INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 24, 987–1017 (1997)

THE PRESSURE GRADIENT FORCE IN SIGMA-CO-ORDINATE OCEAN MODELS

LEIV HÅVARD SLØRDAL[1]*

[1]*Department of Geophysics, University of Oslo, PO Box 1022 Blindern, N-0315 Oslo, Norway*

SUMMARY

The error in computing the horizontal pressure gradient force near steep topography is investigated in a primitive equation, σ-co-ordinate, numerical ocean model (Blumberg and Mellor, in *Three-Dimensional Coastal Ocean Models*, Vol. 4, *American Geophysical Union*, Washington D.C., 1987, pp. 1–16). By performing simple test experiments where the density field is allowed to vary in both the vertical and the horizontal direction, severe errors are detected in the areas where the isopycnals hit the sloping bottom. An alternative method of computing the pressure force (Stelling and van Kester, *Int. j. numer. methods fluids*, **18**, 915–935 (1994) is adopted, resulting in substantial reduction of the errors. However, a systematic underestimation of the calculated quantities is revealed, leading to erroneous depth-mean values of the pressure force. In this study a modification of the Stelling and van Kester method is proposed which seems to improve the overall performance of the method.

(No. of Figures: 17. No. of Tables: 1. No. of Refs: 26.)

KEY WORDS: pressure gradient force; steep topography; terrain-following coordinates

1. INTRODUCTION

Ocean currents appear basically to be geostrophically balanced. An accurate treatment of the horizontal pressure gradient (HPG) forces is therefore of vital importance when numerical models are applied to study the oceanic circulation. Assuming a hydrostatic pressure distribution in the vertical, the HPG force (per unit mass) can be written as

$$\frac{1}{\rho}\nabla p = \frac{g}{\rho}\nabla \int_z^{\zeta} \rho \mathrm{d}z', \tag{1}$$

where p is the pressure, ∇ is the horizontal gradient operator, ρ is the density and g is the constant acceleration due to gravity. To get the HPG force in the form (1), we have assumed a constant atmospheric pressure at the free surface $z = \zeta(x, y, t)$. The numerical model employed in the present study is a version of the rather well-merited Blumberg–Mellor (BM) ocean model,[1–3] frequently also referred to as the Princeton ocean model. This is a primitive equation (hydrostatic) model making use of a terrain-following vertical co-ordinate, i.e. the so-called σ-co-ordinate, and the finite difference equations are solved on a three-dimensional (x, y, σ) grid. Terrain-following co-ordinates have been applied in numerical weather prediction models for many years.[4] Furthermore, it is well known that such a co-ordinate system may increase the truncation errors in the finite difference calculations of the HPG forces, especially when there is strong vertical stratification in areas with steep bottom

*Current address: Norwegian Institute for Air Research, PO Box 100, N-2007 Kjeller, Norway

CCC 0271–2091/97/100987–31$17.50 *Received March 1996*

slopes.[5,6] A related problem with this type of co-ordinate system is that it can render the numerical scheme 'hydrostatically inconsistent'.[7,8] This means that the difference between the discrete and the analytical version of the HPG force does not decrease when the vertical resolution is increased. This happens when a certain condition, involving depth, bottom slope and horizontal and vertical grid distances, is violated.[9] A discussion of these problems in numerical *ocean* models was given by Haney.[10] His conclusion was that one should use as many σ-levels as possible, but at the same time comply with the condition for hydrostatic consistency.

However, as pointed out by Mellor *et al.*,[11] the hydrostatic consistency requirement is quite difficult to fulfil in oceanic applications. Furthermore, they show that when the density is a function of the vertical co-ordinate (z) only, the truncation errors of the HPG terms are eliminated for those particular locations within the fluid where the limiting value of the condition for hydrostatic consistency is reached. Using the same numerical model as employed here, Mellor *et al.*[11] presented some numerical results from which they argued that the inherent problems of the σ-co-ordinates seemed to be surprisingly unimportant. This conclusion, however, was based mainly on model comparisons of depth-integrated quantities. As will be shown in this study, these quantities are not as much affected as the three-dimensional velocity structure. This is because the original BM model produces large errors in the HPG forces which more or less cancel when integrated in the vertical.

Recently, a new method for computing HPG forces was presented by Stelling and van Kester.[12] This method can be used both in the calculation of horizontal pressure gradients as well as for evaluating horizontal diffusive fluxes (to avoid artificial diapycnal mixing). In their paper, Stelling and van Kester[12] mainly focus on the treatment of the diffusive fluxes. Consequently, their method is constructed so as to ensure that gradients of concentrations are always reduced for every time step, i.e. fulfilment of the so-called min-max condition. Even though this is an important condition for diffusion problems, our experimental results indicate that this requirement leads to a systematic underestimation when the method is applied to calculate HPG forces. Furthermore, Stelling and van Kester[12] only present test results where the correct solution is equal to zero. This means that if their method underestimated the HPG force, the test experiments described in their paper would not have revealed this fact.

In Section 2 a description of the σ-co-ordinate transform will be given. Here we also present the two finite difference formulations of the HPG forces applied in the original BM model and in the Stelling and van Kester (SvK) method respectively. In addition, a slightly modified version of the SvK method will be presented. It has been introduced to reduce the previously mentioned underestimation that seems to be inherent in the original SvK method. In Section 3 these methods are applied to calculate the HPG force in some idealized cases where the analytical (true) solution is known. In Section 4 the differences in *model* performance of the various numerical methods are demonstrated. Finally, in Section 5 some concluding remarks are given.

2. TERRAIN-FOLLOWING CO-ORDINATES IN NUMERICAL OCEAN MODELS

There are several reasons for using a terrain-following (σ) co-ordinate. One of the main benefits is that the bottom profile does not need to be approximated by a staircase form, as is the case when depth, i.e. z, is used as the vertical co-ordinate. When constructing a discrete depth field for use in a z-co-ordinate model, gentle variations in the bottom topography are easily lost. Instead of a slight bottom slope, the discrete depth field will consist of large horizontal areas separated by more or less severe jump discontinuities. Associated with these abrupt depth changes, large vertical velocities are very often found.[13] Another advantage of the σ-system is that the kinematic bottom boundary condition becomes very simple. This is due to the fact that the bottom is a co-ordinate surface. Furthermore, for processes where the dynamics of the bottom boundary layer is important, a good

resolution of the bottom boundary layer is achieved in a σ-system simply by increasing the number of sigma layers close to the bottom. Alternatively, in a z-co-ordinate model an increased resolution would be needed over the whole depth interval spanned by the bottom. In addition, the terrain-following co-ordinate system is also advantageous when considering computer resources. This is because the grid points are always kept within the fluid, thereby automatically increasing the vertical resolution in shallow areas.

Unfortunately, there are also some inherent disadvantages associated with the use of σ-co-ordinates in numerical models. One important problem that has evoked considerable concern is related to the size of the truncation error when computing the discrete version of the HPG force. Another undesirable side-effect of the co-ordinate transformation is that the horizontal diffusion terms may introduce artificial diapycnal mixing. A method to reduce this diapycnal mixing has already been implemented in the BM model.[14]

In this presentation we will focus on the problem with the HPG force. However, both the difficulties with the diapycnal mixing and the increased truncation in the calculation of the pressure force are related to the two-term representation of horizontal derivatives in the σ-co-ordinate system.

2.1. The σ-co-ordinate transform

For oceanic applications the σ-co-ordinate is generally defined by

$$\sigma = \frac{z - \zeta}{H + \zeta}. \tag{2}$$

Here z is the Cartesian vertical co-ordinate, $\zeta(x, y, t)$ denotes the position of the sea surface and $H(x, y)$ is the undisturbed ocean depth. The total depth is then $D = H + \zeta$. With this definition there is a one-to-one correspondence between the allowed values of $z \in [-H, \zeta]$ and $\sigma \in [-1, 0]$. For a rigid-lid ocean the terrain-following transformation is reduced to

$$\sigma = z/H. \tag{3}$$

For simplicity reasons the transformation (3) is preferred in error analysis, although the transformation (2) is usually applied in the numerical models. As long as $|\zeta| \ll H$, the difference between the two transformations is of minor importance.

In Figure 1 we have depicted a possible distribution of sigma levels when there is a linear slope in the bottom topography. Each curve represents a constant σ-value.

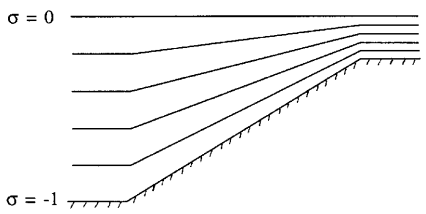

Figure 1. Sketch of possible distribution of sigma levels in an ocean model with linear slope in bottom topography

Application of the chain rule gives the relations between derivatives in the ordinary z-system and the σ-system. If σ is defined by (2), these relations can be expressed as

$$\left.\frac{\partial}{\partial\xi}\right|_z = \left.\frac{\partial}{\partial\xi}\right|_\sigma - \left(\frac{1}{D}\frac{\partial\zeta}{\partial\xi} + \frac{\sigma}{D}\frac{\partial D}{\partial\xi}\right)\frac{\partial}{\partial\sigma}, \tag{4}$$

$$\frac{\partial}{\partial z} = \frac{1}{D}\frac{\partial}{\partial\sigma}. \tag{5}$$

In (4) the variable ξ denotes either horizontal co-ordinates x, y or time t. With σ defined by (3), these relations become

$$\left.\frac{\partial}{\partial\xi}\right|_z = \left.\frac{\partial}{\partial\xi}\right|_\sigma - \frac{\sigma}{H}\frac{\partial H}{\partial\xi}\frac{\partial}{\partial\sigma}, \tag{6}$$

$$\frac{\partial}{\partial z} = \frac{1}{H}\frac{\partial}{\partial\sigma}. \tag{7}$$

In (6) the variable ξ denotes only horizontal co-ordinates, since σ now is time-independent. The subscripts in the equations above indicate that the x- and y-derivatives are to be performed along lines of constant values of z or σ in the respective co-ordinate system. The second term on the right-hand side of (4) and (6) therefore represents a correction which compensates for the vertical change due to the tilt of the σ-levels.

Applying (2), (4) and (5), the HPG force from (1) becomes in the σ-system

$$\frac{1}{\rho}\nabla_z p = g\nabla\zeta + \frac{gD}{\rho}\int_\sigma^0 \left(\nabla_\sigma\rho - \frac{\sigma'}{D}\nabla D\frac{\partial\rho}{\partial\sigma'}\right)\mathrm{d}\sigma'. \tag{8}$$

Here we have introduced ∇_z and ∇_σ to denote 'horizontal' gradients along surfaces of constant z and constant σ respectively. Since isopycnal surfaces are almost horizontal in the ocean, horizontal gradients of density are relatively small quantities compared with vertical gradients. However, in areas of steep bottom slopes the density gradients along the tilting surfaces of constant σ may become large owing to significant vertical changes in density. Accordingly, in such areas the integrand of (8) becomes the difference between two large terms of approximately the same size. Analytically, this represents no problem, but when the two terms in the integrand are discretized for use in a numerical model, the truncation error of each term may become of the same size as or greater than the horizontal gradient we are seeking.

In order to reduce the magnitude of each term in the integrand of (8), we express the density as

$$\rho = \bar{\rho}(z) + \hat{\rho}(x, y, z, t), \tag{9}$$

Here $\bar{\rho}(z)$ is a horizontal mean of the density field. Within the frame of the Boussinesq approximation the HPG force then becomes

$$\frac{1}{\rho_\mathrm{r}}\nabla_z p = g\nabla\zeta + \frac{gD}{\rho_\mathrm{r}}\int_\sigma^0 \left(\nabla_\sigma\hat{\rho} - \frac{\sigma'}{D}\nabla D\frac{\partial\hat{\rho}}{\partial\sigma'}\right)\mathrm{d}\sigma'. \tag{10}$$

where ρ_r is a constant reference density. As pointed out by several investigators,[15,16] the removal of the z-dependent density distribution $\bar{\rho}(z)$ may improve the results in those situations where there are only small horizontal density differences within the model domain. However, when there are large horizontal differences, i.e. situations where the integral of (10) contributes significantly to the depth variation of the HPG force, the removal technique may be of less help.[10]

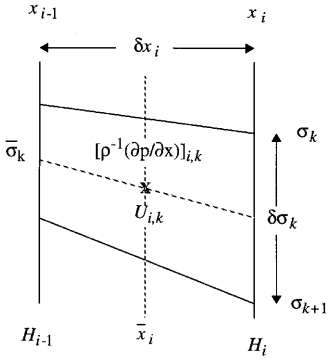

Figure 2. Grid point numbering and position of calculated HPG force

2.2. The discrete formulation of the HPG force

Since the main problem with the HPG force in three-dimensional numerical ocean models are related to the treatment of the integral in (10), we will here focus on the discretization of this expression. In the rest of this subsection we will therefore assume that the surface is level everywhere, i.e. $\zeta \equiv 0$, and $D = H(x, y)$.

The grid point numbering is depicted in Figure 2. Here the quantities δx_i, $\delta\sigma_k$ and $\bar{\sigma}_k$ are defined as $\delta x_i = x_i - x_{i-1}$, $\delta\sigma_k = \sigma_k - \sigma_{k+1}$ and $\bar{\sigma}_k = 0{\cdot}5(\sigma_{k+1} + \sigma_k)$ respectively. Density (i.e. salinity and temperature) and pressure are defined at the points $(x_i, \bar{\sigma}_k)$ and $(x_{i-1}, \bar{\sigma}_k)$. The vertical velocities are defined at the same x-points but at the σ_k-levels. An Arakawa C grid is employed and therefore the x-component of the horizontal velocity, $U_{i,k}$, is defined at the points $\bar{x}_i = 0{\cdot}5(x_{i-1} + x_i)$ and $\bar{\sigma}_k$. Since the HPG force is needed in the calculation of the horizontal velocity, this force is calculated at the same grid points as $U_{i,k}$.

For σ-co-ordinate models, Haney[10] emphasized, as mentioned in Section 1, the importance of specifying grid distances such that the 'hydrostatic consistency' condition is always fulfilled. This requires that

$$\left| \frac{\sigma}{H} \frac{\partial H}{\partial x} \right| \delta x < \delta\sigma. \tag{11}$$

From (11) it is seen that by increasing the number of σ-levels and thereby decreasing $\delta\sigma$, the consistency condition will sooner or later be violated.

In Figure 3(a) we have shown an example of a hydrostatically consistent distribution of two neighbouring σ-levels, whereas Figure 3(b) depicts an inconsistent distribution. It is seen that the geometrical interpretation of this criterion is that a horizontal displacement of one grid distance from a given σ-level should not result in a crossing of the σ-level immediately above or below.

2.2.1. The discrete HPG force in the original BM model. In our version of the BM model the discrete form of the x-component of (10) was formulated as

$$XPG_{i,k} = g \frac{H_i + H_{i-1}}{2\rho_r} \left(\frac{\hat{\rho}_{i,1} - \hat{\rho}_{i-1,1}}{\delta x_i} \frac{\delta\sigma_1}{2} + \sum_{k'=1}^{k-1} F_{i,k'} \delta\bar{\sigma}_{k'} \right). \tag{12}$$

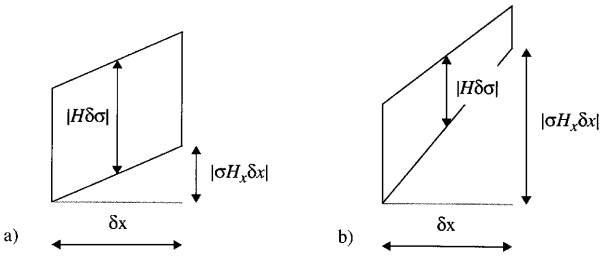

Figure 3. Hydrostatically (a) consistent and (b) inconsistent spacing between two neighbouring σ-levels. Index x denotes partial differentiation with respect to this variable

Here

$$F_{i,k} = \frac{\hat{\rho}_{i,k+1} + \hat{\rho}_{i,k} - \hat{\rho}_{i-1,k+1} - \hat{\rho}_{i-1,k}}{2\delta x_i} - \frac{\bar{\sigma}_k + \bar{\sigma}_{k+1}}{H_i + H_{i-1}} \frac{H_i - H_{i-1}}{\delta x_i} \frac{\hat{\rho}_{i,k} + \hat{\rho}_{i-1,k} - \hat{\rho}_{i,k+1} - \hat{\rho}_{i-1,k+1}}{2\delta \bar{\sigma}_k}, \tag{13}$$

where we have defined $\delta\bar{\sigma}_k = \bar{\sigma}_k - \bar{\sigma}_{k+1}$. (When the surface elevation is non-zero, the depth values H in (12) and (13) are simply replaced by the total depth values D.) The integration is thus performed with an ordinary midpoint formula, where we assume that the density is uniform with depth between the surface and the depth of the uppermost values of density. As long as the bottom is level, i.e. $H_i = H_{i-1}$, then naturally (12) and (13) are identical with the equations in a Cartesian co-ordinate system.

As shown by Haney,[10] it is possible to do some calculations in finite difference form on (12) and (13) if we take $\hat{\rho}$ to be either linearly dependent on z or totally independent of z. It should be mentioned that Haney's expressions were slightly different from (12) and (13) since he was analysing the departure of the HPG force from its depth-mean value. Following Haney's procedure, we find that for a density perturbation which is uniform with depth, our discrete version (12) of the HPG force becomes

$$XPG_{i,k} = -g \frac{\bar{H}^x}{\rho_r} \frac{\delta\hat{\rho}}{\delta x} \bar{\sigma}_k. \tag{14}$$

Here we have used the standard expressions for horizontal differencing and averaging. If we calculate the analytical expression (10), we find that (14) is an exact finite difference form of the true answer. As can be seen from (14), the error in the calculation of the HGP force depends solely on the horizontal differencing.

If, on the other hand, the density perturbation is a linear function of z only, i.e.

$$\hat{\rho}_{i,k} = \hat{\rho}_0 - \frac{\rho_r}{g} N_0^2 \bar{\sigma}_k H_i, \tag{15}$$

where $\hat{\rho}_0$ and N_0^2 are constants, then the HPG force should be identically zero. Substitution of (15) into (13) reveals that $F_{i,k}$ vanishes completely in this case. The only error contribution therefore comes from equation (12) and we have

$$XPG_{i,k} = \left(\frac{\delta\sigma_1}{2}\right)^2 \bar{H}^x \frac{\delta H}{\delta x} N_0^2. \tag{16}$$

Equation (16) shows that as long as the bottom is sloping, the difference version of the HPG force will not be equal to the true (zero) value. Furthermore, reducing the spacing of the uppermost σ-level

to half its original value will reduce the error by a factor of one-quarter. This has also been confirmed in numerical tests.

When the density perturbation is a *general* function of z only, it can be shown[11] that the truncation error of $F_{i,k}$ in (13) can be written as

$$E(F_{i,k}) = \frac{\bar{H}^x}{4}\frac{\delta H}{\delta x_i}\left[(\delta\bar{\sigma}_k)^2 - \sigma_{k+1}^2\left(\frac{\delta H}{\bar{H}^x}\right)^2\right]\left(\frac{\partial^2\hat{\rho}}{\partial z^2} + \frac{\sigma_{k+1}\bar{H}^x}{3}\frac{\partial^3\hat{\rho}}{\partial z^3} + \dots\right), \tag{17}$$

where $\delta H = H_i - H_{i-1}$. From this expression it is seen that the error vanishes locally if

$$(\delta\bar{\sigma}_k)^2 = \sigma_{k+1}^2\left(\frac{\delta H}{\bar{H}^x}\right)^2. \tag{18}$$

It is interesting to note that (18) is exactly the limiting value for maintaining hydrostatic consistency; see (11). This indicates that it actually can be beneficial to be on the verge of breaking this condition. On the other hand, it also demonstrates that when this condition is severely violated, the errors are increased in proportion.

For more general density distributions it is difficult to find simple expressions for the finite difference HPG force. Instead, we will in the next section compute this force numerically and compare the results with analytical results for different density distributions and bottom slopes.

2.2.2. The discrete HPG force as calculated by the method of Stelling and van Kester. Recently, Stelling and van Kester[12] presented a new method for the calculation of horizontal gradients in σ-co-ordinate models. Here we will only give a brief summary of the method; the reader is referred to the original paper for a more thorough discussion.

The basic idea is to change the grid system before calculating the integral term on the right-hand side of (10). Sketches of the original σ-co-ordinate system and the redefined grid system are given in Figure 4(a) and 4(b) respectively.

The positions of the density points are identical in the two systems. The new rectangular grid is then defined so that the density points are located centrally within each grid box. If there are a total of N σ-levels ($N = 4$ in the example shown in Figure 4), the horizontal sides of the grid boxes divide the

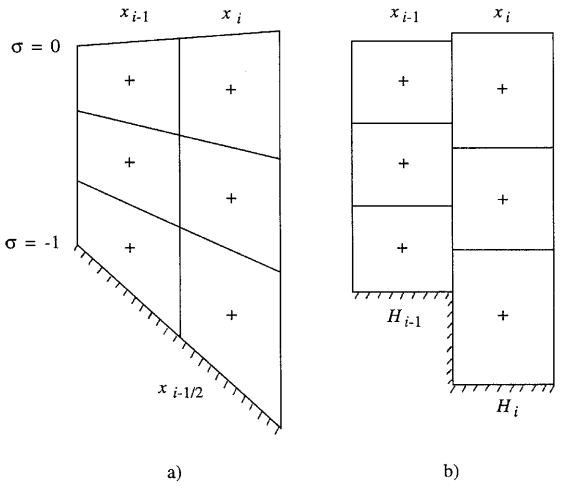

Figure 4. Grid cell structure of (a) original σ-co-ordinate system and (b) redefined system used by Stelling and van Kester[12]

vertical $x_{i-1/2}$ into a maximum of $2N-1$ subintervals, as seen in Figure 4(b). Except for the ones at the top and at the bottom, each of these intervals connects a grid box with the densities $\rho_{i-1,m}$ on the left-hand side and $\rho_{i,n}$ on the right-hand side. The aim is then to calculate a representative value of $\partial\rho/\partial x$ for each of the intervals. In order to accomplish this task, Stelling and van Kester propose the following stepwise procedure. First, the horizontal density gradient is set directly to zero for those intervals which only have grid boxes on one side, i.e. near the surface and the bottom. For the remaining intervals along the line $x_{i-1,2}$, two estimates of the horizontal density gradient, $\delta_m\rho/\delta x$ and $\delta_n\rho/\delta x$, are computed. To help explain how these two values are found for a specific interval, we refer to Figure 5. For the hatched interval in Figure 5 we define

$$\frac{\delta_m\rho}{\delta x}=\frac{\bar{\rho}^z(x_i,z_m)-\rho_{i-1,m}}{x_i-x_{i-1}}, \tag{19}$$

$$\frac{\delta_n\rho}{\delta x}=\frac{\rho_{i,n}-\bar{\rho}^z(x_{i-1},z_n)}{x_i-x_{i-1}}, \tag{20}$$

where $\bar{\rho}^z(x,z)$ is given by a linear interpolation formula between the two nearest density values on either side of z. Among those intervals separating two neighbouring grid cells, the uppermost and lowermost ones will have only one (not two) density value available for the computation of $\bar{\rho}^z(x_i,z_m)$ and $\bar{\rho}^z(x_{i-1},z_n)$; see Figure 4. In this case the single density value will be used in (19) or (20) instead of the interpolated value. The final value of the density gradient for each interval is then defined according to

$$\frac{\delta\rho}{\delta x}=\begin{cases}\min\left(\dfrac{\delta_m\rho}{\delta x},\dfrac{\delta_n\rho}{\delta x}\right), & \dfrac{\delta_m\rho}{\delta x}>0 \quad\wedge\quad \dfrac{\delta_n\rho}{\delta x}>0,\\ \max\left(\dfrac{\delta_m\rho}{\delta x},\dfrac{\delta_n\rho}{\delta x}\right), & \dfrac{\delta_m\rho}{\delta x}<0 \quad\wedge\quad \dfrac{\delta_n\rho}{\delta x}<0,\\ 0, & \dfrac{\delta_m\rho}{\delta x}\dfrac{\delta_n\rho}{\delta x}<0.\end{cases} \tag{21}$$

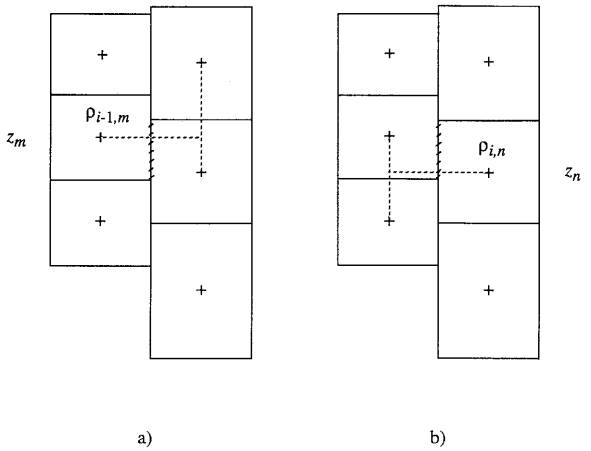

Figure 5. Illustration of density points involved in estimation of two values of horizontal density gradient which might represent value for interval marked by hatched line: (a) points used in computation of $\delta_m\rho/\delta x$; (b) points used for $\delta_n\rho/\delta x$

We are thus minimizing the value of the density gradient. As mentioned above, one should also keep in mind that the density gradient is zero along those intervals where there are no neighbouring rectangular grid cells.

The HPG force is then given by

$$\frac{1}{\rho_r}\nabla_z p(x_{i-1/2}, z) = \frac{g}{\rho_r}\sum_{n'=1}^{n-1}\left[(\delta z)_{i-1/2,n'}\left(\frac{\delta\rho}{\delta x}\right)_{i-1/2,n'}\right] + \frac{g}{\rho_r}\left[(z_{i-1/2,n} - z)\left(\frac{\delta\rho}{\delta x}\right)_{i-1/2,n}\right] + g\frac{\zeta_i - \zeta_{i-1}}{x_i - x_{i-1}}. \tag{22}$$

Here

$$(\delta z)_{i-1/2,n} = z_{i-1/2,n+1} - z_{i-1/2,n}, \quad n \in \langle 1, 2N-1\rangle, \tag{23}$$

is the length of each interval. In (22), z is the depth where the HPG force is to be computed. The value of n is chosen such that $z \in [z_{i-1/2,n}, z_{i-1/2,n+1}]$.

According to Stelling and van Kester,[12] one of the main benefits of this new method is that convergence is maintained even if the condition (11) for hydrostatic consistency is violated.

Since the process leading to (22) is rather complex, the counterparts of (14) and (16) have not been found for this method. Note, however, that for a rigid-lid ocean both the original BM method and the new SvK method become identical (and equal to a Cartesian method) when there are no bottom slopes. Furthermore, for all practical purposes they can be regarded as identical for a constant depth and free surface ocean as well, as long as $|\zeta| \ll H$. The important contribution from the free surface is through the last term on the right-hand side of (22).

It should be pointed out here that the SvK method is quite costly compared with the BM method when computer time is considered. This at least is so for the test experiments that have been performed with our computer code. Stelling and van Kester[12] claim that 'the increase in computational cost is relatively small'. This may suggest that our code is somewhat inefficient. Anyway, it is important that the implementation is made such that the particular computer at hand is utilized in the most cost-efficient way.

2.2.3. A modified version of the SvK method. As will be shown in Section 3, the original SvK method seems to underestimate the HPG force. However, promising results have been obtained by modifying the method slightly. Instead of choosing the minimum value of the density gradient as outlined in (21), we compute the value of $\delta\rho/\delta x$ simply by interpolating linearly between the two values $\delta_m\rho/\delta x$ and $\delta_n\rho/\delta x$ into the midpoint of their mutual interval. Along those intervals where there are no neighbouring grid cells, a zero value is still used for the density gradient. Apart from this modification the method is as described in the previous subsection. It should be noted that this slight change in fact makes the whole method equivalent to first interpolating the densities linearly in the vertical and then computing the gradients in the required depths!

3. TESTING OF PERFORMANCE

In this section we will compare the previously described methods by computing numerically the HPG forces for some simple analytical density distributions in which the density varies in both the vertical and the horizontal direction.

3.1. Experimental set-up

A three-layered system is considered (Figure 6) with constant densities ρ_s and ρ_b in the upper and lower layers respectively. Through the intermediate layer (the pycnocline) the density increases linearly with depth. The horizontal dependence is achieved by a linear tilt of the pycnocline layer. This layer is thus bounded by the lines $z = h_1(x)$ and $z = h_2(x)$ defined as

$$h_1(x) = -h_0 - \alpha\left(x - \frac{L}{2}\right) + \frac{\Delta h}{2}, \tag{24}$$

$$h_2(x) = -h_0 - \alpha\left(x - \frac{L}{2}\right) - \frac{\Delta h}{2}. \tag{25}$$

Here L is the total width of the testing area and h_0 is the depth of the midpoint of the pycnocline layer at $x = L/2$. Furthermore, α and Δh are the slope and vertical thickness of the intermediate layer respectively. The bottom topography is specified as a linear slope according to

$$z = \frac{H_{\max} - H_{\min}}{L}(x - L) - H_{\min}. \tag{26}$$

From this density field and utilizing the Boussinesq approximation, the HPG force (per unit mass) can be calculated according to

$$\frac{1}{\rho_b}\nabla_z p = g\left(\nabla\zeta + \rho_b^{-1}\int_z^0 \nabla\rho \, dz\right). \tag{27}$$

For simplicity we have explicitly defined the surface elevation to be zero and the HPG force is therefore zero everywhere in the surface layer. We have also converted the value of the HPG force to the equivalent geostrophic velocity. This is done simply by dividing (27) by the local Coriolis parameter f. In the results to be presented below, this velocity is shown instead of the actual HPG force. The analytical solution for the situation sketched in Figure 6 is therefore given as zero geostrophic velocity in the surface layer and a constant negative (directed out of the paper) geostrophic velocity in the lower layer. In the intermediate layer the velocity changes linearly with depth from the upper to the lower layer value.

A layered system with a sloping bottom as depicted in Figure 6 is typically found off the western coast of Norway; see e.g. Reference 17. Values representative of the Norwegian Sea area have therefore been chosen for the different parameters. Seven of these parameters have been kept constant through all the experiments presented below. The parameters and their values are given in Table I.

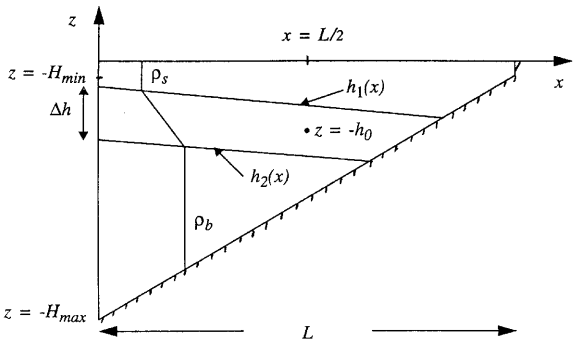

Figure 6. Sketch of density field applied in numerical calculation of HPG forces

Table I. Constant parameters

Parameter	Value
f	$1{\cdot}3\times10^{-4}$ s^{-1}
ρ_s	1026·7 kg m^{-3}
ρ_b	1028·1 kg m^{-3}
h_0	800 m
L	80 km
α	$2{\cdot}43\times10^{-3}$
H_{max}	3000 m

The value of the gravitational acceleration g is taken to be 9·81 m s^{-2}. With these values the true HPG force should give a geostrophic velocity of -25 cm s^{-1} in the bottom layer. The width of the testing area, L, was chosen to fit with the horizontal dimensions of the steeper parts of the Norwegian continental slope.

A large number of experiments have been performed. The experimental variables which have been varied are the bottom slope (achieved by modifying the value of H_{min}) and the value of the pycnocline thickness. Since the surface is horizontal in these experiments, the different methods are identical when the depth is constant, i.e. when $H_{min} = H_{max}$. Values ranging from zero to 1000 m have been used for the pycnocline thickness Δh. In addition, the influence of different horizontal and vertical grid resolutions has been tested. For obvious reasons, only a fraction of the numerous test results will be presented here. Nevertheless, the selected material should reveal the most essential differences between the various methods.

3.2. Experimental results

In order to make the comparison between the different methods as easy as possible, the corresponding results from the BM method, the original SvK method and the modified SvK method are presented together as plots (a)–(c) respectively in each of the figures to be shown in this section.

In accordance with the findings of earlier investigators,[10,11] our experiments have revealed that the results from the BM method are significantly improved by subtracting the horizontal mean value from the actual density field before computing the HPG forces. Consequently, the presented results from the BM method have all been obtained after subtraction of this mean value. For this experimental set-up the results from the two versions of the SvK method did not show any significant improvement when the horizontal mean density was removed, so the material from the calculations with these methods has been presented *without* any such subtraction. However, it should be noted that for other density distributions this removal technique might be beneficial for the SvK methods as well. Tests have shown that this is at least the case for the rather idealized situation where the density depends only on the vertical co-ordinate z.

In Figure 7, examples of the error fields produced by the different methods are presented. The errors are defined as the computed value minus the true analytical solution. These results were obtained with a bottom slope $\partial H/\partial x = 3{\cdot}74 \times 10^{-2}$, i.e. with $H_{min} = 10$ m. In the vertical, 31 equally spaced σ-levels were used, while the horizontal grid distance was 10 km. The positions were the HPG forces are calculated are marked by plus signs in the plots. It should be noted that in this experimental set-up the condition for hydrostatic consistency (11) is severely broken. Nevertheless, if one is to model areas covering the continental slopes of the Norwegian Sea, bottom slopes of this magnitude must be considered.

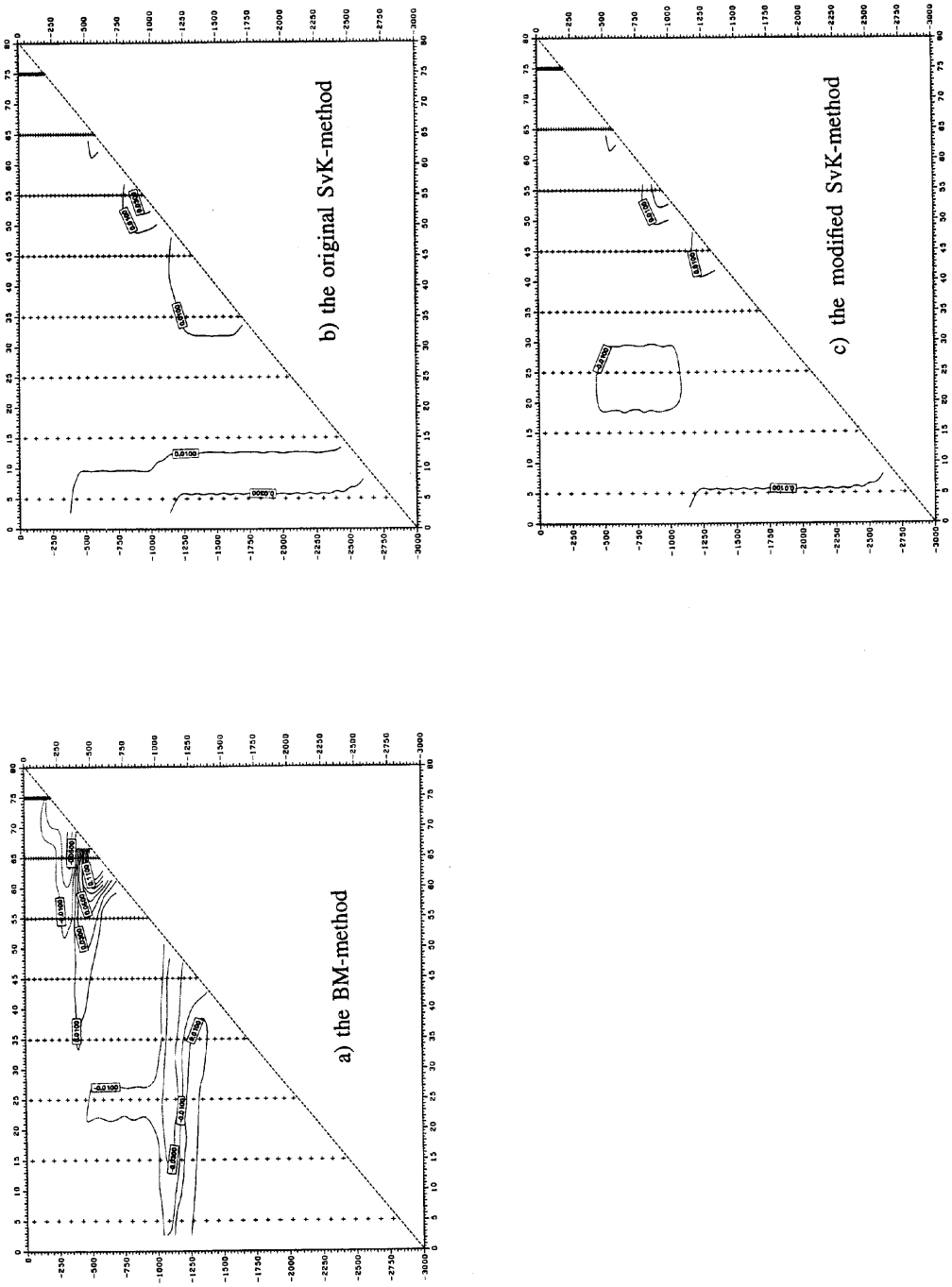

Figure 7. Departure from analytical solution (converted to equivalent geostrophic velocity) when applying (a) BM method, (b) original SvK method and (c) modified SvK method. Thickness of pycnocline is 750 m and $H_{\min} = 10$ m. Contour interval is 2 cm s^{-1}

The error field produced by the BM method (Figure 7(a)) reveals the most characteristic features of the problems associated with a straightforward computation of HPG forces in σ-co-ordinate models. Large errors of both positive and negative signs are found in the areas where the pycnocline intersects the bottom slope. The corresponding results from the original SvK method (Figure 7(b)) and the modified SvK method (Figure 7(c)) clearly demonstrate that application of these methods leads to a considerable reduction in the computed errors. Maximum errors of about 20 cm s^{-1} were found when the BM method was applied. This amounts to 80% of the maximum bottom value of 25 cm s^{-1}. In the case of the original SvK method the errors were bounded by a value of 5·4 cm s^{-1}, and when the modified SvK method was employed, only errors below 5·0 cm s^{-1} were found.

In the numerical model where these methods are to be applied, the depth-averaged currents (and hence the volume transports) are calculated from the depth-averaged momentum equations. In these equations the depth-integrated values of the HPG forces represent important forcing functions. It is therefore essential that the computational method is able to produce accurate integral values of the HPG forces. To get some information about these integral properties, the volume transports of the three error fields in Figure 7 were computed. Surprisingly enough, the large errors of the BM method contributed only to a small transport of $-0{\cdot}065$ Sv, while the errors of the original and modified SvK methods amounted to 0·81 and $-0{\cdot}16$ Sv respectively. Even though all these transports are small compared with the true transport of the analytical solution of 16·5 Sv, the differences are significant. As will be shown in the following, this tendency of producing the largest erroneous transports is a characteristic feature of the original SvK method. This is due to the fact that this method systematically underestimates the HPG forces and thereby produce errors which tend to be of the same sign. The errors of the other two methods have no preferred sign. Accordingly, they add up to lower flux values.

The results shown in Figure 7 represent one particular experiment with a very steep bottom slope. In order to investigate the significance of the bottom topography, similar experiments to those shown in Figure 7 have been performed with different bottom slopes. As mentioned earlier, the bottom slope has been changed by varying $H_{\min}$ between the values of 3000 and 10 m. This implies that the slope spans the interval from zero to $3{\cdot}74 \times 10^{-2}$. For each bottom slope the maximum error (cm s^{-1}) and the error-induced transport (Sv) have been stored. In all experiments the thickness of the pycnocline was 750 m. With these values the bottom reaches the lower border of the pycnocline at a slope value of $2{\cdot}16 \times 10^{-2}$ and it breaks through the upper part of the pycnocline at a slope value of $3{\cdot}10 \times 10^{-2}$. These experiments have been performed with different vertical and horizontal resolutions as well.

In Figure 8 the maximum error as a function of bottom slope is shown for the three methods. For each method we have depicted three curves representing vertical resolutions of 16, 31 and 61 σ-levels respectively. Here each increase in level leads to a doubling of the σ-'layers' in the vertical. The horizontal resolution is 10 km.

For the experiments with 16 σ-levels the hydrostatic consistency condition is violated for bottom slopes steeper than about $1{\cdot}3 \times 10^{-2}$. For the other two choices of vertical resolution this requirement is broken at the smaller slope of $7{\cdot}9 \times 10^{-3}$ and $4{\cdot}4 \times 10^{-3}$. This means that the schemes are hydrostatically inconsistent long before the bottom reaches the pycnocline layer.

The plots of Figure 8 show that the three methods give identical results when the bottom slope is equal to zero. Furthermore, the increase in vertical resolution leads to substantial reductions in the maximum errors for all the methods as long as the bottom is not too steep. It should be noted that the increase in vertical resolution improves the results even for slopes which far exceed the limiting values for maintaining hydrodystatic consistency. The results from the BM method (Figure 8(a)), for example, are improved for all slopes less than about $1{\cdot}5 \times 10^{-2}$, while for the two SvK methods (Figures 8(b) and 8(c)) the errors are reduced until the bottom reaches the pycnocline layer at the value of $2{\cdot}16 \times 10^{-2}$. The largest errors are produced by the BM method when the bottom penetrates

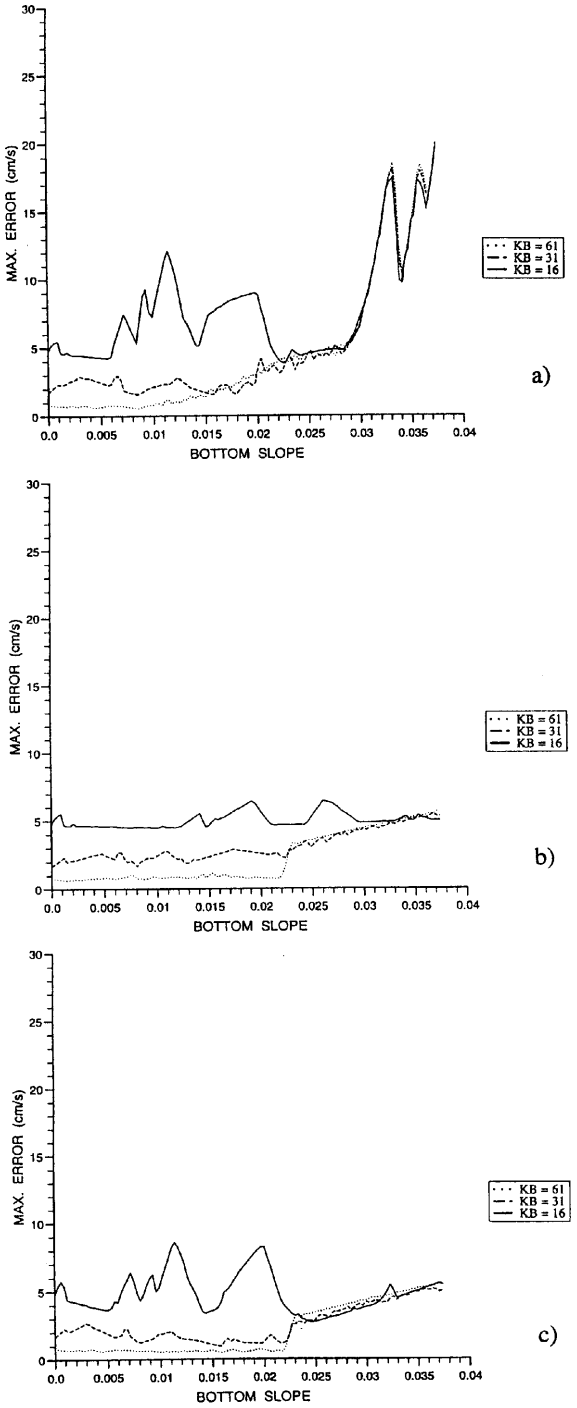

Figure 8. Maximum error as a function of bottom slope. Vertical thickness of pycnocline is 750 m and horizontal grid distance is 10 km. (a) BM method, (b) original SvK method, (c) modified SvK method. Results are shown for three different vertical resolutions

the pycnocline layer, i.e. for bottom slopes greater than about $3{\cdot}1 \times 10^{-2}$. For these slope values, errors of 15–20 cm s^{-1}, i.e. 60%–80% of the analytical bottom value, are reached. Errors of this magnitude are not found in the results obtained from the SvK methods, but at least for the high-resolution experiments a significant *increase* in the errors is found for bottom slopes greater than about $2{\cdot}2 \times 10^{-2}$. This increase is caused by bottom-trapped errors which are produced by the SvK method's interpretation of the sloping bottom as a staircase-shaped form; see Figure 4. These errors are created whenever isopycnals intersect the sloping bottom, and in order to reduce their magnitude, both the vertical and horizontal resolutions must be improved in proportion. Nevertheless, compared with the results from the BM method, both the SvK methods clearly represent a remarkable improvement as far as reduction of errors is concerned. The overall impression is that the modified SvK method (Figure 8(c)) produces the smallest errors. This is most clearly demonstrated by the high-resolution curves.

In Figure 9 the computed error-induced transports as a function of bottom slope are given for the different vertical resolutions. From these plots it is seen that both the BM method (Figure 9(a)) and the modified SvK method (Figure 9(c)) are capable of producing quite accurate transport values. The transport calculations of these methods are also distinctly improved by increasing the number of σ-levels from 16 to 31. Increasing the vertical resolution to 61 σ-levels, however, does not lead to further improvements. The results obtained from the original SvK method (Figure 9(b)) clearly demonstrates that the major drawback of this method is associated with the integral properties. There is a tendency to produce systematic errors in the integrated values of the HPG forces and thereby create significant artificial transports. These erroneous transports are reduced as the vertical resolution is increased, but even with 61 σ-levels in the vertical the original SvK method results are still no better than the results obtained from the modified SvK method when only 16 σ-levels were employed.

In Figures 10 and 11 the maximum error and the error-induced transport as functions of bottom slope are presented for different *horizontal* resolutions. In all these experiments a vertical resolution of 31 σ-levels has been applied. The results from each of the numerical methods have then been computed with horizontal grid distances of 20, 10 and 5 km. With these resolutions the bottom slopes of $1{\cdot}3 \times 10^{-2}$, $7{\cdot}3 \times 10^{-3}$ and $4{\cdot}4 \times 10^{-3}$ still apply as the limiting values for maintaining hydrostatic consistency. In contrast with the situation where the vertical resolution was altered, it is now the high-resolution (5 km) grid which allows the steepest bottom slope.

Many of the features seen in Figures 8 and 9 are repeated in the curves of Figures 10 and 11, including the large maximum errors produced by the BM method when the bottom reaches the pycnocline layer (Figure 10(a)) and the large artificial transports created by the original SvK method (Figure 11(b)). However, the different horizontal resolutions leads to quite interesting modifications of the results. It was naively expected that the errors would be reduced when the horizontal resolution was improved. As seen in Figure 10, this is only the case for rather steep bottom slopes. For more moderate slopes there is an *increase* in the maximum errors when the horizontal grid distance is reduced! The reason for this increase is not fully understood, but our experimental results indicate that it is linked to the grid's ability to discern the gentle tilt of the pycnocline. To accomplish this properly, it appears that the ratio between the vertical and the horizontal grid distance must be approximately of the same magnitude as the slope of the isopycnals. For the results in Figure 10 this grid ratio is closest to the value of the pycnocline slope when the horizontal grid distance is 20 km. When considering the error-induced transports, the results from both the BM method (Figure 11(a)) and the modified SvK method (Figure 11(c)) are improved by an increased horizontal resolution. This desirable property is not evident in the corresponding results from the original SvK method (Figure 11(b)).

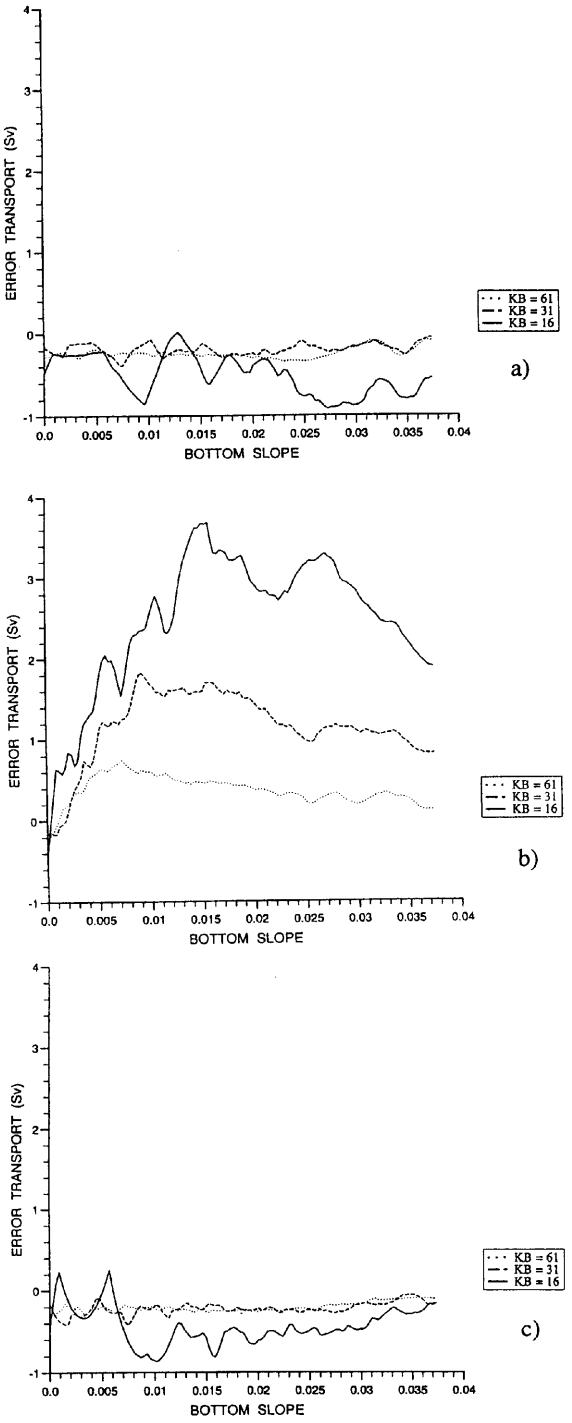

Figure 9. Volume transport computed from error field as a function of bottom slope. Vertical thickness of pycnocline is 750 m and horizontal grid distance is 10 km. (a) BM method, (b) original SvK method, (c) modified SvK method. Results are shown for three different vertical resolutions

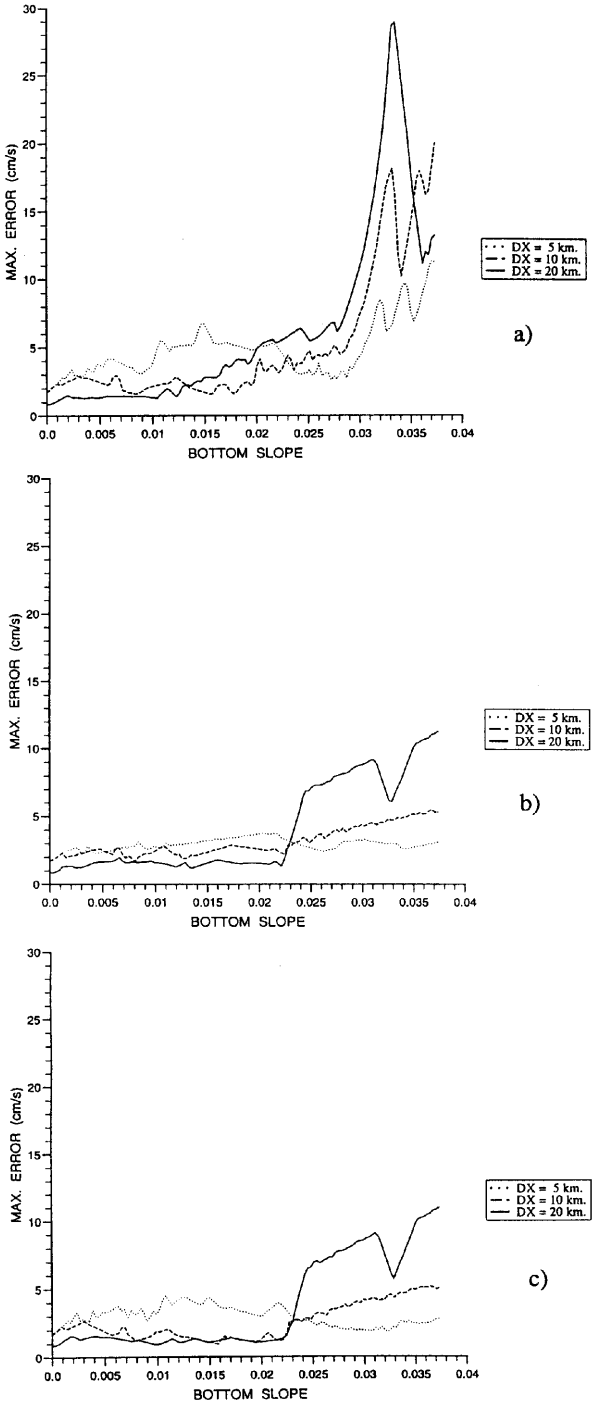

Figure 10. Maximum error as a function of bottom slope. Vertical thickness of pycnocline is 750 m and 31 σ-levels are used in vertical. (a) BM method, (b) original SvK method, (c) modified SvK method. Results are shown for three different horizontal resolutions

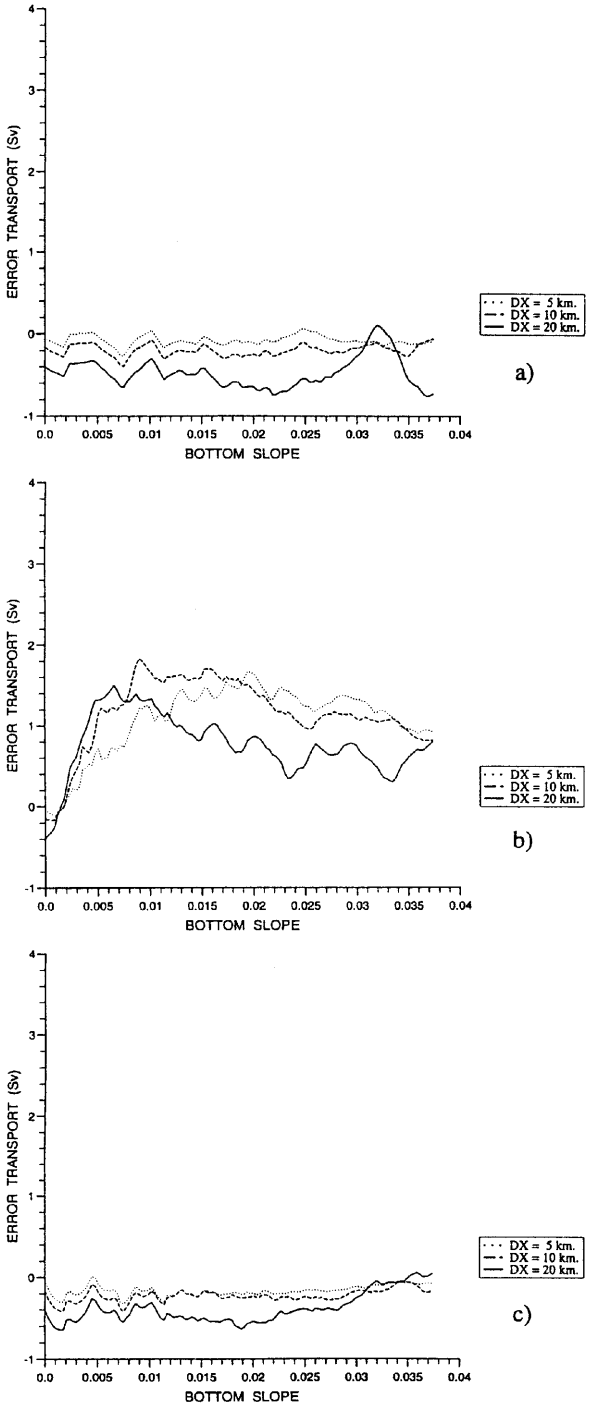

Figure 11. Volume transport computed from error field as a function of bottom slope. Vertical thickness of pycnocline is 750 m and 31 σ-levels are used in vertical. (a) BM method, (b) original SvK method, (c) modified SvK method. Results are shown for three different horizontal resolutions

In the results presented so far, a relatively thick pycnocline of vertical extent 750 m has been considered. Since the maximum depth is 3000 m and the σ-levels are equally spaced in the vertical, this means that the pycnocline is rather well resolved, at least in the experiments where 31 and 61 σ-levels are applied. In order to investigate the importance of the vertical width of the stratified layer, this experiment has also been performed with other values of the pycnocline thickness. In Figure 12 the maximum error as a function of bottom slope (i.e. the equivalent of Figure 8) is shown for the case where the pycnocline was only 250 m thick. It should be noted that the scaling of the vertical axis has been doubled in Figure 12 as compared with the plots in Figure 8. The curves displayed in Figure 12 clearly demonstrate that the errors increase substantially as the pycnocline narrows. This is most apparent in the results produced by the BM method (Figure 12(a)). A pronounced increase in the errors is also found in the results from the original (Figure 12(b)) and the modified (Figure 12(c)) SvK method. Apart from the size of the maximum errors, most of the other characteristics of the curves in Figure 8 are still recognized in Figure 12. These include the improvement achieved at moderate bottom slopes when the number of σ-levels is increased and the striking lack of such an improvement, at least for the BM method (Figure 12(a)), for steeper slopes. This is a general trend that has been seen in all the experimental results: when the pycnocline thickness is reduced, the magnitude of the errors and the corresponding erroneous volume transports increase, but the spatial distribution and the relative dependence upon vertical and horizontal resolutions are almost unaltered.

Only the *maximum* errors from a large number of individual experiments are given in Figure 12. As an example of the actual solution which is found by each of the methods for a particular bottom slope, the calculated geostrophic velocities along a chosen test section are presented in Figure 13. These are the results from an experiment with a bottom slope of $3{\cdot}74 \times 10^{-2}$ (i.e. $H_{\min} = 10$ m), a horizontal grid distance of 10 km and with 31 equally spaced σ-levels. To facilitate the comparison, the analytical (true) solution is displayed in Figure 13(d). The weakness of the BM method is clearly demonstrated in Figure 13(a), where the errors overshadow the true solution almost completely. The characteristic changes in the signs of the errors are also conspicuous. Results of much higher quality are pronounced by the two SvK methods. Nevertheless, in the contour plot of the original SvK method (Figure 13(b)) the previously mentioned tendency of underestimation is obvious. Instead of the true value of -25 cm s^{-1}, absolute values below 22 cm s^{-1} are computed in large parts of the bottom layer. Accordingly, this velocity field integrates to a volume transport of $-13{\cdot}9$ Sv, which is significantly less (in absolute value) than the true value of $-16{\cdot}4$ Sv. The modified SvK method (Figure 13(c)) exhibits the strongest resemblance to the analytical solution. The most distinct errors appear as a slight underestimation in the area where the stratified layer meets the bottom. Based on the velocity field produced by this method, a volume flux of $-16{\cdot}8$ Sv was found, i.e. within 0·5 Sv of the true answer. The most puzzling feature of the results depicted in Figure 3, however, is the fact that the rather chaotic velocity distribution of the BM method adds up to the true volume transport of -16.4 Sv! Even though this extraordinary agreement is accidental, the general picture is that the BM method produce transport values which are approximately of the same accuracy as those found by the modified SvK method.

The results presented in Figures 7–13 clearly illustrate that there are serious problems associated with the use of the original BM method. Even when the horizontal mean density is subtracted, this method introduces large errors in the calculated HPG forces in areas of steep topography. These problems are not pronounced at locations where the isopycnals intersect the sloping bottom. Our results also emphasize that when the condition for hydrostatic consistency is severely broken, an increase in vertical resolution does not reduce the errors. The effect of increasing the horizontal resolution is more obscure. Generally the errors are reduced in areas of steep bottom slopes and in most cases they are enhanced in areas of moderate steepness. Nevertheless, as a consequence of the characteristic alternating signs in the error field produced by the BM method, the computed volume

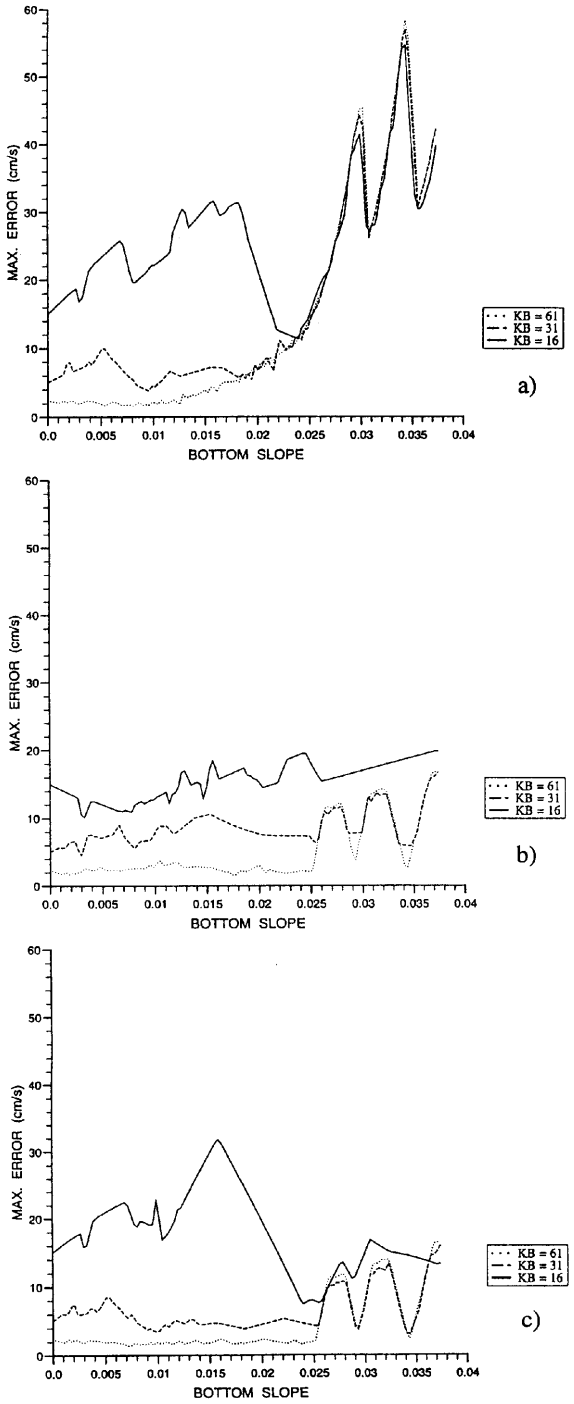

Figure 12. Maximum error as a function of bottom slope. Vertical thickness of pycnocline is 250 m and horizontal grid distance is 10 km. (a) BM method, (b) original SvK method, (c) modified SvK method. Results are shown for three different vertical resolutions

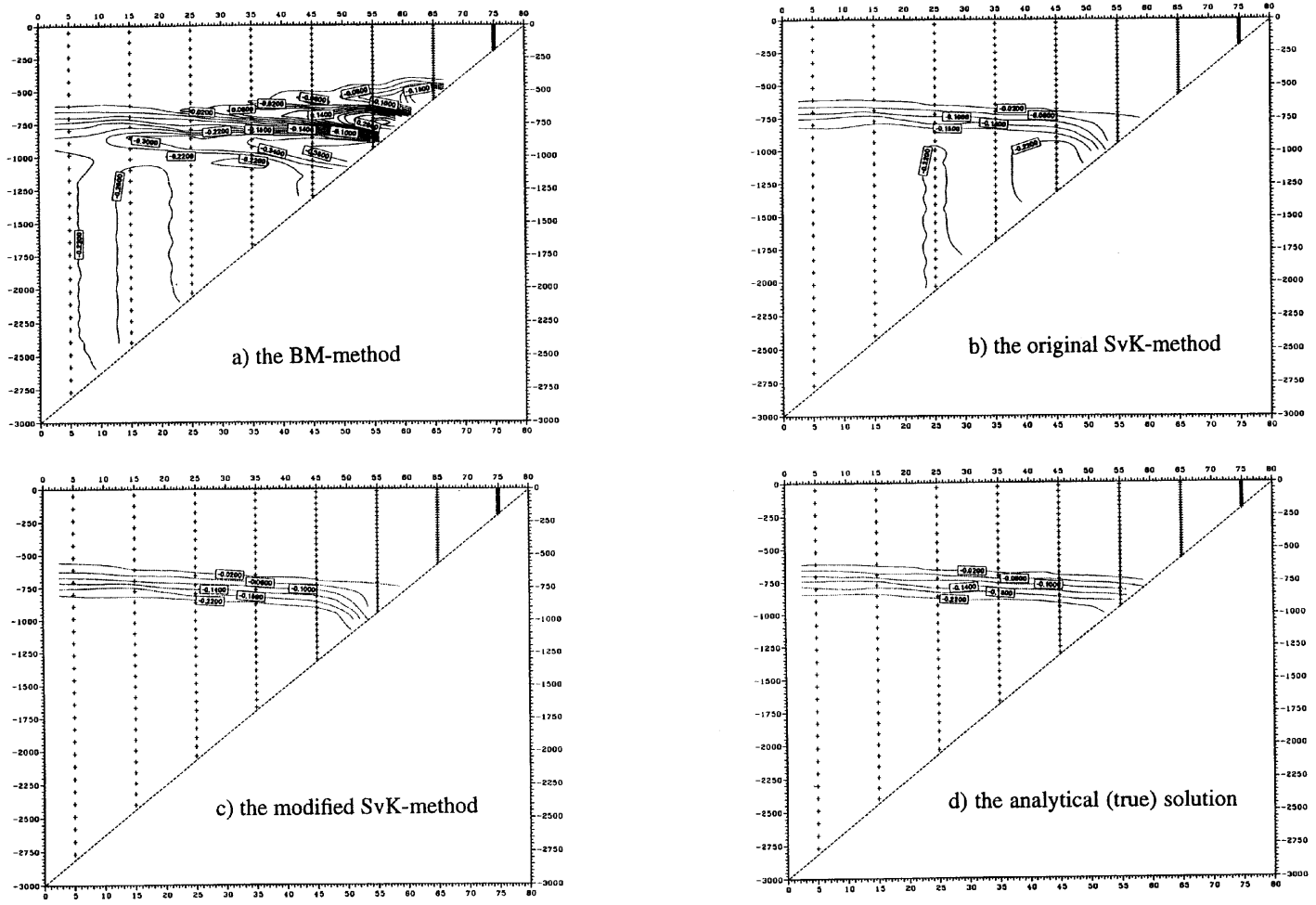

Figure 13. Contour plots of computed HPG forces (converted to equivalent geostrophic velocity) when applying (a) BM method, (b) original SvK method and (c) modified SvK method. In (d) analytical (true) solution is given for comparison purposes. Thickness of pycnocline is 250 m and $H_{\min} = 10$ m. Contour interval is 4 cm s^{-1}

transports through the test section are much better than would be expected. The results also indicate that these flux calculations are more or less systematically improved when both the vertical and horizontal resolutions are increased. Since fulfilment of the hydrostatic consistency condition is presently unrealistic for numerical studies involving both shelf and abyssal areas, this straightforward method of calculating the HPG force is clearly not satisfactory.

There is no doubt that the original SvK method represents an improvement over the BM method. The errors are drastically reduced. However, the experimental results reveal quite clearly that a new problem, associated with the calculations of the integrated HPG forces, is introduced by this method. This problem, which seems to be linked to a systematic underestimation of the computed HPG forces, is reduced by increasing the vertical resolution. On the other hand, improving the horizontal resolution tends to enhance the problem, at least for the steeper bottom slopes.

Based on the presented material (and a large number of additional experiments), we conclude that the modified SvK method represents a combination of the best qualities of the other two methods. The errors are small compared with those created by the BM method and the integral properties of the method are superior to the original version of the SvK method. Even though these results are promising, we emphasize that the errors and the corresponding artificial transports are still of considerable size, especially when the pycnocline is poorly resolved in the vertical.

As mentioned in Section 2.2.3, the modified SvK method is equivalent to performing a simple linear interpolation of the densities in the vertical prior to the calculation of the density gradients. To reveal the influence of the interpolant, experiments with higher-order interpolation (*cubic splines*) have also been performed. Apart from increasing the computer costs, this led only to minor improvements in the results.

4. NUMERICAL SIMULATIONS WITH THE DIFFERENT METHODS

In the previous section the HPG forces were computed by the various methods and direct comparisons were made. In this section we will demonstrate the influence on overall model performance when the different methods are applied in the numerical model. To do so, a simple geostrophic flow experiment has been performed. As mentioned in Section 1, the time-dependent, three-dimensional, estuarine and coastal circulation model of Blumberg and Mellor[1,2] has been adopted.

The prognostic variables of this model are the free surface elevation, the three components of velocity, temperature, salinity (hence density through an equation of state) and two quantities which characterize the turbulence, namely the turbulent kinetic energy and a turbulent length scale. The free surface elevation is calculated prognostically with only some sacrifice in computational time so that tides and storm surge events also can be simulated. This is accomplished by use of a time-splitting procedure whereby the depth-mean currents and vertical velocity shear are solved separately. The depth-mean currents and the surface slope are found from the depth-integrated equations, whereas the velocity shear is computed directly from the primitive equations. It should be noted that the calculated three-dimensional velocity field is corrected after each time step so as to give the same depth-integrated values as found from the depth-integrated equations. The vertical mixing processes are calculated by applying an analytical turbulence closure model of small-scale turbulence.[18,19] At the open boundaries the flow relaxation scheme (FRS) of Martinsen and Engedahl[20] is applied.

The equations of the model, together with their boundary conditions, are solved by finite difference techniques. The computational grid arrangement of points is of the Arakawa C type. The time differencing is leap-frog and an implicit numerical scheme in the vertical direction has been adopted for computational efficiency. At every time step the solution is filtered to remove any residual time splitting due to the leap-frog method.[21] For advection a variant of the Lilly[32] method is used. The

Holland and Lin[23] scheme is used for the Coriolis force. Horizontal friction is lagged in time for computational stability. The numerical scheme conserves mass. The total energy is also conserved except for the time filter and explicit sources and sinks due to friction. For a more thorough description of this model we refer to Reference 2 and 3.

4.1. Model set-up

The experimental set-up is chosen as a straight channel of length 300 km and with a cross-sectional geometry and density distribution as in the experiment described in Section 3; see Figure 6. The tilting pycnocline, in which the density increases linearly with depth, is thus bounded by the lines defined in expressions (24) and (25), while the bottom topography is given by (26). Neither the density nor the topography depends on the along-channel co-ordinate y. The values of the physical parameters are given in Table I. Furthermore, the minimum depth is 300 m, leading to a bottom slope of $3{\cdot}375 \times 10^{-2}$. The thickness of the pycnocline has been set to 750 m. Apart from the inherent time filter,[21] the model has been run without frictional influence.

A geostrophically balanced surface elevation and velocity field can easily be constructed from the density field defined above. We have here chosen a velocity field with zero currents in the bottom layer. By using a Coriolis parameter with a constant value of $1{\cdot}3 \times 10^{-4}\ \mathrm{s}^{-1}$ and a gravitational acceleration of $9{\cdot}81\ \mathrm{m\ s}^{-2}$, this leads to an along-channel velocity of 25 cm s^{-1} in the constant density surface layer and a surface slope

$$\frac{\partial \eta}{\partial x} = \frac{f}{g} v_{\mathrm{surf}} = 3{\cdot}3 \times 10^{-6}. \tag{28}$$

Through the pycnocline layer there is a linear decrease in velocity with depth. This current field and surface elevation were given initially everywhere in the model domain. At the open boundaries, i.e. within the FRS zones, the geostrophic solution was specified throughout the simulations. To make the test experiment as simple and straightforward as possible, the model was run in the so-called *diagnostic* mode, i.e. the density field was kept fixed throughout the simulation. It should therefore be stressed that the results below apply to such simulations only. *Prognostic continuation*, as described by Mellor *et al.*,[11] has not been considered in this work.

In the vertical direction we used 31 σ-levels. Since the maximum depth was 3000 m, this means that the physical spacing between the σ-levels was equal to or less than 100 m. The pycnocline of thickness 750 m is therefore relatively well resolved everywhere. A horizontal grid spacing of 10 km was used in both the x- and the y-direction. The width of the FRS zones was 100 km, i.e. 10 grid points across.

4.2. Results

Based on the test results discussed in Section 3, it was expected that the deviations from the analytical solution would be most pronounced when the original BM method was applied. Moreover, it was believed that the underestimation of the HPG forces in the original SvK method would lead to some detectable errors in the depth-averaged solutions.

The calculated HPG forces (converted to equivalent geostrophic currents) for this model set-up are shown in Figure 14. The results obtained with the original BM method are shown in Figure 14(a), while the original and modified SvK methods produced the results in Figures 14(b) and 14(c) respectively.

Comparing the data from the BM method (Figure 14(a)) with the analytical (true) values, errors in the interval from $-6{\cdot}6$ to $10{\cdot}5$ cm s^{-1} are found. This means that the maximum errors amount to 42% of the true surface velocity. For the other two methods, smaller but still significant errors were

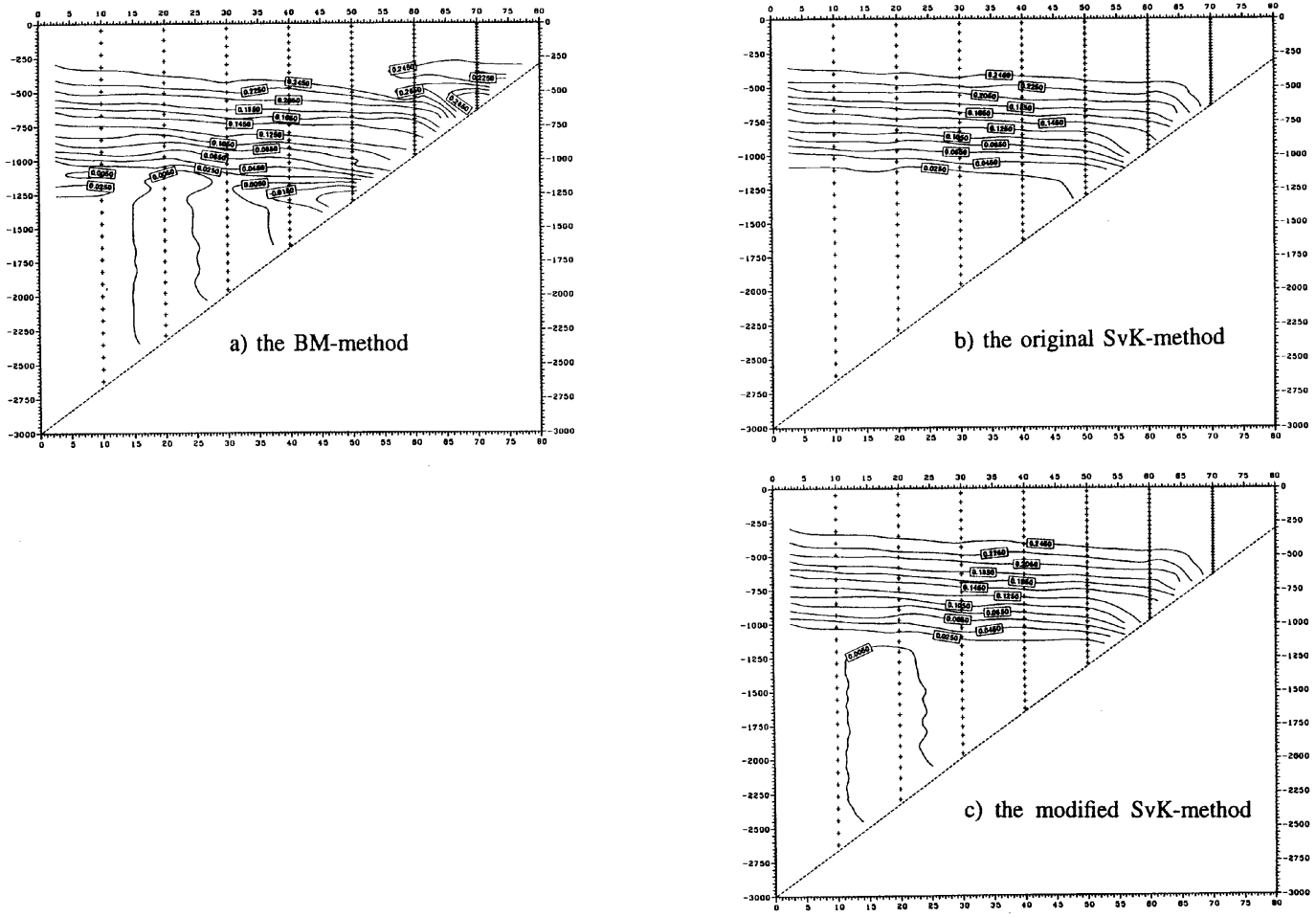

Figure 14. Contour plots of calculated HPG force (converted to equivalent geostrophic velocity) when applying (a) BM method, (b) original SvK method and (c) modified SvK method. Thickness of pycnocline is 750 m and $H_{\min} = 300$ m. Contour interval is 2 cm s^{-1}

detected with maximum values up to about 22% and 17% of the surface value. Here the lowest value is obtained with the modified version. For all methods the maximum errors are found close to the bottom in the region where the pycnocline intersects the bottom slope.

Since the density is not changed during a diagnostic model run, the calculated internal HPG forces shown in Figure 14 will remain practically constant during the simulation. If the free surface is deflected, the position of the σ-levels will be changed in proportion and this might lead to a slight modification of the calculated HPG forces. However, as long as the depth is much larger than the surface elevation, which indeed is the case in our experiment, these modifications are insignificant. Consequently, the internal HPG forces need only be calculated once, i.e. initially, and the increased computational costs of the SvK methods are therefore hardly detectable in diagnostic simulations.

Ideally, the initial solution should be in dynamic balance with the model equations, so the prescribed surface elevation and velocity field should remain as a steady solution throughout the simulation period. However, as a consequence of truncation errors in the numerical model, the prescribed solution will not be in perfect balance and an artificial dynamic adjustment will take place. The size of the truncation errors will be reflected by the strength of this adjustment process. To illustrate the degree of adjustment in our experiments when the different methods are applied, time series of the total kinetic energy within the model domain are presented in Figure 15. When the original BM method is applied (Figure 15(a)), large fluctuations are found in the total kinetic energy. Values which are more than 25% larger than the initial value are occurring at intervals of 60–70 h. These low-frequency oscillations are most likely caused by topographic (Rossby) waves which are generated as a result of cross-isobath movements initiated by the errors in the internal HPG terms. In addition to this long periodic rise and fall in kinetic energy, more rapid fluctuations with the inertial period, here 13–14 h, are also evident in the time series. When the original and modified versions of the SvK method are applied (Figures 15(b) and 15(c) respectively), the kinetic energy remains closer to the specified initial value throughout the 12 day simulation period, signalling that the adjustment processes are less energetic. Nevertheless, oscillations with near-inertial frequency are readily seen. These are most clearly revealed in the results obtained with the original SvK method. Based on the displayed curves, we therefore conclude that application of the original BM method induces the strongest artificial response. Of the two SvK methods the modified version gives the least response. This finding supports our claim that the modified SvK method provides the best overall fit to the correct distribution of HPG forces.

Examples of model-generated velocity distributions are presented in Figure 16, where the along-channel velocity through a centrally located cross-section of the channel is shown. As a standard of comparison the initial (analytical) velocity field is given in Figure 16(d), while Figures 16(a)–16(c) give the velocity after 48 h of simulation with the original BM method, the original SvK method and the modified SvK method respectively. As expected, the errors are most pronounced when the original BM method is applied (Figure 16(a)). Inspection of the data shows that these errors range from -28 to 37 cm s^{-1}. These errors are considerably larger than those found in Figure 14(a). One should also note the peculiar pattern of alternating directions of the currents along the bottom in Figure 16(a). This is a characteristic feature of the current field when the original BM method is used. Substantial deviations from the analytical solution are also found when the two versions of the SvK method are applied. The errors in Figure 16(b) are bounded by the values $-5{\cdot}0$ and $11{\cdot}7$ cm s^{-1}, while the error bounds in Figure 16(c) are $-5{\cdot}1$ and $10{\cdot}8$ cm s^{-1}. The results produced by the two versions of the SvK method are thus quite similar, and even though they are not offering a completely satisfactory solution, they definitely lead to improved results as compared with the original BM method.

Somewhat to our surprise, the calculated volume transports in this specific experiment were not changed significantly by any of the different HPG methods. The calculated depth-mean currents were

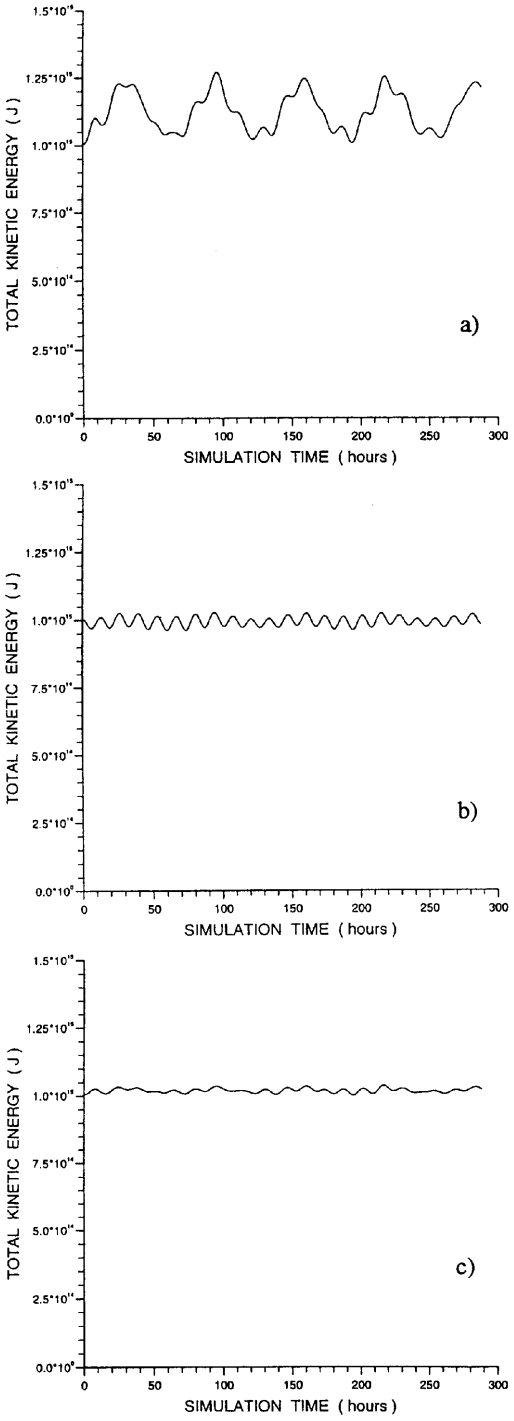

Figure 15. Time series of kinetic energy within the model domain (FRS zones not included) when applying (a) original BM method, (b) original SvK method and (c) modified SvK method

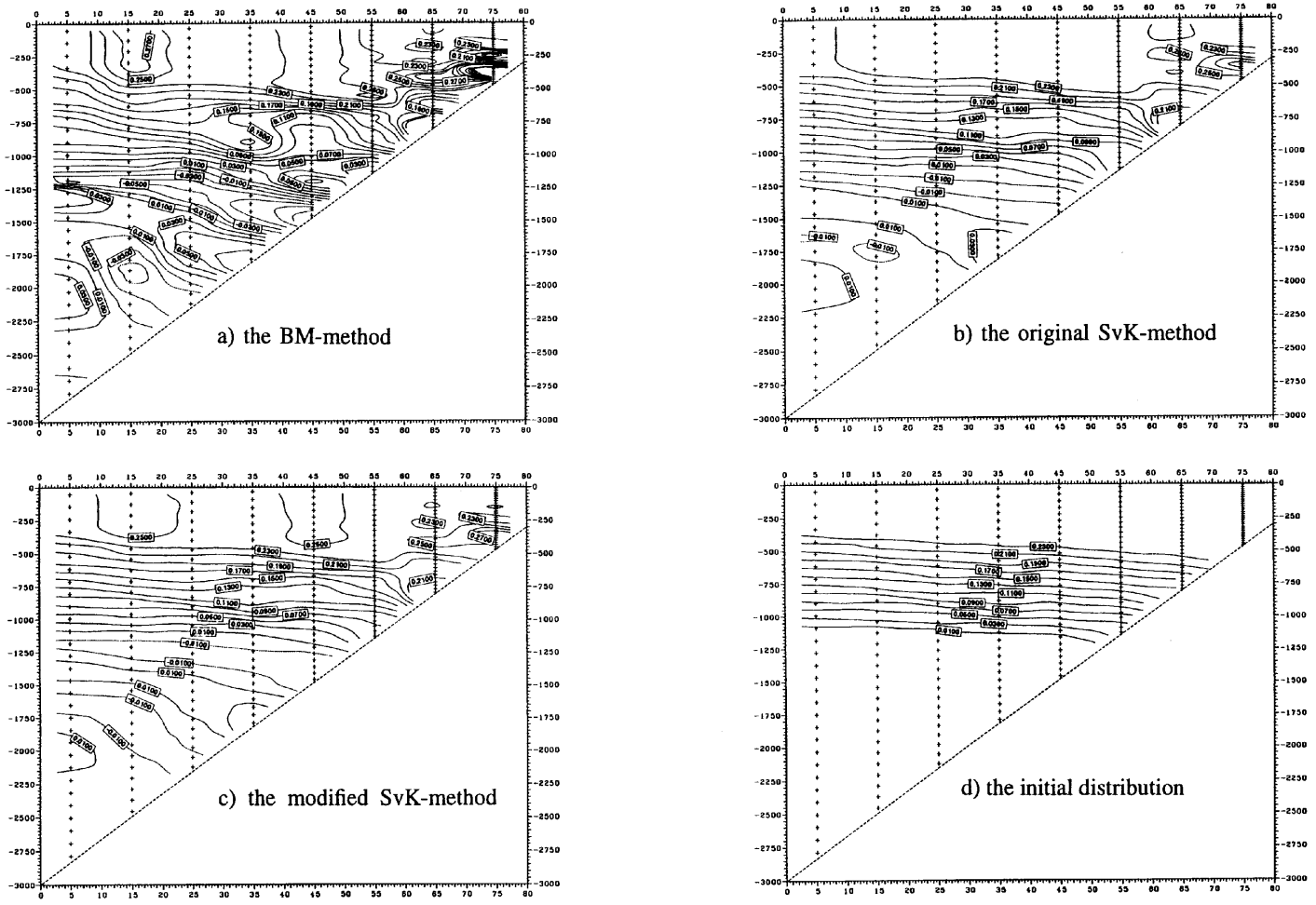

Figure 16. Contour plots of along-channel velocity after 48 h when applying (a) original BM method, (b) original SvK method and (c) modified SvK method. In (d) initial distribution (i.e. analytical solution) is given for comparison purposes. Thickness of pycnocline is 750 m and $H_{\min} = 300$ m. Contour interval is 2 cm s^{-1}

in fact almost identical throughout the simulation and equal to the prescribed analytical solution. Instead of altering the depth-mean currents, however, the effect of the different HPG methods on the integrated solution was to modify the surface elevation. Time series of the simulated surface elevation are shown in Figure 17. The data are taken from four stations which are located symmetrically along the central cross-section of the channel. For all the methods a slight adjustment of the surface is seen during the first 10 h. This rapid adjustment, which must be brought about by fast-moving gravity waves, is most pronounced when the original SvK method is applied (Figure 17(b)). This is in agreement with the results of Section 3, where it was found that this method to some extent underestimated the integral value of the HPG force. The puzzling fact is that the surface is modified in such a way that it exactly compensates the errors in the integrated HPG terms! Based on these results, it is not possible to conclude that the transport calculations will be equally good in more general cases. A reason for this is that the channel is very narrow in our experiment, i.e. the channel width is much less than the barotropic Rossby radius of deformation. This means that rotational effects will be unimportant for this initial adjustment.[24] Consequently, since the HPG forces are identically zero in the along-channel direction, only the depth-averaged current in the cross-sectional direction and the surface elevation will be influenced by the errors in the depth-integrated HPG forces. However, the depth-mean currents in the cross-channel direction are effectively forced towards zero by the channel walls and therefore only the surface elevation is modified. In more open and complex ocean regions this adjustment may influence the transports as well.

5. CONCLUDING REMARKS

The main goal of this study has been to investigate the inherent problems associated with the calculation of the internal HPG forces of the primitive equation, sigma-co-ordinate, Blumberg–Mellor ocean model. In σ-co-ordinate models these problems are most pronounced in areas where the pycnoclines intersect steep bottom slopes. By performing simple test experiments, where the density field is allowed to vary in both the horizontal and the vertical direction, severe errors have been detected. In order to reduce the errors, the alternative method of Stelling and van Kester[12] has been implemented in the BM model. This method basically transforms the σ-grid back to a Cartesian system before the HPG forces are calculated. Our results clearly demonstrate that the errors are reduced when the SvK method is applied. As shown by the test experiments, however, this method tends to underestimate the absolute value of the HPG force, which in turn leads to noticeable errors in the depth-integrated values of the HPG terms. These errors were actually larger than those obtained with the original BM method. This unwanted effect was avoided by introducing a slight modification of the SvK method. The modified version was found to retain the best assets of the other two methods in that it produced small errors as well as accurate integral values. It should be noted, however, that when there are strong vertical gradients in the density field, substantial errors are also produced by the modified SvK method. Moreover, the different methods yield practically identical results in areas of constant depth as long as the deflection of the free surface is negligible compared with the total depth.

A major drawback with the modified SvK method (also the original SvK method), compared with the straightforward and simpler BM method, is that it is rather time-consuming on the computer. This is especially the case in so-called prognostic simulations, i.e. simulations where the density field varies in time. In such situations the internal HPG forces have to be recomputed whenever the mass field is updated, and the increase in computer time can then add up considerably. This technical problem should be given priority in future work so as to improve model efficiency.

In this study, simple pressure distributions and bottom topographies have been considered. This has been done in order to reveal the artificial solutions caused by the numerical scheme. Nevertheless, the dimensions and physical parameters have been carefully chosen to fit the characteristic values of the

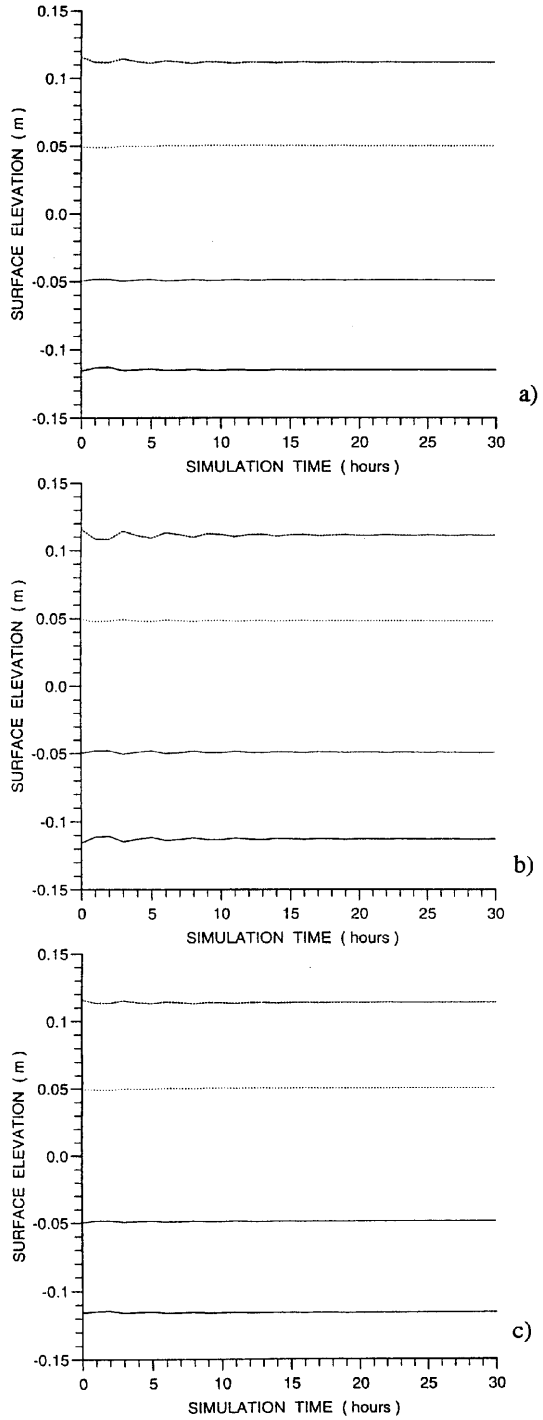

Figure 17. Time series of surface elevation for four stations symmetrically spaced along central cross-section of channel. (a) BM method, (b) original SvK method, (c) modified SvK method

continental slope west of Norway. It is therefore likely that the insight gained from these idealized experiments is applicable to more realistic simulations covering this geographical region. The particular importance of an accurate treatment of the HPG force in the areas of steep bottom slopes stems from the fact that these are regions of strong currents.[25,26] It is, for example, well established that the Norwegian Current, consisting of Atlantic water which enters the Norwegian/Greenland Sea through the Faroe–Shetland Channel, closely follows the isobaths of the continental slope as it flows northward. The core of this current leans onto the continental slope and the main vertical shear is associated with the transition between the saline Atlantic water and the cold Norwegian Sea water below. However, large temporal variability is observed in this current, indicating that instability processes are of importance.[17] Since such processes originate from the basic current, the success of modelling this system numerically depends obviously on the model's capability of accurately simulating the physical state of this current. In order to determine whether the modified SvK method complies with these demands, more realistic numerical experiments must be performed. In association with these experiments, comparison with field measurements should be made.

ACKNOWLEDGEMENTS

I express my gratitude to Professor J. E. Weber and Professor J. Berntsen for their valuable comments and suggestions. The help from Dr. Morten Skogen in getting rid of a serious 'bug' in the computer code is also greatly appreciated. This work has received support from The Research Council of Norway (Programme for Supercomputing) through a grant of computing time.

REFERENCES

1. A. F. Blumberg and G. L. Mellor, 'Diagnostic and prognostic numerical circulation studies of the South Atlantic Bight,' *J. Geophys. Res.*, **88**, 4579–4592 (1983).
2. A. F. Blumberg and G. L. Mellor, 'A description of a three-dimensional coastal ocean circulation model,' in N. Heaps (ed.), *Three-Dimensional Coastal Ocean Models*, Vol. 4, American Geophysical Union, Washington D.C., 1987, pp. 1–16.
3. G. L. Mellor, 'User's guide for a three-dimensional, primitive equation, numerical ocean model', *Program in Atmospheric and Oceanographic Sciences*, Priceton University, Princeton, NJ, 1993.
4. N. A. Phillips, 'A coordiante system having some special advantages for numerical forecasting,' *J. Meteorol.*, **14**, 184–185 (1957).
5. H. Sundqvist, 'On truncation errors in sigma-system models', *Atmosphere*, **13**, 81–95 (1975).
6. H. Sundqvist, 'On vertical interpolation and truncation in connexion with use of sigma-system models', *Atmosphere*, **14**, 37–52 (1976).
7. Z. I. Janjic, 'Pressure gradient force and advection scheme used for forecasting with steep and small scale topography,' *Contrib. Atmos. Phys.*, **50**, 186–199 (1977).
8. F. Mesinger, 'On the convergence and error problems of the calculation of the pressure gradient force in sigma coordinate models,' *Geophys. Astrophys. Fluid Dyn.*, **19**, 105–117 (1982).
9. F. Mesinger and Z. I. Janjic, 'Problems and numerical methods of the incorporation of mountains in atmospheric models', *Lect. Appl. Math.*, **22**, 81–120 (1985).
10. R. L. Hanely, 'On the pressure gradient force over steep topography in sigma coordinate models', *J. Phys. Oceanogr.*, **21**, 610–619 (1991).
11. G. L. Mellor, T. Ezer and L.-Y. Oey, 'The pressure gradient conundrum of sigma coordinate ocean models', *J. Atmos. Oceanic, Technol.*, **11**, 1126–1134 (1994).
12. G. S. Stelling and J. A. Th. M. van Kester, 'On the approximation of horizontal gradients in sigma co-ordinates for bathymetry with steep bottom slopes', *Int. j. numer. methods fluids*, **18**, 915–935 (1994).
13. R. Gerdes, 'A primitive equation ocean circulation models using a general vertical coordinate transformation. 1. Description and testing of the model', *J. Geophys. Res.*, **98**, 14,683–14,701 (1993).
14. G. L. Mellor and A. F. Blumberg, 'Modelling vertical and horizontal diffusivities with the sigma coordinate system', *Mon. Weath. Rev.*, **113**, 1379–1383 (1985).
15. J. M. Gary, 'Estimate of truncation error in transformed coordinate, primitive equation atmospheric models,' *J. Atmos. Sci.*, **30**, 223–233 (1973).
16. D. R. Johnson and L. W. Uccellini, 'A comparison of methods for computing the sigma-coordinate pressure gradient force for flow over sloped terrain in a hybrid theta–sigma model', *Mon. Weather Rev.*, **111**, 870–886 (1983).

17. L. A. Mysak and F. Schott, 'Evidence for baroclinic instability of the Norwegian current', *J. Geophys. Res.*, **82**, 2087–2095 (1977).
18. G. L. Mellor and T. Yamada, 'Development of a turbulence closure model for geophysical fluid problems,' *Rev. Geophys. Space Phys.*, **20**, 851–875 (1982).
19. B. Galperin, L. H. Kantha, S. Hassid and A. Rosati, 'A quasi-equilibrium turbulent energy model for geophysical flows', *J. Atmos. Sci.*, **45**, 55–62 (1988).
20. E. A. Martinsen and H. Engedahl, 'Implementation and testing of a lateral boundary scheme as an open boundary condition in a barotropic ocean model', *Coastal Eng.*, **11**, 603–627 (1987).
21. R. Asselin, 'Frequency filters for time integrations', *Mon. Weather Rev.*, **100**, 487–490 (1972).
22. D. K. Lilly, 'On the computational stability of numerical solutions of time-dependent non-linear geophysical fluid dynamical problems,' *Mon. Wehater Rev.*, **93**, 11–26 (1965).
23. W. R. Holland and L. L. Lin, 'On the generation of mesoscale eddies and their contribution to the oceanic general circulation', *J. Phys. Oceanogr.*, **5**, 642–669 (1975).
24. A. E. Gill, 'Adjustment under gravity in a rotating channel,' *J. Fluid Mech.*, **77**, 603–621 (1976).
25. J. M. Huthnance, 'Extensive slope currents and the ocean-shelf boundary', *Prog. Oceanogr.*, **29**, 161–196 (1992).
26. G. T. Csanady, 'Ocean currents over the continental slope', *Adv. Geophys.*, **30**, 95–203 (1988).

