# The Contour-Advective Semi-Lagrangian Algorithm: Keeping the Balance

D. G. Dritschel & A. R. Mohebalhojeh
Mathematical Institute, University of St Andrews,
St Andrews KY16 9SS, UK

**Abstract**

We review recent efforts to improve the accuracy of both the vortical and gravity-wave components in primitive equation simulations. The present study focuses on the shallow-water equations, in an idealised doubly-periodic geometry, where it is at least conceivable to obtain accurate results over a wide range of flow regimes. We discuss three main topics: (1) the importance of accurately representing potential vorticity, (2) the advantage of using different sets of prognostic variables for maintaining balance and for minimising erroneous gravity-wave production, and (3) the sensitivity to spatial resolution.

# 1 Introduction

Global atmospheric modelling continues to demand ever increasing resolution and computer power. With increasing resolution, improvements are made in forecast skill, but often these improvements are not dramatic given the additional computer cost. Is there an alternative? Here, we discuss a possibility, a promising one in the idealised shallow-water context studied so far, that suggests that increasing resolution is far less worthwhile than simply changing the prognostic dynamical variable set.

The first step in this work was to use the potential vorticity (PV) explicitly, and to ensure its accurate representation by evolving it in a fully-Lagrangian manner, as in Contour Advection (Dritschel, Polvani & Mohebalhojeh 1999; hereafter DPM). The explicit use of PV was shown in DPM to greatly improve the accuracy of shallow-water simulations in comparison to standard pseudo-spectral and semi-Lagrangian treatments. Moreover, it was shown to be hugely cost effective: for a given accuracy, contour advection is a factor of 100 to 1000 times more efficient than these other numerical methods.

The next step was to replace the other shallow-water prognostic variables commonly used (the height and divergence) by other variables (Mohebalhojeh & Dritschel 2000a; hereafter MDa). This idea was motivated by the observation that the divergence fields in the (undamped) Contour-Advection (CA) simulations of DPM contained an unrealistically high level of gravity wave activity. This noise diminished only gradually with increasing numerical resolution. The principal source of this noise was found to come from the height-tendency equation; here, small errors in computing the nonlinear term generate significant erroneous gravity waves. This suggested that using a different set of prognostic variables which avoids the numerical integration of the height-tendency equation might lead to improvements in accuracy. The first such scheme used the divergence $\delta$ and its first time derivative $\delta_t$ as prognostic variables in place of height $h$ and $\delta$. This leads then to a *diagnostic* equation for $h$, given $\delta$ and $\delta_t$. The nice feature of this scheme is that, in the limit of vanishing Froude number, it reduces to quasi-geostrophy, a property not shared by the standard schemes in use. More than that, this scheme does not filter gravity waves, and therefore it is not a balanced model, yet importantly it recovers the underlying balance expected to be there at small Froude

numbers. The numerical results for this scheme show a dramatic improvement in solution accuracy, and the virtual elimination of erroneous gravity waves.

Here we discuss this scheme and its generalisations (further details may be found in the submitted paper: Mohebalhojeh & Dritschel 2000b). All of these schemes are based on hierarchies of balance conditions, yet do not filter the gravity waves. We emphasise that there are great advantages to underpinning a numerical model with variables that respect physical relations between the flow variables, as in balance. The key result is that, in doing this, one improves directly on the accuracy of the large and intermediate scales. It turns out that, as a consequence, there is a dramatic improvement at the small scales as well. That is, our results point to the need for improvements in the modelling of "well-resolved" scales, rather than for finding fixes (like diffusion) that control the behaviour of small scales.

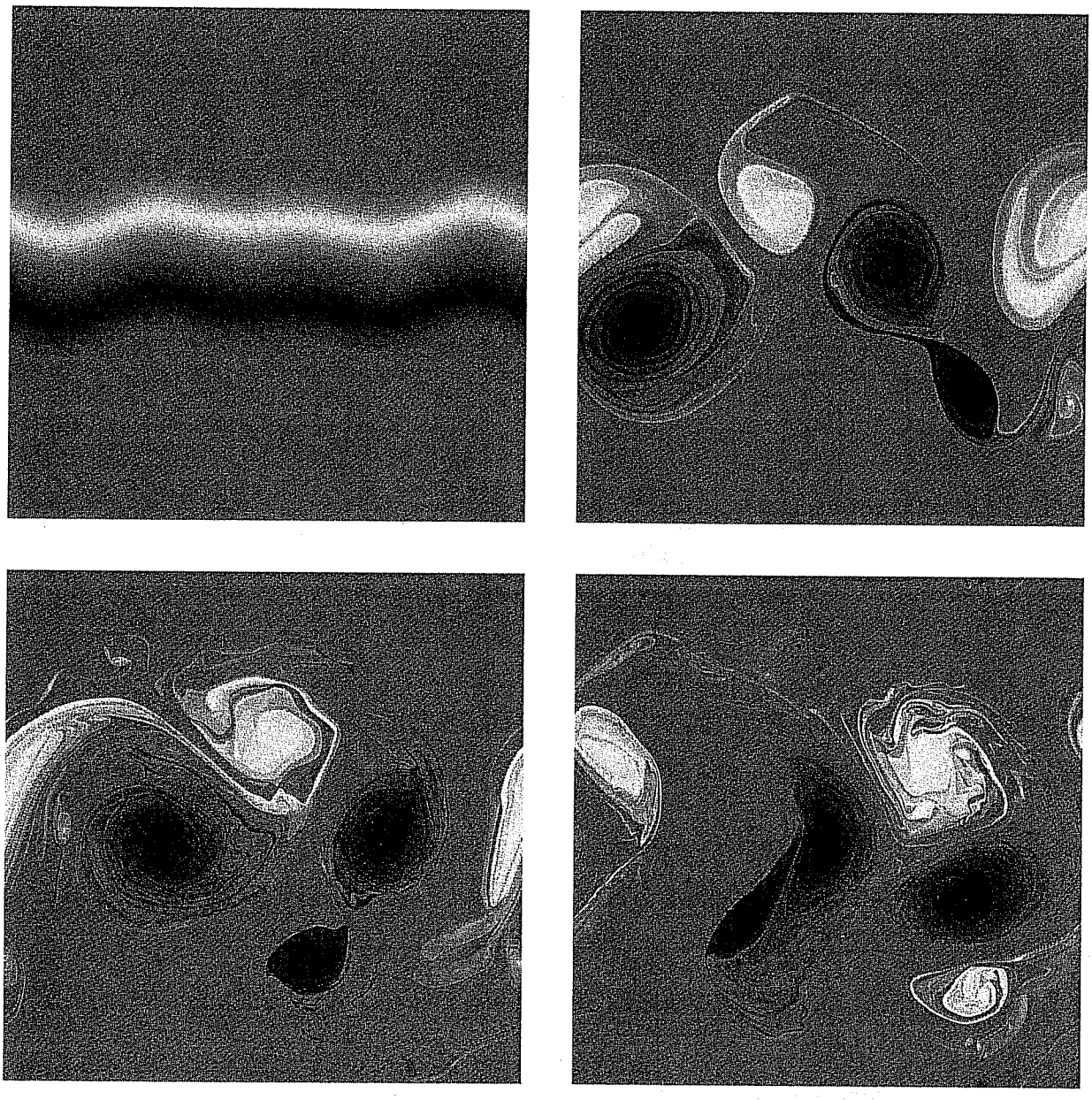
In the next section, we briefly review contour advection and present the basic simulation test case. In §3, we describe the various balance hierarchies that generate new variable sets for shallow-water simulation. In §4 we examine the convergence of these schemes with increasing resolution and summarise the results over a wide range of Froude numbers. In §5, we show results for a modified pseudo-spectral method which uses the prognostic variables $\delta$ and $\delta_t$ (in addition to the relative vorticity $\zeta$), and demonstrate that significant improvements are also possible in such a method. Finally, some conclusions are made in §6, including remarks on the extension of these ideas to the three-dimensional primitive equations.

## 2 The example

We consider the same test case used by DPM, namely that of a jet instability, with maximum initial Rossby number $R_o \approx 0.95$ and Froude number $F_r \approx 0.45$. This is a challenging test case which starts with a smooth distribution of PV yet rapidly develops sharp PV gradients and (arguably typical) abundant fine-scale structure. The PV evolution, computed using the basic Contour Advection code (hereafter $CA_0$) employing PV, height and divergence as the prognostic variables, is illustrated in figure 1 up to $t = 15$ days. This simulation was done on a $512^2$ grid, and is highly accurate as regards the PV evolution, as discussed in DPM. In all the simulations discussed below, we employ a time step that strongly satisfies the CFL condition; we will not discuss the sensitivity to time step in the paper, except to say that larger time steps reduce the accuracy of not only the imbalanced part of the flow but also the balanced part. The two are not independent.

The jet destabilises by forming a vortex street, with a clear asymmetry between cyclonic and anticyclonic vorticity. A detailed description of this case can be found in DPM, along with comparisons with other numerical methods. We show it here only as a point of reference for the results below.

[next page] Snapshots of the PV field at times $t = 0,\ 5,\ 10,\ 15$.



# 3 Balance Hierarchies and New Algorithms

Generally, balance means that there exists a pair of independent functional relations between the three prognostic variables. In view of the advective character of PV, it is most natural to express two of the variables in terms of functionals of the PV, i.e.

$$h = \mathcal{F}[q] \quad \& \quad \delta = \mathcal{G}[q]$$

where $q$ is the PV. Equivalently, one could demand such relations for the time derivatives of $h$ and/or $\delta$, so long as the relations are independent. For instance, a particularly simple choice is to set $\delta = 0$ and $\delta_t = 0$. This is the first member of the *$\delta$-hierarchy*, and its $N$th member is just $\delta^{(N)} = 0$ and $\delta^{(N+1)} = 0$ (where $\delta^{(N)} = \partial^N \delta / \partial t^N$; see McIntyre and Norton 2000). These balance conditions require one to solve a set of nonlinear implicit equations,

$$\begin{aligned}
\mathcal{H}h &= H(fq_\ell + 2J(u,v) - \nabla.(\mathbf{v}\delta) \\
&\quad -\delta^{(1)}) \qquad (1)\\
\mathcal{H}\delta &= \{f\nabla.(\mathbf{v}\zeta) - g\nabla^2\nabla.(\mathbf{v}h)\} \\
&\quad -\{2(J(u,v) - \nabla.(\mathbf{v}\delta)\}^{(1)} \qquad (2)\\
&\quad +\delta^{(2)}\\
\mathcal{H}\delta^{(n)} &= \{f\nabla.(\mathbf{v}\zeta) - g\nabla^2\nabla.(\mathbf{v}h)\}^{(n)} \\
&\quad -\{2(J(u,v) - \nabla.(\mathbf{v}\delta)\}^{(n+1)} \qquad (3)\\
&\quad +\delta^{(n+2)}\\
&\quad (n = 1, \ldots, N-1)
\end{aligned}$$

$$\delta^{(n)} = 0 \ (n = N, N+1)$$

$$\zeta = (h + H)q - f$$

$$\zeta^{(n)} = \{-f\delta - \nabla.(\mathbf{v}\zeta)\}^{(n-1)} \ (n = 1, \ldots, N)$$

$$h^{(n)} = \{-H\delta - \nabla.(\mathbf{v}h)\}^{(n-1)} \ (n = 1, \ldots, N-1)$$

$$\mathbf{v}^{(n)} = \mathbf{k} \times \nabla\nabla^{-2}\zeta^{(n)} + \nabla\nabla^{-2}\delta^{(n)} \ (n = 0, \ldots, N)$$

where $\mathcal{H} = gH\nabla^2 - f^2$ is the modified Helmholtz operator, $H$ is the mean fluid depth, and $q_\ell \equiv \zeta - fh/H$ is the linear PV. This procedure is generally referred to as *PV inversion.*

Two other hierarchies considered in this paper involve the *ageostrophic vorticity* $\gamma \equiv f\zeta - g\nabla^2 h$ and its time derivatives. This variable is sometimes (inappropriately) called the "imbalance". The three hierachies considered are then

- $\delta$-$\gamma$ Balance
$$\delta^{(N)} = 0 \quad , \quad \gamma^{(N)} = 0$$
- $\delta$ Balance
$$\delta^{(N)} = 0 \quad , \quad \delta^{(N+1)} = 0$$
- $\gamma$ Balance
$$\gamma^{(N)} = 0 \quad , \quad \gamma^{(N+1)} = 0$$

How can these ideas help in the numerical solution of the full shallow-water equations? If we return to the original divergence equation, take its time derivative, one can simply derive the following *inertia-gravity wave equation* for the divergence:

$$(\partial^2/\partial t^2 - \mathcal{H})\delta = S$$

where

$$S = g\nabla^2\nabla.(\mathbf{v}h) - f\nabla.(\mathbf{v}\zeta) + 2J(u,v)_t - \nabla.(\mathbf{v}\delta)_t$$

Now, the first member of the $\delta$ hierarchy sets $\delta = 0$ and $\delta_t = 0$; however, we need not do that here but instead directly solve the above "wave equation" as a second-order in time differential equation for $\delta$ , or equivalently as a set of two first-order in time differential equations for $\delta$ and $\delta_t$. That is, *the balance conditions are relaxed.* We still have to solve a *diagnostic* equation for $h$, namely

$$\mathcal{H}h = H(fq_\ell + 2J(u,v) - \nabla.(\mathbf{v}\delta) - \delta_t)$$

and it is this equation, numerically, that helps to keep the balance intrinsic in many shallow-water flows. This is demonstrated in MDa,b and again below.

The above scheme can be generalised to any member of the $\delta$ hierarchy as follows:

$$(\partial^2/\partial t^2 - \mathcal{H})\delta^{(N)} = S^{(N)} \qquad (N \geq 1)$$

where

$$S^{(N)} = \{g\nabla^2\nabla.(\mathbf{v}h) - f\nabla.(\mathbf{v}\zeta) + 2J(u,v)_t - \nabla.(\mathbf{v}\delta)_t\}^{(N)}$$

We solve the above equation for $\delta^{(N)}$ and $\delta^{(N+1)}$ and replace the balance conditions with them. Then, we recover, *diagnostically*, $h$ and $\delta^{(n)}$, $n = 0,\ 1,\ ...,\ N-1$ from Eqs.(1-3).

To summarise, the above demonstrates that there are hierarchies of shallow water CASL algorithms, all mathematically equivalent but having possibly significant numerical differences. Table 1 gives the names, the prognostic, and the diagnostic variables used in the CASL and pseudo-spectral (PS) algorithms examined in DPM, MDa, and here.

| algorithm | prognostic variables | diagnostic variables |
|---|---|---|
| $\mathrm{CA}_0$ | $(q\,,\,h\,,\,\delta)$ | $\zeta$ |
| $\mathrm{CA}_{1,\delta-\gamma}$ | $(q\,,\,\delta\,,\,\gamma)$ | $(h\,,\,\zeta)$ |
| $\mathrm{CA}_{1,\gamma}$ | $(q\,,\,\gamma\,,\,\gamma_t)$ | $(h\,,\,\delta\,,\,\zeta)$ |
| $\mathrm{CA}_{1,\delta}$ | $(q\,,\,\delta\,,\,\delta_t)$ | $(h\,,\,\zeta)$ |
| $\mathrm{CA}_{2,\delta}$ | $(q\,,\,\delta_t\,,\,\delta_{tt})$ | $(h\,,\,\delta\,,\,\zeta)$ |
| $\mathrm{PS}_0$ | $(\zeta\,,\,h\,,\,\delta)$ | |
| $\mathrm{PS}_{1,\delta}$ | $(\zeta\,,\,\delta\,,\,\delta_t)$ | $h$ |

Table 1: Various CASL and PS (pseudo-spectral) algorithms examined in DPM, MDa, and here.

We next present some results indicating these differences in the context of the test case presented in figure 1.

# 4 Balance in CASL algorithms

For details of balance and gravity waves in $CA_0$, $CA_{1,\delta}$, and $CA_{2,\delta}$ we refer the reader to MDa. The computationally desirable feature of $CA_{1,\delta-\gamma}$ and $CA_{1,\gamma}$ is the linearity of the diagnostic equations needed to be solved at every time step, as opposed to the nonlinear equations for $CA_{1,\delta}$ and $CA_{2,\delta}$. The difference between $CA_{1,\delta-\gamma}$ and $CA_{1,\gamma}$ was found, however, to be marginal. For this reason, we omit $CA_{1,\gamma}$ from consideration and concentrate on $CA_{1,\delta-\gamma}$.

In Fig. 2, the ratio $||\delta_{\rm imb}||/||\delta_{\rm b}||$ is shown against resolution for the CASL algorithms $CA_0$, $CA_{1,\delta-\gamma}$, and $CA_{1,\delta}$. In this paper, for any quantity $X$, $||X||$ denotes its $l^2$ norm as estimated by $||X|| = \frac{\pi}{n_g}(\sum_{i,j=1}^{n_g}(X)_{i,j}^2)^{1/2}$, $n_{\rm g}$ being the number of grid points in each direction. Further, $\delta_{\rm b}$ and $\delta_{\rm imb}$ refer to balanced and imbalanced divergence fields as obtained by inverting the instantaneous PV field by means of the third-order $\delta$ balance that employs $\delta^{(2)} = \delta^{(3)} = 0$. For the measure shown, a four-fold increase in resolution is needed for $CA_0$ in order to compete with $CA_{1,\delta}$. At the low-resolution end ($n_{\rm g} = 16$), $CA_0$ is seriously in danger of over-estimating imbalance. It is worth mentioning that for the lowest resolution, we have nearly one grid point per Rossby radius. Such poor resolution of the Rossby radius is common for the high vertical modes in 3-dimensional primitive-equation models.

In Fig. 3, we present the quadratic energy spectra in the *imbalanced* part at $t = 10$ and for very high resolution ($n_{\rm g} = 512$) for $CA_0$, $CA_{1,\delta-\gamma}$, and $CA_{1,\delta}$. The quadratic energy is evaluated according to $(2\pi/n_{\rm g})^2 \sum_{i,j=1}^{n_{\rm g}} 1/2(|{\rm v}|^2 + gh^2)$. One can see that the imbalance at all scales is sensitive to the numerical algorithm. In other words, it is not exclusively the small scale part of the spectrum which is affected by a poor representation of the waves.

In Fig. 4, the results presented in MDa for the behaviour of CASL algorithms $CA_0$, $CA_{1,\delta}$, $CA_{2,\delta}$ against Froude number are complemented by including the results from $CA_{1,\delta-\gamma}$. At small Froude number limit, $CA_{1,\delta-\gamma}$ largely circumvents the false generation of imbalance manifested by $CA_0$. Further, even at the highest Froude number covered (0.73), $CA_{1,\delta-\gamma}$ still outperforms, albeit marginally, $CA_0$. This is important because of the need for an algorithm capable of covering a greater range of Froude numbers without loss of accuracy at small Froude numbers limit.

This property of $\mathrm{CA}_{1,\delta-\gamma}$ comes at minimal cost. The resulting diagnostic relation for $h$ is linear and can be solved for example with a fast multi-grid solver. Overall, the main conclusion is that $\mathrm{CA}_{1,\delta-\gamma}$ provides a good compromise between accuracy, efficiency, and robustness.

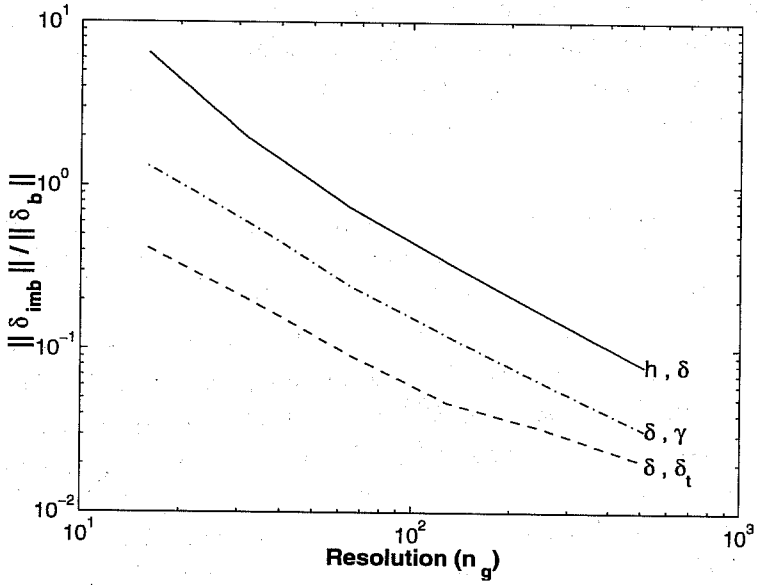



Figure 2: The ratio $||\delta_{\mathrm{imb}}||/||\delta_{\mathrm{b}}||$ for the CASL algorithms with prognostic variables $(q\,,\,h\,,\,\delta)$ (solid), $(q\,,\,\delta\,,\,\gamma)$ (dash-dotted), and $(q\,,\,\delta\,,\,\delta_t)$ (dashed).

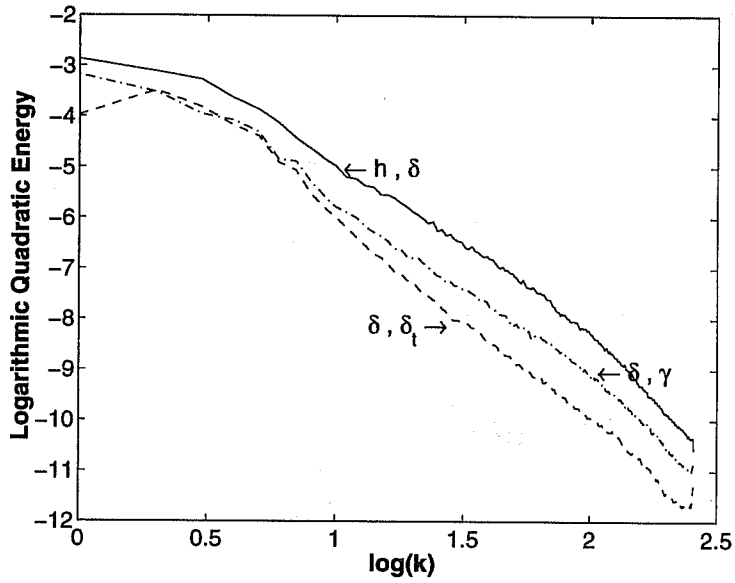



Figure 3: The quadratic energy spectra in the *imbalanced* part for the CASL algorithms with prognostic variables $(q\,,\,h\,,\,\delta)$ (solid), $(q\,,\,\delta\,,\,\gamma)$ (dash-dotted), and $(q\,,\,\delta\,,\,\delta_t)$ (dashed).

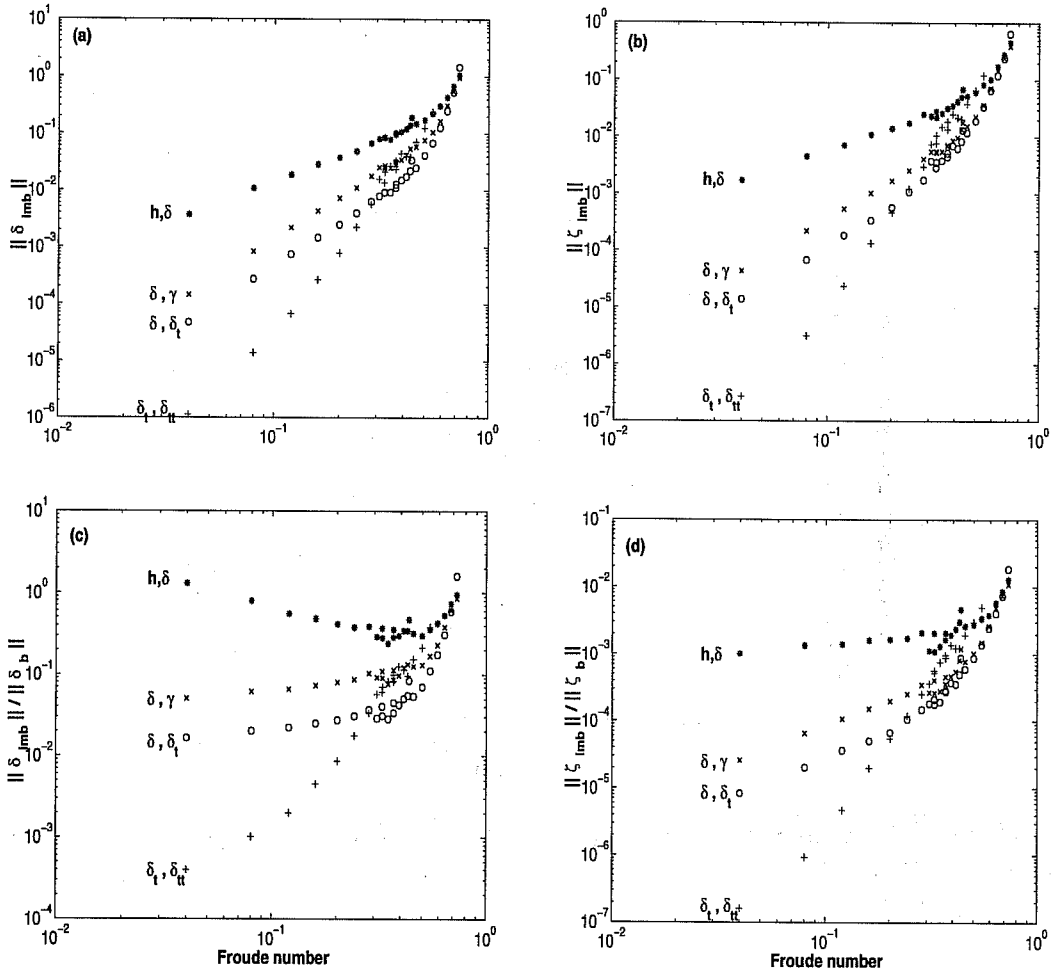



Figure 4: Measures of imbalance for $CA_0$ (∗), $CA_{1,\delta}$ (∘), $CA_{2,\delta}$ (+), and $CA_{1,\delta-\gamma}$ (×). The measures are (a) $||\delta_{\rm imb}||$, (b) $||\zeta_{\rm imb}||$, (c) $||\delta_{\rm imb}||/||\delta_{\rm b}||$, and (d) $||\zeta_{\rm imb}||/||\zeta_{\rm b}||$. See also figure 9 of MDa.

# 5 Balance in pseudo-spectral algorithms

The foregoing ideas on maintenance of balance can be applied to conventional pseudo-spectral algorithms as well. As a preliminary comparison, in Fig. 5 we show the time evolution of $||\delta_{\rm imb}||$ and $||\zeta_{\rm imb}||$ in the pseudo-spectral algorithms $\rm PS_0$ and $\rm PS_{1,\delta}$ together with the corresponding results for the CASL algorithms $\rm CA_0$ and $\rm CA_{1,\delta}$. The resolution is $n_{\rm g} = 128$. The low level of imbalance in the pseudo-spectral algorithms relative to their CASL counterparts results from the action of explicit diffusion of small-scale vorticity (see MDa). In terms of $||\delta_{\rm imb}||$, there is little gain by going from $\rm PS_0$ to $\rm PS_{1,\delta}$. But the difference in $||\zeta_{\rm imb}||$ is substantial. The almost continuous growth of imbalanced vorticity in $\rm PS_0$ in conjuntion with a continuous decay of balanced vorticity due to diffusion is clearly due to the mishandling of the balanced dynamics. In fact, for much longer simulations (200 days) we have not observed any bound on the growth of imbalanced vorticity. The same patern of growth happens for imbalanced height (not shown).

What are the consequences of a poor representation of these small amplitude gravity waves on the balanced dynamics itself? This is an important question for which we can provide at least a partial answer. Let us examine for this purpose the broad scale field of perturbation height $h$, i.e. the pressure field for the SW equations. In Fig. 6, we present $h$ at $t = 15$ for $\rm CA_0$, $\rm CA_{1,\delta}$, $\rm PS_0$, and $\rm PS_{1,\delta}$, all using $n_{\rm g} = 128$. Notice the merging of two anticyclonic vortices and also the filling of the cyclonic centres in $\rm PS_0$ simulation. Remarkably, $\rm PS_{1,\delta}$ shows better agreement with $\rm CA_0$ and $\rm CA_{1,\delta}$. In other words, the elimination of the spurious growth of imbalanced vorticity can improve the performance of conventional pseudo-spectral algorithms. To better understand the problem with $\rm PS_0$, we have performed two more $\rm PS_0$ simulations: (a) 5-day and (b) 10-day simulations starting from the fully-developed states given by $\rm CA_0$ at $t = 5$ and $t = 10$, respectively (Fig. 7). One can see clearly the cumulative effect of errors due to the misrepresentation of small amplitude gravity waves in the 10-day simulation. The significance or otherwise of the associated loss of predictability compared with other sources of errors needs a separate study.

The same effect can be observed in the CASL algorithms. The height field at $t = 15$ for $\rm CA_0$ at two higher resolutions $n_{\rm g} = 256$ and $n_{\rm g} = 512$ are shown in Fig. 8 (cf. top panel of Fig. 6). Although it is hard to judge by direct inspection, in certain

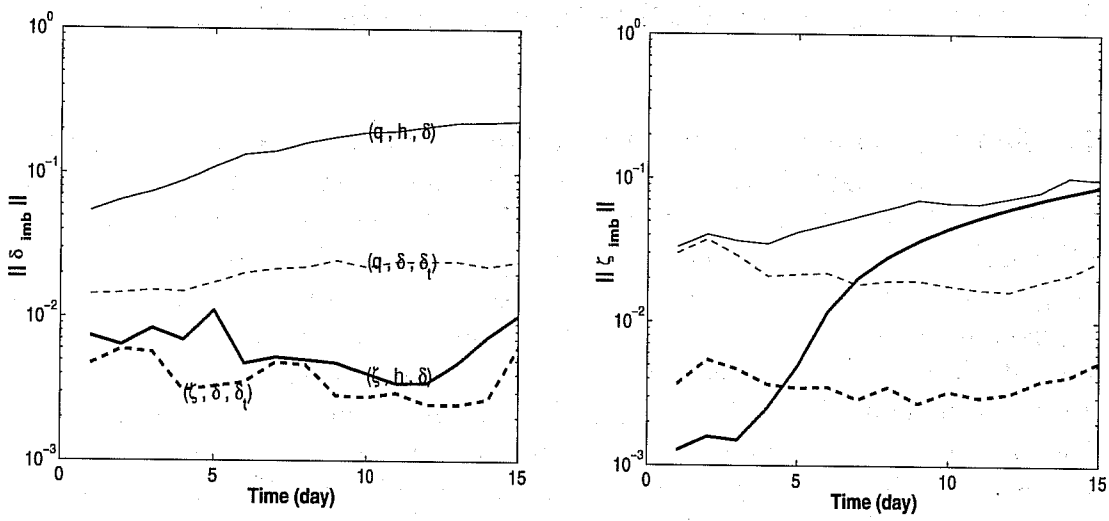



Figure 5: The $l^2$ norms of imbalance versus time for the divergence (left) and the vorticity (right) fields. The results shown are for CASL algorithms $\mathrm{CA}_0$ (solid), $\mathrm{CA}_{1,\delta}$ (dashed); the pseudo-spectral algorithms $\mathrm{PS}_0$ (thick solid), and $\mathrm{PS}_{1,\delta}$ (thick dashed).

features $\mathrm{CA}_{1,\delta}$ at $n_g = 128$ clearly outperforms $\mathrm{CA}_0$ at the same resolution. These include the position of the centre of the bottom-right cyclone and the downward trough of the top-left cyclone. Compared with non-PV based pseudo-spectral $\mathrm{PS}_0$, however, the overall effect of the erroneous gravity waves is less dramatic. This is not surprising, given the fact that PV is the dynamical quantity that is least affected by the presence of gravity waves, either physical or numerical. This is another advantage of PV-based numerical algorithms.

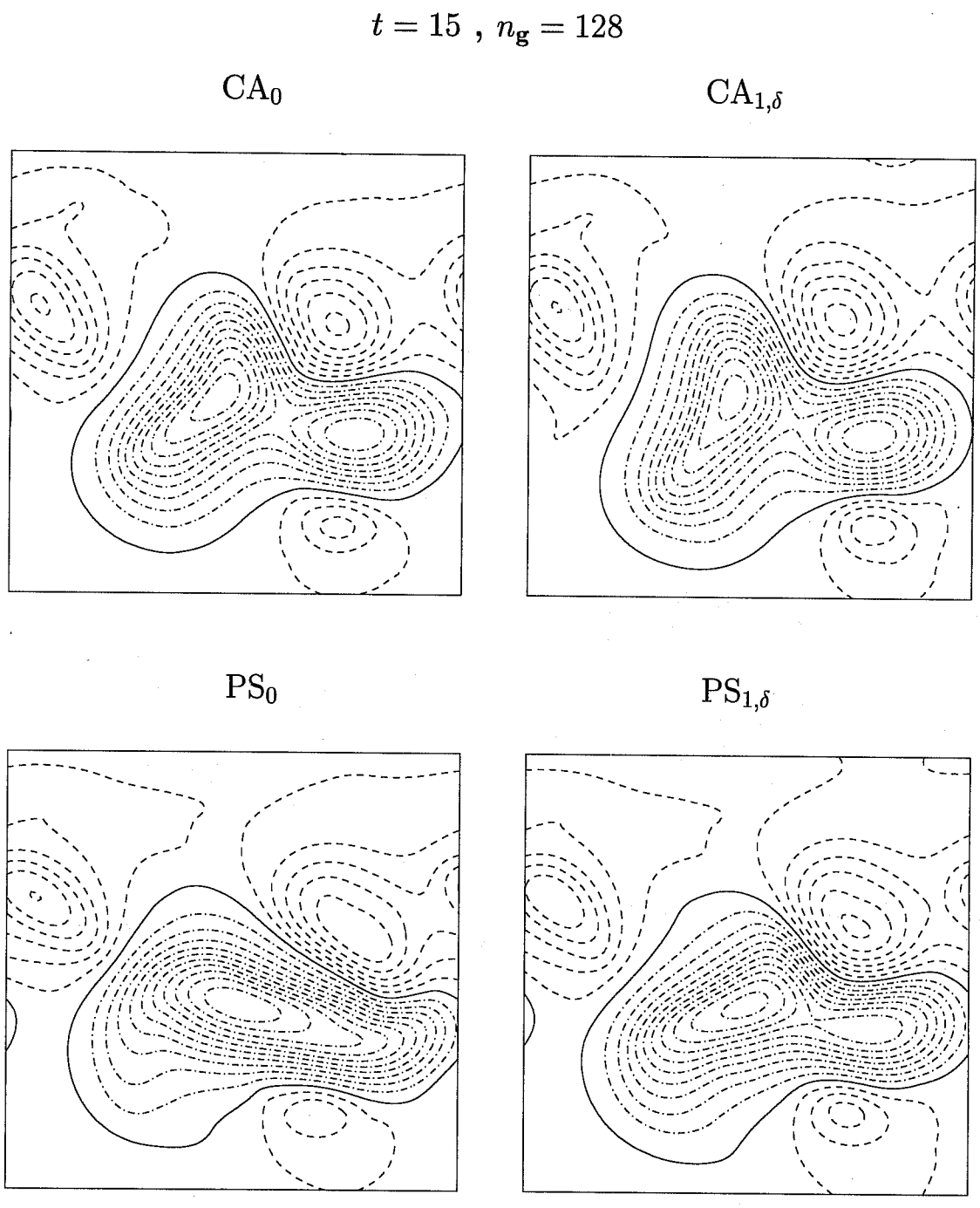



Figure 6: The perturbation height $h$ at $t = 15$ for (top-left) $CA_0$, (top-right) $CA_{1,\delta}$, (bottom-left) $PS_0$, and (bottom-right) $PS_{1,\delta}$. The resolution is $n_g = 128$. The contour inerval is 0.05. The solid line is the zero contour, dashed and dashed-dotted lines are, respectively, for negative and positive values.

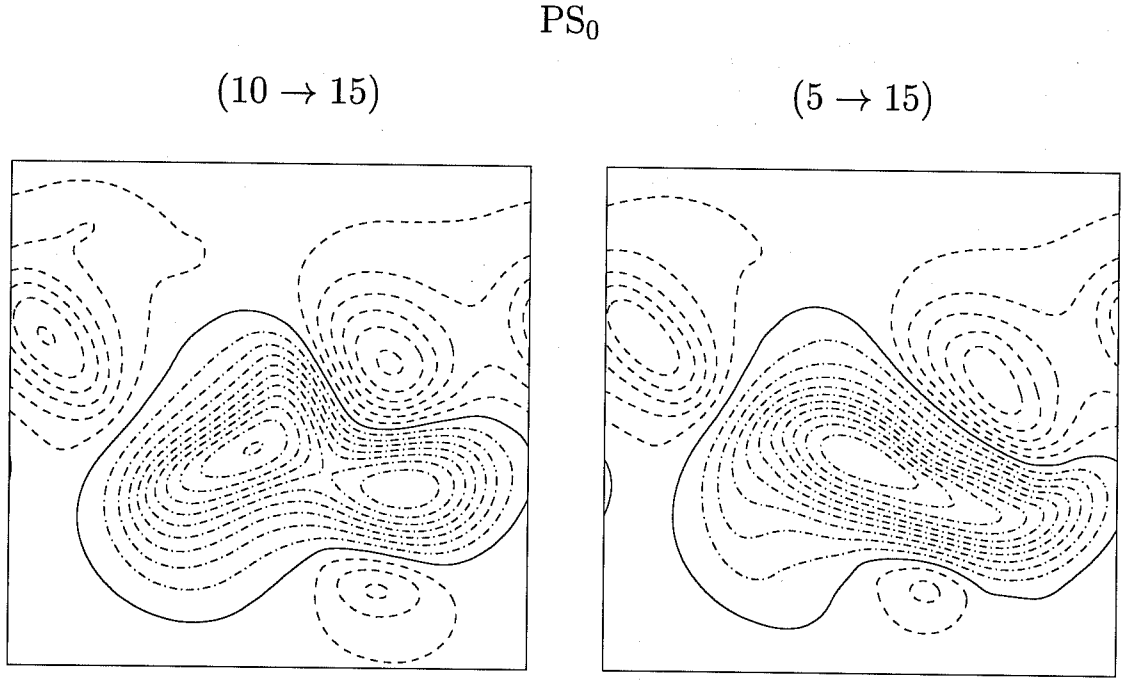



Figure 7: The perturbation height $h$ for (left) 5-day and (right) 10-day integrations by $\mathrm{PS}_0$ starting from the fields given by $\mathrm{CA}_0$ at $t = 10$ and $t = 5$, respectively. The contour interval and line style is the same as fig. 6.

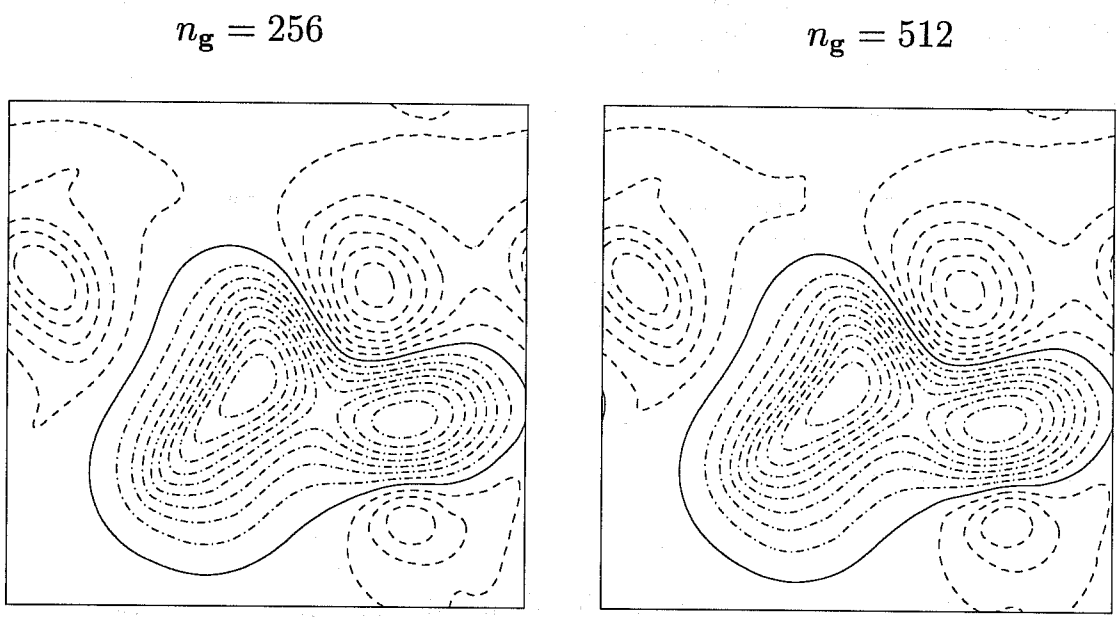



Figure 8: The height field as given by the CASL algorithm $CA_0$ for $t = 15$ and resolutions (left) $n_g = 256$ and $n_g = 512$. The contour interval and line style is the same as fig. 6.

# 6 Conclusion

We have described here recent efforts to improve the dynamical accuracy of SW simulations without compromising computational efficiency. Indeed, our results suggest that going to ever higher resolution with the currently used models is not optimal. A more optimal and viable approach, it appears, is to employ different sets of prognostic variables that respect the underlying physical flow relations, like near balance, that normally exist under typical flow regimes. In particular the use of PV explicitly, as in CA, gives a significant improvement in solution accuracy. But also, changing the other prognostic variables, e.g. $(h\,,\,\delta) \to (\delta\,,\,\gamma)$, has been shown to give significant further improvement.

Currently we are implementing these ideas in the 3D PE context. This is challenging in part due to the difficulty in resolving high vertical modes. However, recent work is encouraging, and we hope to announce new efficient and accurate models in the near future.

# 7 References

Dritschel, D. G., L. M. Polvani, and A. R. Mohebalhojeh, 1999: The contour-advective semi-Lagrangian algorithm for the shallow water equations. *Mon. Wea. Rev.*, **127**, 1551–1565.

McIntyre, M. E., and W. A. Norton, 2000: Potential vorticity inversion on a Hemisphere. *J. Atmos. Sci.*, **57**, 1214–1235.

Mohebalhojeh, A. R., and D. G. Dritschel, 2000: On the representation of gravity waves in numerical models of the shallow water equations. *Q. J. Roy. Meteorol. Soc.*, **126**, 669–688.

Mohebalhojeh, A. R., and D. G. Dritschel, 2000: Hierarchies of balance conditions for the $f$-plane shallow water equations. *J. Atmos. Sci.*, submitted.