Aspects of the numerics of the ECMWF model

Mariano Hortal
European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, UK

1. INTRODUCTION

The forecast model at ECMWF is a spectral model since 1983, when the triangular T63 resolution was introduced substituting the previous grid-point representation. In 1985 the resolution was changed to T106 in the horizontal and in 1986 the vertical resolution was increased from 16 to 19 levels. All these versions of the model used the vorticity-divergence representation and the Eulerian leapfrog semi-implicit time integration scheme.

Further increases in resolution needed a more efficient integration scheme in order to be able to run a 10-day forecast within the operational constraint of about 1 hour elapsed time. Therefore a start was made in 1988 at the development of the semi-Lagrangian integration scheme following the approach of *Ritchie* (1987). The semi-Lagrangian scheme became operational at ECMWF in 1991 using a resolution of T213 in the horizontal and 31 levels in the vertical. The dynamic fields represented in grid-point space (reduced Gaussian grid) are the wind components, the temperature, the specific humidity and the logarithm of the surface pressure. For a full description of this first version of the semi-Lagrangian model see *Ritchie el al* (1995). The reduced Gaussian grid is described in *Hortal and Simmons* (1991) and the basis of the present implementation is as in *Courtier and Naughton* (1994).

Further improvements in efficiency were achieved by the introduction in 1996 of the two-time-level version of the semi-Lagrangian scheme (*Temperton et al* 1999) and the linear Gaussian grid (*Hortal* (1999)) in 1998. The first version of the two-time-level scheme was in fact slightly unstable and therefore the time step size had to be reduced with the consequence that the promised increase in efficiency was not realized until a new, more stable version was introduced at the same time as the linear Gaussian grid.

Even then, the first-order accuracy of the time integration of the physical parameterizations inside the forecast model prevents the usage of the time step which would be allowed by the dynamical (second order accurate) part of the model. A new, second order accurate in space and time treatment of the physical parameterizations is being tested and will become operational soon (*Wedi* (1999)).

On the other hand, a new version of the model using 50 levels in the vertical to span the atmosphere up to 0.1 hPa will become operational at the beginning of 1999 (*Untch et al* (1998), *Untch* (1998b)). In this version of the model the thickness of the vertical layers becomes very small in terms of a pressure-based vertical coordinate and therefore the CFL condition for stability of the Eulerian scheme in the vertical becomes much more restrictive than with the 31-level version, which is not the case for the semi-Lagrangian treatment.
Also the wind at the polar night stratospheric jet (which is not captured by the 31-level version of the model) is much stronger than the wind at the part of the atmosphere covered by the 31-level version and this limits even further the stability of the Eulerian scheme but not the stability of the semi-Lagrangian scheme.

In section 2 the general form of the evolution equations is revised from the point of view of both the Eulerian and the semi-Lagrangian schemes. The reduced and the linear Gaussian grids are described in section 3 while in section 4 the original and the improved two-time-level semi-Lagrangian schemes are presented. All these developments led to an increase of efficiency of the model which is estimated in section 5 as the number of 10-day forecasts which could be run in one day using the available computer resources at the time the improvements were (or will be) implemented and compared with what would have been possible if these improvements had not been made. Section 6 presents some improvements to the semi-Lagrangian scheme which do not lead to efficiency increases but to a more accurate and noise-free behavior of the scheme. Finally some conclusions are drawn in section 7.

2. GENERAL FORM OF THE EVOLUTION EQUATIONS

Any of the evolution equations in the forecast model can be expressed as:

$$\frac{dX}{dt} = R \equiv N + L \tag{2.1}$$

where X is the variable to be forecasted and R is the right hand side of the equation which is split into a linearized part to be treated implicitly in time (L) and the remainder (N) which we will call non-linear term. This form of the evolution equations, where the total time derivative is used, reflects the Lagrangian point of view in which, given a parcel of air, the only independent variable is time and the position of that parcel is a function of time to be determined by the evolution equations themselves. From this perspective the solution procedure is to choose a certain set of air parcels in the atmosphere and follow their evolution during a time step of the model during which a parcel starts at a "departure point" D at the beginning of the time step and arrives at an "arrival point" A at the end of the time step. The evolution equations then tell us how much the quantity X has changed within that parcel during that time step. In the operational procedure, the set of arrival points is chosen to be the grid-points of the Gaussian grid.

In contrast with this procedure, the Eulerian point of view considers the space variables as independent variables and therefore Eq.(2.1) has to be recast in the form:

$$\frac{\partial X}{\partial t} = -U\frac{\partial X}{\partial x} - V\frac{\partial X}{\partial y} - W\frac{\partial X}{\partial z} + N + L \tag{2.2}$$

In this approach the change of the quantity X is computed at each of the fixed Gaussian grid points.

In order to solve Eq. (2.2) we discretize in the time dimension by means of a (second order accurate) centered scheme as follows:

$$\frac{X^{(t+\Delta t)} - X^{(t-\Delta t)}}{2\Delta t} = \left[-U\frac{\partial X}{\partial x} - V\frac{\partial X}{\partial y} - W\frac{\partial X}{\partial z} + N\right]^t + \frac{L^{(t+\Delta t)} + L^{(t-\Delta t)}}{2} \tag{2.3}$$

where all the values correspond to the same grid point and the superindex refers to the moment in time.

A similar discretization can be applied to the semi-Lagrangian Eq. (2.1) as follows:

$$\frac{X_A^{(t+\Delta t)} - X_D^{(t-\Delta t)}}{2\Delta t} = N_M^t + \frac{L_A^{(t+\Delta t)} + L_D^{(t-\Delta t)}}{2} \tag{2.4}$$

Now the quantities referring to time level t+Δt should be evaluated at the arrival point of the trajectory "A", the quantities referring to time level t-Δt should be estimated at the corresponding departure point "D" and the quantities referring to time step t should be given at the middle of the trajectory "M". This is the three-time-level semi-Lagrangian scheme.

As we know the positions of the arrival points, which are chosen to be the Gaussian grid points, we need to find out the positions of the corresponding departure points and the middle points of each trajectory by tracking back the movement of the corresponding air parcel through two time intervals. This can be accomplished by solving the motion equation:

$$\frac{d}{dt}\vec{R} = \vec{V} \tag{2.5}$$

which can be discretized as:

$$\frac{A^{t+\Delta t} - D^{t-\Delta t}}{2\Delta t} = \vec{V}_M^t \tag{2.6}$$

using the same discretization as for Eq. (2.4). In order to solve this discretized equation, the velocity at the present time step has to be interpolated to the middle point of the trajectory which in turn is defined by the unknown position of the departure point D, therefore an iterative procedure, to be described later, is applied.

The main advantage of the semi-Lagrangian procedure over the Eulerian procedure in terms of efficiency comes from the existence of an upper limit in the size of the time step which can be used in the solution of Eq. (2.3) so that this solution is stable, the so called CFL condition. An analysis of the linearized advection equation (Eq. (2.2) with N=L=0) shows that the solution is stable provided the information travels during one time step a space smaller than the space interval used for the computation of the space derivatives, that is to say that the information which reaches a certain grid-point at the future time step t+Δt comes from inside the interval used for the computation of the space derivatives. In the semi-Lagrangian procedure, the information used is the value of the quantity interpolated at the departure point of the trajectory. If in this interpolation we use the values of the grid-points surrounding the computed departure point, the above mentioned stability criterion is always fulfilled no matter how large the size of the time step is and therefore there is no limit on the size of the time step we may use coming from CFL stability considerations. This result is confirmed by a proper stability analysis of the method.

At a horizontal resolution of T213 and 31 levels in the vertical (T213L31), the limit of stability of the Eulerian scheme is Δt~ 3 minutes and even smaller than this for the 50-level version. In contrast, the three-time-level semi-Lagrangian version of the model runs stably using a time step of Δt=15 minutes and more at both vertical resolutions.

In comparing the efficiency of the semi-Lagrangian with the Eulerian scheme one has to take into account that the solution procedure in the case of the semi-Lagrangian involves the computation of the trajectories and the interpolations to the departure and middle point of different fields. This amounts to some 20% of the cost of the model itself but this increased cost is more than compensated by the allowed increase of the size of the time step by a factor of 5 in the case of T213L31, but may not be the same for other (lower) resolutions.

As mentioned above, the computation of the departure points involves an iterative procedure which can be described as follows: Let V_0 be the vector wind at the present time step t at the arrival point "A" of the trajectory of a parcel of air (the Gaussian grid points by definition). If that parcel of air is travelling at that constant speed during the previous and the future time steps it would have been at position D_0 at the time t-Δt, with $(A-D_0)=2V_0.\Delta t$. We may then interpolate the velocity at the present time to the middle point between A and D_0 and recompute the trajectory using this more accurate approximation to V_M. The new departure point is D_1 and we may repeat the procedure until it hopefully converges. It can be shown that the procedure converges if the time step is small enough so that the trajectories do not cross and this depends on the size of the spatial variation of the wind.

According to the above, although the size of the time-step in the semi-Lagrangian procedure is not limited by the CFL stability criterion it is still limited by the (Lifshitz) criterion that the iterative trajectory computation converges. In the conditions given in the atmosphere this criterion is much less restrictive than the CFL condition which applies to the Eulerian scheme.

3. SPECTRAL TECHNIQUE AND THE GAUSSIAN GRID

In a spectral model the fields to be forecasted are represented as a finite (truncated) linear combination of basis functions which, in the case of spherical geometry, are the spherical harmonics $Y_n^m = P_n^m(\mu)e^{im\lambda}$ where λ is the longitude and μ the sine of the latitude. The $P_n^m(\mu)$ are the associated Legendre functions of the first kind. Therefore field $F(\lambda,\mu)$ is represented as:

$$F(\lambda, \mu) = \sum_{(m=-M)}^{M} \sum_{(n=|m|)}^{N} F_n^m P_n^m(\mu) e^{im\lambda} \tag{3.1}$$

Here the spectral coefficients F_n^m can be computed as:

$$F_n^m = \frac{1}{4\pi} \int_{-1}^{1} \int_{0}^{2\pi} F(\lambda, \mu) P_n^m(\mu) e^{-im\lambda} d\lambda d\mu \tag{3.2}$$

given the orthogonality properties of the Legendre functions and of the Fourier functions $e^{im\lambda}$.

In the transform technique, which is the practical way to apply the spectral method to a non-linear problem, the non-linear terms are computed in the physical (or grid-point) space by evaluating every spectral field at a set of locations, doing the non-linear computations at each of these locations or grid points and transforming the result back to spectral space.

The truncation applied in the ECMWF model is a triangular truncation which means that $M=N\neq f(m)$ in Eq. (3.1). Superindex "m" is the zonal wavenumber and subindex "n" is the total wavenumber. The latter is related with the wavelength of a wave in its direction of propagation and the former to its projection on the longitudinal direction. The fact that the upper limit in the summation over the total wavenumber is the same for all the values of the zonal wavenumber means that the shortest wave represented is the same whatever its direction of propagation and therefore its zonal wavenumber. The resolution is therefore isotropic.

The waves resolved in the triangular truncation are the ones inside the grey triangle in wavenumber space of Fig. (3.1)

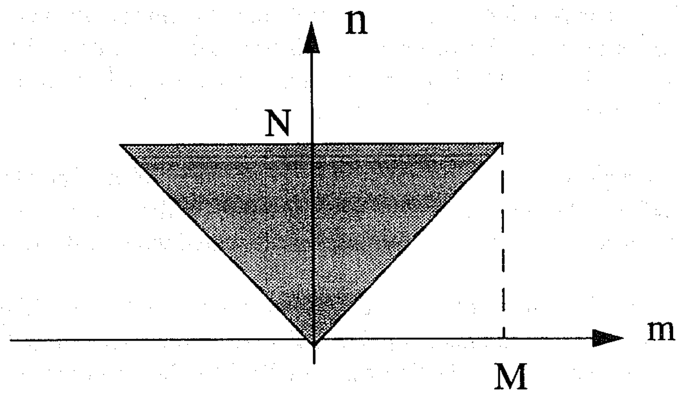

Fig. (3.1) schematic of the waves represented by Eq. (3.1) using triangular truncation.

The direct and inverse spectral transforms represented by Eq. (3.2) and Eq. (3.1) resp. are the exact inverse of each other due to the orthogonality property of the Legendre functions:

$$\int_{-1}^{1} P_n^m(\mu) P_{n'}^m(\mu) d\mu = \delta_{n,n'} \tag{3.3}$$

Now, due to the fact that the amplitude of the Legendre functions tends to zero when the latitude approaches the poles and when the zonal wavenumber increases, there are values of the zonal wavenumber for which:

$$\max_{n,n'} \left| \int_{\mu_m}^{1} P_n^m(\mu) P_{n'}^m(\mu) d\mu \right| \le \varepsilon \tag{3.4}$$

This means that, if we choose ε small enough, for the latitudes $|\mu| \geq \mu_m$ for which Eq. (3.4) holds we can neglect the corresponding zonal wavenumbers m$\geq m$ and still Eq. (3.3) will be nearly true and the direct and inverse spectral transforms be nearly the inverse of each other. This is the basis of the reduced Gaussian grid as described by *Courtier and Naughton* (1994) and implemented operationally at ECMWF with ε=1E-12.

The direct spectral transform Eq. (3.2) can be computed exactly as:

$$F_n^m = \frac{1}{4\pi} \sum_{NG} \sum_{NF} F(\lambda_F, \mu_G) P_n^m(\mu_G) e^{-im\lambda_F} \omega(\mu_G) \tag{3.5}$$

i.e. the integration with respect to the longitude can be substituted by a discrete sum over NF points equally distributed along a longitude row and the integral over latitude can be computed exactly by means of a Gaussian quadrature formula with weights $\omega(\mu_G)$ over NG (approximately equally-spaced) latitude rows which satisfy $P^0_{NG}(\mu)=0$, latitudes called Gaussian latitudes. The set of NFxNG points is called the Gaussian grid.

These discrete sums are exact if NF≥2N and NG≥N and give alias-free computations of quadratic terms using the transform technique if NF≥3N and NG≥3N/2. The first distribution of points is called the linear Gaussian grid and the second one the quadratic Gaussian grid. This quadratic grid is necessary when using the Eulerian form of the equations to avoid aliasing in the quadratic advection terms but it is not needed in the semi-Lagrangian treatment because the advection is not a quadratic term in that version of the equations.

The linear Gaussian grid for a given spectral truncation has therefore a number of points smaller than the quadratic grid by a factor of 3/2x3/2=9/4=2.25 and, as most of the computations in the model are performed in grid-point space, allows a performance of the model better than a factor of 2 as compared with the quadratic grid.

For latitudes close to the poles, the number of wave-numbers retained with the use of the reduced Gaussian grid is smaller than M and correspondingly the number of points NF used for the direct Fourier transform (sum over NF in Eq. (3.5)) can be also less than either 2N for the linear grid or 3N/2 for the quadratic grid. At T213 this amounts to a saving of 33% in the total number of points in the grid. This also means that the distance between grid-points in the longitudinal direction is larger than with the full Gaussian grid and therefore more similar in terms of physical distance to the distance between Gaussian latitudes. This makes the resolution more isotropic in physical space than the full Gaussian grid and therefore more consistent with the isotropy of the triangular spectral truncation. This was the starting idea behind the first implementation of the reduced Gaussian grid as described in *Hortal & Simmons* (1991).

The reduced (quadratic) Gaussian grid for T42 is shown in Fig(3.1)

Reduced T42 Gaussian grid

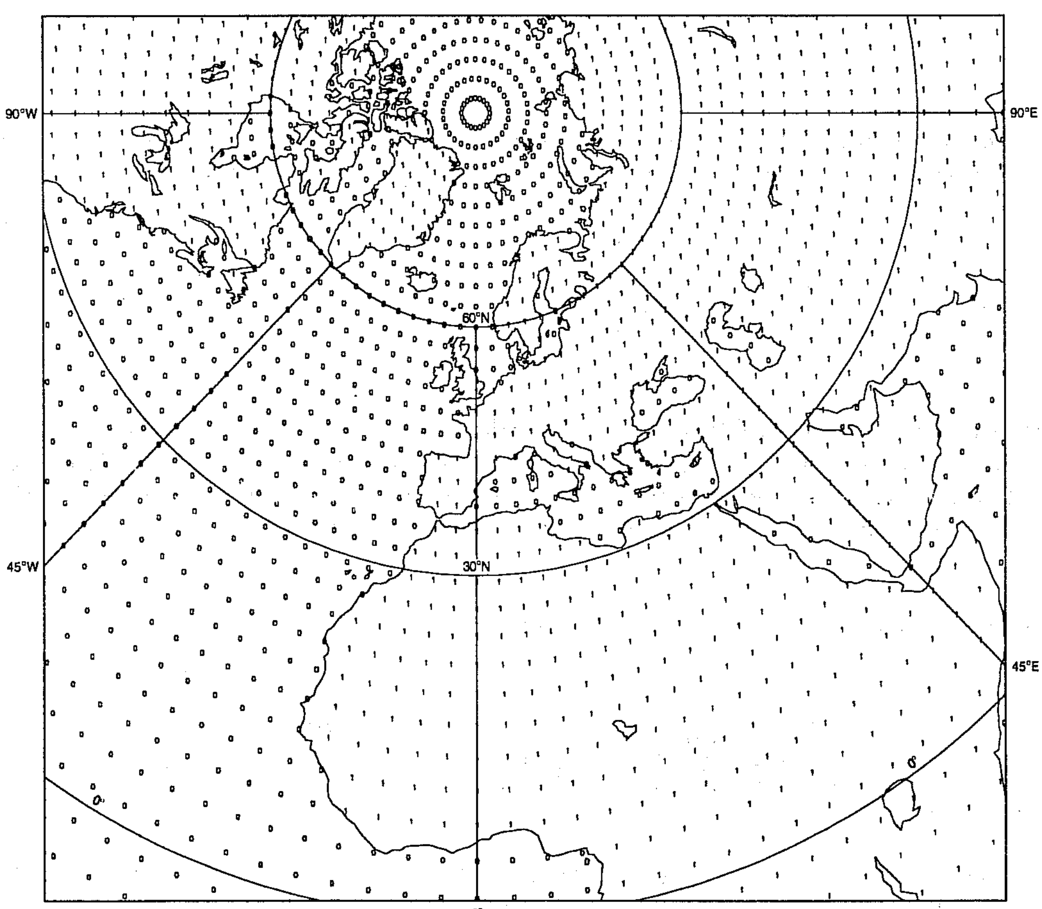

Fig. (3.1) Distribution of points in the T42 (quadratic) reduced Gaussian grid. Symbol 0 indicates a sea point and symbol 1 are land points

4. TWO-TIME-LEVEL SEMI-LAGRANGIAN SCHEME

In principle, a two-time-level semi-Lagrangian scheme provides a doubling of efficiency as compared with a three-time-level scheme, through a procedure which is usually referred to as "doubling the timestep". This is slightly misleading, as in a three-time-level leapfrog scheme the length of each timestep in the usual notation is 2Δt, but successive timesteps overlap by Δt. More precisely, the two-time-level scheme doubles the efficiency by eliminating this overlapping, so that only half the number of timesteps is needed to complete the forecast.

Schematically this can be represented as follows:

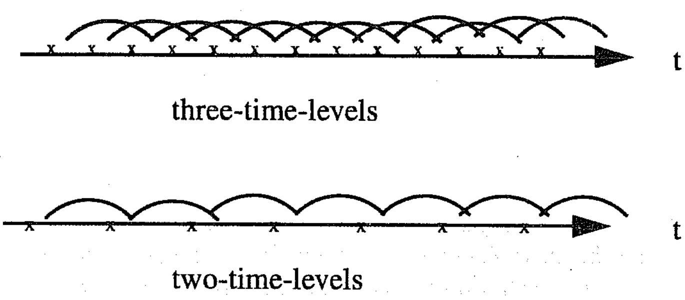

Fig. (4.1) Graphical view of the difference between a three-time-level and a two-tim-level scheme

Viewed in this light, it is clear that the time truncation error can be the same for a three-time-level scheme and for the corresponding two-time-level scheme with a "doubled" timestep. For a two-time-level scheme to be accurate as well as efficient, it is important that second-order accuracy in time be maintained in the trajectory calculations. A simple way to achieve this was independently suggested by *McDonald & Bates* (1987) and *Temperton & Staniforth* (1987), and formed the basis for later developments.

In this first version of the scheme, Eq. (2.4) corresponding to the three-time-level discretization is replaced by

$$\frac{X_A^{(t+\Delta t)} - X_D^t}{\Delta t} = \frac{N_A^{\left(t+\frac{\Delta t}{2}\right)} + N_D^{\left(t+\frac{\Delta t}{2}\right)}}{2} + \frac{L_A^{(t+\Delta t)} + L_D^t}{2} \tag{4.1}$$

This means that the arrival points are still the positions of the selected air parcels at the future time step t+Δt but the departure points are now the positions of these same parcels at the present time-step t. The interpolation of the non-linear terms to the middle point of the trajectory has been replaced by an averaging "along the trajectory" in the space dimension alone. This can also be done in the three-time-level scheme. In order to retain the second-order accuracy in time for the non-linear terms, these are evaluated at the "central time" t+Δt/2 by means of a time extrapolation using the values at times t and t-Δt:

$$N^{\left(t+\frac{\Delta t}{2}\right)} = \frac{3}{2}N^t - \frac{1}{2}N^{(t-\Delta t)} \tag{4.2}$$

Similarly, the discretized displacement equation Eq. (2.6) becomes

$$\frac{A^{(t+\Delta t)} - D^{t}}{\Delta t} = \frac{\vec{V}_A^{\left(t+\frac{\Delta t}{2}\right)} + \vec{V}_D^{\left(t+\frac{\Delta t}{2}\right)}}{2} \tag{4.3}$$

and the velocity at the central time is evaluated by means of a linear extrapolation in time:

$$\vec{V}^{\left(t+\frac{\Delta t}{2}\right)} = \frac{3}{2}\vec{V}^{t} - \frac{1}{2}\vec{V}^{(t-\Delta t)} \tag{4.4}$$

The iterative scheme and first-guess for solving Eq. (4.3) are exactly analogous to those for solving Eq. (2.6) and the choices for the variables X and for the interpolation schemes remain exactly as for the three-time-level scheme. The semi-implicit equations to be solved in spectral space have the same form as for the three-time-level scheme, except that Δt is replaced by $\Delta t/2$. In principle a two-time-level scheme should have no $2\Delta t$ computational mode, and the time-filtering procedure needed in a three-time-level scheme is no longer needed.

Nevertheless, the extrapolation in time given by Eq. (4.4) (and also the one given by Eq. (4.2)) is inconsistent with the semi-Lagrangian point of view in what it treats the time dimension as independent of the space dimensions. This leads to a slightly unstable scheme as shown by *Durran* (personal communication) who also shows that an extrapolation "along the semi-Lagrangian trajectory" gives a stable scheme.

The instability mentioned in the previous paragraph led to some cases of noise in the operational forecast when it was using a time step of 1800 sec. An example of these noisy forecasts can be seen if Fig. (4.2) which shows the forecast at 5 days into the forecast of level 9 temperature from initial data corresponding to the 4th of January 1997

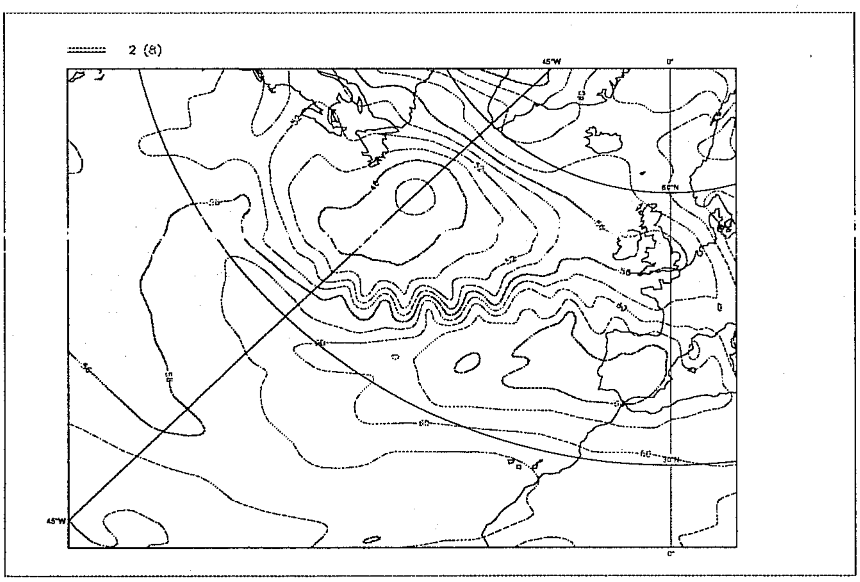

Fig. (4.2) Level 9 temperature forecast for D+5 from 970104

It shows also, even more clearly, in the upper levels of the 50-level version of the model, an example of which is shown in Fig. (4.3)

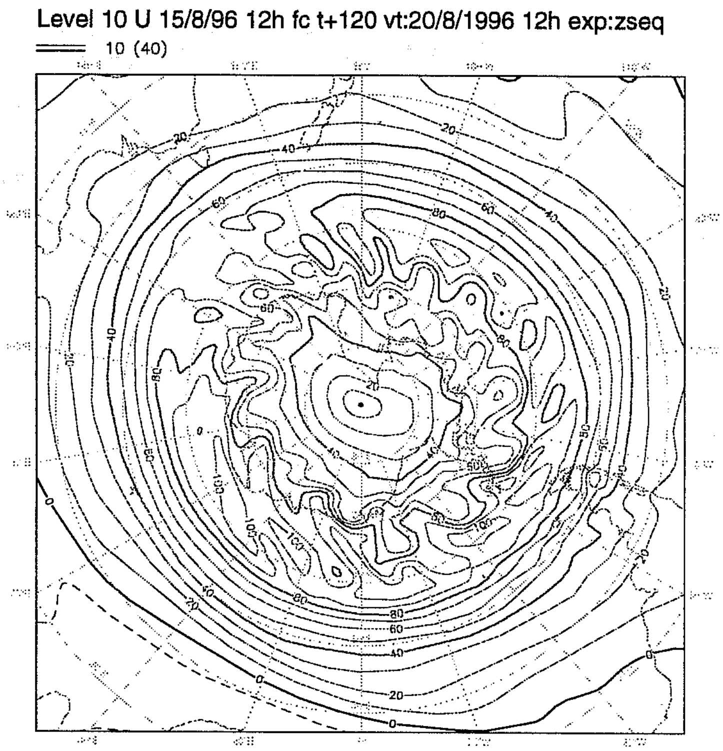

Fig. (4.3) Level 10 (~5 hPa) forecast zonal wind over the southern hemisphere for D+5 from initial data of 15th August 1996 using the 50-level version of the model

These kinds of noise do not appear in the corresponding three-time-level forecasts (not shown) and therefore are attributable to instability in the two-time-level version.

On the other hand the scheme proposed by *Durran* of extrapolating along the semi-Lagrangian trajectory does get rid of the noise but produces forecasts which score much worse than the operational forecasts.

A new scheme had then to be developed which avoids the time extrapolations by using an approach more consistent with the semi-Lagrangian point of view.

Lets expand both the unknown X of Eq.(2.1) and the position vector of Eq.(2.5) by means of Taylor series to second order around the departure point D (corresponding to the present time-step t) of the trajectory of the parcel of air arriving at point A at the next time step t+Δt

$$X_A^{(t+\Delta t)} = X_D^t + \Delta t \cdot \left(\frac{dX}{dt}\right)_D^t + \frac{(\Delta t)^2}{2} \cdot \left(\frac{d^2 X}{dt^2}\right)_{AV} \qquad (4.5)$$

$$\vec{R}_A^{(t+\Delta t)} = \vec{R}_D^t + \Delta t \cdot \left(\frac{d}{dt}\vec{R}\right)_D^t + \frac{(\Delta t)^2}{2} \cdot \left(\frac{d^2}{dt^2}\vec{R}\right)_{AV} \tag{4.6}$$

in both expansions the semi-Lagrangian point of view is adopted that, given the air parcel, the only independent variable is the time dimension. Here subindex AV means some average value along the interval t -> t+Δt and therefore also along the trajectory.

If we keep for the linear terms the same discretization as in Eq. (4.1) and substitute the remainder of the total time derivative in Eq.(4.5) we get:

$$X_A^{(t+\Delta t)} = X_D^t + \Delta t \cdot N_D^t + \frac{(\Delta t)^2}{2} \cdot \left(\frac{dN}{dt}\right)_{AV} + \Delta t \cdot \left(L_D^t + L_A^{(t+\Delta t)}\right) \tag{4.7}$$

and similarly from (4.6)

$$\vec{R}_A^{(t+\Delta t)} = \vec{R}_D^t + \Delta t \cdot \vec{V}_D^t + \frac{(\Delta t)^2}{2} \cdot \left(\frac{d}{dt}\vec{V}\right)_{AV} \tag{4.8}$$

All the quantities on the r.h.s. of Eq. (4.7) are known except the time derivative of the non-linear terms (the linear terms at the future time step make (4.7) implicit exactly the same as in the previous scheme). In the same way, the only unknown on the r.h.s. of Eq. (4.8) is the time derivative of the velocity.

Notice that Eq. (4.8) is the equation of an uniformly accelerated motion with initial speed V^t_D and acceleration $(dV/dt)_{AV}$. The trajectory is no longer a straight line on a plane or the arc of a great circle over the sphere and the middle point of the trajectory is not the average between the departure and the arrival points.

In order to make an estimate of the time derivative of N in Eq. (4.7) it is important to realize that, in the case on the momentum equations, the contribution to the pressure gradient term coming from the orography is arbitrarily distributed between the linear and the non-linear terms depending on the value chosen for the reference temperature profile for the semi-implicit scheme. A proper equilibrium of the terms involved in the pressure-gradient computation can only be achieved if the linear and the non-linear terms are evaluated at the same geographical points. This suggests the following approximation for dN/dt:

$$\left(\frac{dN}{dt}\right)_{AV} \approx \frac{N_A^t - N_D^{(t-\Delta t)}}{\Delta t} \tag{4.9}$$

and similarly we can approximate dV/dt in Eq. (4.8) as:

$$\left(\frac{d}{dt}\vec{V}\right)_{AV} \approx \frac{\vec{V}_A^{t} - \vec{V}_D^{(t-\Delta t)}}{\Delta t} \tag{4.10}$$

Substituting these approximations in Eq. (4.7) and (4.8) we get:

$$X_A^{(t+\Delta t)} = X_D^{t} + \frac{\Delta t}{2}\left(\left(2N_D^{t} - N_D^{(t-\Delta t)}\right) + N_A^{t}\right) + \Delta t \cdot \left(L_D^{t} + L_A^{(t+\Delta t)}\right) \tag{4.11}$$

and

$$\vec{R}_A^{(t+\Delta t)} = \vec{R}_D^{t} + \frac{\Delta t}{2}\left(\left(2\vec{V}_D^{t} - \vec{V}_D^{(t-\Delta t)}\right) + \vec{V}_A^{t}\right) \tag{4.12}$$

The expression in Eq. (4.12) looks at first sight very similar to the scheme used by some authors to integrate Eq. (2.5) using the trapezoidal rule, i.e.

$$\vec{R}_A^{(t+\Delta t)} = \vec{R}_D^{t} + \frac{\Delta t}{2}\left(\vec{V}_A^{t+\Delta t} + \vec{V}_D^{t}\right) \tag{4.13}$$

and estimating $\vec{V}_A^{t+\Delta t}$ by means of a forward extrapolation in time as:

$$\vec{V}_A^{t+\Delta t} \approx 2\vec{V}_A^{t} - \vec{V}_A^{t-\Delta t} \tag{4.14}$$

but in this case the extrapolated value is used at the arrival point. As this last scheme also involves extrapolation in time, it should also be unstable as (4.3)-(4.4). When tested on the same case of Fig. (4.3) it gave a forecast even noisier than our original scheme. This is to some extent to be expected as the extrapolation in time projects to one timestep into the future while ours projected only to half that far.

Notice that, although $2V^t_D - V_D^{t-\Delta t}$ looks like an extrapolation in time of the velocity, the way we arrived at expression (4.12) only involved a transposition in time of the time derivative of the velocity and not an extrapolation in time. Keeping the total time derivative constant is the idea behind the "extrapolation along the semi-Lagrangian trajectory" of *Durran* and therefore it should give a stable scheme. For that reason it has been called Stable Extrapolation Two-Time-Level Scheme (SETTLS). Nevertheless, the application of the linear stability analysis to this scheme does not give a stability independent of the CFL number as *Durran's* scheme does.

The new scheme was tested at T213 to produce a forecast from the fourth of January 1997. The resulting level 9 temperature forecast for day 5 is shown in Fig. (4.4)

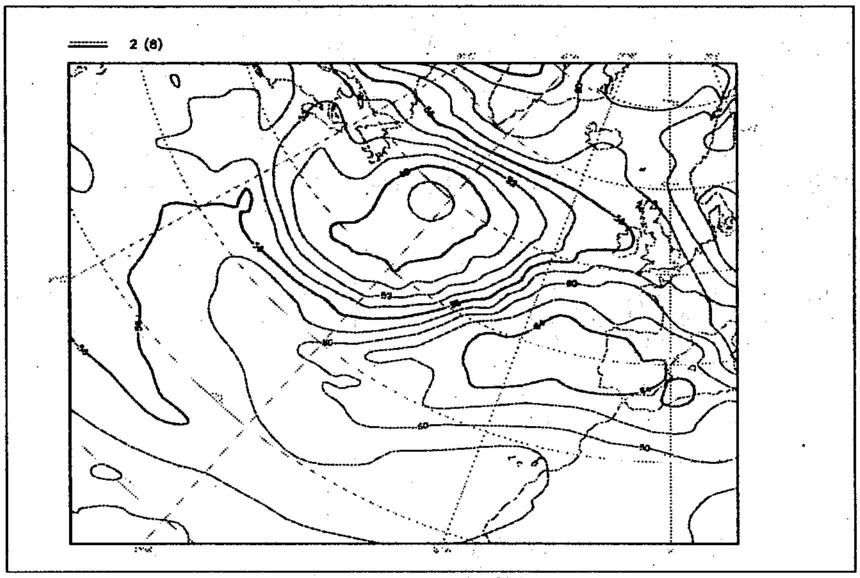

Fig. (4.4) Level 9 temperature forecast for D+5 from 970104 using the SETTLS scheme

The noise seen in Fig. (4.2) has disappeared and the shape of the isotherms looks now identical to the ones obtained with the three-time-level scheme using half the size of the time step (not shown).

The scheme was also applied to produce a forecast with the 50-level version of the model using the initial data which produced the forecast of Fig. (4.3). The corresponding forecast of zonal wind at level 10 is shown in Fig. (4.5)

The noise in Fig. (4.3) does not show any more indicating that the instability which produced it in the original two-time-level scheme has been removed.

The scheme has been tested also on a series of T159 forecasts using 1 hour time step and on a series of T213 forecasts distributed throughout the year using 15 minutes time step. The scores of the first series as compared with the control forecast of the Ensemble Prediction System (EPS) (T159 using 45 min. time step) and of the second as compared with the operational forecast (T213 with 15 minutes time step) give a neutral impact of the new scheme.

With a view to its operational implementation the scheme was tested in conjunction with the linear Gaussian grid at T319 resolution (T_L319) using 20 minutes time step. The positive results obtained led to its operational implementation at ECMWF on April the 1st 1998.

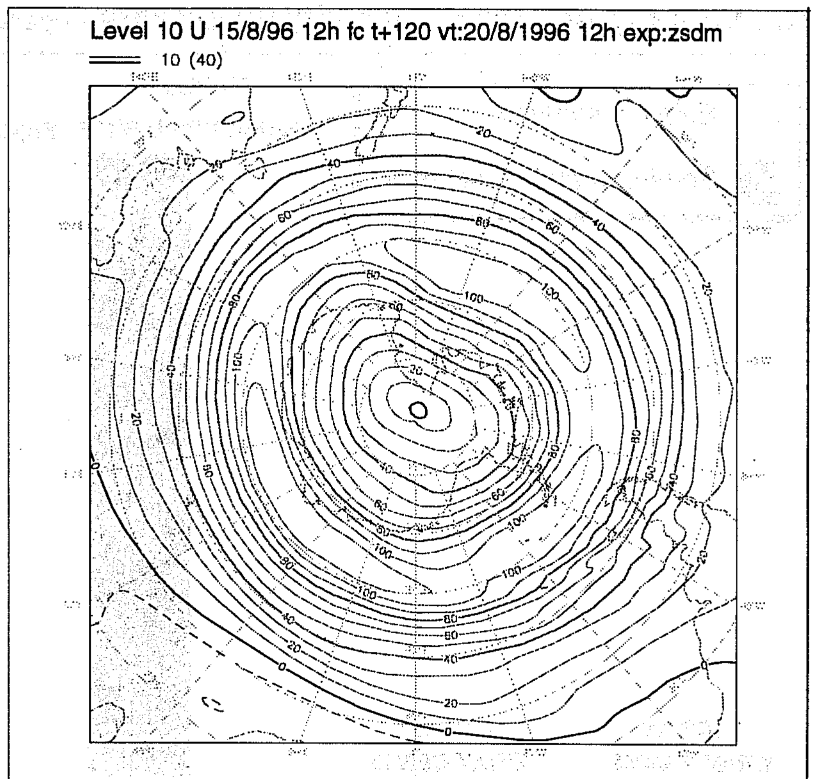

Fig. (4.5) Level 10 (~5 hPa) forecast zonal wind over the southern hemisphere for D+5 from initial data of 15th August 1996 using the 50-level version of the model and the SETTLS scheme.

5. EFFICIENCY INCREASE OF THE FORECAST MODEL

Fig. (5.1) represents the efficiency of the forecast model at the operational resolutions as number of forecast days which could be produced per day of computer time using the whole computer available at the time (dark bars). As a reference to show the improvements brought about by the algorithmic changes introduced in the model, the light bars represent the efficiency which could have been achieved by a model at the operational resolution, using the numerical algorithm used in the operational model in 1987.

The improvement (ratio between the dark and the light bars) from 1987 to 1992 represents only the contribution of the semi-Lagrangian scheme which at T213L31 resolution can be estimated as a factor of four. The further improvement in the efficiency from 1992 to 1997 accumulates the effect of the reduced Gaussian grid and the two-time-level version of the semi-Lagrangian scheme which can be estimated as a factor of three, giving a total factor of 12 as the efficiency increase of the operational model with respect to an Eulerian model at the same resolution in 1997.

The 50-level version of the model which has been under test during 1998 needs a reduction of the size of the time-step in its Eulerian form due to the stronger winds present in the part of the stratosphere resolved by the 50-level model which was not resolved in the 31-level version, as compared with the winds present in the troposphere. The semi-Lagrangian version does not need this reduction of the time-step. This produces a further relative increase of efficiency of the semi-Lagrangian model with respect to its Eulerian counterpart. The same can be applied to the spectral resolution increase from T213 to T319 introduced operationally in 1998. These relative increases, together with the improvement brought about by the introduction of the linear Gaussian grid in 1998, can be estimated as a factor of 6 in efficiency increase, giving in 1999 a total improvement by a factor of about 72 brought about by the algorithmic changes introduced in the forecast model since 1987.

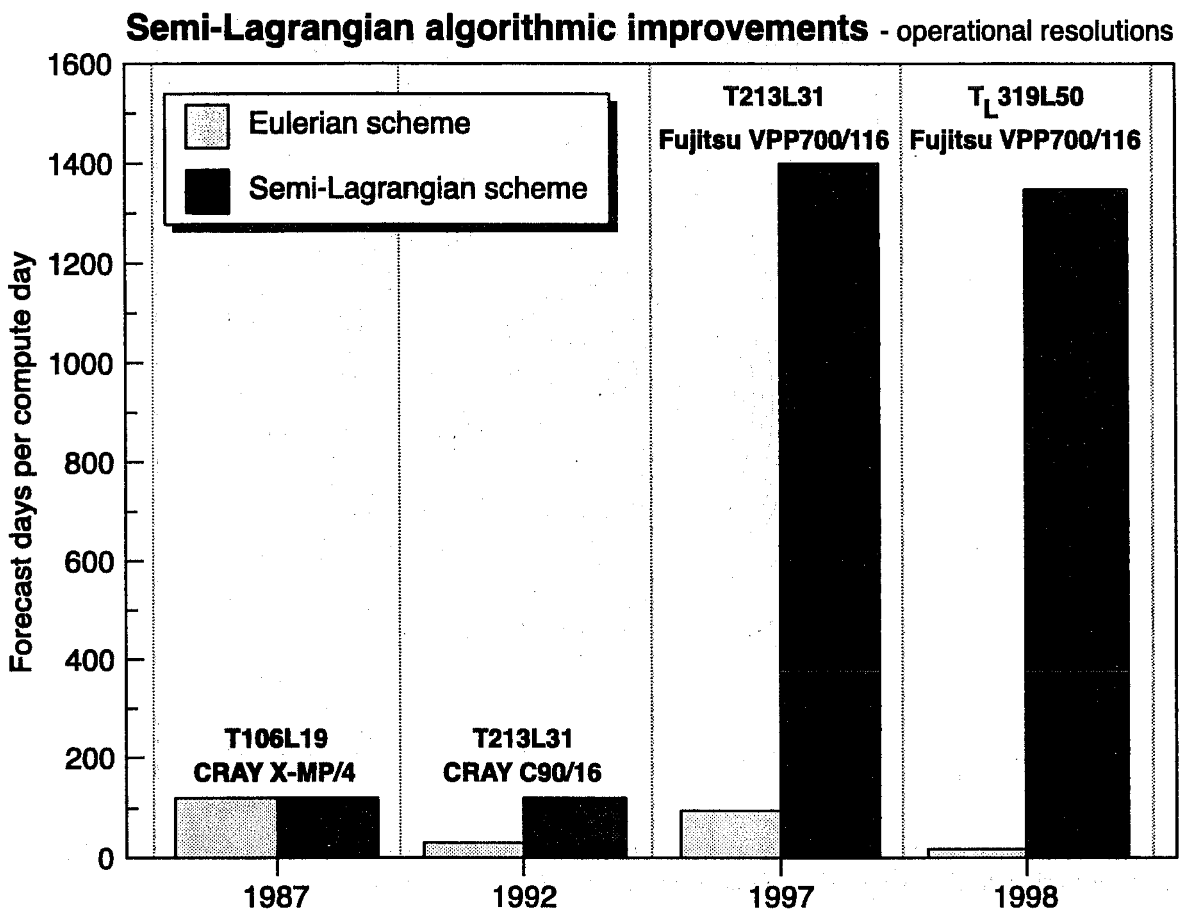

Fig. (5.1) Efficiency improvements in the forecast model. Dark bars represent the number of forecast days per day using the whole computer with the operational model version. Light bars represent what could be achieved with the version of the model used in 1987, run at the same resolution as the operational model.

6. REFINEMENTS TO THE FORMULATION OF THE SEMI-LAGRANGIAN SCHEME

6.1) Quasi-monotone interpolation

In the semi-Lagrangian scheme (Eq. 2.4) the unknown field X corresponding to the previous time step (in a three-time-level scheme) or to the present time step (in a two-time-level scheme) has to be interpolated to the departure point of the semi-Lagrangian trajectory. In order to avoid an excessive smoothing of the field coming from the interpolation, it is customary to use cubic interpolation for this purpose. At ECMWF a 32-point interpolation is used which includes cubic Lagrange polynomial interpolation in the longitudinal direction, in the latitudinal direction and in the vertical direction.

For smooth enough fields a cubic interpolation does not produce any problems, however if the field is insufficiently smooth the use of high-order interpolation procedures may be counterproductive, the advected field developing oscillations in the vicinity of low-smoothness points. The solution proposed by *Bermejo and Staniforth* (1992) blends low- and high-order approximate interpolations in such a way that in regions where the field is smooth, the blended solution is exactly a cubic polynomial interpolation and in regions where the field is not sufficiently smooth to justify a high-order interpolation, the blended solution represents a low-order (monotonic) interpolation.

In *Bermejo and Staniforth's* proposal, illustrated in their paper with a bi-dimensional interpolation, the interpolated value of the field obtained by a bicubic interpolation is subsequently limited to lie inside the interval defined by the values of the field at the four grid-points surrounding the departure point. At ECMWF it was found more effective to use the quasi-monotone procedure at each one-dimensional cubic interpolation by limiting the interpolated values to lie inside the interval defined by the values used for the interpolation at the two grid-points surrounding the departure point. Therefore, if in a certain one-dimensional cubic interpolation the interpolated value is larger than both the values at the surrounding grid-points, the interpolated value is substituted by the larger of the two values at the grid-points and similarly by the smaller of the two if it is smaller than both of them.

The use of cubic interpolations in the horizontal within the semi-Lagrangian procedure at ECMWF forecast model without the quasi-monotone limiter produces an excessive gravity-wave activity which affects negatively the verification scores of the forecast. On the other hand the use of the quasi-monotone limiter in the vertical interpolation of the temperature field produces an error in the tropopause by smoothing out the temperature minimum.

The quasi-monotone interpolation is therefore used operationally in the ECMWF model in the horizontal direction only to interpolate the wind components and the temperature and in the three dimensions to interpolate the humidity, avoiding in this way the production of negative humidities.

6.2) New form of the continuity equation

Ritchie and Tanguay (1996) discuss the problems associated with the incorporation of orographic forcing into a semi-implicit semi-Lagrangian model and examine a possible modification of the form of the continuity equation to alleviate these problems.

Their proposal results basically in substracting on both sides of the continuity equation the time derivative of the approximate contribution to the logarithm of the surface pressure coming from the presence of orography.

Let us write the continuity equation as

$$q \equiv \ln p_s \Rightarrow \frac{dq}{dt} = [RHS]$$

and split the field q in two parts

$$q = q^* + q' \qquad \text{where} \quad q^* = -\frac{\Phi_s}{R\bar{T}}$$

here q* represents the contribution to q coming from the presence of the orography Φ_s in an isothermal atmosphere with temperature $\bar{T}$, and it does not depend explicitly on time. The continuity equation can then be written as

$$\frac{d}{dt}(q - q^*) = [RHS] + \frac{1}{R\bar{T}} \cdot \vec{V} \cdot \nabla\Phi_s$$

and the extra term appearing on the right hand side can be treated together with the rest of the terms [RHS].

Ritchie and Tanguay show that a forecast made with this version of the continuity equation has much better noise characteristics than the original model. This result can be understood as follows: the quantity advected in the new form of the continuity equation and which therefore has to be interpolated cubically to the departure point of the semi-Lagrangian trajectories is (q-q*) which is much smoother than the original quantity q. The cubic interpolation is therefore more accurate and free from overshoot and undershoot than in the original version.

6.3) New form of the thermodynamic equation

The same idea leading to the new form of the continuity equation can be applied to the thermodynamic equation. For this purpose let us define the following quantity:

$$T_C = \left(p_s \cdot \frac{\partial p}{\partial p_s} \cdot \frac{\partial T}{\partial p}\right)_{ref} \frac{\Phi_s}{R\bar{T}}$$

where the quantity in brackets is computed for a reference atmosphere for each of the vertical levels of the model.

Tc represents approximately the dependence in the real atmosphere of the variation of the temperature in a model level at points with an underlying orography Φs.

The thermodynamic equation at a model level can be written as

$$\left(\frac{dT}{dt}\right)_k = [RHST]_k$$

if we substract from both sides the total time derivative of the quantity Tc computed for that model level and we take into account that Tc does not depend explicitly on time but it depends on the vertical coordinate we get

$$\left[\frac{d}{dt}(T - T_C)\right]_k = [RHST]_k -$$

$$\frac{1}{R\bar{T}}\left[\left(p_s\frac{\partial p}{\partial p_s}\frac{\partial T}{\partial p}\right)_{ref} \vec{V} \cdot \nabla\Phi_s + \Phi_s\dot{\eta}\frac{\partial}{\partial\eta}\left(p_s\frac{\partial p}{\partial p_s}\frac{\partial T}{\partial p}\right)_{ref}\right]_k$$

The quantity to be advected is now (T-Tc) which is much smoother than T over orography and therefore allows a more accurate cubic interpolation at the departure points of the semi-Lagrangian trajectories.

7. CONCLUSIONS

- The semi-Lagrangian scheme together with the reduced linear Gaussian grid has allowed us to increase the efficiency of the forecast model by a factor larger than 70 since 1991.

- The limited accuracy of cubic interpolation obliges to take special care when a field is not smooth. Possible improvements include a quasi-monotone or shape-preserving limiter and the transformation of the forecast equation to make the field to be advected by the semi-Lagrangian procedure smoother than the original field.

- The extrapolations used for the two-time-level semi-Lagrangian scheme should be made keeping the total time derivative constant rather than the partial time derivative, corresponding to the Lagrangian point of view of considering the space variables as variables dependent on time (for a given air parcel) rather than as independent variables.

REFERENCES

Bermejo R., Staniforth A., (1992): The conversion of semi-Lagrangian advection schemes to quasi-monotone schemes. *Mon. Wea. Rev.* **120**, 2622-2632

Courtier P., and Naughton M. (1994): A pole problem in the reduced Gaussian grid. *Q. J. Roy. Meteor. Soc.* **120**, 1389-1407

Hortal M., and Simmons A. J., (1991): Use of reduced Gaussian grids in spectral models. *Mon. Wea. Rev*, **119**, 1057-1074

Hortal M., (1999): The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. *Q. J. Roy. Meteor. Soc.* (sent for publication)

McDonald A., and Bates J. R., (1987): Improving the estimate of the departure point in a two-time-level semi-Lagrangian and semi-implicit model. *Mon. Wea. Rev*, **115**, 737-739

Ritchie H. (1987): Semi-Lagrangian advection on a Gaussian grid. *Mon. Wea. Rev.*, **115**, 608-619

Ritchie H., Temperton C., Simmons A. J., Hortal M., Davies T., Dent D., Hamrud M., (1995): Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. *Mon. Wea. Rev*, **123**, 489-514

Ritchie H. and Tanguay M., (1996): A comparison of spatially-averaged Eulerian and semi-Lagrangian treatments of mountains. *Mon. Wea. Rev.* **124**, 167-181

Temperton C. and Staniforth A., (1987): An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. *Q. J. Roy. Meteor. Soc.* **113**, 1025-1039

Temperton C., Hortal M. and Simmons A.J., (1999): A two-time-level semi-Lagrangian global spectral model. *Q. J. Roy. Meteor. Soc* (sent for publication)

Untch A., Simmons A.J., Hortal M., Jakob C. and colleagues (1998): Increased stratospheric resolution in the ECMWF forecasting system. *Proc. workshop on chemical data assimilation* KNMI (DeBilt)

Untch A., (1998b): Aspects of stratospheric modelling at ECMWF. *This volume*

Wedi N. P., (1999): The numerical coupling of the physical parameterizations to the "dynamical" equations in a forecast model. *RD technical memorandum No* **274**, ECMWF