. This term is made up of an energy conserving coefficient,  $\alpha$ , and a pattern,  $\delta F$ , that is a function of the preliminary solution, where  $F$  is one of  $u$ ,  $v$  and  $h$ . Denoting variables at the beginning of the time step with the superscript  $n$ , the preliminary scheme with the superscript  $p$ , and the variables at the end of the time step (after the application of the restoration scheme) with the superscript  $n + 1$ , the energy is added back as

$$\begin{aligned} u^{n+1} &= u^p + \alpha \delta u, \\ v^{n+1} &= v^p + \alpha \delta v, \\ h^{n+1} &= h^p + \alpha \delta h. \end{aligned} \quad (15)$$

### 3.1. Restoration Variable Pattern

As shown by Thuburn *et al.*, (2014) for the vorticity equation, the optimal choice of variable pattern is dependent on the preliminary scheme. The energy conserving restoration scheme replaces the energy that is dissipated by the preliminary scheme, and the energy is generally removed from the small, near grid scales. This implies that the variable pattern should be an approximation of these scales or of the diffusion in the preliminary scheme. There are a number of choices to use for the variable pattern. The difference between the preliminary solution and a large scale average, for example  $\delta F = F - \overline{F}^r$  where  $r \in \mathbb{N}$  and  $\overline{F}^r$  is the average of  $F$  over a  $r\Delta\lambda \times r\Delta\varphi$  region in  $(\lambda, \varphi)$  space, is an approximation that will add energy to the small scales. Numerically, the  $r\Delta\lambda \times r\Delta\varphi$  average is **an area average located at the centre of the current grid cell. This averaging extends a distance of  $r\Delta\lambda/2$  and  $r\Delta\varphi/2$  in  $\lambda$  and  $\varphi$  respectively from the centre point, and can be thought of as an averaging from cell**

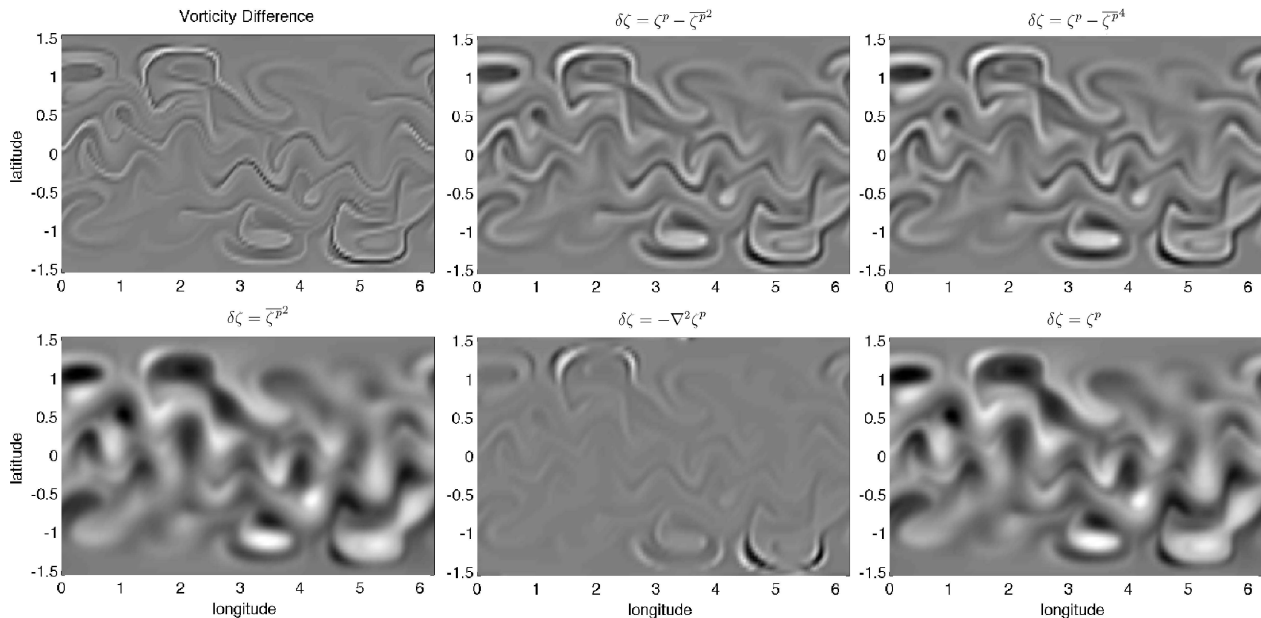
**centres to vertices and then a further averaging of these back to the cell centre.** For example, for  $r = 2$  the average at the grid cell with indices  $i$  and  $j$  (in the zonal and meridional direction respectively) would be

$$\begin{aligned} \overline{F}_{i,j}^2 &= \frac{1}{4} \left[ F_{i,j} + \frac{1}{2} (F_{i+1,j} + F_{i-1,j} + F_{i,j+1} + F_{i,j-1}) \right. \\ &\quad \left. + \frac{1}{4} (F_{i+1,j+1} + F_{i+1,j-1} + F_{i-1,j+1} + F_{i-1,j-1}) \right]. \end{aligned}$$

Note that on a latitude-longitude grid this averaging procedure is not an isotropic filter as the filtered out scales depend on the size of each grid cell, which decreases towards the poles. An area weighted average could also be used for the averaging procedure, although for  $r = 2$  and  $r = 4$  **no noticeable difference is found** between the weighted and unweighted average. Using the large scale average  $\delta F = \overline{F}^r$  would add energy into the large scales. The Laplacian operator would approximate the diffusion in the preliminary scheme, hence using the negative Laplacian,  $\delta F = -\nabla^2 F$ , would replace energy as an anti-diffusion term. Finally, setting  $\delta F = F$  would add energy to the range of scales that are already present in  $F$ .

**As large-scale atmospheric flows are dominated by balance, it is desirable that any changes to the flow by the energy conserving restoration scheme minimise any introduction of imbalance. Therefore we follow Bowler *et al.*, (2009) and use variable patterns that are non-divergent. To do this we make use of the relative vorticity and the stream function, and use a depth pattern that is related to the vorticity patterns.** Using  $\delta\zeta$ , a variable pattern for relative vorticity, the velocity patterns can be recovered by solving for the stream function (7) of  $\delta\zeta$ , and the depth pattern can be recreated from geostrophic balance. This will create the variable pattern for velocity that is based on the vorticity dynamics.

Figure 1 provides a comparison of different variable patterns for relative vorticity,  $\delta\zeta$ . The freely decaying vortices test (see Section 4.3) is simulated to day 5 on a high-resolution latitude-longitude grid of  $1024 \times 512$  grid points using the semi-implicit semi-Lagrangian (SISL) shallow water model by Thuburn *et al.*, (2010) and Zerroukat *et al.*, (2009b) (see also Section 4). The day 5 snapshot of this simulation then provides the initial conditions for a one time step run using the same model in



**Figure 1.** Comparison of vorticity patterns,  $\delta\zeta$  with the difference between the reference relative vorticity and the SISL scheme on the  $128 \times 64$  latitude-longitude grid (top left). We use  $\delta\zeta = \zeta^p - \bar{\zeta}^2$  (top centre),  $\delta\zeta = \zeta^p - \bar{\zeta}^4$  (top right),  $\delta\zeta = \bar{\zeta}^2$  (bottom left),  $\delta\zeta = -\nabla^2 \zeta^p$  (bottom centre), and  $\delta\zeta = \zeta^p$  (bottom right). All plots are normalized, with negative vorticity shown as black and positive vorticity white.

a low-resolution ( $128 \times 64$ ) configuration. We also use these initial conditions for a simulation to the same time, i.e. one time step of the low-resolution simulation, using the high-resolution grid. The difference between the high-resolution and coarse-resolution runs shows what should be replaced by the energy conserving restoration scheme (Figure 1 top left). The remaining plots of Figure 1 show 5 examples of  $\delta\zeta$ : i)  $\delta\zeta = \zeta^p - \bar{\zeta}^2$ , ii)  $\delta\zeta = \zeta^p - \bar{\zeta}^4$ , iii)  $\delta\zeta = \bar{\zeta}^2$ , iv)  $\delta\zeta = -\nabla^2 \zeta^p$ , and v)  $\delta\zeta = \zeta^p$ . Each plot in Figure 1 is normalized to allow a comparison of the vorticity patterns. Using  $\delta\zeta = \zeta^p - \bar{\zeta}^2$  and  $\delta\zeta = \zeta^p - \bar{\zeta}^4$  produces the closest match to the actual difference. Section 5 performs tests using the variable patterns above, and error norm analysis indicates that  $\delta\zeta = \zeta^p - \bar{\zeta}^2$  produces the lowest error norms. For the rest of this paper, unless specified otherwise, we use

$$\delta\zeta = \zeta^p - \bar{\zeta}^2. \quad (16)$$

The velocity patterns,  $\delta u$  and  $\delta v$ , are then calculated by solving an elliptic equation numerically for the stream function in terms of  $\delta\zeta$ . The fluid depth  $\delta h$  is calculated from a numerical approximation to geostrophic balance,

$$fv = \frac{g}{a \cos \varphi} \frac{\partial H}{\partial \lambda}, \quad (17)$$

$$fu = -\frac{g}{a} \frac{\partial H}{\partial \varphi}. \quad (18)$$

Setting  $\delta h$  to be in geostrophic balance to  $\delta u$  and  $\delta v$  and then rearranging gives an elliptic equation

$$\nabla^2 \delta h = \frac{1}{a^2 \cos^2 \varphi} \frac{\partial}{\partial \lambda} \left( \frac{a \cos \varphi f \delta v}{g} \right) - \frac{1}{a^2 \cos \varphi} \frac{\partial}{\partial \varphi} \left( \frac{a \cos \varphi f \delta u}{g} \right), \quad (19)$$

which can be solved numerically for  $\delta h$ . Note that the restoration patterns will only be in geostrophic balance provided that the energy conserving coefficient  $\alpha$  is the same for the velocity and depth patterns (see Section 5.1). Finally, we subtract the mean of the variable pattern (so that the restored energy does not change the mean of the preliminary variables). For the depth variable pattern this ensures that any conservation of mass from the preliminary scheme is not affected. The effect of the choice of variable pattern on the potential entrophy is discussed in the Appendix.

Note that the vorticity patterns shown in Figure 1 are not an exhaustive list of all possible variable pattern, but are based on

those tested in Thurnburn *et al.*, (2014). Another set of variable patterns would be obtained by using nonlinear functions of  $u^p$  and  $v^p$  to generate  $\delta u$  and  $\delta v$  (not shown here). This would have the disadvantage that  $(\delta u, \delta v)$  may not be divergence free, but in terms of computational cost this approach would be cheaper as an elliptic solver would not be required to calculate  $\delta u$  and  $\delta v$  from  $\delta \zeta$ .

### 3.2. Calculating the Energy Conserving Coefficient

To ensure energy conservation, we equate the energy at the beginning of the time step with the energy at the end of the time step

$$E^n = E^{n+1}. \quad (20)$$

As energy is calculated using equation (12), it follows that

$$E^n = \int_A \frac{1}{2} h^{n+1} \left( (u^{n+1})^2 + (v^{n+1})^2 \right) + \frac{1}{2} g \left( (H^{n+1})^2 - h_0^2 \right) dA, \quad (21)$$

which, for a given variable pattern, can be rewritten as

$$E^n = \int_A \frac{1}{2} (h^p + \alpha \delta h) \left( (u^p + \alpha \delta u)^2 + (v^p + \alpha \delta v)^2 \right) + \frac{1}{2} g \left( (h^p + \alpha \delta h)^2 + 2h_0(h^p + \alpha \delta h) \right) dA. \quad (22)$$

For most applications the approximation that the energy conserving coefficient  $|\alpha| \ll 1$  is valid and therefore higher powers of  $\alpha$  can be considered negligible. Rearranging equation (22) gives, to a first-order approximation in  $\alpha$ ,

$$E^n = E^p + \alpha \int_A (u^p \delta u + v^p \delta v) h^p + g \delta h (h^p + h_0) \quad (23)$$

$$+ \frac{1}{2} \delta h \left( (u^p)^2 + (v^p)^2 \right) dA \quad (24)$$

which we can rearrange to give

$$\alpha = \frac{E^n - E^p}{\mathcal{I}}, \quad (25)$$

where the denominator is given as

$$\mathcal{I} = \int_A (u^p \delta u + v^p \delta v) h^p + g \delta h (h^p + h_0) + \frac{1}{2} \delta h \left( (u^p)^2 + (v^p)^2 \right) dA. \quad (26)$$

Note that due to the magnitude of the terms in equation (26) and our choice of the variable patterns from Section 3.1 ( $\delta u$ ,  $\delta v$  and  $\delta h$  are positively correlated with  $u^p$ ,  $v^p$  and  $h^p$ ), this integral does not become zero during our tests. Therefore, for any pattern of  $\delta u$ ,  $\delta v$  and  $\delta h$  we can calculate  $\alpha$  to conserve energy.

### 3.3. Available and Unavailable Potential Energy

We now consider how much energy the scheme restores as available and unavailable potential energy. The global integral of the available potential energy is defined as

$$APE = \int_A \frac{g(H - H_m)^2}{2} dA, \quad (27)$$

for the reference height  $H_m = \int_A H dA / \int_A dA$ . The total height field is updated by the energy conserving restoration scheme as

$$H^{n+1} = h^{n+1} + h_0 = h^p + h_0 + \alpha \delta h = H^p + \alpha \delta h, \quad (28)$$

which gives an update equation for the available potential energy as

$$\begin{aligned} APE^{n+1} &= \int_A \frac{g(H^p + \alpha \delta h - H_m)^2}{2} dA, \\ &= \int_A \frac{g(H^p - H_m)^2}{2} dA + \alpha g \int_A \delta h (H^p - H_m) dA, \\ &= APE^p + \alpha g \int_A \delta h (H^p - H_m) dA. \end{aligned} \quad (29)$$

Again, this makes use of the assumption that  $|\alpha| \ll 1$  and therefore higher powers of  $\alpha$  are neglected. Similarly, the unavailable potential energy becomes

$$UPE^{n+1} = UPE^p + \alpha g H_m \int_A \delta h dA. \quad (30)$$



Note that  $gH_m \int_A \delta h dA = 0$ , as  $\int_A \delta h dA = 0$  by design. Therefore  $UPE^{n+1} = UPE^p$  and the total unavailable potential energy is due to the preliminary scheme. The unavailable potential energy is conserved by the continuous equations, and will be conserved in the model provided that the preliminary scheme is mass conserving. For the available potential energy, equation (29) simplifies to

$$APE^{n+1} = APE^p + \alpha g \int_A \delta h H^p dA. \quad (31)$$

This shows that the restored potential energy goes into the available potential energy.

### 3.4. ECRS: The Energy Conserving Restoration Scheme

For clarity, this section describes the default specification used for the energy conserving restoration scheme (ECRS) in this paper.

First a preliminary scheme is used to calculate  $u^p$ ,  $v^p$  and  $h^p$ . The vorticity pattern is then calculated as

$$\delta\zeta = \zeta^p - \overline{\zeta^p}, \quad (32)$$

and is inverted to calculate the stream function pattern,  $\delta\psi$ . From here, the velocity patterns are calculated as

$$\delta u = -\frac{1}{a} \frac{\partial \delta\psi}{\partial \varphi}, \quad \delta v = \frac{1}{a \cos \varphi} \frac{\partial \delta\psi}{\partial \lambda}. \quad (33)$$

The depth pattern is then calculated by solving an elliptic equation (19) that approximates geostrophic balance. The spatial means of the variable patterns are then subtracted, to ensure that the variable patterns have a mean of zero.

The variables at the new time step are updated as in equation (15). A single  $\alpha$  is used for both the velocity patterns and the depth pattern, and is calculated using equation (25).

## 4. Numerical Testing

The preliminary scheme used for all tests is the mass conserving semi-implicit semi-Lagrangian scheme of Thuburn *et al.*, (2010), which is based upon that of Zerroukat *et al.*, (2009b). This scheme is denoted SISL. SISL uses the Semi-Lagrangian Inherently Conserving and Efficient (SLICE) scheme for the continuity

equation (Zerroukat *et al.*, 2002, 2009a) coupled to a semi-implicit semi-Lagrangian discretization of the momentum equations. SISL solves the governing equations on a latitude-longitude grid, and makes use of the Arakawa C-grid (Arakawa, 1977). The geopotential,  $\phi = gh$ , is used as the mass variable instead of the depth,  $h$ . For all tests we use a regular size earth with radius  $a = 6371220$  m, regular rotation rate  $2\Omega = 1.4584 \times 10^{-4} \text{ s}^{-1}$ , and gravity  $g = 9.80616 \text{ ms}^{-2}$ .

We run the preliminary scheme with the energy conserving restoration scheme and compare the results with the preliminary scheme on its own. In this testing we use the energy conserving restoration scheme as described in section 3.4. This scheme is denoted ECRS.

To use the energy conserving restoration scheme on the staggered C grid, we first calculate all the non-linear terms in the energy equation at their native points. We then average all the values to cell centres (height points). We then calculate energy at the cell centres, and use this to calculate the value of  $\alpha$  for depth and the staggered velocities.

### 4.1. Williamson Test Cases

The use of the Williamson test cases (Williamson *et al.*, 1992) is to make sure that the restoration scheme does not introduce any errors into simple flows. The first test is the steady state geostrophic flow test (TC2) of Williamson *et al.*, (1992). We consider the case of zonal flow, where the velocities and depth field are initialized as

$$u = u_0 \cos \varphi, \quad (34)$$

$$v = 0, \quad (35)$$

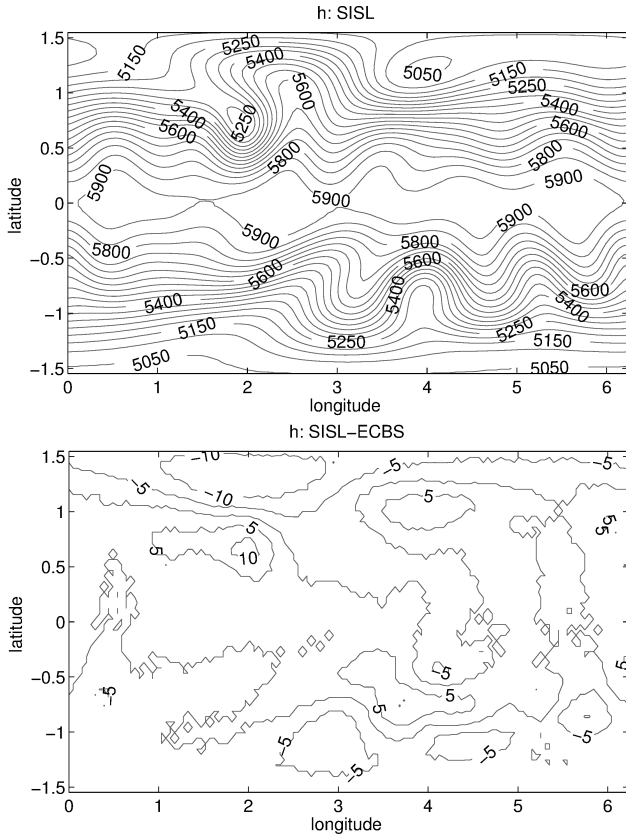
$$gh = gh_C - \left( a\Omega u_0 + \frac{u_0}{2} \right) \sin^2 \varphi, \quad (36)$$

where  $gh_C = 2.94 \times 10^4 \text{ m}^2 \text{ s}^{-2}$  and  $u_0 = 2\pi a / \tau$  where  $\tau = 1036800$  s. We calculate the normalized  $\ell_2$  error norm,

$$\ell_2(h) = \sqrt{\frac{\int (h - h_T)^2 dA}{\int h_T^2 dA}}, \quad (37)$$

after 5 days, where  $h_T$  is the true solution (the initial condition). The error norms for SISL and ECRS are virtually identical for

both the  $64 \times 32$  and  $128 \times 64$  resolution latitude-longitude grids, indicating that ECRS does not introduce any additional errors into the solution. After 15 days the SISL scheme loses  $5.82 \times 10^{-5}\%$  of the energy compared to a change of  $2.52 \times 10^{-11}\%$  for ECRS on the  $128 \times 64$  grid. The energy change for ECRS is not exactly zero due to the approximation that higher powers of  $\alpha$  are negligible. The application of ECRS does not affect the conservation of mass.



**Figure 2.** The total height field,  $H$  (m), for the Mountain Test (TC5) at day 15 for the SISL scheme (top), and the difference plot,  $SISL - ECRS$ , (bottom) on the  $128 \times 64$  resolution latitude-longitude grid.

The second test is the zonal flow over an isolated mountain (TC5) test of Williamson *et al.*, (1992). The initial velocity and height fields are as in the steady state geostrophic flow (TC2), except  $h_C = 5960$  m and  $u_0 = 20 \text{ ms}^{-1}$ . The mountain is given as

$$h_0 = h_1 \left( 1 - \frac{r}{\pi/9} \right), \quad (38)$$

where  $r^2 = \min [(\pi/9)^2, (\lambda - 3\pi/2)^2 + (\varphi - \pi/6)^2]$  and  $h_1 = 2000$  m.

The testing shows that ECRS conserves mass to the same degree of accuracy as SISL. The height field for SISL and the difference plot ( $SISL - ECRS$ ) of the height fields are shown

at day 15 on the  $128 \times 64$  grid in Figure 2, demonstrating that SISL and ECRS produce very similar solutions. After 30 days on the  $64 \times 32$  grid, SISL loses 0.13% of the total energy, whereas the percent change for ECRS is of the order  $10^{-5}$ . After 30 days both schemes have dissipated potential enstrophy: SISL dissipates 1.57% of the potential enstrophy compared to 0.63% for ECRS.

#### 4.2. Unstable Barotropic Wave

The unstable barotropic wave test (BW) is described by Galewsky *et al.*, (2004). A perturbation is added to a balanced, barotropically unstable, mid-latitude jet to initiate an instability. We calculate a reference solution using the SISL scheme on a  $1024 \times 512$  latitude-longitude grid. This reference solution is used to calculate the normalized  $h$  and  $q$  error norms after 6 and 10 days for the schemes on the  $128 \times 64$  latitude-longitude grids.

Figure 3 shows the potential vorticity at day 6 for the barotropic wave test. The left plot shows the reference solution whereas the right plot shows the SISL scheme on the  $128 \times 64$  resolution grid. The reference solution is able to capture the small scale potential vorticity filaments as they stretch out and wrap up. The coarse resolution simulation is unable to represent these small scale features. The normalized  $\ell_2$  and  $\ell_{\text{inf}}$  error norms for  $h$  and  $q$  at days 6 and 10 are presented in Table 1. The normalized  $\ell_{\text{inf}}$  norm is calculated as

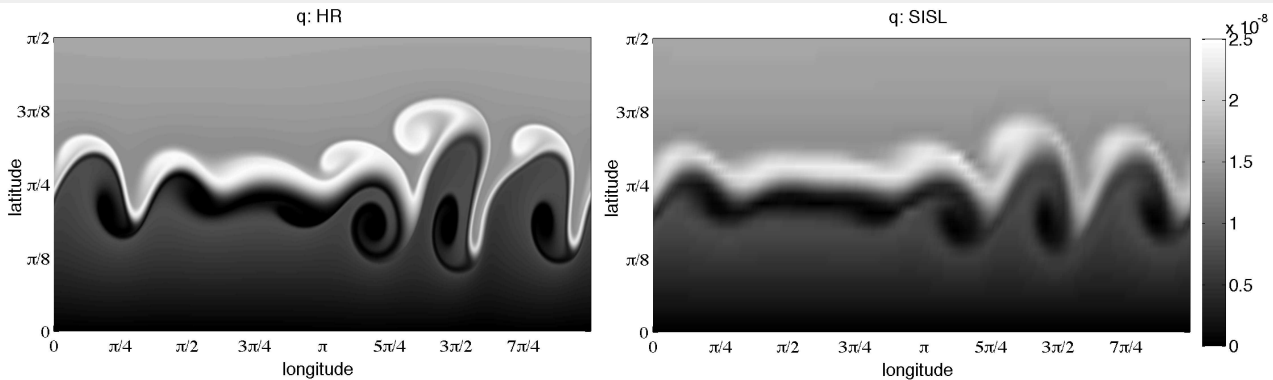
$$\ell_{\text{inf}}(h) = \frac{\max |h - h_T|}{\max |h_T|}, \quad (39)$$

where  $h_T$  is taken from the reference solution. The error norms for  $h$  and  $q$  are similar for SISL and ECRS. **This is because the error norms are dominated by the truncation error of SISL. However, these results do show that using ECRS does not introduce significant errors to the flow.** The percent energy change after 10 days is  $-0.059$  for SISL and  $2.12 \times 10^{-5}$  for ECRS, whereas the percent potential enstrophy change after 10 days is  $-5.21$  for SISL and  $-3.82$  for ECRS. This shows that while both schemes are dissipating potential enstrophy, with SISL dissipating more than ECRS, only SISL is dissipating significant energy.

#### 4.3. Field of Vortices

The final test is the freely decaying field of vortices based upon the vorticity described on the plane by Kent *et al.*, (2012). The

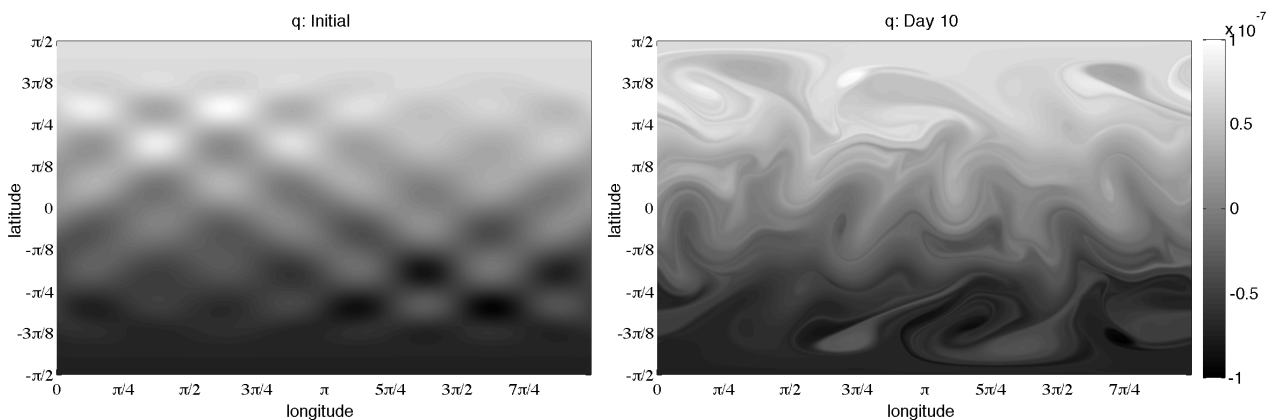




**Figure 3.** The potential vorticity ( $\text{m}^{-1}\text{s}^{-1}$ ) for the unstable barotropic wave test at day 6 for the high-resolution (HR) reference solution (left) and the SISL scheme on the  $128 \times 64$  resolution grid (right).

Table 1. The normalized  $\ell_2$  and  $\ell_{\text{inf}}$  error norms for  $h$  and  $q$  at days 6 and 10 for the unstable barotropic wave test case on the  $128 \times 64$  resolution latitude-longitude grid.

	$\ell_2(h)$	$\ell_{\text{inf}}(h)$	$\ell_2(q)$	$\ell_{\text{inf}}(q)$
SISL (day 6)	0.0066	0.0452	0.2501	0.8310
ECRS (day6)	0.0063	0.0447	0.2560	0.8587
SISL (day 10)	0.0101	0.0693	0.2591	0.6407
ECRS (day10)	0.0102	0.0799	0.2705	0.7344



**Figure 4.** The potential vorticity ( $\text{m}^{-1}\text{s}^{-1}$ ) for the field of vortices test at day 0 (left) and at day 10 (right) using the SISL scheme on a  $1024 \times 512$  resolution grid.

test has analytical initial conditions for the velocities, vorticity and

divergence. The velocities are given as

$$u = \frac{3S}{4} \left[ 32 \sin(4\lambda) \cos(8\varphi) - 19.2 \cos(3\lambda) \sin(6\varphi) - 6 \cos(5\lambda) \sin(10\varphi) + \cos(2\varphi) \right], \quad (40)$$

$$v = -\frac{3S}{4} \left[ 16 \cos(4\lambda) \sin(8\varphi) - 9.6 \sin(3\lambda) \cos(6\varphi) - 3 \sin(5\lambda) \cos(10\varphi) + 0.5 \cos(\lambda) \right], \quad (41)$$

$$\zeta = \frac{3S}{4a \cos \varphi} \left\{ \left( \sin(\varphi) + 4 \cos(\varphi) \varphi^7 \right) \left[ 32 \sin(4\lambda) \cos(8\varphi) - 19.2 \cos(3\lambda) \sin(6\varphi) - 6 \cos(5\lambda) \sin(10\varphi) + \cos(2\varphi) \right] + \left( 1 + 4 \cos(\varphi) \right) \left[ 64 \sin(4\lambda) \sin(8\varphi) + 28.2 \cos(3\lambda) \cos(6\varphi) + 15 \cos(5\lambda) \cos(10\varphi) \right] + 0.5 \sin(\lambda) + 2 \cos(\varphi) \sin(2\varphi) \right\}, \quad (42)$$

resulting in the vorticity as

and divergence as

$$D = \frac{3S}{4a \cos \varphi} \left\{ \left( \sin(\varphi) + 4 \cos(\varphi) \varphi^7 \right) \right. \quad (43)$$

$$\left[ 16 \cos(4\lambda) \sin(8\varphi) - 9.6 \sin(3\lambda) \cos(6\varphi) \right.$$

$$\left. - 3 \sin(5\lambda) \cos(10\varphi) + 0.5 \cos(\lambda) \right]$$

$$+ \left( 1 - \cos(\varphi) \right) \left[ 128 \cos(4\lambda) \cos(8\varphi) + 57.6 \sin(3\lambda) \sin(6\varphi) \right.$$

$$\left. + 30 \sin(5\lambda) \sin(10\varphi) \right\},$$

where the shape function,  $S = S_0 \exp(-\frac{1}{2}\varphi^8)$  with  $S_0 = 1 \text{ ms}^{-1}$ , is to remove the dependence on  $\lambda$  at the poles. The depth field is calculated from a numerical approximation to geostrophic balance, i.e (17) and (18). As  $h_0 = 0 \Rightarrow H = h$ , the equation for geostrophic balance can be rewritten as

$$\nabla^2 h = \frac{1}{a^2 \cos^2 \varphi} \frac{\partial}{\partial \lambda} \left( \frac{a \cos \varphi f v}{g} \right) - \frac{1}{a^2 \cos \varphi} \frac{\partial}{\partial \varphi} \left( \frac{a \cos \varphi f u}{g} \right), \quad (44)$$

where the velocity derivatives can be calculated exactly or, as for the results in this paper, approximated using centred differences. This equation can then be solved numerically using an elliptic solver. We then add a constant to the depth field such that the mean of  $h$  is 2000 m. Again, we use the SISL scheme on a  $1024 \times 512$  grid to calculate the reference solution. This reference solution is used to calculate error norms up to day 10 of the simulation. Figure 4 shows the initial and day 10 potential vorticity,  $q$ , calculated using the reference solution. After day 10 the test becomes unpredictable due to the turbulent nature of the flow, and therefore we focus on the energy and potential enstrophy statistics after this point.

Figure 5 shows the normalized  $\ell_2$  and  $\ell_{\text{inf}}$  error norms for  $h$  and  $q$  up to day 10 for SISL and ECRS on the  $128 \times 64$  grid for the freely decaying vortices test. As with the previous tests the error norms are very similar for both schemes, showing that ECRS does not affect the accuracy of the preliminary scheme (in this case SISL). Another metric that is examined is the global maximum and minimum relative vorticity values for the schemes. The difference between the maximum and minimum vorticity values is initially  $1.2 \times 10^{-4} \text{ s}^{-1}$ . For SISL this becomes  $1.5 \times$

$10^{-4} \text{ s}^{-1}$  at day 10 and  $3.7 \times 10^{-5} \text{ s}^{-1}$  at day 365, and for ECRS this is  $1.6 \times 10^{-4} \text{ s}^{-1}$  at day 10 and  $5.3 \times 10^{-5} \text{ s}^{-1}$  at day 365. As the maxima and minima are larger for ECRS, these values suggest that ECRS reduces the diffusion from SISL.

The total energy, kinetic energy and potential enstrophy are plotted as time series (up to 300 days) for SISL and ECRS on the  $128 \times 64$  grid, in Figure 6. As with the previous tests total energy is dissipated by SISL, with a loss of 0.43% after 365 days, and effectively conserved by ECRS, with a change of  $1.39 \times 10^{-5}\%$ . As stated above, the change in energy by ECRS is not exactly zero due to the first order approximation of the  $\alpha$  values. Repeating these tests on the very coarse  $64 \times 32$  grid gives total energy changes after 365 days of  $-0.95\%$  for SISL and  $1.89 \times 10^{-4}\%$  for ECRS, demonstrating that the accuracy of the energy conservation in ECRS improves as number of grid points increases. For the kinetic energy and potential enstrophy statistics in Figure 6 we present results from the reference solution averaged onto the coarser grid. Although the high resolution reference solution is not used to compute error norms for long simulations (over 10 days) due to the turbulent nature of the flow, the kinetic energy and potential enstrophy statistics are used for guidance. The reference kinetic energy fluctuates because of the turbulent nature of the flow, and gradually decreases because there is a systematic conversion of kinetic energy to potential energy. Throughout the run the SISL kinetic energy is considerably smaller than the reference value and decreases faster, indicating excessive dissipation. The ECRS kinetic energy time series, on the other hand, is similar to the reference time series, indicating that the ECRS is able to maintain the energy of the turbulent eddies against numerical dissipation. Potential enstrophy is dissipated by both schemes, although the amount of potential enstrophy dissipated by ECRS is less than that dissipated by SISL and is similar to that of the reference solution.

Figure 7 shows the relative vorticity,  $\zeta$ , and the depth,  $h$ , for SISL and ECRS at day 365. The small vortices of the initial state have merged over time to leave two large vortex pairs. For both relative vorticity and depth, ECRS has the largest maximum values and the steepest gradients. However, the plots show that in general the structure of the flow after 365 days for SISL is similar to ECRS.

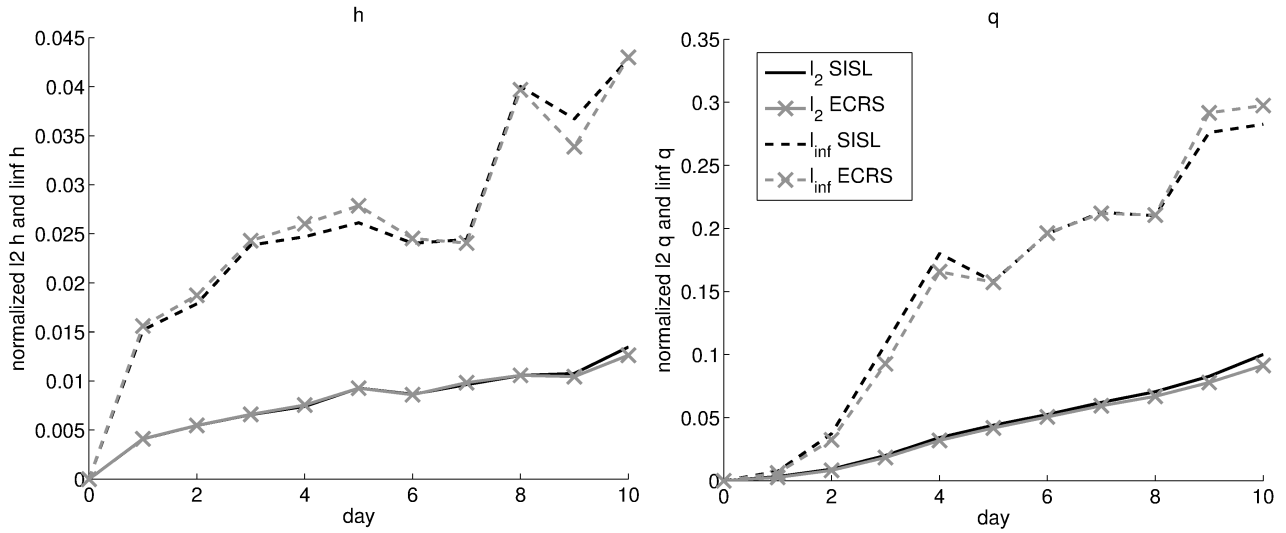


Figure 5. The normalized  $\ell_2$  and  $\ell_{\text{inf}}$  error norms for  $h$  and  $q$  plotted against time on the  $128 \times 64$  grid for the freely decaying vortices test.

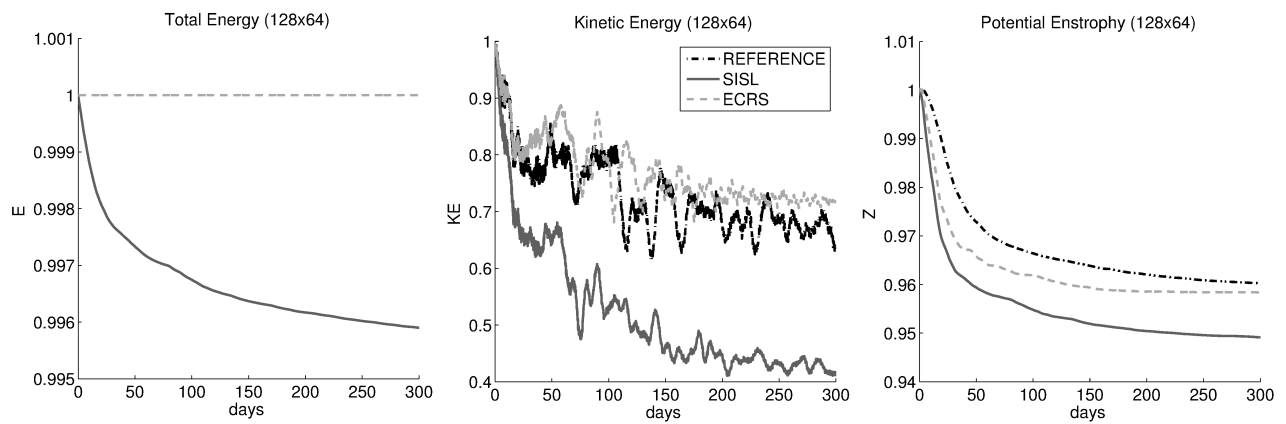


Figure 6. The normalized total energy (left), kinetic energy (centre) and potential enstrophy (right) against time on the  $128 \times 64$  grid for the freely decaying vortices test. The reference solution is calculated on the  $1024 \times 512$  grid.

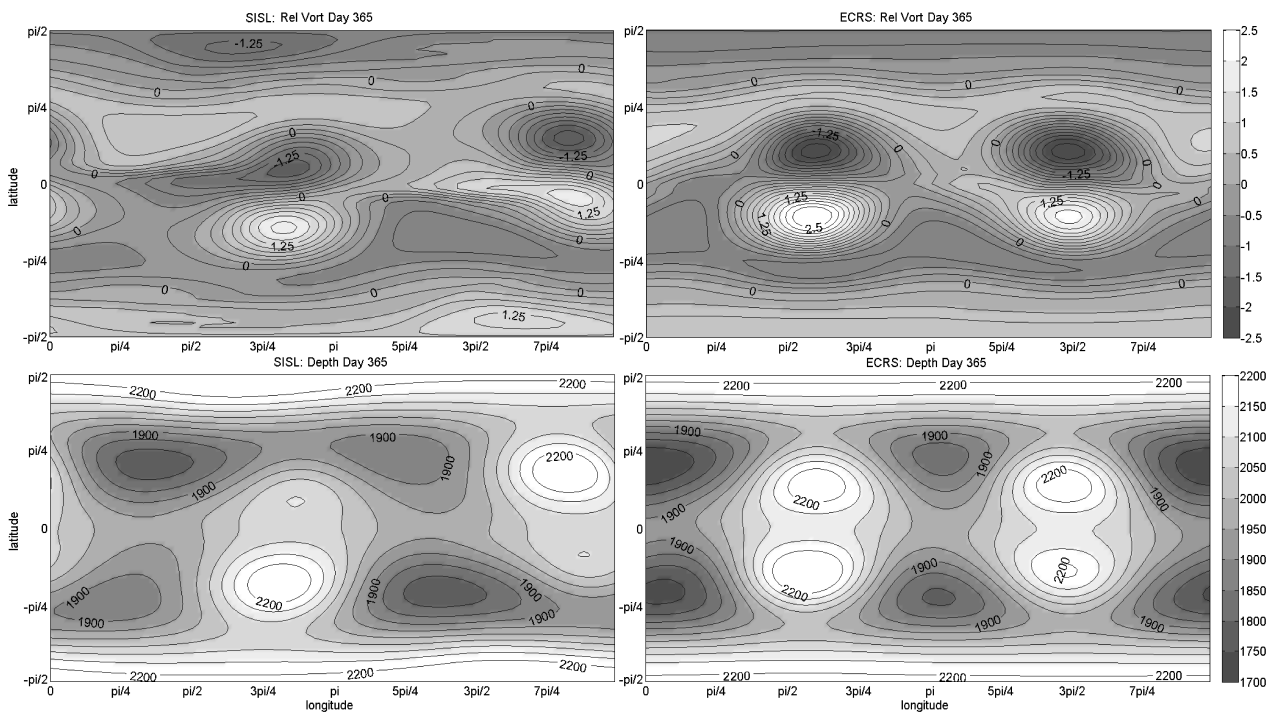


Figure 7. Day 365 plots of relative vorticity (top), expressed in units of  $10^{-5} \text{ s}^{-1}$ , and depth (bottom), m, for SISL (left) and ECRS (right).

## 5. Sensitivity to Parameters

In this section we discuss the sensitivity of the energy conserving restoration scheme to the choice of variable pattern,  $\delta F$ , and to changes in  $\alpha$  that allow a choice in the partition of the replaced energy into potential and kinetic forms.

### 5.1. Potential and Kinetic Energy Partition

The aim of this section is to investigate the effect of the different choices of potential and kinetic energy partition. For some situations it may be preferable to specify how much energy is replaced in the form of kinetic or potential energy (for example, if it is known how much kinetic energy dissipation is physical then this can be used to judge the restoration of kinetic energy).

This is also useful when considering how the energy conserving restoration scheme will restore energy via the velocities and the temperature field when applied to atmospheric dynamical cores.

If, for example, we only require energy to be added back in the form of additions to the depth field we can set  $\delta u = \delta v = 0$ , giving the following equation for  $\alpha$

$$\alpha = \frac{E^n - E^p}{\int_A g \delta h (h^p + h_0) + \frac{1}{2} \delta h ((u^p)^2 + (v^p)^2) dA}. \quad (45)$$

Similarly, if we only require energy to be added back as additions to the velocities we can set  $\delta h = 0$  to give

$$\alpha = \frac{E^n - E^p}{\int_A (u^p \delta u + v^p \delta v) h^p dA}. \quad (46)$$

Alternatively, we can specify a fraction of the energy dissipated by the preliminary scheme to add back in the form of additions to the depth and velocities respectively. Note, however, that the kinetic energy contribution to the total energy, the first term in (12), is dependent on the depth  $h$ . This means that modifying the depth field will affect the kinetic energy as well as the potential energy.

We denote  $\Delta E = E^n - E^p$  as the energy lost by the preliminary scheme over one time step. We then specify the partition fraction  $C_{PE}$  to be the proportion of energy lost to be added back strictly to the potential energy term. The partition fraction  $C_{KE} = 1 - C_{PE}$  is the proportion of energy lost to be

replaced in the depth multiplied by the kinetic energy term. The energy conserving coefficient is now split into velocity and height parts such that

$$\begin{aligned} u^{n+1} &= u^p + \alpha_v \delta u, \\ v^{n+1} &= v^p + \alpha_v \delta v, \\ h^{n+1} &= h^p + \alpha_h \delta h. \end{aligned} \quad (47)$$

For a given  $C_{PE}$  and  $C_{KE}$ , the energy conserving coefficients  $\alpha_h$  and  $\alpha_v$  are calculated as

$$\alpha_h = \frac{C_{PE} \Delta E}{\int_A g \delta h (h^p + h_0) dA}, \quad (48)$$

$$\alpha_v = \frac{C_{KE} \Delta E - \frac{1}{2} \alpha_h \int_A \delta h ((u^p)^2 + (v^p)^2) dA}{\int_A (u^p \delta u + v^p \delta v) h^p dA}. \quad (49)$$

The partition fractions  $C_{PE}$  and  $C_{KE}$  can be either specified and fixed for the simulation, or they can be variable and based upon the amount of potential energy that is dissipated by the preliminary scheme. We consider  $C_{PE} = 0.2$  and  $C_{KE} = 0.8$ , as empirical testing using the freely decaying field of vortices test shows that approximately 20% of the energy lost by SISL is in the form of potential energy, and  $C_{PE} = 0.5$  and  $C_{KE} = 0.5$ , to show an equal partition. For the variable fractions we proceed as follows. At each time step the energy lost due to the preliminary scheme,  $\Delta E = E^n - E^p$ , and the potential energy change due to the preliminary scheme,  $\Delta PE = PE^n - PE^p$ , are calculated. The partition fractions  $C_{PE}$  and  $C_{KE}$  are then calculated as

$$C_{PE} = \min \left[ \max \left( \frac{\Delta PE}{\Delta E}, 0 \right), 1 \right], \quad (50)$$

$$C_{KE} = 1 - C_{PE}. \quad (51)$$

The change in potential energy is not necessarily a loss of potential energy, therefore the max and min functions are used to ensure that  $C_{PE} \in [0, 1]$ .

## 5.2. Parameter Testing

In each case  $\delta u$  and  $\delta v$  are calculated from the relative vorticity pattern, and  $\delta h$  is calculated to be in geostrophic balance with  $\delta u$  and  $\delta v$  using (19). We consider the following relative vorticity patterns, most of which were shown in Figure 1:

$$\begin{aligned} \delta\zeta &= \zeta^p - \overline{\zeta^{p^2}}, & \delta\zeta &= \zeta^p - \overline{\zeta^{p^4}}, & \delta\zeta &= \overline{\zeta^{p^2}}, \\ \delta\zeta &= \overline{\zeta^{p^4}}, & \delta\zeta &= -\nabla^2 \zeta^p, & \delta\zeta &= \zeta^p. \end{aligned}$$

We consider six cases of energy replacement partition fraction. The first set uses  $\alpha = \alpha_v = \alpha_h$ . Note that using this with  $\delta\zeta = \zeta^p - \overline{\zeta^{p^2}}$  is the version of ECRS used in the previous sections. The second set uses the variable partition fractions  $C_{PE}$  and  $C_{KE}$  from equations (50) and (51). The next case sets  $C_{PE} = 0.2$  and  $C_{KE} = 0.8$ , and the next uses  $C_{PE} = C_{KE} = 0.5$ . The final two cases consider only adding energy back to the depth field, and only adding energy back to the velocity. We look at the effect of these parameters on both short and long term simulations. We consider the field of vortices test from Section 4.3 and discuss error norms in the short term (after 10 days) and energy and potential enstrophy statistics in the long term (365 days).

As with the previous testing, the error norms are dominated by the truncation error of the preliminary scheme, SISL. For the  $128 \times 64$  resolution latitude-longitude grid the error norms are very similar for each case, showing that the choice of parameters in ECRS has little impact on the potential vorticity errors. However, for each choice of energy partition apart from only adding to the depth field, using  $\delta\zeta = \overline{\zeta^{p^2}}$ ,  $\delta\zeta = \overline{\zeta^{p^4}}$  or  $\delta\zeta = \zeta^p$  results in the largest error norms, indicating that these variable patterns introduce more error than the others. The pattern  $\delta\zeta = \zeta^p - \overline{\zeta^{p^2}}$  consistently has the lowest error norms, supporting the choice of using  $\delta\zeta = \zeta^p - \overline{\zeta^{p^2}}$  as the default in our paper. The energy partition also has an effect on the error norm. The results suggest that adding a smaller amount of energy to the potential energy, e.g.  $C_{PE} = 0.2$  or  $C_{PE} = 0$ , improves the accuracy of scheme. Using the variable partition fractions of  $C_{PE}$  and  $C_{KE}$  from equations (50) and (51) and the variable pattern  $\delta\zeta = \zeta^p -$

$\overline{\zeta^{p^2}}$ , the mean values for this test are  $C_{PE} = 0.46$  and  $C_{KE} = 0.54$ .

For each choice of  $\delta\zeta$  and energy partition fraction, total energy is conserved to the same order-of-accuracy. However, the potential enstrophy statistics are significantly affected by the parameter choices. The potential enstrophy difference from the initial conditions after 365 days for the energy conserving restoration scheme with different  $\delta\zeta$  patterns and energy partitions on the  $64 \times 32$  resolution latitude-longitude grid are given in Table 2. The choice of  $\delta\zeta$  pattern does have an effect on the potential enstrophy. Each choice of energy partition using  $\delta\zeta = \overline{\zeta^{p^r}}$  or  $\delta\zeta = \zeta^p$  results in a loss of potential enstrophy after 365 days, yet some choices of energy partition using  $\delta\zeta = \zeta^p - \overline{\zeta^{p^r}}$  or  $\delta\zeta = -\nabla^2 \zeta$  results in an increase in potential enstrophy. The possible increase in potential enstrophy when using ECRS with certain parameters is discussed in detail in the Appendix. The Appendix shows that both the variable pattern and the energy partition fraction affect the possible increase in potential enstrophy. If energy is restored at smaller scales than those from which it is dissipated, then this can lead to an increase in potential enstrophy. This explains why in the numerical testing for some choices of  $C_{PE}$  and  $C_{KE}$  the variable patterns of the form  $\delta F = \overline{F^{p^r}}$ , which restore energy at the larger scales, do not increase potential enstrophy with time, yet the patterns of the form  $\delta F = F^p - \overline{F^{p^r}}$ , which approximates the smaller scales, may lead to an increase in potential enstrophy. Choosing an energy partition fraction that restores more energy in the form of potential energy, e.g. using variable partition fractions or only adding back to  $h$  or using  $C_{PE} = C_{KE} = 0.5$  can also result in an increase in potential enstrophy for the above mentioned schemes.

## 6. Conclusions and Discussion

We have presented an energy conserving restoration scheme that can be applied to shallow water models on the sphere. The total energy is conserved to a good approximation by restoring the energy that is dissipated by a preliminary scheme, and this is achieved by adding a variable pattern to a combination of the velocities (kinetic energy) and the depth field (potential energy). The variable pattern that is added back is a function of the preliminary solution.



Table 2. The percent potential enstrophy difference from the initial conditions after 365 days for the field of vortices test case on the on the  $64 \times 32$  resolution latitude-longitude grid. Negative values show a decrease in potential enstrophy after 365 days. As a comparison, the SISL scheme produces a  $-6.05\%$  change after 365 days.

	$\zeta^p - \overline{\zeta^p}^2$	$\zeta^p - \overline{\zeta^p}^4$	$\overline{\zeta^p}^2$	$\overline{\zeta^p}^4$	$-\nabla^2 \zeta^p$	$\zeta^p$
$\alpha_v = \alpha_h = \alpha$	-3.33	-3.20	-5.71	-5.43	-2.38	-5.64
$C_{PE}, C_{KE}$ variable	2.12	1.93	-3.86	-3.28	1.86	-3.90
$C_{PE} = 0.2, C_{KE} = 0.8$	-3.32	-3.10	-5.63	-5.64	-2.49	-5.62
$C_{PE} = 0.5, C_{KE} = 0.5$	1.00	0.78	-4.60	-4.16	1.68	-4.93
Only $h$ (i.e. $\alpha_v = 0$ )	0.44	1.15	-4.34	-3.08	4.03	-2.88
Only $u, v$ (i.e. $\alpha_h = 0$ )	-5.51	-5.64	-4.79	-4.47	-5.77	-4.88

Numerical testing is performed using the mass conserving semi-implicit semi-Lagrangian scheme (SISL) as the preliminary scheme (Thuburn *et al.*, 2010). For the tests in this paper the error norms for SISL and the energy conserving restoration scheme (ECRS) are very similar. **As the error norms are dominated by the truncation error of the preliminary scheme**, this indicates that using ECRS does not introduce any significant errors into the solution, and in some cases may improve the accuracy of the solution. Using ECRS produces better energy statistics than SISL, and for some tests better potential enstrophy statistics. For longer simulations ECRS conserves energy to a good approximation, and the accuracy of this approximation increases as the grid resolution increases. Energy conservation is not exact due to the linear approximation of the energy conserving coefficient. In general, both SISL and ECRS dissipate potential enstrophy, although the dissipation rate is smaller for ECRS.

For the testing presented here we use a variable pattern based on an approximation of the small scale vortical flow. Comparing with other variable patterns, for example those based on an approximation of the large scales or a diffusion term, shows that the approximation of the small scales is a suitable variable pattern. We have also investigated the use of energy partition fractions, which specify how much of the replaced energy is restored as either kinetic or potential energy. The energy partition fractions can be held constant or they can be variable. Note, however, that for some parameter choices potential enstrophy may increase for ECRS. Results from further testing (not presented here) show that the increases in potential enstrophy for ECRS decrease as the grid resolution is increased. The possible increase in potential enstrophy is discussed in the Appendix, and is dependent on the variable pattern that is used to restore the energy and the partition fractions that partition the restored energy into potential and

kinetic energy parts. Future research will investigate the optimum choice of the energy partition fractions and variable patterns.

The energy conserving restoration scheme for the shallow water equations is designed as a first step towards creating a similar model for atmospheric dynamical cores. **An important consideration for atmospheric models is the computational cost of the energy conserving method. The shallow water testing in this paper was performed on a serial machine, and the runtime of ECRS was approximately 10% more expensive than SISL. For an atmospheric model on parallel processors solving an elliptic equation is undesirable, and so a different choice of  $\delta u$ ,  $\delta v$  and  $\delta h$  may be advantageous.** Additionally there are other significant considerations that would need to be taken into account to make the model suitable for atmospheric dynamical cores. For example, in the atmosphere the potential energy can be split into gravitational and internal energy, and a variable pattern can be added to the temperature field to restore total energy. These could be incorporated into the restoration scheme by changing the equations for  $\alpha$  to include these terms. The amount of energy to restore as kinetic energy or through additions to temperature in the restoration scheme could be tuned depending on the flow. **This may be desirable because the kinetic energy that is lost due to the numerical scheme should be restored in its original form. If it is instead restored as thermodynamic energy then energy conservation is obtained but at the price of spuriously enhancing or reducing an energy conversion term.**

#### Acknowledgements

We would like to thank the two reviewers for their detailed comments that helped improve the manuscript.

The work was supported by the Office of Science, U.S. Department of Energy, Award No. DE-SC0006684.



## Appendix

In this appendix we discuss why the potential enstrophy may increase for certain parameter choices when using the energy conserving restoration scheme (ECRS). We use both grid point and Fourier analysis.

First we consider the dependence on the partition fractions. The potential enstrophy added by the energy conserving restoration scheme is

$$\alpha_Z \delta Z = \alpha_q h^p q^p \delta q + \frac{1}{2} \alpha_h \delta h (q^p)^2. \quad (52)$$

Potential vorticity is defined as  $q = (\zeta + f)/h$ , giving

$$q^p + \alpha_q \delta q = \frac{\zeta^p + \alpha_v \delta \zeta + f}{h^p + \alpha_h \delta h}, \quad (53)$$

which can be approximated as

$$\alpha_q \delta q = \frac{\alpha_v \delta \zeta - \alpha_h \delta h q^p}{h^p}. \quad (54)$$

This results in the change in potential enstrophy due to ECRS as

$$\alpha_Z \int \delta Z dA = \alpha_v \int \delta \zeta q^p dA - \frac{1}{2} \alpha_h \int \delta h (q^p)^2 dA. \quad (55)$$

Consider the second integral in this equation. The  $(q^p)^2$  term will have a large equator to pole gradient due to the rotating sphere, and if a similar pattern (with opposite sign) appears in  $\delta h$  then there will be a large contribution to the potential enstrophy from this integral. This may lead to an increase in potential enstrophy. For the case of  $\alpha_v \neq \alpha_h$ , the magnitude of the integral is dependent on  $\alpha_h$ , which in turn is dependent on the choice of  $C_{PE}$ . Therefore larger values of  $C_{PE}$  can result in the increase in potential enstrophy, as shown in Table 2. A choice of  $\delta h$  that is not negatively correlated with  $(q^p)^2$ , or subtracting the zonal mean from  $\delta h$  to remove the equator to pole gradient in  $\delta h$ , will prevent the increase in potential enstrophy due to this term

Now consider the dependence on the shape of the variable pattern. We wish to express the changes in energy and potential enstrophy in Fourier space, so we use a doubly periodic  $f$ -plane with no orography. Also, we need to express the energy and potential enstrophy as quadratic quantities, so we make the

assumption that  $Fr^2 \ll Ro$ , where  $Fr$  is the Froude number and  $Ro$  is the Rossby number (defined as  $Ro = V/(Lf)$  and  $Fr = V/c$ , where  $V$  is the characteristic velocity,  $L$  is a characteristic length, and  $c = \sqrt{gh}$  is the characteristic wave propagation velocity). Under this assumption equations (12) and (13) may be approximated by

$$\begin{aligned} E &= const + \frac{1}{2} \int_A \bar{h} (u^2 + v^2) + g (h')^2 dA \\ Z &= const + \frac{1}{2} \int_A \frac{1}{\bar{h}} \left( \zeta - f \frac{h'}{\bar{h}} \right)^2 dA \end{aligned}$$

where  $\bar{h}$  is the global mean of  $h$ ,  $h' = h - \bar{h}$ , and  $const$  indicates constant terms.

We let  $\hat{F}$  indicate the Fourier transform of a variable  $F$ . Then, by Parseval's theorem,

$$\begin{aligned} E &= const + \frac{C}{2} \int \bar{h} (|\hat{u}|^2 + |\hat{v}|^2) + g |\hat{h}'|^2 d\mathbf{k} \\ Z &= const + \frac{C}{2} \int \frac{1}{\bar{h}} \left| \hat{\zeta} - f \frac{\hat{h}'}{\bar{h}} \right|^2 d\mathbf{k} \end{aligned}$$

where  $C$  is a normalization constant,  $\mathbf{k} = (k, l)$  is the wave number vector, and  $k$  and  $l$  are the wave numbers in  $x$  and  $y$ . If we further assume that the flow is geostrophically balanced then

$$\begin{aligned} \hat{v} &= ik\hat{\psi}, & \hat{u} &= -il\hat{\psi}, \\ \hat{\zeta} &= -K^2\hat{\psi}, & \hat{h}' &= \frac{f}{g}\hat{\psi}, \end{aligned} \quad (56)$$

where  $\psi$  is the stream function,  $i$  is the imaginary unit, and  $K = \sqrt{k^2 + l^2}$  is the total wave number. The expressions for energy and potential enstrophy become

$$\begin{aligned} E &= const + \frac{C}{2} \bar{h} \int \left( K^2 + \frac{1}{\bar{a}^2} \right) |\hat{\psi}|^2 d\mathbf{k} \\ Z &= const + \frac{C}{2} \frac{1}{\bar{h}} \int \left( K^2 + \frac{1}{\bar{a}^2} \right) |\hat{\psi}|^2 d\mathbf{k} \end{aligned}$$

where  $\bar{a} = \sqrt{gh}/f = c/f$  is the Rossby radius.

Now let  $\Delta$  indicate the change in any variable due to dissipation plus the effects of the ECRS. If these changes are small then, using the fact that  $\psi$  is real so that  $\hat{\psi}(-\mathbf{k}) = \hat{\psi}(\mathbf{k})^*$ , where  $*$  indicates complex conjugate, the resulting changes in energy and potential

enstrophy are

$$\Delta E = C\bar{h} \int \left( K^2 + \frac{1}{a^2} \right) \hat{\psi} \Delta \hat{\psi}^* d\mathbf{k} \quad (57)$$

$$\Delta Z = C \frac{1}{h} \int \left( K^2 + \frac{1}{a^2} \right)^2 \hat{\psi} \Delta \hat{\psi}^* d\mathbf{k} \quad (58)$$

The ECRS should compensate the dissipation of energy by the preliminary scheme to ensure that  $\Delta E = 0$ . Because of the additional factor  $(K^2 + 1/a^2)$  in (58), potential enstrophy will decrease provided energy is restored at larger scales (smaller  $K$ ) than those at which it is dissipated. If energy is restored at smaller scales than the dissipation then potential enstrophy may increase with time.

## References

- Arakawa, A. and Lamb, V. R. 1977. Computational design of the basic dynamical processes of the UCLA general circulation mode. *Methods of Computational Physics*, **17**, New York: Academic Press. 173-265.
- Bates, J. R., Li, Y., Brandt, A., McCormick, S. F. and Ruge, J. 1995. A global shallow-water numerical model based on the semi-Lagrangian advection of potential vorticity. *Quart. J. Roy. Meteor. Soc.*, **121**, 1981-2005.
- Berner, J., Shutts, G. J., Leutbecher, M. and Palmer, T. N. 2009. A Spectral Stochastic Kinetic Energy Backscatter Scheme and Its Impact on Flow-Dependent Predictability in the ECMWF Ensemble Prediction System. *J. Atmos. Sci.*, **66**, 603-626.
- Boville, B. A. 2000. Toward a complete model of the climate system. in: Mote, P. and O'Neill, A. (Eds), *Numerical Modeling of the Global Atmosphere in the Climate System*, NATO Science Series C: Mathematical and Physical Sciences, *Kluwer Academic Publishers*, pp. 419-442.
- Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R. and Shutts, G. J. 2009. The local ETKF and SKED: Upgrades to the MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **135**, 767-776.
- Burkhardt, U. and Becker, E. 2006. A consistent diffusion dissipation parameterization in the ECHAM climate model. *Monthly Weather Review*, **134**, 1194-1204.
- Galewsky, J., Scott, R. K. and Polvani, L. M. 2004. An initial-value problem for testing numerical models of the global shallow-water equations. *Tellus*, **56A**, 429-440.
- Jablonowski, C. and Williamson, D. L. 2011. The pros and cons of diffusion, filters and fixers in atmospheric general circulation models, in: Lauritzen, P. H., Jablonowski, C., Taylor, M. A. and Nair, R. D. (Eds), *Numerical Techniques for Global Atmospheric Models*, *Springer*, pp. 381-493.
- Kent, J., Thuburn, J. and Wood, N. 2012. Assessing Implicit Large Eddy Simulation for Two-Dimensional Flow. *Quart. J. Roy. Meteor. Soc.*, **138**, 365-375.
- Lorenz, E. N. 1955. Available Potential Energy and the Maintenance of the General Circulation. *Tellus*, **7**, 157-167.
- Shutts, G. 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079-3102.
- Thuburn, J. 2008. Some Conservation Issues for the Dynamical Cores of NWP and Climate Models. *Journal of Computational Physics*, **227**, 3715-3730.
- Thuburn, J., Kent, J. and Wood, N. 2014. Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. *Quart. J. Roy. Meteor. Soc.*, **140**, 626-638.
- Thuburn J., Zerroukat M., Wood N., Staniforth A. 2010. Coupling a mass-conserving semi-Lagrangian scheme (SLICE) to a semi-implicit discretization of the shallow-water equations: Minimizing the dependence on a reference atmosphere. *Quart. J. Roy. Meteor. Soc.*, **136**, 146-154.
- Williamson, D. L., Olson, J. G. and Jablonowski, C. 2009. Two Dynamical Core Formulation Flaws Exposed by a Baroclinic Instability Test Case. *Mon. Weather Rev.*, **137**, 790-796.
- Williamson, D. L., Drake, J. B., Hack, J. J., Jakob, R. and Swarztrauber, P. N. 1992. A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.*, **102**, 211-224.
- Zerroukat, M., Wood, N. and Staniforth, A. 2002. SLICE: A Semi-Lagrangian Inherently Conserving and Efficient scheme for transport problems. *Quart. J. Roy. Meteor. Soc.*, **128**, 2801-2820.
- Zerroukat, M., Wood, N. and Staniforth, A. 2009. An improved version of SLICE for conservative monotonic remapping on a C-grid. *Quart. J. Roy. Meteor. Soc.*, **135**, 541-546.
- Zerroukat, M., Wood, N., Staniforth, A., White, A. and Thuburn, J. 2009. An inherently mass-conserving semi-implicit semi-Lagrangian discretisation of the shallow-water equations on the sphere. *Quart. J. Roy. Meteor. Soc.*, **135**, 1104-1116.