Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

Numerical Discretization of Rotated Diffusion Operators in Ocean Models

J.-M. BECKERS*, H. BURCHARD+, E. DELEERSNIJDER#1 AND P. P. MATHIEU#

**University of Liège, Liege, Belgium*
+Joint Research Centre, Ispra, Italy
#Unité ASTR, Université Catholique de Louvain, Louvain-La-Neuve, Belgium

Abstract

A method to improve the behavior of the numerical discretization of a rotated diffusion operator such as, for example, the isopycnal diffusion parameterization used in large-scale ocean models based on the so-called *z*-coordinate system is presented. The authors then focus exclusively on the dynamically passive tracers and analyze some different approaches to the numerical discretization. Monotonic schemes are designed but are found to be rather complex, while simpler, linear schemes are shown to produce unphysical undershooting and overshooting. It is suggested that the choice of an appropriate discretization method depends on the importance of the rotated diffusion in a given simulation, whether the field to be diffused is dynamically active or not.

1. INTRODUCTION

Ocean models typically no longer rely on parame-terizations of subgrid-scale processes involving only second derivatives taken along the coordinates underlying the numerical grid. As recognized by several authors (e.g., Cox 1987; McDougall 1984; Gent and McWilliams 1990), these parameterizations are physically unsound and should be replaced by more sophisticated formulations.

In large-scale ocean models, a widely accepted parameterization is based on the concepts of Gent and McWilliams (1990) and McDougall and Church (1986). Their arguments are based on the fact that mixing preferentially occurs along isopycnal surfaces and that bar-oclinic instabilities tend to flatten these surfaces. The parameterizations that are generally retained to consider these two aspects are then diffusion acting on isopycnal surfaces to describe the mixing and an additional ad-vection velocity for tracer fields that has a tendency to flatten the isopycnals. This approach, combining isopycnal diffusion and the so-called bolus velocity (Gent et al. 1995), will subsequently be called improved isopycnal mixing. The parameterized additional bolus velocity field depends on the shape of the isopycnal surfaces and can be calculated from the prognostic variables in a primitive equation model.

Though this idea was initially developed for coarse-resolution models, Roberts and Marshall (1998) showed that the use of horizontal diffusion rather than isopycnal diffusion still induces diapycnal transfers when resolution is increased beyond the deformation radius and they concluded that *adiabatic subgrid dissipation schemes are required, even in eddy-resolving models.* Adiabatic dissipation schemes rely on the concept of mixing along isopycnals, which in their high-resolution model is parameterized with more scale-selective operators as the biharmonic diffusion operator. Here, however, we will work on the original versions of the diffusion part of the improved isopycnal mixing based on Laplacian diffusion, since they are the most widely used.

The improved isopycnal mixing is easily implemented into layer models (e.g., Bleck and Boudra 1981; Oberhuber 1993; Chassignet et al. 1996; Paiva et al. 1999). This is because in such models the natural coordinates of the isopycnal mixing are also the coordinates in which the model operates. The improved isopycnal mixing has also been introduced into level models by several authors. For the diffusion part $\mathcal{D}$, this leads to a rotated tensor $\mathbf{K}$ as shown in Redi (1982):

[1] Research associate at the National Fund for Scientific Research, Brussels, Belgium.

$$\mathcal{D} = \nabla \cdot (\mathcal{A}\mathbf{K} \cdot \nabla\Psi), \qquad (1)$$

where $\mathbf{K}$ is defined as

$$(\rho_x^2 + \rho_y^2 + \rho_z^2)\mathbf{K} = \begin{pmatrix} \rho_y^2 + \rho_z^2 & -\rho_x\rho_y & -\rho_x\rho_z \\ -\rho_x\rho_y & \rho_x^2 + \rho_z^2 & -\rho_y\rho_z \\ -\rho_x\rho_z & -\rho_y\rho_z & \rho_x^2 + \rho_y^2 \end{pmatrix}. \qquad (2)$$

Here, x and *y* are horizontal coordinates, while *z* is the vertical one—*x, y, z* being a Cartesian coordinate system; $\nabla\cdot$ and ∇ denote the divergence and gradient operators, respectively; and $\mathcal{A}$ is the isopycnal diffusion coefficient relevant to the scalar variable Ψ.

It is also possible to include parameterizations of isopycnal diffusion as a rotated diffusion operator into ocean models based on terrain-following coordinates (see, e.g., Hedström 1994; Blumberg and Mellor 1987; Beckers 1991). The inclusion of additional velocity that flattens isopycnals is also possible but will not be analyzed here. Since these models are often applied to shallow seas, traditional diffusion along the coordinate surfaces of the numerical grid is commonly retained to account for the impact of the topography on the subgrid-scale motions to be parameterized (see, e.g., Mellor and Blumberg 1985). But, when stratification effects are dominant, this strategy may be questionable, at least for those regions where strong stratification occurs over sloping bottom topography. In this case, some rotated diffusion laws could account for the preferred direction of mixing. Another situation in which rotated diffusion in terrain-following coordinates is advantageous occurs when isopycnal surfaces are identical to geopotential surfaces. In this case, no baroclinic pressure gradient is generated. For a departure from such a situation, when applying a rotated diffusion of density back into geopotential surfaces, it can then be ensured that the system tends toward a standstill when left to itself (see Stelling and van Kester 1994), contrary to a diffusion along terrain-following coordinates, which tends to create a thermal wind.

The attractive physical features of parameterizations in terms of mixing along isopycnals and flattening of the isopycnals have led to numerical implementations in several *z*- or terrain-following coordinate models. It was then found, however, that the straightforward numerical discretizations of the parameterization in terms of diffusion and advection were not without problems. The first attempts by Cox (1987) and Gerdes et al. (1991) needed the addition of a background diffusion along grid coordinate surfaces, the limitation of the slope of the computed isopycnals, or the application of intermittent filtering. Despite these controlled numerical inconsistencies, improvements of the simulated circulation were generally found (Danabasoglu et al. 1994).

Several authors (e.g., Griffies et al. 1998; Beckers et al. 1998; Mathieu and Deleersnijder 1997) then analyzed why the discretized diffusion term (normally a smoothing operator) needs such additional damping, potentially masking the desired rotated diffusion. Others analyzed the best way to compute the additional advection (e.g., Gerdes 1993), or how to efficiently combine the advection and diffusion part into a general tensor formalism (Griffies 1998; Gnanadesikan 1999). Some improvements to the classical discretizations were introduced. In particular, Griffies et al. (1998) and Griffies (1998) designed a scheme that eliminated some problems for dynamically active tracers. They showed how to ensure that along isopycnal surfaces, diffusion fluxes for temperature and salinity combine to give a zero density flux. This ensures that the numerical formulation of the rotated diffusion applied to the density field itself is zero as does the mathematical formulation of Eqs. (1) and (2) when applied to the density field. The violation of this constraint was a major reason for instabilities in the first implementations of the isopycnal diffusion. Griffies et al. (1998) also introduced new discretizations of the diffusion tensor that allow a more accurate diffusion where the density field varies rapidly in space. However, none of these methods solves the problem identified in Beckers et al. (1998), who showed that there is no linear, first- or second-order discretization of the rotated diffusion operator that is mono-tonic. This means that new local extema may be created by the numerical version of the diffusion operator. This can be disastrous when the diffusion of a dynamically passive tracer[2] is performed: without a positive definite scheme, specially when the tracer interacts with other tracers as in biological models, the model behavior may lead to completely unrealistic fields. Typically, local negative concentrations are both unphysical and difficult to manage in nonlinear biological laws. This was also recognized by Gnanadesikan (1999), who showed the disastrous effects of spurious minima and maxima of a given tracer field due to discretized rotated diffusion, which in principle should smooth the fields rather than introduce new extrema.

[2] Hydrodynamically speaking; they could be active in the sense of a biological growth, for example.

In the present paper we investigate in detail some of the remedies proposed in Beckers et al. (1998) to deal with the violation of the monotonicity in classical linear schemes. This article is organized as follows: First, we formulate the mathematical problem in a general framework (section 2), then we summarize the known numerical problems associated with the discretizations of the rotated diffusion operator (see section 3). In section 4, some remedies will be described and then compared in section 5. Finally, their application in a general 3D model is discussed.

2. PROBLEM FORMULATION

In this work, we investigate only the case of a vertical 2D section, by assuming that the 3D problem can be treated as two 2D problems. This implies that the small-slope approximation (Cox 1987) is adopted, leading to simplifications of the tensor $\mathbf{K}$ [see Eq. (2)] such that the fluxes in the x direction are not influenced by the gradients of the field in the y direction.

Our understanding is that isopycnal diffusion is often used with small-slope approximation and that it is reasonable to follow a step-by-step procedure. This consists of first designing numerical schemes that perform well in 2D and then attempting to generalize them to 3D with respect to the full tensor.

We will also restrict our analysis to the case of a dynamically passive tracer, since the problem of dynamically active tracers leads to a complex coupling and feedback with circulation, and depending on the scales at which the model is applied, the conclusions drawn from the study of dynamically active tracers may be different. In any case, since studying dynamically passive tracers is the final purpose of an increasing number of applications, a proper discretization of their diffusion terms is necessary (e.g., Gnanadesikan 1999).

Though the main objective of this study is the design of appropriate discrete algorithms of isopycnal diffusion, a slightly more general problem can also be easily addressed.

In order to encompass a wide range of model implementations and coordinate systems (not necessarily orthogonal), we take the general case in which diffusion of a field Ψ along the coordinate line $s(\xi, \eta)$, on which the other coordinate $n(\xi, \eta)$ remains constant, can be written as

$$\mathcal{D} = \frac{\partial}{\partial s}\left(\mathcal{A}' \frac{\partial \Psi}{\partial s}\right). \qquad (3)$$

Here, (ξ, η) is a local coordinate system along the numerical grid [see Fig. 1, in which the coordinate system in which the parameterization is formulated is referenced by (s, n)].

In the case of isopycnal diffusion, the coordinate s varies along the isopycnal, and the generalized (positive) diffusion coefficient $\mathcal{A}'$ depends on the isopycnal slopes and the distance between isopycnals.

For the purpose of a conservative formulation, the diffusion term can be reformulated as follows:

$$\mathcal{D} = \frac{\partial}{\partial s}(-\Phi), \qquad (4)$$

where Φ is the diffusion flux, which can be rewritten in terms of the derivatives along the numerical coordinates:

$$\Phi = -\mathcal{A}' \frac{\partial \Psi}{\partial s} = -\mathcal{A}'\left(\frac{\partial \xi}{\partial s}\frac{\partial \Psi}{\partial \xi} + \frac{\partial \eta}{\partial s}\frac{\partial \Psi}{\partial \eta}\right). \qquad (5)$$

Since (e.g., Anderson 1995)

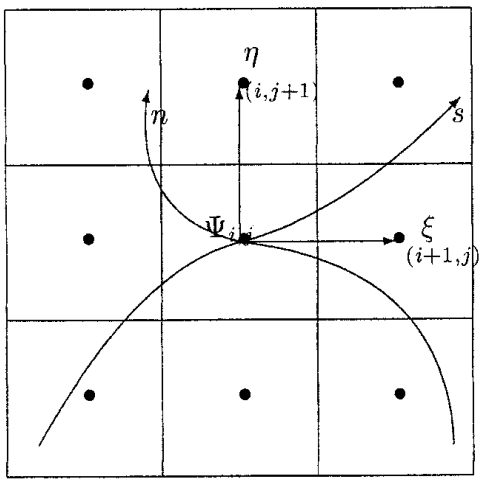
$$\frac{\partial}{\partial s} = \frac{\partial \xi}{\partial s}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial s}\frac{\partial}{\partial \eta}, \qquad (6)$$

$$\frac{\partial \eta}{\partial s} = -\frac{1}{\mathcal{J}}\frac{\partial n}{\partial \xi}, \quad \text{and} \qquad (7)$$

$$\frac{\partial \xi}{\partial s} = \frac{1}{\mathcal{J}}\frac{\partial n}{\partial \eta}, \qquad (8)$$

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 1. *Grid and naming convention.*



we can easily find an equation that will enables us to formulate a conservative equation from Eq. (3):

$$\frac{\partial}{\partial \xi}\left(\mathcal{J}\frac{\partial \xi}{\partial s}\right) + \frac{\partial}{\partial \eta}\left(\mathcal{J}\frac{\partial \eta}{\partial s}\right) = 0. \qquad (9)$$

From this general tensor analysis (e.g., Aris 1962) Eq. (5) can be transformed to

$$-\mathcal{J}\mathcal{D} = \frac{\partial}{\partial \xi}\left(\mathcal{J}\frac{\partial \xi}{\partial s}\Phi\right) + \frac{\partial}{\partial \eta}\left(\mathcal{J}\frac{\partial \eta}{\partial s}\Phi\right), \qquad (10)$$

where $\mathcal{J} = \partial(s, n)/\partial(\xi, \eta)$ is the Jacobian of the coordinate transformation. Evidently this formulations easily leads to a conservative type of discretization.

By using a classical integration over the finite volume box in (ξ, η) space, Eq. (10) can be translated into a conservative finite-difference scheme, provided that the integrated fluxes $\mathcal{J}(\partial\eta/\partial s)\Phi$ and $\mathcal{J}(\partial\xi/\partial s)\Phi$ are known at the interfaces.

As we will see below, an important factor of the formulation is

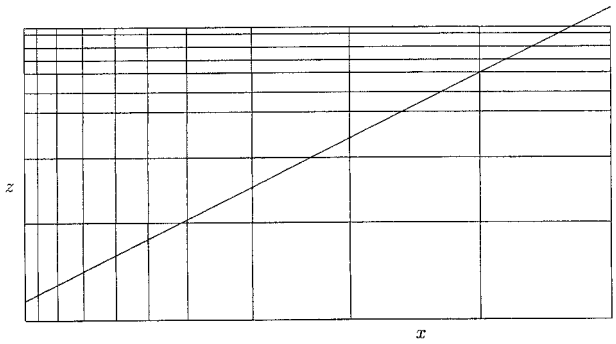
$$r = \frac{\partial \eta/\partial s}{\partial \xi/\partial s} = -\frac{\partial n/\partial \xi}{\partial n/\partial \eta}, \qquad (11)$$

which measures the ratio of the slope of the lines on which diffusion occurs (e.g., isopycnals) compared to the numerical grid aspect ratio.

From here on, it is clear that even if the small-slope approximation holds in the physical space, we cannot assume that the coordinate s along which the diffusion takes place is almost flat in the numerical coordinates ξ, η. Those coordinates take into account the anisotropy of the numerical grid. If the line were flat in the discrete space, then that would mean that in no event can the numerical grid correctly resolve the isopycnal slope. This is analogous to the z-level models, which cannot correctly resolve topographic slopes that are below the grid aspect ratio. In other words, if the isopycnals are almost flat in the discrete coordinates, then there is no point in trying to use isopycnal diffusion since the numerical scheme will not be able to resolve the slope correctly. We thus have to assume that the numerical grid is such that the relative height of the isopycnal may vary significantly from one horizontal position to the next. This translates mathematically as parameter r being at least of 0(1). A graphical representation of this is given in Fig. 2, where the physical slope may be weak, but where depending on the horizontal and vertical discretization, the isopycnal vertical position changes significantly from one horizontal grid point to the next.

To illustrate how a specific rotated diffusion may be retrieved from the generic formulation, we shall show two classical applications.

FIG. 2. *Representation of an isopycnal line in physical space, with a numerical grid superimposed. Depending on the discrete grid, the isopycnal line can significantly change its vertical position relative to the numerical grid from one horizontal grid point to the next.*



a. Horizontal diffusion in uniformly discretized σ coordinates

In this case, the σ coordinate ranges from $\sigma = 0$ at the bottom, $z = -d$, to $\sigma = 1$ at the surface, $z = \zeta$. Assuming uniform discretizations along the x and σ coordinate with increments Δx and $\Delta\sigma$, respectively, we may define the coordinates ξ, η to vary in the discrete space exactly as the indices of the grid points so that

$A\xi = \Delta\eta = 1$[3] (see Fig. 1):

$$\xi\Delta x = x \quad \text{and} \tag{12}$$

$$\eta\Delta\sigma = \frac{z + d}{\zeta + d} = \sigma. \tag{13}$$

The assumption made here that the grid spacing is constant is not required but simplifies the presentation.

The coordinates that define the line on which diffusion takes place are thus (s, n), where the diffusion takes place for constant n along the s line:

$$s = x = \xi\Delta x \quad \text{and} \tag{14}$$

$$n = z = \sigma(\zeta + d) - d. \tag{15}$$

From these definitions, we can easily compute the metric coefficients of the transformations,[4]

$$\frac{\partial\xi}{\partial s} = \frac{1}{\Delta x} \quad \text{and} \tag{16}$$

$$\frac{\partial\eta}{\partial s} = \frac{1}{\Delta\sigma}\left\{\frac{d_x}{\zeta + d} - \sigma\frac{\zeta_x + d_x}{\zeta + d}\right\}, \tag{17}$$

and the Jacobian is

$$\mathcal{J} = \frac{\partial(s, n)}{\partial(\xi, \eta)} = \left[\frac{\partial(\xi, \eta)}{\partial(s, n)}\right]^{-1} = (d + \zeta)\Delta x\Delta\sigma. \tag{18}$$

[3] This does not mean of course that the physical grid space is isotropic, but that we made the arbitrary choice of using a numerical coordinate system varying by unit steps from one grid point to the next.

[4] Here, d_x stands for $\partial d/\partial x$.

For small surface slopes (neglecting the sea surface slope compared to the bottom slope), parameter r reads

$$r = \frac{\Delta x \frac{d_x}{d}}{\Delta \sigma \frac{1}{1 - \sigma}} \qquad (19)$$

and can be readily interpreted as the measure for the hydrostatic consistency in terrain-following coordinates (e.g., Blumberg and Mellor 1987; Haney 1991; De-leersnijder and Beckers 1992). This gives the interesting interpretation of the hydrostatic consistency requirement $|r| \leq 1$ in terms of variation of the z coordinates when seen in discrete space. These lines of constant z should never vary more than one discrete vertical unit when moving one discrete horizontal unit.

b. Isopycnal diffusion in uniformly discretized *z*-coordinates

If again, we assume a uniformly discretized space with constant horizontal grid size Δx and vertical grid size Δz, the numerical grid is defined by the coordinates ξ, η, which vary as the gridpoint indices:

$$\xi = \frac{x}{\Delta x} \quad \text{and} \qquad (20)$$

$$\eta = \frac{z}{\Delta z}. \qquad (21)$$

The coordinates that define the direction of the diffusion are those for which n is constant:

$$s = x \quad \text{and} \qquad (22)$$

$$n = \rho(x, z). \qquad (23)$$

The vertical coordinate is density p in this case, and the metrics of the transformations can be calculated using properties of Eqs. (7) and (8), so that they read

$$\frac{\partial \xi}{\partial s} = \frac{1}{\Delta x} \quad \text{and} \qquad (24)$$

$$\frac{\partial \eta}{\partial s} = -\frac{\rho_x}{\rho_z}\frac{1}{\Delta z}, \qquad (25)$$

with the Jacobian

$$\mathcal{J} = \rho_z \Delta x \Delta z. \qquad (26)$$

In this case, it may be shown that we retrieve the formulation of Redi (1982) [see Eq. (2)] by using the generalized diffusion coefficient $\mathcal{A}''$ given by

$$\mathcal{A}' = \frac{1}{\rho_z}\frac{\rho_z^2}{\rho_x^2 + \rho_z^2}\mathcal{A}. \qquad (27)$$

Here, the parameter r can be identified as the physical isopycnal slope $S = -\rho_x/\rho_z$ multiplied by the grid aspect ratio.

3. PROBLEMS IDENTIFIED

As already mentioned in the introduction, various authors realized that the straightforward discretization of the reformulated diffusion in the grid coordinates leads to numerical problems, which are summarized in the following sections.

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

a. Dynamically active tracers

The major problems for large-scale applications were the noncancellation of density flux contributions of temperature and salinity on isopycnal surfaces (Griffies et al. 1998). This led to inaccuracies that could explain some of the instabilities observed in the studies of Gough and Welch (1994) and Gough (1997). This problem is of a dynamic nature, but any additional problem that exists for dynamically passive tracers is also a potential problem for temperature and salinity, although the dynamical effect may be controlled by the method of Griffies et al. (1998). Therefore we focus now on the problems identified for dynamically passive tracers.

b. Dynamically passive tracers

For dynamically passive tracers, one problem identified in Griffies et al. (1998) as well is the inappropriate computation and averaging of products of grid slopes and gradients in situations where the density field exhibits rapid horizontal variations with a spatially high wavenumber. Griffies et al. (1998) show how this problem appears in the classical discretization (Cox 1987) and how it can be solved by using different averaging techniques.

The focus of the present paper is on a problem that remains unsolved even for very slowly varying or constant slopes: as shown inBeckers et al. (1998), no mono-tonic scheme can be obtained when classical consistent and linear schemes are used, unless the slope and the grid is such that $r = 0$ or $r = \pm$ l or r = $\pm \infty$. These are cases in which diffusion occurs on a line that crosses the grid points. In these cases a direct classical diffusion discretization along these grid points works well. But if the diffusion direction does not coincide with the grid, as shown in Beckers et al. (1998), the numerical stencil, which gives the contribution of the surrounding point to the evolution of the central point, always contains negative coefficients. This feature is responsible for the possibility of creating new local extrema.

Even for constant slopes, uniform grid spacing, and constant diffusion coefficient, no well-behaved linear scheme can be found. Furthermore, it is also clear that any method of a priori limiting the isopycnal slope to a prescribed maximum amplitude is not appropriate to eliminate this monotonicity violation. Anyway, it is not the slope that is the important control factor of the "neg-ativeness," but the slope compared to the coordinate slope, as reflected by the parameter r.

Griffies (1998) and Gnanadesikan (1999) argue in the case of isopycnal diffusion that the problem of monotonicity may be overcome by the concurrent use of the Gent-McWilliams advection parameterization. They showed how the isopycnal diffusion and the layer thickness diffusion can be cast into a single asymmetric tensor formulation. In the numerical experiment of Gnanadesikan (1999), this helped to stabilize a coupled biological model. However, the asymmetric tensor formulation modifies the flux computation in such a way that negative coefficients in the discrete stencil appear only in the vertical fluxes. The vertical direction is where the biological gradients and fluxes are most important and the stabilization seems to result from a better combination of the Gent-McWilliams advection-dif-fusion fluxes resulting from the asymmetric tensor formulation with these strong diapycnal exchanges and local production terms. Since the stabilization effect probably depends on the biological model, it might not work in all situations. And there are further reasons why this stabilizing effect might fail. The Gent-McWilliams advection parameterization is not always relevant, for example, in the case of regional σ-coordinate models with geopotential diffusion. In other words, the tracer diffusion coefficient A used in the diffusion part and the thickness diffusion coefficient κ used in the advection part may differ significantly from one situation to another, such that a compensation effect can generally not be expected. Typically, A is taken as a constant parameter, whereas the isopycnal thickness diffusivity used in the bolus velocity calculation is calculated depending on the local Richardson number (Visbeck et al. 1997). Furthermore, an advection operator generally introduces positive and negative coefficients into the nine-point stencil unless a numerically very diffusive scheme is used. Since the Gent-Mc Williams advection velocity is related to second-order derivatives of the density field, whereas the rotated diffusion is related to first-order derivatives, there is no particular reason why the negative coefficients of one operator are canceled out by a positive coefficient of the other operator. In the skew flux formulation of Griffies (1998), this is somehow hidden, since this formulation is introduced as a "diffusive" formulation with a (nonsymmetric) diffusion tensor depending on the isopycnal slopes only and not on their second derivatives. Because of the asymmetry of the tensor, the actual derivatives of the fluxes lead to an advection part and a diffusion part depending on different derivatives of the density field. A striking example is the locally relevant case of a uniform slope in the density field. This leads to isopycnal diffusion without any Gent-Mc Williams advection both in the skew formulation and the classical formulation. In this case, there is no cancellation between the two effects. Since the locally uniform slope is very likely to occur, we certainly should correctly represent the pure diffusion part, because the advection part vanishes in this

case. The development of a well-behaved discretization of the rotated diffusion operator is thus discussed in the next section.

4. DISCRETIZATION METHODS

A well-behaved scheme should be monotonic, and in order to ensure such a method, one has to eliminate one of the requirements that were shown (Beckers et al. 1998) to lead to the impossibility of having a monotonic scheme: the scheme was assumed to be based on a nine-point stencil (in the vertical plane), a linear method (discretization being not a function of the solution), and a consistent scheme. At least one of these conditions cannot be satisfied by a monotonie algorithm. On the other hand, a conservative scheme could be essential for long-term climatic calculations and tracer dispersion. Otherwise, a simple, nonconservative method is given by clipping overshooting and undershooting values such that monotonicity is obtained. However, here we do not concentrate on nonconservative schemes (easily forced to be monotonie), we focus rather on the behavior of classical linear schemes compared to certain new conservative discretizations. Before introducing the new schemes aiming at achieving a monotonie behavior, some classical linear discretizations are presented.

a. Linear consistent schemes

Here, we will not explain in detail the linear consistent discretizations used classically, because they may differ strongly in the way the isopycnal slopes are computed, averaged, and combined with the gradients of the fields to be diffused. Griffies et al. (1998) show that a modification in this kind of averaging for variable slopes may result in drastic changes when diffusing in a spatially rapidly varying density field. Since all these different averaging techniques for the isopycnal slopes lead to the same scheme when the slope is constant, we will focus on this case. This is also justified by the fact that locally uniform slopes are likely to be present in real situations, and that such a situation is the basic situation any scheme should be able to deal with. For the linear schemes, we therefore assume a constant slope and refer to Griffies et al. (1998) for a generalization to cases with variable slopes.

For constant slopes and uniform grids $s = \Delta s\xi$, the expression for the diffusion flux Φ of Eq. (5) simplifies to

$$\Phi = -\mathcal{A}'\frac{\partial \Psi}{\partial s} = -\frac{\mathcal{A}'}{\Delta s}\left(\frac{\partial \Psi}{\partial \xi} + r\frac{\partial \Psi}{\partial \eta}\right), \qquad (28)$$

where the r coefficient, which measures the relative slope compared to the aspect ratio, is now a constant. The diffusion operator $\mathcal{D}$ given in Eq. (3) can then also be simplified:

$$\frac{\Delta s^2}{\mathcal{A}'}\mathcal{D} = \frac{\partial}{\partial \xi}\left(\frac{\partial \Psi}{\partial \xi} + r\frac{\partial \Psi}{\partial \eta}\right) + \frac{\partial}{\partial \eta}\left(r\frac{\partial \Psi}{\partial \xi} + r\frac{\partial \Psi}{\partial \eta}\right). \qquad (29)$$

This simplification for constant slopes and uniform grid allows for understanding of the nonmonotonic behavior more easily by analyzing the stencil obtained by the discretization: the stencils shown in the following figures provide the contribution of the surrounding points to the evolution of the central point when performing flux differencing in the linear schemes; any negative coefficient except the central point leads to a nonmonotonic scheme (Beckers et al. 1998).

1) LINEAR

As an example of the classical Cox discretization of Redi's rotated diffusion, as well as the adaptation of Griffies et al. (1998), the numerical stencil is shown in Fig. 3 when the slope is constant. This stencil is obtained by calculating the fluxes at the interfaces by using classical algebraic averages of gradients at the interface.

To illustrate the standard way of discretization of Eq. (28), we take the calculation of the flux at the interface between points i, j and $i + 1, j$:

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

$$-\frac{\Delta s}{\mathcal{A}'}\Phi_{i+1/2,j} = \Psi_{i+1,j} - \Psi_{i,j} + r/4 \times (\Psi_{i,j+1} - \Psi_{i,j} + \Psi_{i+1,j+1} - \Psi_{i+1,j} + \Psi_{i,j} - \Psi_{i,j-1} + \Psi_{i+1,j} - \Psi_{i+1,j-1}). \quad (30)$$

Similar expressions are easily obtained for the other interfaces and thus involve the calculation of averages for gradients that are not naturally defined by a single finite difference.

By using a very straightforward flux differencing of Eq. (29),

$$\frac{\Delta s^2}{\mathcal{A}'}\mathcal{D} = -\Phi_{i+1/2,j} + \Phi_{i-1/2,j} + r(-\Phi_{i,j+1/2} + \Phi_{i,j-1/2}), \quad (31)$$

a discretization is obtained that is simply a standard discretization of

$$\frac{\Delta s^2}{\mathcal{A}'}\mathcal{D} = \frac{\partial^2\Psi}{\partial\xi^2} + 2r\frac{\partial^2\Psi}{\partial\eta\partial\xi} + r^2\frac{\partial^2\Psi}{\partial\eta^2}, \quad (32)$$

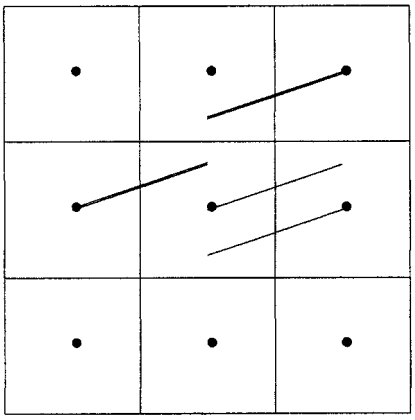
leading to the stencil of Fig. 3.

This stencil clearly shows that the nonmonotonic tendencies stem from the cross-derivative terms, since these are the derivatives that introduce the negative coefficients into the stencil and include the possibility of introducing new extrema into the solution. The simplest situation for the generation of a new extrema is a constant field with a positive perturbation at the points where the coefficients in the stencil are negative. This will lead to a time tendency that will create a negative (mathematically incorrect) perturbation in the center.

FIG. 3. *Stencil for the standard LINEAR discretization in the case of constant slopes and grid spacings. For the sake of clarity, no multiplication constant was included. We assume that we reference fluxes and points compared to the central point of the control volume, whose coordinates are thus i, j. For fluxes, when they are not referenced by coordinates, they correspond to either the interface i + ½, j or i, j + ½.*

$-r/2$	r^2	$r/2$
1	$-2-2r^2$	1
$r/2$	r^2	$-r/2$

FIG. 4. *Fluxes involved by the modified method LINEAR1. At the right interface, only two vertical derivatives are used (single lines), whereas at the top interface, only two horizontal derivatives are used (double lines). For the fluxes at the right interface, the discretization of Eq. (33) indeed uses the vertical derivatives* $\Psi_{i+1,j+1} - \Psi_{i+1,j}$, *and* $\Psi_{i,j} - \Psi_{i,j-1}$ *only.*



2) LINEAR1

Another linear scheme, LINEARl, can be constructed by computing the averages of the four triads defined by Griffies et al. (1998), not as a simple average, but taking into account the direction of the slope, so as to consider only the triads in the corresponding direction. This means that if we have to calculate the gradients at an interface, we do not compute the average of the four surrounding gradients, but only the average of the two gradients that are crossed by the isopycnal line.

For a positive slope, this would lead to the following evaluation of the flux:

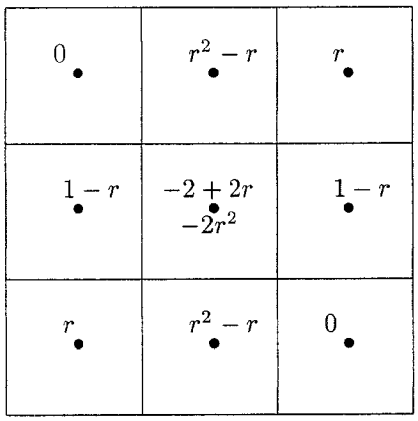
$$-\frac{\Delta s}{\mathcal{A}'}\Phi_{i+1/2,j} = \Psi_{i+1,j} - \Psi_{i,j} + r/2 \times (\Psi_{i+1,j+1} - \Psi_{i+1,j} + \Psi_{i,j} - \Psi_{i,j-1}). \quad (33)$$

This discretization clearly involves the fluxes shown in Fig. 4. From this and analogous formulations for the other interfaces, one obtains the stencil for the case shown in Fig. 5.

This stencil potentially leads to a smaller monoto-nicity violation when the slope is close to one, since the negative coefficients are lower, but the monotonicity problem remains. Interestingly enough, for the stencil shown in Fig. 5, the method can be rendered monotonic not by adding a horizontal background diffusion as done usually, but by adding vertical diffusion proportional to the slope-dependent parameter $r - r^2$. This also suggests that this scheme may behave better when a dia-pycnal diffusion is present, because such a diapycnal diffusion could cancel the negative coefficients, which would not be the case for the classical scheme with a horizontal background diffusion.

If we add however a vertical background diffusion that always cancels exactly the negative weighting coefficients, then we simply obtain the inconsistent scheme COMBI presented in section 4b(2).

FIG. 5. *Stencil for the modified linear method LINEAR1. For the sake of clarity, no multiplication constant was included and a positive slope parameter r was assumed in this case.*



3) LINEAR2

Other slope-dependent choices of the interface flux weighting can be envisaged, as for example a linear combination of the two vertical differences as a function of the slope parameter:

$$-\frac{\Delta s}{\mathcal{A}'}\Phi_{i+1/2,j} = \Psi_{i+1,j} - \Psi_{i,j} + r/4 \times [(1 - r)(\Psi_{i,j+1} - \Psi_{i,j}) + (1 + r)(\Psi_{i+1,j+1} - \Psi_{i+1,j}) + (1 + r)(\Psi_{i,j} - \Psi_{i,j-1}) + (1 - r)(\Psi_{i+1,j} - \Psi_{i+1,j-1})]. \quad (34)$$

The ensuing stencil (Fig. 6) can also be interpreted as being obtained by using the classical stencil of Fig. 3 in which the vertical diffusion part of Eq. (32), rather than being discretized on the central vertical line, has been distributed on the surrounding vertical lines (a consistent truncation error).

b. Inconsistent linear schemes

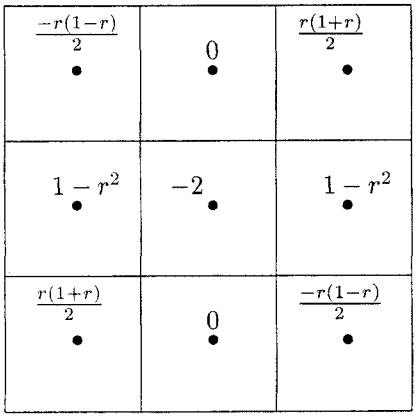
A first approach to render linear schemes monotonie is to relax the consistency.

1) CLASSIC

An inconsistent scheme, hereafter called CLASSIC, which is already used currently, is the scheme in which a background diffusion in the "horizontal" numerical grid is maintained. Typically, the diffusion coefficient associated with this background diffusion is 20% of the isopycnal diffusion coefficient. But when looking at the stencils of the linear methods, this procedure is not likely to reduce the nonmonotonicity problem, since it does not influence the cross-derivative terms.

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 6. *Stencil for a second modified linear method LINEAR2. For the sake of clarity, no multiplication constant was included.*



2) COMBI

Since the problem of nonmonotonicity stems from the cross-derivative terms, one may try to eliminate them by using a combination of kinds of diffusion along the grid lines, which mimic the directional effect of the rotated diffusion: in a nonflux form this can be written as

$$\frac{\Delta s^2}{\mathcal{A}'}\mathcal{D} = \alpha\mathcal{D}^{|} + \beta\mathcal{D}^{-} + \gamma\mathcal{D}^{/} + \delta\mathcal{D}^{\backslash}, \tag{35}$$

$$\mathcal{D}^{|} = \Psi_{i,j+1} + \Psi_{i,j-1} - 2\Psi_{i,j}, \tag{36}$$

$$\mathcal{D}^{-} = \Psi_{i+1,j} + \Psi_{i-1,j} - 2\Psi_{i,j}, \tag{37}$$

$$\mathcal{D}^{/} = \Psi_{i+1,j+1} + \Psi_{i-1,j-1} - 2\Psi_{i,j}, \quad \text{and} \tag{38}$$

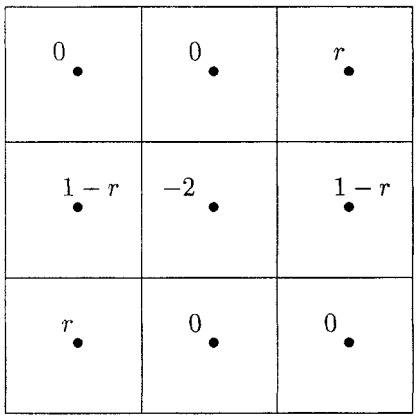
$$\mathcal{D}^{\backslash} = \Psi_{i+1,j-1} + \Psi_{i-1,j+1} - 2\Psi_{i,j}. \tag{39}$$

This COMBI discretization is generally *not* consistent, but if the coefficients $\alpha, \beta, \gamma, \delta$ are nonnegative, then the monotonicity is easily satisfied for small time steps. Such coefficients can then at least be chosen so that a discretization mimics the real diffusion as well as possible. Here we used a linear combination of two $\mathcal{D}^{*}$'s depending on the slope parameter r, such that when $r = 1$ for example, $\mathcal{D}^{/}$ is retrieved (for $0 \le r \le 1$ one uses for example $\alpha = 0, \delta = 0, \beta = 1 - r, \gamma = r$). The problem with this formulation is that the conservative form is more complicated due to $\mathcal{D}^{/}$ and $\mathcal{D}^{\backslash}$ Those terms written in conservative form require some averaging of Ψ involving at each interface more points than those of the nonconservative form. When slopes and diffusion coefficients are constant, those contributions cancel out when flux differencing is performed, but for nonuniform grids and slopes, this cannot be guaranteed and the purely three-point stencils combination may be lost.

Another problem is the clear diapycnal diffusion that will develop when using only positive weightings. If we assume for example a moderate positive slope, the operators $\mathcal{D}^{/}$ and $\mathcal{D}^{-}$ would be combined. This will however introduce diffusion in the diagonal direction. A signal will thus propagate between the horizontal line in the discrete space and the diagonal direction in the discrete space, rather than propagating only in the direction of the slope. This will be observed later in the test cases.

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 7. *Stencil for the inconsistent linear method COMBI. A slope parameter r e [0, 1] was assumed. For the sake of clarity, no multiplication constant was included.*



Of course some more complicated and nonlinear weighting of the two operators could be used, but the fact remains that there will always be only pure diffusion in both the diagonal and horizontal direction. One of these diffusions could be kept small locally, but in order to have some kind of diffusion, at least the other one must be present, and we will thus always tend to diffuse away from the slope direction.

In the case of a constant slope, the implementation of the COMBI method simply combines some discrete diffusion along oblique, horizontal, and vertical directions weighting them in the function of the slope parameter *r.* In this case, the scheme COMBI shown in Fig. 7 clearly ensures monotonicity but is inconsistent. Compared to the consistent discretization LINEAR 1 of Fig. 5, we have added a permanent vertical diffusion proportional to $r - r^2$.

c. Nonlinear computations

1) AMPMIN

Stelling and van Kester (1994) presented a monotonic method based on a nonlinear flux minimization. In their work, the authors discuss the problem of diffusion along geopotentials in a σ-coordinate model. Their approach is based on a transformation of the numerical grid, where the finite volumes (normally defined in the σ space) are first rotated into rectangular horizontal boxes. Then, since the box interfaces do not generally horizontally match their neighbors, a *z* interpolation of sca-lars is needed to compute the fluxes at interfaces. For small slopes this involves only the classical nine points and can be efficient. However, when slopes are arbitrary, the interpolation method requires the scanning of the whole water column for each flux computation.

On the other hand, the authors prove their scheme to be monotonic, if the fluxes are computed by means of a nonlinear minimization. Unfortunately, this method is by and large time consuming. The computational burden of a hopefully small effect ("horizontal" diffusion) should not penalize the whole ocean model. Therefore, we could adopt the approach of Stelling and van Kester (1994) if relative slopes are small (which could possibly be enforced by slope limiting in the code) or if we find another similar nonlinear interpolation method limited to the local stencil rather than the whole water column.

The idea of limiting fluxes at the interfaces can be used here in the following approach. Similar to Stelling and van Kester (1994), when the slopes are weak $|r| \leq 1$, two consistent flux calculations can be performed at each interface. At the interface between *i, j* and $i + 1, j,$ two possibilities to compute Φ from Eq. (28) are defined:

$$\Phi^{+}_{i+1/2,j} = -\frac{\mathcal{A}'}{\Delta s}[(1-|r|)\Psi_{i+1,j} + |r|\Psi_{i+1,j+\epsilon} - \Psi_{i,j}] \quad \text{and} \tag{40}$$

$$\Phi^{-}_{i+1/2,j} = -\frac{\mathcal{A}'}{\Delta s}[\Psi_{i+1,j} - (1-|r|)\Psi_{i,j} - |r|\Psi_{i,j-\epsilon}], \tag{41}$$

with

$$\epsilon = \text{sign}(r). \tag{42}$$

For convenience, the AMPMIN function is defined by

$$\text{AMPMIN}(a, b) = 0.5[\text{sign}(a) + \text{sign}(b)]\min(|a|, |b|), \tag{43}$$

which is an interpolating weighting function selecting the flux whose amplitude is minimal or zero if fluxes have different signs.

In the case $|r| \leq 1$, the flux at the interface between i, j and $i + 1, j$ is then chosen as

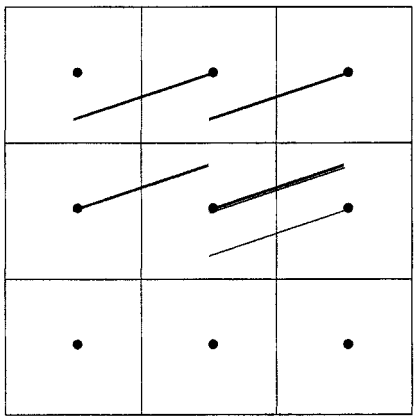
$$\Phi = \text{AMPMIN}(\Phi^{+}, \Phi^{-}), \tag{44}$$

with the r parameter being computed as the vertical average of the slopes at the grid corners above and below the interface.

For the interface between i, j and $i, j + 1$ the method reads

$$\jmath\frac{\partial\eta}{\partial s}\Phi = 0.5\left(\jmath\frac{\partial\eta}{\partial s}\Phi|_{i+1/2,j+1/2} + \jmath\frac{\partial\eta}{\partial s}\Phi|_{i-1/2,j+1/2}\right). \tag{45}$$

FIG. 8. *Fluxes involved in the nonlinear method AMPMIN. At the right interface, two fluxes are involved (single and triple lines), whereas at the top interface, four (double and triple lines) are taken into account for $|r| \leq 1$.*



The fluxes at the corners are then again chosen by the AMPMIN function for the two nearest available flux computations.

$$\Phi|_{i+1/2,j+1/2} = \text{AMPMIN}(\Phi^{+}_{i+1/2,j}, \Phi^{-}_{i+1/2,j+1}), \tag{46}$$

in the case r is positive.

In this case, the r parameter is again computed at the corners and is used for the two fluxes that are surrounding the corners.

For the case of slopes $|r| \leq 1$ the fluxes involved in the calculation of the interface flux are those depicted in Fig. 8. Two evaluations of gradients are needed at the right interface, whereas four are involved in the flux at the top interface.

It is easily shown that the flux calculation itself is consistent, because fluxes are always interpolated (non-linearly). As shown in the appendix, the flux differencing for a finite volume may however lead to inconsistencies for the diffusion operator. This is because each of the consistent fluxes has a different type of truncation errors (due to the presence of the AMPMIN functions selecting different flux discretizations and thus truncation errors at the different interfaces) that, when performing the flux differencing, introduce the inconsistency.

Several tests with grids satisfying $|r| \leq 1$ everywhere showed that the method behaved correctly concerning the monotonicity properties, though we were not able to demonstrate that the scheme satisfies the monotonicity principle for slopes $|r| \leq 1$. But even if the scheme seems monotonie for $|r| \leq 1$, the problem of larger slopes must be tackled. The generalization for larger slopes should solely be based on the local stencil in order to avoid the expensive scanning of the water column. This problem arises when the relative slope increases to the value where an extrapolation rather than an interpolation is performed during the flux calculation according to (44). In the original version of Stelling and van Kester (1994), a search of the two points that really surround the s coordinate line is carried out. But this is very expensive and in some cases even impossible (near the bottom boundaries, for example).

If the slope is steeper, $|r| \geq 1$, we suggest computing

$$\Phi^{+}_{i+1/2,j} = -\frac{\mathcal{A}'}{\Delta s}\left[\left(1 - \frac{1}{|r|}\right)\Psi_{i,j+\epsilon} + \frac{1}{|r|}\Psi_{i+1,j+\epsilon} - \Psi_{i,j}\right], \tag{47}$$

and similarly for the other flux.

While in the case of $|r| \leq 1$, two fluxes are involved at the interface between i, j and $i + 1, j$ and four fluxes were involved at the interface between i, j and $i, j + 1$, we now have the inverse: Two fluxes must be used at the interface above the grid point and four laterally. Not only does the number of fluxes involved in the minimization process vary depending on the slopes, but also the decision of which two fluxes are relevant depends on the slope (via the sign of the slope). This thus leads to important testing sections in the algorithm.

An advantage of the method is that if formulated as a weighting of different fluxes at the interfaces, linear schemes may be recovered by the appropriate choice of the weighting functions. This would for example allow for switching from the Griffies et al. (1998) scheme to the AMPMIN scheme merely by changing the weighting functions.

A major practical problem of the present method is however the treatment of the vertical part of the fluxes, since those can lead to restrictions on the time step (Mathieu et al. 1999), because the nonlinear algorithm is not easily implemented in the framework of an existing implicit treatment of vertical fluxes.

In addition, if any slope limitation is desired (for other reasons than numerical), then this limitation can only be imposed at the moment when the flux is computed. However, some tests will then indicate that the method does not behave correctly, because the finite volume is not changed by the slope limiting, and the differencing of the fluxes creates problems.

As we will show later in examples, if we use the minimum amplitude flux approach proposed here for both smaller and larger slopes, the method seems to behave monotonically, but we are not able to prove this mathematically.

2) EQUIVALENT NONLINEAR DIFFUSION ALONG GRID LINES

Another discretization already used is a fully implicit scheme as in Harvey (1995), but this is neither easily implemented into existing GCMs nor very efficient in terms of CPU resources. But Harvey (1995) suggests a different approach. One could rewrite

$$\mathcal{JD} = \mathcal{JD}_\xi + \mathcal{JD}_\eta, \tag{48}$$

$$\mathcal{JD}_\xi = \frac{\partial}{\partial \xi}\left[\mathcal{J}\left(\frac{\partial \xi}{\partial s}\right)^2 \mathcal{A}'\left(1 - \frac{r}{R}\right)\frac{\partial \Psi}{\partial \xi}\right], \quad \text{and} \tag{49}$$

$$\mathcal{JD}_\eta = \frac{\partial}{\partial \eta}\left[\mathcal{J}\left(\frac{\partial \eta}{\partial s}\right)^2 \mathcal{A}'\left(1 - \frac{R}{r}\right)\frac{\partial \Psi}{\partial \eta}\right], \tag{50}$$

where

$$R = -\frac{\partial \Psi/\partial \xi}{\partial \Psi/\partial \eta} \tag{51}$$

is the relative slope of the field to be diffused, compared to the aspect ratio of the numerical coordinate grid.

The problem is formally equivalent to diffusion along the grid lines, with diffusion coefficients that depend on the solution and can be negative locally in time and space. A sufficient condition to ensure a monotonic solution for each small time step is the use of a positive apparent diffusion coefficient along the grid lines. Using only positive apparent diffusion coefficients is in principle not necessary to ensure that the monotonicity principle is satisfied, since (48) is just a reformulation of the monotonic physical diffusion. This equivalence shows that "negative" diffusion along the grid lines may be necessary. By limiting the apparent (solution dependent) diffusion coefficients to positive values, one can clearly satisfy the monotonicity principle if the chosen time step is short enough (otherwise one can also impose an upper limit on the apparent diffusion coefficient). This scheme is also conservative but inconsistent when any apparent negative equivalent diffusion coefficient is set to zero. The advantages of the scheme are promising and, in addition, the practical implementation of such a scheme is almost immediate in a model already including a diffusion module along grid lines with varying diffusion coefficients. Since on the vertical, time stepping is generally implicit, time step restrictions associated with the vertical flux in the case of strong slopes can be dealt with automatically. Furthermore, it is relatively easy to ensure that for a system in which density depends linearly only on temperature, for example, the isopycnal diffusion of temperature is zero. This can be easily achieved by computing r and R in an identical way so that $r = R$ when temperature is constant on s lines. There are thus numerous advantages but, unfortunately, when the s line and the solution have a slope of the same sign, the apparent diffusion coefficient must always be limited, because one of the two equivalent diffusion coefficients is always negative in this case. This is because $1 - r/R$ and $1 - R/r$ have different signs when $r/R > 0$. For slopes of opposite sign ($r/R \leq 0$), the method is however consistent and simply leads to a strong, physically correct diffusion in both ξ and η directions.

Another problem of this equivalent diffusion is that sometimes an upper bound limit for the equivalent diffusion must be set. This is due to the appearance of gradients of the isopycnals and the tracer fields in the denominator. In principle, this denominator should cancel out during the final flux computation, but since some averaging of the equivalent diffusion may be necessary, this cancellation cannot be guaranteed numerically.

We will not show results of a method in which the equivalent diffusion coefficients are always forced to be nonnegative, since the next method encompasses this possibility.

3) FLUX LIMITER APPROACH FLUXCORR[5]

The preceeding method of using zero equivalent diffusion whenever it is negative has the disadvantage that it introduces an inconsistent scheme, whenever the slope of the solution and the diffusion direction have the same sign. Limiting the apparent diffusion coefficient to positive values is however only a sufficient condition to ensure a monotonie scheme, since downgradient fluxes at interfaces are not necessary to ensure a positive definite scheme. What matters is the compensation of some upgradient fluxes at an interface by sufficiently strong downgradient fluxes at another interface of the grid box. In other words, upgradient fluxes (i.e., negative

[5] As is sometimes the case in the literature, we will also call this method the "flux corrected method," although this terminology is somewhat incorrect since such methods are generally based on two-stage approaches.

equivalent diffusion coefficients) may be allowed as long as the budget over the grid box remains positive definite. One possibility is thus to add positive diffusion at the interfaces, but just the minimal quantity necessary to guarantee that the next time step does not create extrema outside the range around the grid point.

This leads naturally toward flux limiting approaches. Though these methods were developed for advection problems, the basic idea is to add just as much numerical diffusion as is necessary locally to ensure a monotonie scheme. This is, superficially at least, similar to our problem. Similarly to the advection problem, we define a high-order flux F^h, which is the flux based on the actual (and thus possibly negative) equivalent diffusion coefficient, and a low-order flux F^l (assuring a positive definite scheme), which is either the high-order flux for positive equivalent diffusion coefficients, or zero for the upgradient case. The low-order flux may at first glance appear to be inconsistent, but as shown in the appendix, the limiter function itself depends upon the resolution, and when grid spacing tends toward zero, the limiter function gives back the high-order flux.

For the fluxes in the ξ direction, this method can be described by the following equations in which the equivalent diffusion $\mathcal{A}^e$ is used:

$$\mathcal{A}^e = \mathcal{J}\left(\frac{\partial \xi}{\partial s}\right)^2 \mathcal{A}'\left(1 - \frac{r}{R}\right). \qquad (52)$$

It should be noted that

$$R - r = \frac{\dfrac{\partial \Psi}{\partial \eta}\dfrac{\partial n}{\partial \xi} - \dfrac{\partial \Psi}{\partial \xi}\dfrac{\partial n}{\partial \eta}}{\dfrac{\partial \Psi}{\partial \eta}\dfrac{\partial n}{\partial \eta}}, \qquad (53)$$

so that in practice the component 1 - *rlR* can easily be translated into a Jacobian in the vertical plane between the solution and the isopycnal lines.

From the conservative formulations (48)-(51) and relation (53), it appears that the most natural way to calculate the equivalent diffusion coefficients $\mathcal{A}^e$ is to compute first corner values (because there *r/R* is most naturally calculated) and then to take an average of the corner values to retrieve the interface values needed for the final flux computation.

The high- and low-order fluxes are computed as

$$F^h_{i+1/2} = -\mathcal{A}^e(\Psi_{i+1,j} - \Psi_{i,j}) \quad \text{and} \qquad (54)$$

$$F^l_{i+1/2} = -\mathcal{A}^e(\Psi_{i+1,j} - \Psi_{i,j})\gamma(\mathcal{A}^e), \qquad (55)$$

where $\gamma(x)$ is the classical Heaviside function, which is zero for negative values of *x,* and takes a unit value otherwise. In classical flux limiter methods, the limiter parameter is computed by taking the difference between the two types of fluxes, and then by taking the ratio of this difference at the interface where the flux is to be computed and a second interface, which is chosen to be the left or right neighbor interface, depending on the sign of the advection velocity.

The classical flux limiter approaches were developed for advection, and the flux differences introduced the gradients of the fields into the limiter function. The ratio of the flux differences was a measure of the ratios of gradients and, thus, the variability of the field (which ultimately is the estimator of the need to increase the diffusion). Here we should base the ratio directly on the gradients rather than on the flux difference, because the fluxes are already based on the gradients.

The choice of the direction in which the ratio is computed depends on the advection direction for an advection problem. Here we base it on the direction of the basic high-order flux, which indicates in which direction the flux should be.

The scheme is then computed as

$$F_{i+1/2} = F^l_{i+1/2} + \phi(q)(F^h_{i+1/2} - F^l_{i+1/2}), \quad (56)$$

with

$$q = \frac{F^h_{i+1/2-m}\mathcal{A}^e_{i+1/2}}{F^h_{i+1/2}\mathcal{A}^e_{i+1/2-m}} \quad \text{and} \quad (57)$$

$$m = \text{sign}(F^h_{i+1/2}). \quad (58)$$

Typical limiters are (e.g., Zijlema 1996)

SMART limiter,

$$\phi(q) = \max[0, \min(4, \tfrac{3}{4}q + \tfrac{1}{4}, 2q)], \quad \text{and} \quad (59)$$

SWEBY limiter (Superbee for $\alpha = 2$, minmod for $\alpha = 1$),

$$\phi(q) = \max[0, \min(\alpha q, 1), \min(q, \alpha)]. \quad (60)$$

In the cases shown hereafter, the SMART limiter was used.

This method needs an additional treatment related to the time discretization itself: an upper limit (in amplitude) must be set for equivalent diffusion coefficients, otherwise the time discretization itself may be unstable. In principle, the limit should be set on the time stepping, but this would penalize the overall model performance very heavily simply because in some occasions, the denominator used in the computation of the equivalent diffusion coefficient vanishes and anyway should cancel out when the flux is calculated. But since the equivalent diffusion coefficients are most naturally computed at the grid-box corner, they need to be averaged at the interfaces. There however, the gradients in the denominator of the equivalent diffusion do not cancel out with the local interface gradient, and for small gradients (where diffusion is anyway small) one has to limit the time step or diffusion coefficient.

It is also noted that from a practical programming point of view, contrary to the AMPMIN method, it is not easy to make the equivalent diffusion coefficient method into a linear method, because the computation of the equivalent diffusion is essentially nonlinear. On the other hand, the method can be easily included in any solver that allows variable horizontal and vertical diffusion coefficients (treated implicitly or not). Practically, one can apply the flux limiter approach to calculate the equivalent diffusion coefficient, since the gradients involved in the flux combination are the same for the two fluxes (lower and higher order), so that one can in fact just combine diffusion coefficients based on the limiter functions. These equivalent diffusion coefficients can then be used very conveniently in existing diffusion solvers (which should not enforce positive diffusion coefficients in the computer code).

The method presented here is based on a treatment that consists of analyzing two one-dimensional problems when it comes to the computation of the limiter functions. In advection problems this is the general procedure. As for advection, the rotated diffusion is also a directional process (1D problem along the velocity or the isopycnal direction), but there is a major difference: for the advection flux correction, when computing the limiters, the advection flux in the other direction will normally not introduce an important diffusion (because this is the goal of the method). This means that in the case of flux limiting in one direction one can assume that the other direction behaves correctly, but will not help in diffusing perturbations. In the case of isopycnal diffusion, however, when looking at the fluxes in one direction, limiting may in fact not be necessary, because the other direction could have diffusive fluxes that render the 2D system monotonic. Therefore, a pure 1D monotone scheme for negative diffusion is too strong (because we neglect the other direction and the feedbacks on diffusion coefficients). For fixed negative diffusion coefficient in one direction and zero diffusion in the other, the scheme is indeed not monotonic. We must thus expect that the twofold monotonie 1D problem will induce unnecessarily high diapycnal diffusion. Indeed, we may add some diffusion in one direction because we do not take into account the positive diffusion that may exist in the other direction. A promising alternative would thus be the use of truly 2D flux limiter approaches (e.g., Thuburn 1996), which however make the numerical scheme increasingly complicated (specially if the 3D generalization is thought of). In addition, truly 2D flux limiter schemes are generally designed for advec-tion schemes (e.g., Thuburn 1997; Drange and Bleck 1997), and their adaptation to diffusive fluxes, as done here in 1D, is not straightforward and some preliminary trials of generalizations were not conclusive. Furthermore, the proof of monotinicity must be two-dimensional and include the feedback onto the diffusion coefficients. For general implementation and further findings in 2D

limiters, the limiter should be parameterized in terms of points in a stencil rather than along a coordinate line since the problem is really two-dimensional.

Though no proof is presented to mathematically ensure that the nonlinear schemes AMPMIN and FLUX-CORR are monotonie, the heuristic explanation of their functioning is similar: Both methods feature flux limiters that use smaller-amplitude fluxes when "problems are expected": this arises when very different values of the different flux computations are encountered (AMPMIN) or rapidly changing gradients (FLUXCORR). In this case, both methods reduce the fluxes and to some extent "freeze" the situation. Typically this is the case when a "plateau" is present before a jump: at the jump, for negative diffusion values, no flux can be allowed from the plateau into the higher value if no fluxes are present in the second direction. So the plateau has a tendency to remain. Similarly, for a local peak and negative diffusion, fluxes must be limited, otherwise the local peak will increase its magnitude.

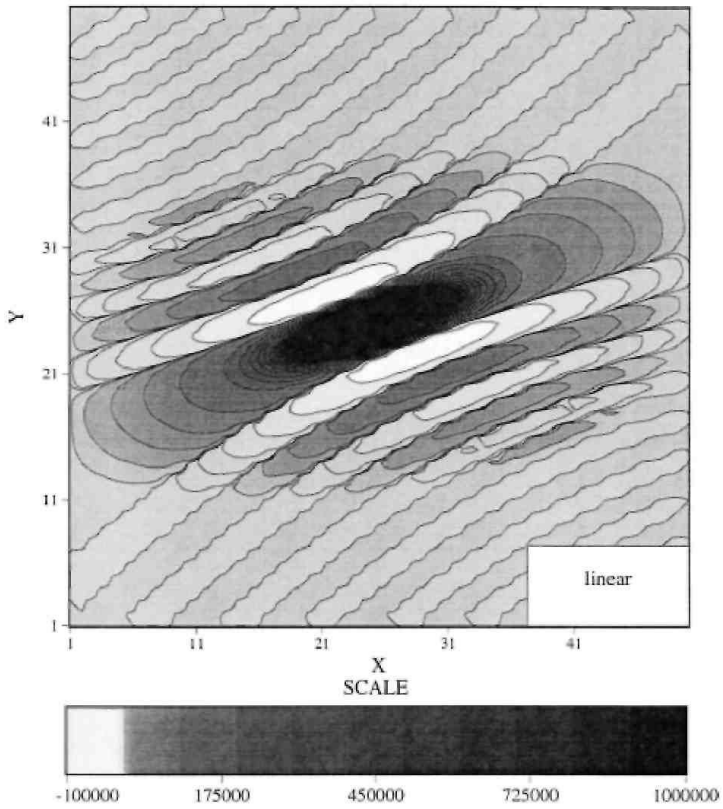
5. COMPARISON OF THE NEW SCHEMES

Since we already have several possible choices for nonstandard discretizations, we will now proceed to a comparison of their behavior with classical linear schemes and exact solutions.

a. Test cases and criteria

Because we have already presented the numerical stencils for the linear methods in the case of a constant slope, we will use this case as a first test for the methods presented here. Only those that are promising in this framework will be further examined in the case of variable slopes.

FIG. 9. *Linear scheme LINEAR. Diffusion of a Dirac signal along* $y = 0.4$ *x after 100 time steps. This standard scheme clearly introduces strong diapycnal diffusion and dispersion with significant undershootings (in white).*



1) LINEAR SLOPE CASE

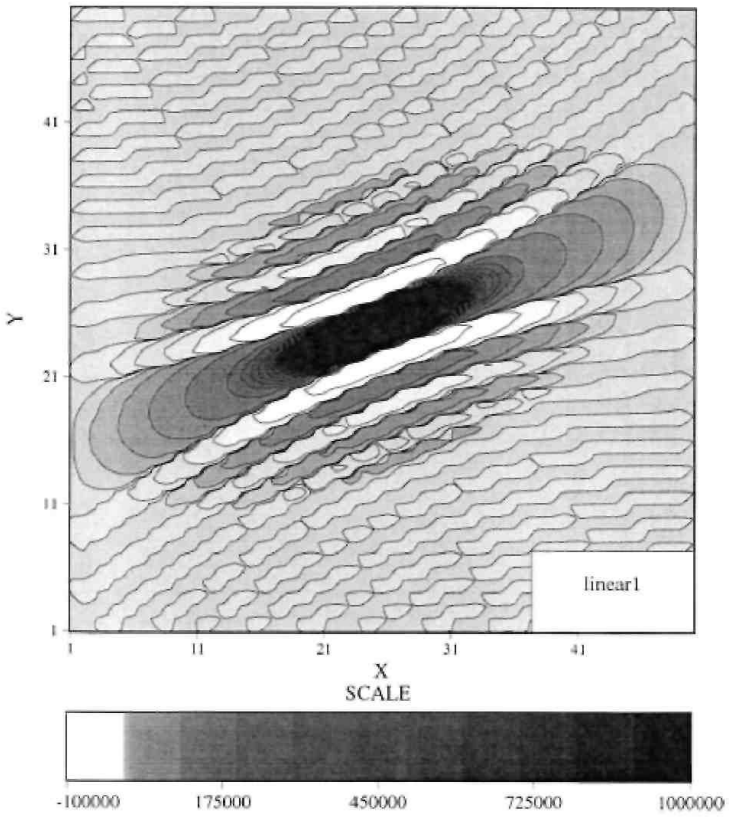
Qualitative information can be gained by diffusing the Dirac function in an uniform slope field with $r = 0.4$ and a time stepping with $\Delta t' = 0.1\,\Delta x^2$ for 100 time steps.

The analytical solution Ψ* for a quantity Q of the tracer initially present at the origin of s is given by

$$\Psi^* = \frac{Q}{\sqrt{4\pi \mathcal{A}' t}} \exp\left(-\frac{s^2}{4\mathcal{A}' t}\right). \qquad (61)$$

It is observed that all linear schemes produce strong diapycnal diffusion dispersion with significant undershootings (see Figs. 9-11). Compared to the standard scheme LINEAR, for the scheme LINEARl (see Fig. 10), the diapycnal diffusion is visible in a narrower band and oriented more "vertically," due to the choice of fluxes involved in the averaging (see Fig. 4). For LIN-EAR2 (Fig. 11) the solution is very similar to the standard solution in this case with a slightly reduced propagation of information into the diapycnal direction. It also seems that adding a 20% background diffusion (see Fig. 12), as it is done classically in GCM models using isopycnal diffusion, does not change the behavior and pattern of the linear scheme, even if the amplitude of the diapycnal dispersion and the associated negative values are slightly reduced. The inconsistent COMBI method (see Fig. 13) clearly shows a strong diapycnal mixing, but remains monotonic. Clearly the diapycnal dispersion of the linear schemes was replaced by a diapycnal diffusion ensuring a monotonic solution that propagates strongly into the diapycnal direction. The solution of the AMPMIN scheme (Fig. 14) is monotonie and presents less diapycnal diffusion than the COMBI method. The staircase pattern results from the nonlinear AMPMIN function and the constant zero concentration in the background. Finally, as in the AMPMIN solution, the solution of the flux-corrected method (see Fig. 15) remains monotonie and presents a propagation into the diapycnal direction that is reduced compared to the stan-dart scheme.

FIG. 10. *Modified linear scheme LINEAR1. Compared to the standard scheme LINEAR, the diapycnal diffusion is visible in a narrower band and oriented more "vertically," due to the choice of fluxes involved in the averaging.*



In Fig. 16, showing the evolution of the maximum value of Ψ in function of time, we can clearly see that the AMPMIN version retains the highest values but is below the exact solution after 100 time steps. There is thus a diapycnal mixing present in the AMPMIN version, but it appears smaller than for the other methods. The most diffusive method is the COMBI method. It should also be noted that the AMPMIN version underestimates the diffusion at the initial stage. Among the linear versions, LINEARl gives the best response in later stages, with LINEAR2 still better than the classical linear version. It is noteworthy that the additional background diffusion does not drastically change the behavior of the maximum. The flux-corrected method initially diffuses strongly,

but then slows down the diapycnal diffusion.

FIG. 11. *Second modified linear scheme LINEAR2. The solution is very similar to the standard solution in this case with a slightly reduced propagation of information into the diapycnal direction.*

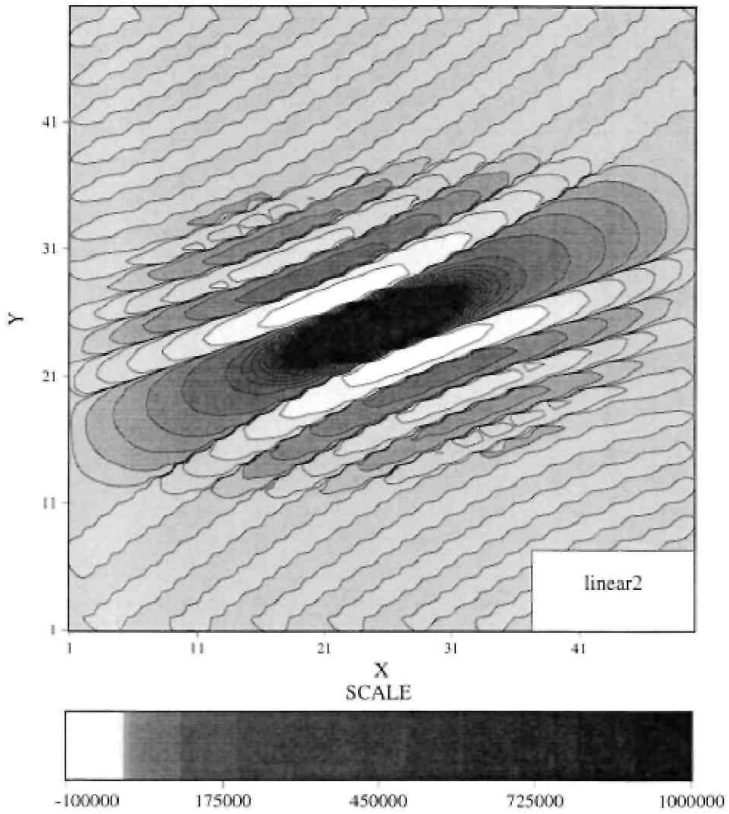
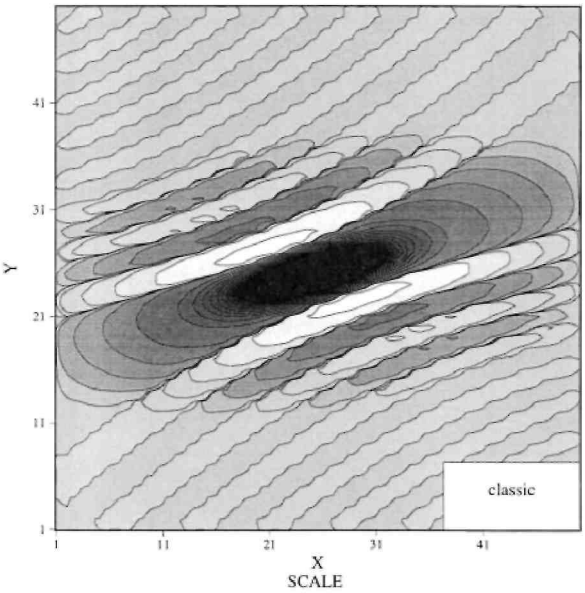


FIG. 12. *Linear discretization CLASSIC with additional 20% horizontal diffusion. This scheme reduces the amplitude of the diapycnal dispersion and the associated negative values, but the pattern is still comparable to the standart scheme LINEAR.*



Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 13. *Inconsistent combination of horizontal and diagonal diffusion COMBI. Clearly the diapycnal dispersion was replaced by a diapycnal diffusion ensuring a monotonic solution that propagates strongly in the diapycnal direction.*

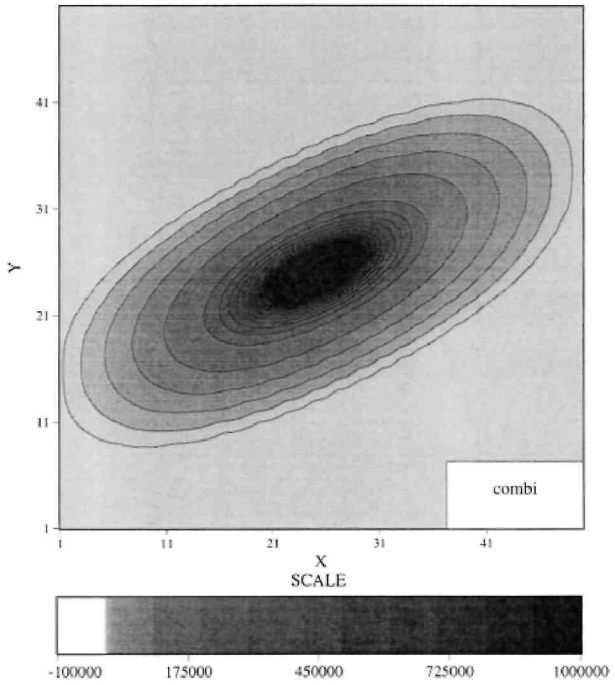
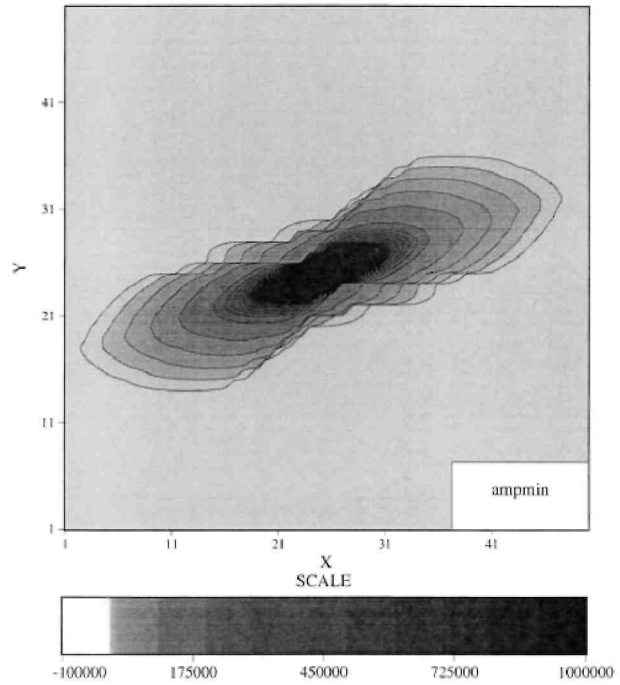


FIG. 14. *Nonlinear AMPMIN method. The solution is monotonic and presents less diapycnal diffusion than the COMBI method. The staircase pattern results from the nonlinear AMPMIN function and constant zero concentration in the background.*



Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 15. *Flux-corrected method FLUXCORR. The solution is monotonie and propagation into the diapycnal direction reduced compared to the standard scheme.*

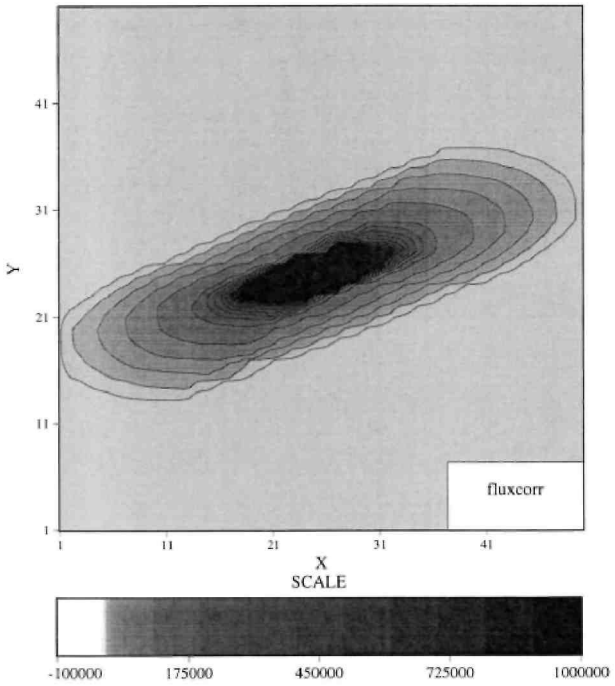


FIG. 16. *Evolution of the maximum value of the field in function of time (unnumbered curves in this figure and the following figures overlay the exact solution; the x axis corresponds to the discrete time step).*

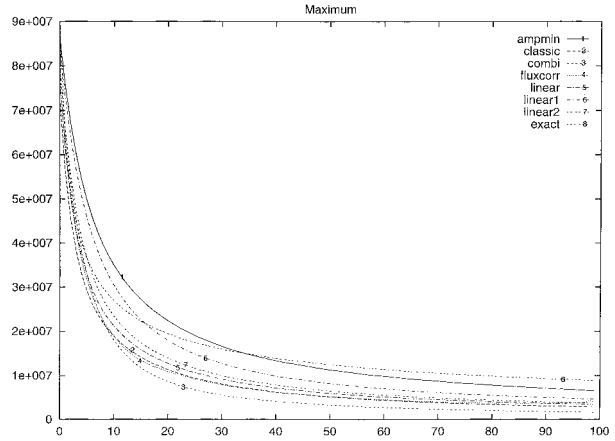
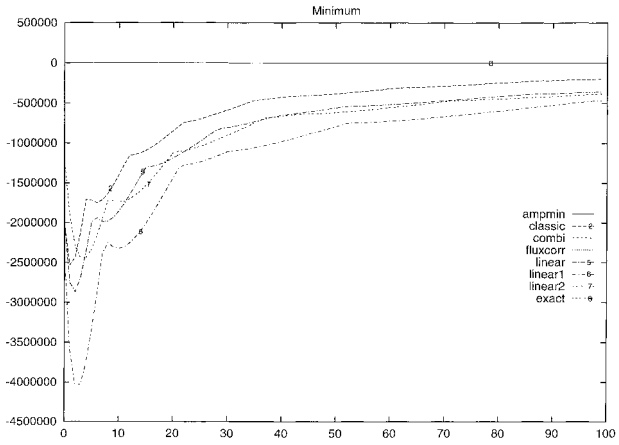


FIG. 17. *Evolution of the minimum value of the field. Due to the Dirac distribution, initially, undershootings are very important but tend to decrease when the field starts to be smoother. LINEARl produces the strongest undershootings and adding a horizontal diffusion (CLASSIC) only slightly reduces the undershooting.*



The evolution of the minimum value of Ψ in function of time (Fig. 17) shows that all the linear methods produce a strong undershooting, especially during the initial phase when the Dirac signal is present and leads to large gradients. Here method LINEAR1 gives the worst result, with the two other methods being similar. An additional background diffusion improves the behavior only slightly. Undershooting tends to decrease once the field is smoothed by the overall diffusion.

In order to quantify the accuracy of the methods, we now use two cost functions I_1 and I_2, which measure the difference between the theoretical solution and the numerical solution:

$$I_1 = \sum_i (\Psi_i - \Psi^*)^2 \quad \text{and} \qquad (62)$$

$$I_2 = \sum_i (\Psi_i - \Psi^{\text{extr}})^2 \gamma, \qquad (63)$$

where Ψ_i represents the computed fields, Ψ^* the analytical solution, Ψ^{extr} are the extrema of the real solution, and γ is a Heaviside function that is zero when the field Ψ_i remains between these extremal values. These cost functions give an idea of the overall truncation error (I_1) and the monotonicity (I_2). We see that the normalized errors in Table 1 give rms errors that are lowest for the linear methods and the AMPMIN method, followed closely by the flux-corrected method. Concerning undershootings, linear methods behave similarly, with an improvement due to the addition of the background horizontal diffusion. We also give a rough estimate[6] on the relative cost of each scheme compared to the classical linear scheme. If the rotation of a diffusion operator does not take an important CPU fraction of a general model, this indicates that all schemes are affordable.

TABLE 1. *Measures of the errors.*

Method	I_1	I_2	Relative cost
AMPMIN	0.53	0.	1.2
CLASSIC	0.69	4.06	1
COMBI	0.81	0	1
FLUXCORR	0.65	0	2
LINEAR	0.65	7.03	1
LINEARl	0.51	8.44	1
LINEAR2	0.61	7.24	1

A way to measure the diffusion of the numerical method is the computation of

[6] The actual value in a GCM will depend on the compiler, the hardware, and the organization of the code. In addition, some of the implementations of the linear schemes were directly based on the assumption of a constant slope, which led to some simplifications in the code.

$$X = \int_D \xi^2 \Psi \, dD \quad \text{and} \tag{64}$$

$$Y = \int_D \eta^2 \Psi \, dD, \tag{65}$$

where D is the total domain. Both X and Y increase linearly in time and proportionally to the diffusion coefficient for the case of the diffusion of a point release whose solution was given in Eq. (61). This means that the quantities

$$x = \frac{2X}{Q\mathcal{A}'t} \quad \text{and} \tag{66}$$

$$y = \frac{2Y}{Qr^2\mathcal{A}'t} \tag{67}$$

should be equal to one for the numerical solution. Also the ratio $p = (X/Y)/r^2$ should be constant and equal to one, as can be seen by the ratio x/y, which is one for the analytical solution. The first two parameters measure the effective diffusion in the two grid directions, whereas the last parameter measures the actual slope on which the method diffuses. If the parameter is lower, this means that the diffusion is too steep, whereas a higher value means that the diffusion is too horizontal.

In Fig. 18, it can be seen that all the linear methods produce a correct average measure of horizontal diffusion. This is because the undershooting and overshooting cancel out when performing the integral. For the CLASSIC scheme, the additional background diffusion clearly displays a 20% increase in effective diffusion, while AMPMIN has considerably reduced the effective diffusion because of its systematic choice of minimal amplitude fluxes (or zero fluxes when fluxes have opposite signs). The flux-corrected method only slightly decreases the effective diffusion. In Fig. 19, it is demonstrated that the linear methods all produce a correct average measure of diffusion in the vertical direction. The additional background diffusion does not show up here for the CLASSIC scheme, since it was added only in the horizontal direction and the scheme is linear. The COMBI method shows a very strong increase in effective diffusion, which is consistent with the interpretation of the stencil given beforehand (Fig. 7). The AMPMIN version has again reduced effective diffusion because of its AMPMIN function. The flux-corrected method only slightly increases effective diffusion. Figure 20 shows that the linear methods all produce a correct average measure of direction of diffusion. The additional background diffusion in CLASSIC leads to a weakening of the effective slope, consistent with Figs. 18 and 19. The other methods increase the slope, with the worst effect in the COMB I and the AMPMIN method.

FIG. 18. *Evolution of the measure of the horizontal diffusion. AMP-MIN clearly reduces diffusion while increasing the horizontal background diffusion by 20% in CLASSIC can clearly be seen on this integral quantity. All linear schemes have the correct average horizontal diffusion while FLUXCORR slightly reduces it.*

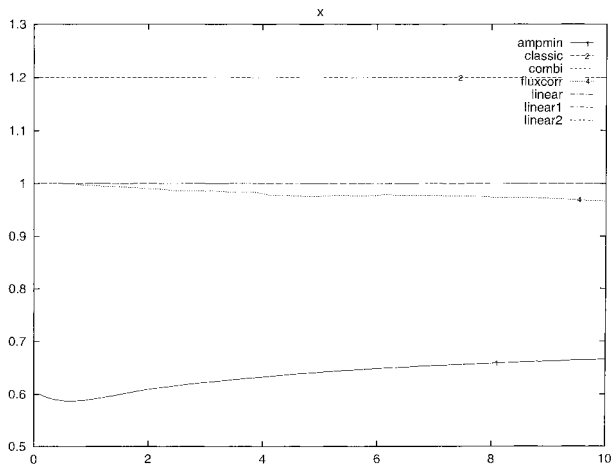
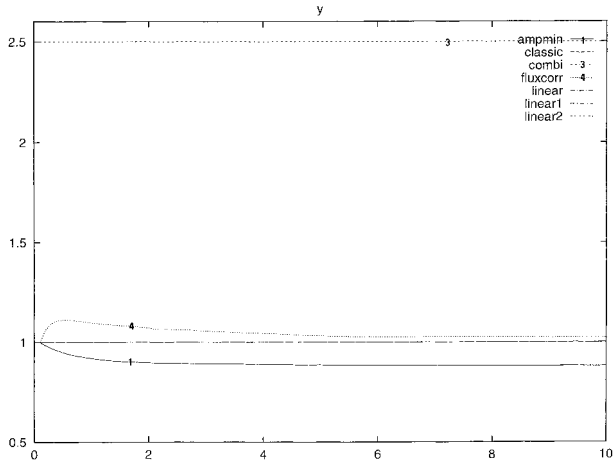


FIG. 19. *Evolution of the measure of the vertical diffusion. As for the horizontal diffusion, AMPMIN reduces the diffusion, and the linear schemes behave correctly. The FLUXCORR method slightly increases the vertical diffusion while the inconsistent COMBI method shows the strong additional vertical diffusion consistent with the analysis of the numerical stencil.*



Another integrated way is to look at the integral of the field in the η direction and show the distribution in the function of ξ after 100 time steps. In principle, this should be exactly the same solution as obtained by a pure horizontal diffusion. An inspection of Fig. 21 shows that the linear methods all produce a "correct" answer, because the discrete summing of the linear schemes in the y direction leads to a discrete equation of the vertical average that is quite simply a horizontal diffusion. For 100 small time steps, the numerical solution to this equation is indistinguishable from the exact solution. The additional background diffusion in CLASSIC leads, as one should expect, to an overdiffusion. The flux-corrected method slightly underestimates the diffusion, while AMPMIN not only strongly underestimates diffusion but also leads to a very particular and unrealistic shape of the diffused field. Though the integral measure seems to indicate that the linear methods are not so bad, the patterns show the problems they can induce. These patterns also show that simple linear integrals are not sufficient to characterize a field, since the wiggles' contributions of the dispersion may cancel in the integrations (unless quadratic measures as in I_2 are used).

In order to verify that the methods also behave correctly with slopes that exceed the aspect ratio of the grid, additional tests with $|r| \geq 1$ were performed (but are not shown here). Both the AMPMIN and FLUX-CORR method behaved similarly to the case presented above, so that the adaptation from the original method of Stelling and van Kester (1994) to a local stencil was successful. After the constant slope case, we will finally test the more promising methods in more general situations.

FIG. 20. *Evolution of the measure of the direction of the diffusion. Consistently with the analysis of the horizontal and vertical diffusion, the average direction of diffusion is correct for the linear schemes, too flat for the CLASSIC scheme and slightly more upward oriented for the FLUXCORR scheme. AMPMIN and COMBI modify the direction very strongly.*

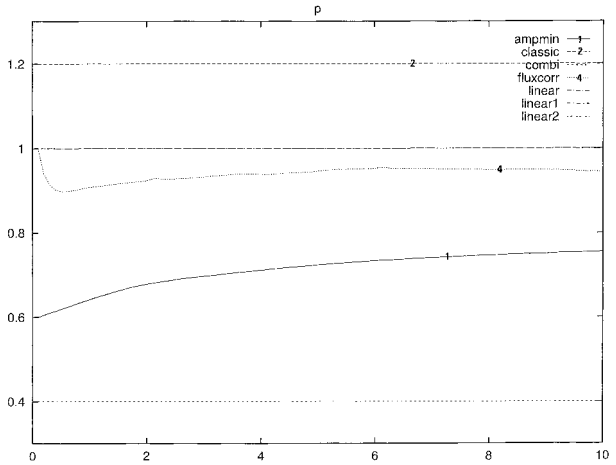


FIG. 21. *Integral on η in the function of the grid point; the exact solution is indistinguishable from the linear solutions. AMPMIN not only shows the reduced diffusion as before but also the effect of the inconsistent flux calculations, which allow an unphysical peak in the integrated field to remain. FLUXCORR has a slightly higher peak than the exact solution, while the CLASSIC method reduces the peak.*

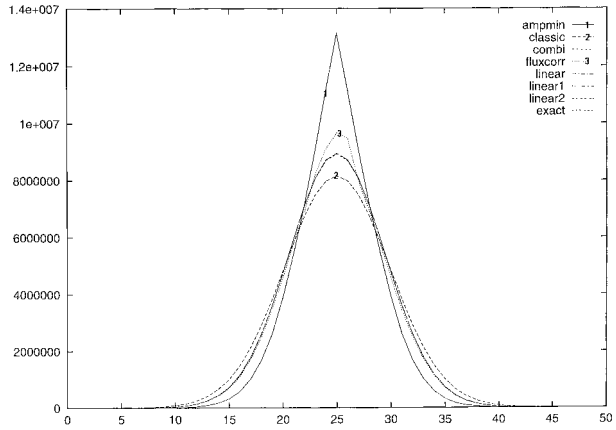
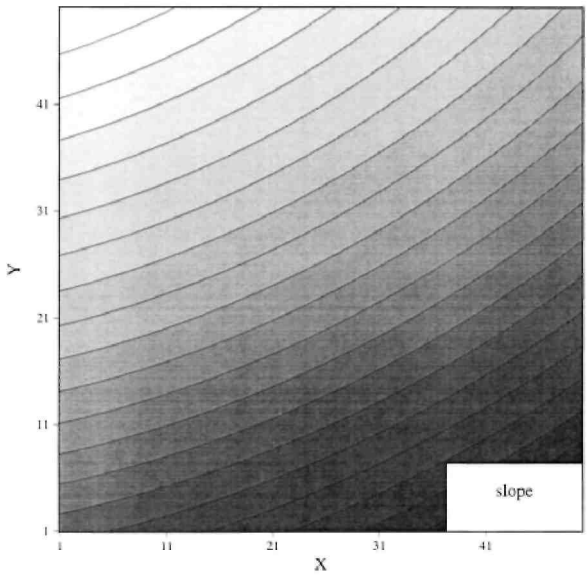


FIG. 22. *Density field used for the variable slope experiment.*



2) VARIABLE SLOPE

In order to verify that the two monotonic nonlinear methods also behave correctly in a situation where the slope is not constant, we performed a simulation in which the lines on which diffusion takes place are curved as shown in Fig. 22

In this situation, both the FLUXCORR and the AMPMIN method (Fig. 23 and 24) diffuse the signal on these lines by bending the patch along the isoline. Both methods remain monotonic, except for very small negative values in the AMPMIN method, which are however in the range of CPU rounding errors and several orders of magnitude lower than the undershooting induced by linear methods.

b. Discussion

The flux-corrected method seems to have some significant diapycnal diffusion, but part of it is due to the fact that the nonlinear scheme needs several points to keep the information contained in a narrow band. Typically, after 100 iterations, the diapycnal signal is spread over seven points containing a signal that is larger than 1% of the central signal. When running the diffusion 10 times longer (Fig. 25) this spreading is only on 10 points, indicating that the initial diapycnal diffusion was necessary to create a large-scale signal. This is very different

from the linear methods: as seen in Fig. 26, the linear scheme has dispersed farther into the diapycnal direction and spreads over 21 points (again for a 1% threshold), while the flux-corrected scheme somehow stabilizes. This indicates that for larger-scale signals than the Dirac function, the flux-corrected method would exhibit less diapycnal mixing. This is also confirmed for the variable slope case integrated over 1000 time steps (see Fig. 27).

FIG. 23. *AMPMIN method for the variable slope experiment showing that the diffused field follows the density field by bending the patch along the isolines.*

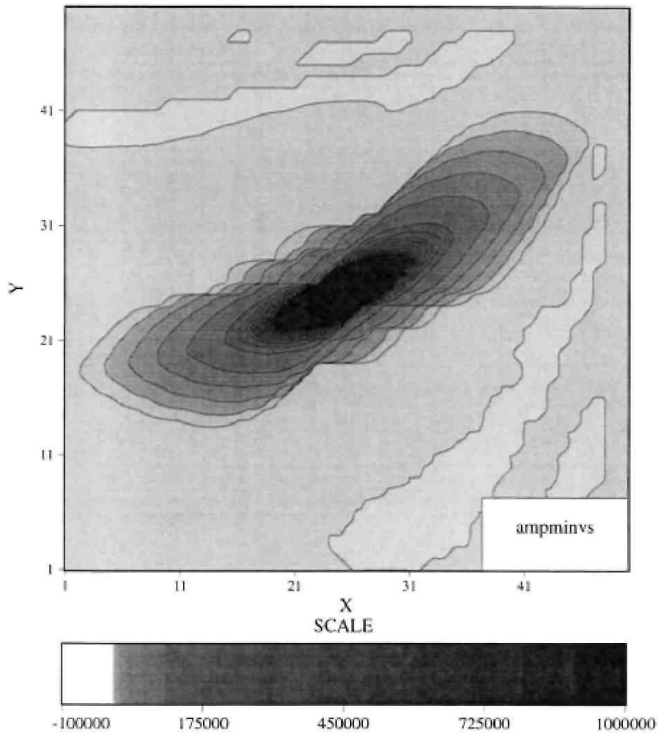
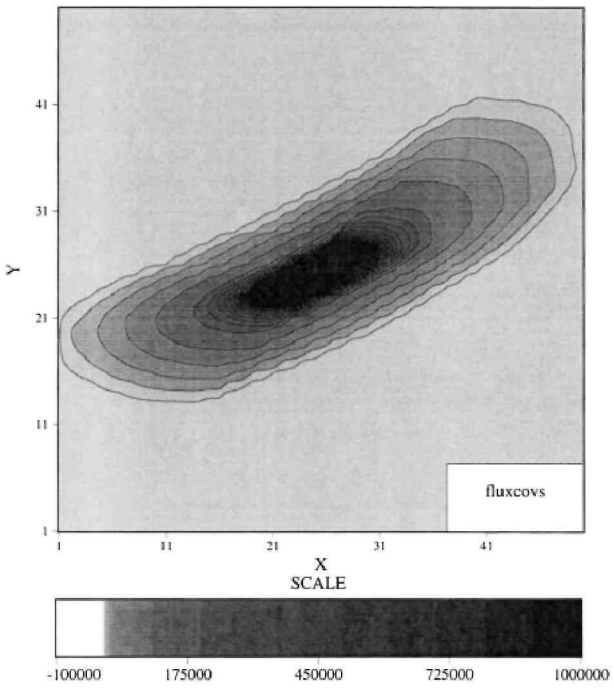
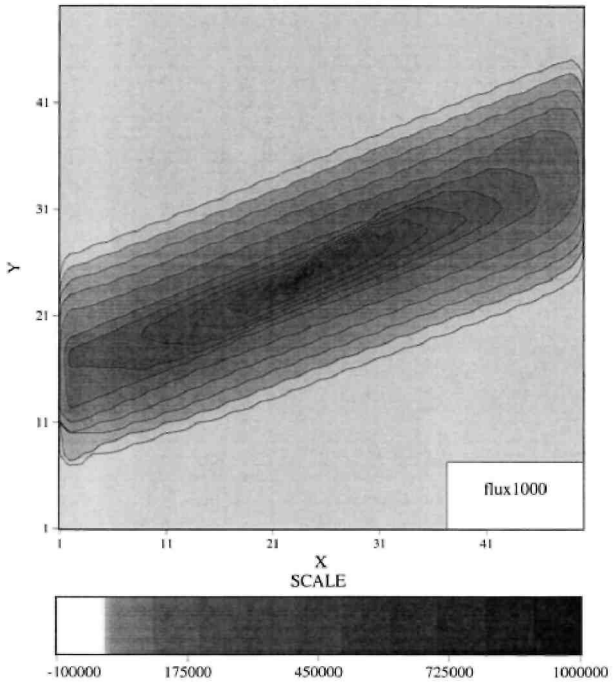


FIG. 24. *FLUXCORR method for the variable slope experiment. As for AMPMIN, the solution is following the density field.*



Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 25. *Flux-corrected method after 1000 time steps, constant-slope case. The method keeps the solution in a rather narrow band only slightly larger than after 100 time steps.*



Presently, if we need to ensure a monotonic scheme, we suggest the use of the flux-corrected method. The COMB I method clearly exhibits too strong a diapycnal diffusion, whereas the generalization of the the AMP-MIN method suggested by Stelling and van Kester (1994) leads to strange patterns in the diffused fields as staircase patterns and non-Gaussian integrals. The inconsistent discretization in this case leads to a numerical solution that is inconsistent with the mathematical properties of diffusion. Since the AMPMIN method is also not easily implemented into vertically implicit schemes (the classical ocean model approach), the flux-corrected method seems to be the indicated choice. However, when such a nonlinear scheme is used, we will face a major problem when diffusing temperature and salinity. Since the nonlinear scheme may behave differently for those two fields, there is no way to ensure that the combined contribution of each of these fields to density fluxes along isopycnal surfaces is nil. This may be acceptable, if the resulting modification in the pressure field does not lead to an unstable coupling between momentum and density. In the case of the classical linear scheme, the incomplete compensation was shown by Griffies et al. (1998) to be the major source of problems in the isopycnal diffusion of the Modular Ocean Model. It is probable, though not certain, that the problem remains also for the nonlinear scheme, unless the instability mechanism is suppressed by avoiding the introduction of new extrema in the *T, S* fields. If, despite the monotonicity, the presence of density fluxes on isopyc-nals still leads to instabilities, certain tricks may perhaps allow us to achieve this compensation even in nonlinear schemes. A simpler approach could thus be to use a linear scheme as in Griffies et al. (1998) (maybe with a more sophisticated averaging technique than the one described in Fig. 9) for dynamically active tracers only, while dynamically passive tracers can be solved by using the new nonlinear schemes. This means that errors for temperature and salinity fields are controlled so that they do not have an influence on isopycnals, the only dynamically important feature. For dynamically passive tracers the use of our monotonic scheme (particularly for biological models or turbulent variables) thus eliminates most problems associated with the rotated diffusion. For further improvements of our scheme, a way to decrease the diapycnal diffusion in the flux-corrected method would be to use a truly 2D flux correction scheme for nonadvective problems. This is, in our opinion, the direction for further research on discretization, rather than searching for larger stencil methods for example; a linear model is not expected to behave better by using more points, since the risk of introducing other negative coefficients increases.

Published in: Monthly Weather Review (2000), vol.128, iss. 8, pp. 2711-2733
Status: Postprint (Author's version)

FIG. 26. *Linear scheme after 1000 steps, constant-slope case. Here the solution has propagated into the overall domain.*

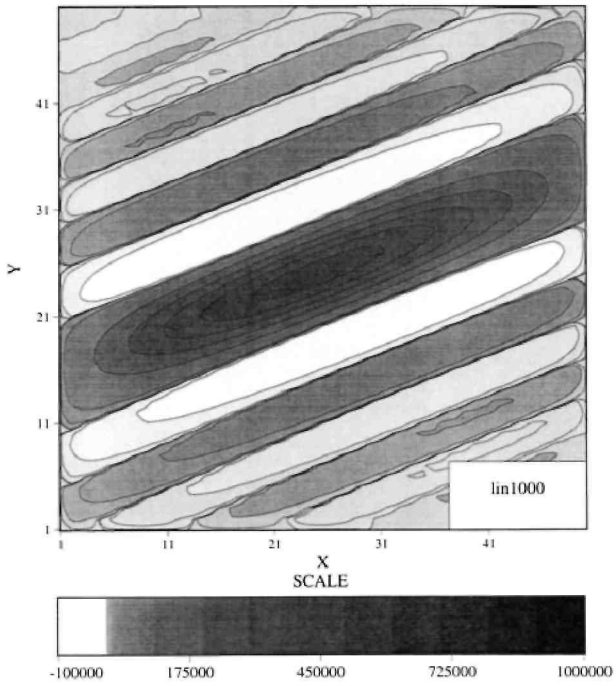
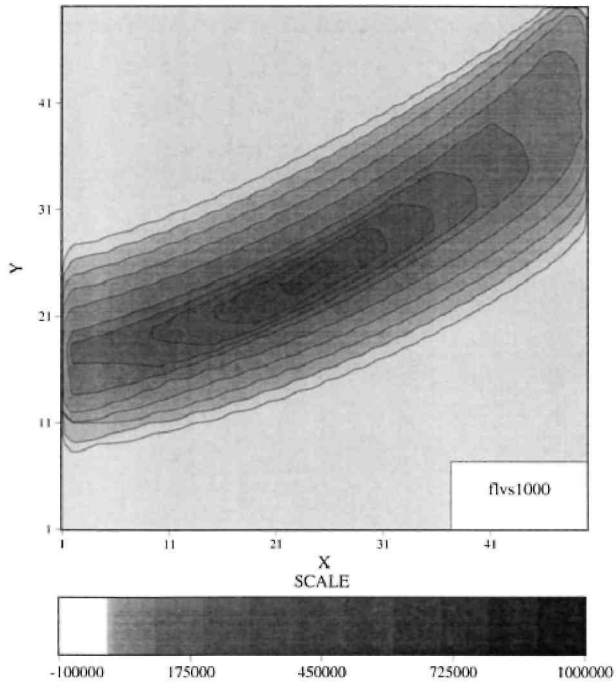


FIG. 27. *Flux-corrected method after 1000 time steps in a variable slope situation. As for the constant-slope case, the solution is maintained in a narrower band.*



6. CONCLUSIONS

In our opinion, the design of a discretized rotated diffusion operator should have the following properties.

- Be conservative;
- satisfy the monotonicity principle;
- extend only over a nine-point stencil [for the (x, z) case] ;
- reduce to the classical horizontal stencil 1, -2,1 when no transformation is present;
- be computationally efficient, since diffusion should be small anyway, and a scheme should not be penalized by a second order term; and
- introduce the smallest possible diapycnal diffusion.

Consistency could be another basic requirement, but we consider it less important than the other properties because the isopycnal diffusion is a parameterization that should disappear when grid spacing decreases. Furthermore, inconsistent schemes are common and sometimes behave even better than consistent schemes (e.g., Cockburn et al. 1999). Nevertheless, if an inconsistent scheme is retained, one should carefully analyze its behavior (see the AMPMIN and COMBI schemes).

We have shown that even when allowing for non-consistent schemes, unfortunately, no scheme analyzed here satisfies all these requirements perfectly.

All linear (and inexpensive) schemes produce diapycnal dispersion and over- and undershootings. As a linear scheme that performs well in situations with strong diapycnal mixing, we presented a slightly modified linear method (LINEARl) that may be sufficient in some cases where monotonicity is not a strong requirement. For linear methods, we showed that adding horizontal background diffusion for purely numerical reasons is not recommended.

We presented a number of new methods to deal with the monotonicity violation of linear schemes. These schemes are either nonlinear, inconsistent, or both nonlinear and inconsistent. Only two schemes are candidates for implementation into general models: the AMPMIN method keeps diapycnal mixing at a lower value, but leads to some strange patterns because of its inconsistent discretization and diffusion, while the FLUX-CORR method has a satisfactory behavior in all the cases we tested, though one might still search for a scheme with even less diapycnal diffusion.

Presently, if we need to ensure a monotonic scheme, we suggest thus the use of the flux-corrected method at least for dynamically passive tracers. For temperature and salinity, it is not clear if the instability mechanism identified in Griffies et al. (1998) is suppressed by the monotonic schemes. In case of the lack of compensation of temperature and salinity fluxes on isopycnals still leading to instabilities, the use of a linear scheme as in Griffies et al. (1998) eliminates the problem for dynamically active tracers.

Though we investigated a large range of methods, other remedies could still be analyzed, but our feeling is that we already reached a level of complexity that is surprisingly high for a diffusive process introduced for parameterization reasons. This shows also that if the isopycnal diffusion is only introduced for numerical reasons, in order to eliminate grid noise, then the use of the rotated operator does not yield an efficient numerical filter. In the case of a complex model in which some filtering of the grid noise is required, one should use a space-selective filter in the numerical grid or, even better, numerical schemes (generally for advection) that do not generate grid noise. Only if the rotated physical parameterization supersedes the numerical filtering requirements should the rotation be discretized. In this case, we proposed monotonie schemes to deal with the monotonicty problem for the given coordinate system. Another way of decreasing the problem of the operators rotation is to design the coordinate system so that it follows closely the direction of interest. A pure iso-pycnal layer model is an example of this, but some hybrid models can adapt their vertical grid positions continuously to density gradients in order to reduce relative slopes.

Finally, we should mention that throughout the whole analysis, we supposed that the mixing paramaterization is adequately performed by a Laplacian formulation, which is the standard version used in isopycnal diffusion parameterizations. This supposes relatively coarse-resolution models, since for higher resolutions, more scale-selective biharmonic diffusions are generally retained. In this case however, even for unrotated operators, the classical scheme is not monotonic, not to mention the problems related to the rotation. One might ask whether

this might be a source of problems in coupled biological models in the future, but since high-resolution models not only use more scale-selective but also much lower values for the diffusion, this may be hidden by other processes involved.

Acknowledgments

A model intercomparison was supported by the European Union (MEDMEX Contract MAS2-CT94-0107 and MATER Contract MAS3-CT96-0051) and led to some questions addressed in the present paper. The development of the numerical code was part of the MEDNET project (MAS3-CT98-0189). E. De-leersnijder and P. P. Mathieu work in the framework of Contract GC/DD/09A with Belgium's Federal Office for Scientific, Technical and Cultural Affairs and Convention N 92/02-208 d'Actions de Recherches Concertées with the Communauté Francaise de Belgique.

The National Fund for Scientific Research (Belgium) has given J. M. Beckers and E. Deleersnijder the opportunity to be active in oceanographic research. P. P. Mathieu was a fellow of Fonds pour la Formation à la Recherche dans l'Industrie et l'Agriculture during this research.

Anand Gnanadesikan and the anonymous reviewers made some helpful comments on the numerical aspects.

Appendix

General Considerations on Consistency

In this appendix we analyze the consistency of some of the numerical schemes, showing that the FLUX-CORR scheme is consistent in contrast to the AMPMIN scheme. In order to do so, we will develop the truncation error of the diffusion operator in terms of the truncation error of the fluxes.

A common way to design discretizations of a divergence such as term *D,*

$$D = -\frac{\partial \mathcal{F}}{\partial x} = -\frac{\mathcal{F}\left(x + \frac{\Delta x}{2}\right) - \mathcal{F}\left(x - \frac{\Delta x}{2}\right)}{\Delta x} + \frac{\Delta x^2}{24}\frac{\partial^3 \mathcal{F}}{\partial x^3}, \quad \text{(A1)}$$

is to calculate consistent discretizations of the flux $\mathcal{F}$ at the interface of a finite volume and then to proceed to a volume integration of the divergence, resulting in the discretized form

$$\tilde{D}\Delta x = -\tilde{\mathcal{F}}\left(x + \frac{\Delta x}{2}\right) + \tilde{\mathcal{F}}\left(x - \frac{\Delta x}{2}\right), \quad \text{(A2)}$$

where $\tilde{}$ stands for the discretized version of the analytical expression. In this case, the truncation error of the scheme is $\tilde{D} - D$ and includes truncation errors in the fluxes and truncation errors due to the flux differencing. Assuming that we use a *n*th-order scheme for the calculation of the fluxes, we have a truncation error of the flux $\mathcal{F} - \tilde{\mathcal{F}}$ defined by

$$\tilde{\mathcal{F}}\left(x + \frac{\Delta x}{2}\right) = \mathcal{F}\left(x + \frac{\Delta x}{2}\right) + \Delta x^n a\left(x + \frac{\Delta x}{2}\right), \quad \text{(A3)}$$

where the function *a* depends on the mathematical formulation of the flux and its discretization. Typically, it consists of some derivatives of the fields on which the flux operator acts.

We can then calculate the truncation error to be

$$D - \tilde{D} = \frac{\Delta x^2}{24}\frac{\partial^3 \mathcal{F}}{\partial x^3} + \Delta x^n \frac{a\left(x + \frac{\Delta x}{2}\right) - a\left(x - \frac{\Delta x}{2}\right)}{\Delta x}. \quad \text{(A4)}$$

This shows that a high-order flux calculation does not necessarily lead to a high-order scheme, since the flux differencing performed here leads to a second-order scheme. This phenomenon, known as degeneracy, is generally discussed when analyzing discretizations on nonuniform grids (e.g., Hoffman 1982) but also applies to nonlinear conservation laws. When a first-order scheme is used for flux calculation, even an inconsistent scheme may result if the function a is not differentiable.

This is the case for the AMPMIN scheme, where the function used to choose the flux involved at the interface is nondifferentiable (since different triads may be involved at the left and right interfaces, the truncation error reads differently on both sides, which in general cannot ensure that the differencing of a converges for $\Delta x \to 0$).

For linear schemes, function a is a differentiable function depending on the slope r and a second derivative of the field that is diffused.

In the case of the FLUXCORR method, the situation is more complicated, since the flux is evaluated by using the equivalent diffusion coefficient modified by the flux limiter function.

First, we show that for $\Delta x \to 0$ the parameter q used in the flux limitation for the flux along the x axis tends in a continuous way to 1 :

$$q_{i+1/2} = \frac{\Psi_{i+1-m} - \Psi_{i-m}}{\Psi_{i+1} - \Psi_i}, \quad \text{(A5)}$$

where m depends on the sign of the high-order local flux. One can develop the solution Ψ in a Taylor series around i, for example, by assuming that the solution is sufficiently smooth (this might not be easily satisfied in regions near the thermocline, since the axes may cross this discontinuity, but all truncation errors are affected by this and our analysis here is therefore consistent with classical analyses):

$$q_{i+1/2} = \frac{2\frac{\partial \Psi}{\partial x} + (1 - 2m)\Delta x \frac{\partial^2 \Psi}{\partial x^2} + O(\Delta x^2)}{2\frac{\partial \Psi}{\partial x} + \Delta x \frac{\partial^2 \Psi}{\partial x^2} + O(\Delta x^2)}, \quad \text{(A6)}$$

where the derivatives are taken at i.

This means that for very small grid sizes, no flux limiting is used anymore, except at extrema, but these occupy a smaller fraction of the domain as resolution increases, since for $q = 1$, all currently used flux limiter functions return the high-order flux, which in our case means that no additional diffusion is added.

Furthermore, for small grid sizes, the equivalent diffusion coefficient clearly converges toward its mathematical counterpart at least as fast as Δx.

This means that the error on the discrete fluxes is

$$\tilde{\mathcal{F}} - \mathcal{F} = \mathcal{A}^e \frac{\Delta x}{24}\frac{\partial^3 \phi}{\partial x^3}, \quad \text{(A7)}$$

which is differentiable and does lead to a consistent scheme.

Recent developments on consistency and convergence can also be found in Cockburn et al.'s (1999) recent paper. Beside general aspects of consistency and convergence of nonlinear conservation laws, they show that in some situations loss of consistency causes su-praconvergence (a cancellation of the loss of consistency by an increased

stability), provided that a conservative flux formulation with consistent fluxes is used in situations with a sufficiently smooth exact solution (in this case the AMPMIN method could be tested again).

References

Anderson, J., 1995: *Computational Fluid Dynamics.* Mechanical Engineering Series, McGraw-Hill, 547 pp.

Aris, R., 1962: *Vectors, Tensors and the Basic Equations of Fluid Mechanics.* Dover, 286 pp.

Beckers, J.-M., 1991. Application of a 3D model to the Western Mediterranean. *J. Mar. Syst.,* 1, 315-332.

------, H. Burchard, J.-M. Campin, E. Deleersnijder, and R-P. Mathieu, 1998: Another reason why simple discretizations of rotated diffusion operators cause problems in ocean models: Comments on "Isoneutral diffusion in a *z*-coordinate ocean model." *J. Phys. Oceanogr.,* 28, 1552-1559.

Bleck, R., and D. Boudra, 1981: Initial testing of a numerical ocean circulation model using a hybrid (quasi-isopycnic) vertical coordinate. *J. Phys. Oceanogr.,* 11, 755-770.

Blumberg, A. E, and G. L. Mellor, 1987: A description of a three dimensional coastal ocean model. *Three-Dimensional Shelf Models,* N. Heaps, Ed., Coastal Estuarine Science, Vol. 5, Amer Geophys. Union, 1-16.

Chassignet, E., L. Smith, R. Bleck, and F. Bryan, 1996: A model comparison: Numerical simulations of the north and equatorial Atlantic oceanic circulation in depth and isopycnic coordinates. *J. Phys. Oceanogr.,* 26, 1849-1867.

Cockburn, B., P.-A. Gremaud, and J. Yang-Xiangrong, 1999: A priori error estimates for numerical methods for scalar conservation laws: Part III: Multidimensional flux-splitting monotone schemes on non-Cartesian grids. *SIAM J. Numer. Anal,* 35, 1775-1803.

Cox, M. D., 1987: Isopycnal diffusion in a *z*-coordinate ocean model. *Ocean Modelling* (unpublished manuscripts), 74, 1-5.

Danabasoglu, G., J. McWilliams, and P. Gent, 1994: The role of mesoscale tracer transports in the global ocean circulation. *Science,* 264, 1123-1126.

Deleersnijder, E., and J.-M. Beckers, 1992: On the use of the σ-coordinate system in regions of large bathymetric variations. *J. Mar. Syst.,* 3, 381-390.

Drange, H., and R. Bleck, 1997: Multidimensional forward-in-time and upstream-in-space-based finite differencing for fluids. *Mon. Wea. Rev.,* 125, 616-630.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.,* 20, 150-155.

------, J. Willebrand, T. McDougall, and J. McWilliams, 1995: Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.,* 25, 463-474.

Gerdes, R., 1993: A primitive equation ocean circulation model using a general vertical coordinate transformation. 1: Description and testing of the model. *J. Geophys. Res.,* 98, 14 683-14 701.

------, C. Köberle, and J. Willebrand, 1991: The influence of numerical advection schemes on the results of ocean general circulation models. *Climate Dyn.,* 5, 211-226.

Gnanadesikan, A., 1999: Numerical issues for coupling biological models with isopycnal mixing schemes. *Ocean Modelling* (unpublished manuscripts), 1, 1-16.

Gough, W., 1997: Isopycnal mixing and convective adjustment in an ocean general circulation model. *Atmos.—Ocean,* 35, 495-511.

------, and W. Welch, 1994: Parameter space exploration of an ocean general circulation model using an isopycnal mixing parama-terization. *J. Mar. Res.,* 52, 773-796.

Griffies, S. M., 1998: The Gent-McWilliams skew-flux. *J. Phys. Oceanogr.,* 28, 831-841.

------, A. Gnanadesikan, R. Pacanowski, V. Larichev, J. Dukowicz, and R. Smith, 1998: Isoneutral diffusion in a *z*-coordinate ocean model. *J. Phys. Oceanogr.,* 28, 805-830.

Haney, R. L., 1991: On the pressure gradient force over steep topography in sigma coordinate ocean models. *J. Phys. Oceanogr.,* 21, 610-619.

Harvey, L. D. D., 1995: Impact of isopycnal diffusion on heat fluxes and the transient response of a two-dimensional ocean model. *J. Phys. Oceanogr.,* 25, 2166-2176.

Hedström, K., 1994: User's manual for a semi-spectral primitive equation ocean circulation model. Institute of Marine and Coastal Sciences, Rutgers University.

Hoffman, J., 1982: Relationship between the truncation error of centered finite difference approximations on uniform and nonuniform meshes. *J. Comput. Phys.,* 46, 469-474.

Mathieu, P., and E. Deleersnijder, 1997: What's wrong with isopycnal diffusion in World Ocean models? *Appl. Math. Model,* 22, 367-378.

------,------, and J.-M. Beckers, 1999: Accuracy and stability of the discretised isopycnal mixing equation. *Appl. Math. Lett.,* 12, 81-88.

McDougall, T. J., 1984: The relative role of diapycnal and isopycnal mixing on subsurface water mass conversion. *J. Phys. Ocean-ogr.,* 14, 1577-1589.

------, and J. A. Church, 1986: Pitfalls with the numerical representation of isopycnal and diapycnal mixing. *J. Phys. Oceanogr.,* 16, 196-199.

Mellor, G., and A. Blumberg, 1985: Modeling vertical and horizontal diffusivities with the sigma coordinate system. *Mon. Wea. Rev.,* 113, 1379-1383.

Oberhuber, J., 1993: Simulation of the Atlantic circulation with a coupled sea ice-mixed layer-isopycnal general circulation model. Part I: Model description. *J. Phys. Oceanogr.,* 23, 808-829.

Paiva, A., J. Hargrove, E. Chassignet, and R. Bleck, 1999: Turbulent behavior of a fine mesh (1/12°) numerical simulation of the North Atlantic. *J. Mar. Syst,* 21, 307-320.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.,* 12, 1154-1158.

Roberts, M., and D. Marshall, 1998: Do we require adiabatic dissipation schemes in eddy-resolving ocean models? J. *Phys. Oceanogr.,* 28, 2050-2063.

Stelling, G. S., and T. van Kester, 1994: On the approximation of horizontal gradients is σ cooordinates for bathymetry with steep bottom slopes. *Int. J. Numer. Methods Fluids,* 18, 915-935.

Thuburn, J., 1996: Multidimensional flux-limited advection schemes. *J. Comput. Phys.,* 123, 74-83.

------, 1997: TVD schemes, positive schemes, and the universal limiter. *Mon. Wea. Rev.,* 125, 1990-1993.

Visbeck, M., J. Marshall, T. Haine, and M. Spall, 1997: On the specification of eddy transfer coefficients in coarse-resolution ocean circulation models. *J. Phys. Oceanogr.,* 27, 381-402.

Zijlema, M., 1996: *Computational Modeling of Turbulent Flow in General Domains.* Cip Data Koninklijke Bibliotheek, 186 pp.