

# A finite-element scheme for the vertical discretization of the semi-Lagrangian version of the ECMWF forecast model

By A. UNTCH and M. HORTAL\*

*European Centre for Medium-Range Weather Forecasts, Reading, UK*

(Received 5 September 2003; revised 19 December 2003)

## SUMMARY

A vertical finite-element (FE) discretization designed for the European Centre for Medium-Range Weather Forecasts (ECMWF) model with semi-Lagrangian advection is described. Only non-local operations are evaluated in FE representation, while products of variables are evaluated in physical space. With semi-Lagrangian advection the only non-local vertical operations to be evaluated are vertical integrals. An integral operator is derived based on the Galerkin method using B-splines as basis functions with compact support. Two versions have been implemented, one using piecewise linear basis functions (hat functions) and the other using cubic B-splines. No staggering of dependent variables is employed in physical space, making the method well suited for use with semi-Lagrangian advection.

The two versions of the FE scheme are compared to finite-difference (FD) schemes in both the Lorenz and the Charney–Phillips staggering of the dependent variables for the linearized model. The FE schemes give more accurate results than the two FD schemes for the phase speeds of most of the linear gravity waves. Evidence is shown that the FE schemes suffer less from the computational mode than the FD scheme with Lorenz staggering, although temperature and geopotential are held at the same set of levels in the FE scheme too. As a result, the FE schemes reduce the level of vertical noise in forecasts with the full model. They also reduce by about 50% a persistent cold bias in the lower stratosphere present with the FD scheme in Lorenz staggering (i.e. the operational scheme at ECMWF before its replacement by the cubic version of the FE scheme described here) and improve the transport in the stratosphere.

**KEYWORDS:** B-splines Charney–Phillips staggering Eigenmodes Galerkin method Hat functions Integral operator Lorenz staggering

## 1. INTRODUCTION

Until recently the vertical discretization in the European Centre for Medium-Range Weather Forecasts (ECMWF) global spectral model had not changed significantly from the finite-difference (FD) scheme of Simmons and Burridge (1981), except for the introduction of the semi-Lagrangian scheme for the vertical advection (Ritchie *et al.* 1995). This vertical FD scheme uses Lorenz staggering of the dependent variables (Lorenz 1960) and is only first order accurate for non-uniform spacing of levels.

The finite-element (FE) method is very attractive for constructing an accurate vertical discretization, since it is a Galerkin-type method, which uses as basis functions piecewise continuous functions with compact support (only locally non-zero), instead of globally defined functions as in spectral methods. Finding suitable, globally defined, orthogonal polynomials as basis functions for the vertical of primitive-equation models has proved difficult (Francis 1972; Machenhauer and Daley 1972; Hoskins 1973). Polynomial functions with compact support are a good alternative and allow for great flexibility in dealing with boundary conditions.

FE-based discretizations for the vertical have been tried successfully in several atmospheric models (Staniforth and Daley 1977; Burridge *et al.* 1986; Steppeler 1987). However, an earlier attempt to replace the Simmons and Burridge FD scheme in the operational version of the ECMWF model by a FE scheme (Steppeler 1986) was unsuccessful because of numerical instability.

\* Corresponding author: European Centre for Medium-Range Weather Forecasts, Shinfield Road, Reading, Berkshire RG2 9AX, UK. e-mail: Mariano.hortal@ecmwf.int

The vertical FE scheme developed here is designed to be used with semi-Lagrangian advection. For this reason, no staggering of dependent variables is employed, so all fields (even vertical velocity) are carried on the same set of levels, unlike in Steppeler (1987). Holding the horizontal winds and the temperature on staggered sets of levels as in Charney–Phillips arrangement (Charney and Phillips 1953) has advantages for treating the gravity-wave terms of the forecast equations, but causes difficulties for semi-Lagrangian advection. In contrast, holding the horizontal winds and temperature at the same set of levels as in Lorenz's arrangement gives rise to a computational (zigzag) mode in the vertical distribution of temperature (Tokioka 1978; Hollingsworth 1995; Arakawa and Konor 1996). In this paper, evidence is shown that the FE discretization damps this computational mode.

Only non-local vertical operations are performed in FE space. Products of dependent variables (and of course, the semi-Lagrangian advection) are evaluated in physical space. This is analogous to the spectral transform method used for the horizontal (Ritchie *et al.* 1995). With semi-Lagrangian vertical advection, the only non-local vertical operations to be evaluated are vertical integrals, since the continuity and the hydrostatic equations are used in their integrated forms; this is unlike Staniforth and Daley (1977) who choose to treat the vertical-derivative form of the evolution equations. An ‘integral operator’ is derived in the FE representation by using the Galerkin method and B-splines as basis functions with compact support. The scheme can, in principle, use piecewise continuous polynomials of any degree with compact support as basis functions. The method has been implemented and tested both using linear elements (linear FE scheme) and using cubic spline elements (cubic FE scheme).

After reviewing the hydrostatic forecast equations in the form used at ECMWF in section 2, the derivation of the FE integral operators with linear and cubic elements is described in section 3, and their order of accuracy, when applied to the integration of a smooth function, is estimated. In section 4 the impact of the FE schemes on the forecast model is shown, beginning with the linearized model, and the results compared with the FD schemes in Lorenz and in Charney–Phillips staggering of variables. Results are presented also from extensive assimilation–forecast experiments using the cubic FE scheme, selected as the more accurate scheme for operational implementation. Finally, a summary of the results and discussion are given in section 5.

## 2. GOVERNING EQUATIONS

The evolution equations used in the semi-Lagrangian version of the ECMWF forecast model are (Ritchie *et al.* 1995):

Momentum equation

$$\frac{d\mathbf{V}_H}{dt} + f\mathbf{k} \times \mathbf{V}_H + \nabla_H \Phi + R_d T_V \nabla_H \ln p = P_V + K_V; \quad (2.1)$$

Thermodynamic equation

$$\frac{dT}{dt} - \frac{\kappa T_V \omega}{\{1 + (\delta - 1)Q\}p} = P_T + K_T; \quad (2.2)$$

Continuity equation

$$\frac{\partial}{\partial t} \left( \frac{\partial p}{\partial \eta} \right) + \nabla_H \cdot \left( \mathbf{V}_H \frac{\partial p}{\partial \eta} \right) + \frac{\partial}{\partial \eta} \left( \dot{\eta} \frac{\partial p}{\partial \eta} \right) = 0; \quad (2.3)$$

Humidity equation

$$\frac{dQ}{dt} = P_Q; \quad (2.4)$$

Hydrostatic equation

$$\nabla_H \Phi = \nabla_H \Phi_S + \nabla_H \int_1^\eta \left( -R_d T_V \frac{\partial}{\partial \eta'} \ln p \right) d\eta'. \quad (2.5)$$

Here  $\mathbf{V}_H$  denotes the ‘horizontal’ velocity,  $\nabla_H$  the horizontal gradient operator,  $f$  the Coriolis parameter,  $T_V$  virtual temperature,  $\Phi_S$  surface geopotential,  $\eta$  the (pressure-based) hybrid vertical coordinate (Simmons and Strüfing 1983) and  $\omega$  the ‘pressure’ vertical velocity ( $dp/dt$ );  $\kappa = R_d/C_{pd}$  and  $\delta = C_{pv}/C_{pd}$ , where  $R_d$  is the gas constant for dry air, and  $C_{pv}$  and  $C_{pd}$  are specific heats of water vapour and dry air at constant pressure.  $P$  and  $K$  are the physical parametrization and the horizontal diffusion contributions, respectively, with appropriate subscripts.

The hybrid vertical coordinate  $\eta(p, p_s)$  is a monotonic function of pressure and also depends on surface pressure such that  $\eta = 1$  for  $p = p_s$  and  $\eta = 0$  for  $p = 0$ . Pressure as a function of  $\eta$  is given by

$$p(\eta) = A(\eta) + B(\eta)p_s, \quad (2.6)$$

where the functions  $A(\eta)$  and  $B(\eta)$  have to fulfil the conditions

$$A(1) = 0, \quad B(1) = 1, \quad A(0) = B(0) = 0, \quad (2.7)$$

which ensure that  $p(1) = p_s$  and  $p(0) = 0$ . An equivalent definition of the vertical coordinate, but one better suited for the construction of the FE-based discretization for the vertical with no staggering of variables presented here, is given by the derivative form of (2.6)

$$\frac{\partial p}{\partial \eta} = \frac{dA}{d\eta} + \frac{dB}{d\eta} p_s, \quad (2.8)$$

together with the conditions

$$\int_0^1 \frac{dA}{d\eta} d\eta = 0, \quad \int_0^1 \frac{dB}{d\eta} d\eta = 1, \quad (2.9)$$

which ensure that the integral of pressure from the top of the atmosphere to the surface yields the surface pressure  $p_s$ .

By integrating the continuity equation in the vertical subject to the boundary conditions  $\dot{\eta} = 0$  at  $\eta = 1$  and at  $\eta = 0$ , the following expressions for the vertical velocity  $\omega$ , for  $\dot{\eta} \equiv d\eta/dt$ , and for the rate of change of surface pressure are obtained:

$$\omega = - \int_0^\eta \left( \frac{\partial p}{\partial \eta'} D + \frac{dB}{d\eta'} \mathbf{V}_H \cdot \nabla_H p_s \right) d\eta' + B \mathbf{V}_H \cdot \nabla_H p_s, \quad (2.10)$$

$$\dot{\eta} \frac{\partial p}{\partial \eta} = - \frac{\partial p}{\partial t} - \int_0^\eta \nabla_H \cdot \left( \mathbf{V}_H \frac{\partial p}{\partial \eta'} \right) d\eta', \quad (2.11)$$

$$\frac{\partial}{\partial t} (\ln p_s) = - \frac{1}{p_s} \int_0^1 \nabla_H \cdot \left( \mathbf{V}_H \frac{\partial p}{\partial \eta} \right) d\eta = - \frac{1}{p_s} \int_0^1 \left( D \frac{\partial p}{\partial \eta} + \frac{dB}{d\eta} \mathbf{V}_H \cdot \nabla_H p_s \right) d\eta. \quad (2.12)$$

Here  $D$  is horizontal divergence.

For the semi-Lagrangian advection scheme an expression for the total time derivative of surface pressure is required:

$$\frac{d}{dt}(\ln p_s) = \int_0^1 \left( \frac{dB}{d\eta} \frac{\partial}{\partial t}(\ln p_s) + \frac{dB}{d\eta} \mathbf{V}_H \cdot \nabla_H \ln p_s \right) d\eta, \quad (2.13)$$

where  $\partial(\ln p_s)/\partial t$  is given by (2.12) above.

### 3. FINITE-ELEMENT SCHEME

#### (a) Vertical discretization

In the Simmons and Burridge (1981) vertical discretization, used operationally at ECMWF until January 2002, when the present FE scheme was implemented, the atmosphere (from the surface to  $p = 0$ ) is divided into  $N$  layers defined by the interface pressures, referred to as half levels. (The list of half levels includes the top of the atmosphere ( $p = 0$ ) and also the surface.) The prognostic variables (with the exception of surface pressure) are represented by their values at full-level pressures (which in the model are defined as the average pressure between the adjacent half levels), while the diagnostic quantities such as  $\omega$  and  $\Phi$  are kept at the half levels, i.e. a Lorenz arrangement of dependent variables (Lorenz 1960) is used. Vertical derivatives are computed by means of centred FD algorithms, and vertical integrals are performed by means of the mid-point rule. This scheme will be referred to as the FD scheme (or as the FD-Lorenz scheme where it needs to be distinguished from a FD scheme with Charney–Phillips staggering of variables).

The values of full-level pressures are not explicitly needed in the FD scheme (see Simmons and Burridge (1981) for details). Only the values of the half-level pressures are required and these are defined by:

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2} \cdot p_s \quad \text{for } 0 \leq k \leq N, \quad (3.1)$$

where  $A_{k+1/2}$  and  $B_{k+1/2}$  are specified constants which include  $A_{1/2} = 0$ ,  $B_{1/2} = 0$ ,  $A_{N+1/2} = 0$ ,  $B_{N+1/2} = 1$  in order to fulfil conditions (2.7).

The values of the vertical coordinate  $\eta$  at full- or half-level positions were not explicitly required by the FD discretization with Eulerian advection (Simmons and Burridge (1981)) but have to be specified for the semi-Lagrangian scheme (Ritchie *et al.* 1995). The  $\eta$  values at full levels are also needed for the FE discretization. The following formal definition of  $\eta$  is used in Simmons and Burridge (1981):

$$\eta_{k+1/2} = A_{k+1/2}/p_0 + B_{k+1/2}, \quad 0 \leq k \leq N, \quad (3.2)$$

where  $p_0$  is a constant pressure chosen to be 1013.25 hPa. At full levels, the values of  $\eta$  are given by

$$\eta_k = 0.5(\eta_{k-1/2} + \eta_{k+1/2}) \quad \text{for } 1 \leq k \leq N. \quad (3.3)$$

In the FE discretization all variables—including pressure—are kept at the same set of levels (at full levels), i.e. no staggering of variables is used. It is therefore convenient to use definition (2.8) of the vertical coordinate and prescribe  $(dA/d\eta)_k$  and  $(dB/d\eta)_k$  at full levels (instead of  $A_{k+1/2}$  and  $B_{k+1/2}$  at half levels) such that

$$\left( \frac{\partial p}{\partial \eta} \right)_k = \left( \frac{dA}{d\eta} \right)_k + \left( \frac{dB}{d\eta} \right)_k p_s, \quad (3.4)$$

together with the conditions (2.9). These conditions have to be fulfilled to a good approximation with the *numerical integration scheme used in the vertical discretization*.

The required pressures at full levels can then be obtained by integrating (3.4) over  $\eta$  from the top of the atmosphere to the full levels.

In the semi-Lagrangian version of the evolution equations in the form given in section 2,  $dB/d\eta$  and  $\partial p/\partial\eta$  are the only vertical derivatives. Since the vertical derivative of pressure can be expressed according to (3.4) in terms of the vertical derivatives of  $A$  and  $B$ , which are constant in time and prescribed, only vertical integrals remain to be evaluated. A vertical integral operator based on a FE representation will now be derived.

(b) *Vertical integral operator in finite-element representation*

The majority of the vertical integrals which need to be evaluated are integrals from the top of the atmosphere to the individual model levels and to the surface. An operator for this type of integral, i.e. an operator which returns the integral from the top of the atmosphere to each of the model levels  $\eta_k$  and to the surface  $\eta = 1$ , will be evaluated using a FE representation. The vertical integral in the hydrostatic equation (i.e. from the surface upwards) can be constructed by taking the integral from the top of the atmosphere to the model level in question minus the integral from the top to the surface.

Let  $\{d_i(\eta)\}_{i=K_1}^{K_2}$  and  $\{e_i(\eta)\}_{i=M_1}^{M_2}$  be two complete sets of linearly independent functions of compact support, which can be used as basis functions to expand any function of the vertical coordinate  $\eta$  given in the domain  $[0, 1]$ .

The vertical integral

$$F(\eta) = \int_0^\eta f(x) dx \quad (3.5)$$

can then be approximated as

$$\sum_{i=K_1}^{K_2} C_i d_i(\eta) = \sum_{i=M_1}^{M_2} c_i \int_0^\eta e_i(x) dx, \quad (3.6)$$

where  $C_i$  are the coefficients of the expansion of  $F(\eta)$  as a linear combination of the basis functions  $d_i(\eta)$ , and  $c_i$  are the coefficients of the expansion of  $f(\eta)$  as a linear combination of the basis functions  $e_i(\eta)$ .

The Galerkin procedure can then be applied to (3.6), by multiplying both sides of this equation by each function from a complete set of ‘test functions’  $\{t_j(\eta)\}_{j=N_1}^{N_2}$  and integrating over the vertical domain:

$$\sum_{i=K_1}^{K_2} C_i \int_0^1 d_i(x) t_j(x) dx = \sum_{i=M_1}^{M_2} c_i \int_0^1 \left( t_j(x) \int_0^x e_i(y) dy \right) dx \quad \text{for } N_1 \leq j \leq N_2. \quad (3.7)$$

In matrix form this can be expressed as  $\mathbf{AC} = \mathbf{Bc} \Rightarrow \mathbf{C} = \mathbf{A}^{-1} \mathbf{Bc}$ , provided the inverse of  $\mathbf{A}$  exists. For the choice of basis and test functions we make in subsection (c) this is the case.

Incorporating into the above expression also the transformations from physical space to FE space and back, i.e.  $\mathbf{c} = \mathbf{S}^{-1} \mathbf{f}$  and  $\mathbf{F} = \mathcal{P} \mathbf{C}$ , one obtains  $\mathbf{F} = \mathcal{P} \mathbf{A}^{-1} \mathbf{B} \mathbf{S}^{-1} \mathbf{f} \equiv \mathbf{If}$ . Here  $\mathbf{f}$  and  $\mathbf{F}$  denote vectors in physical space composed of the values of  $f$  and  $F$ , respectively, at the model levels:  $f_i = f(\eta_i)$ ,  $F_i = F(\eta_i)$ ,  $1 \leq i \leq N$ .  $\mathbf{F}$  also includes the value of the integral at the surface of the model. These vectors may include values of the first and higher order derivatives of  $f$  and  $F$  at the boundaries when higher order than linear spline functions are chosen as basis and/or test functions. Details are given in subsection (c) where it is also shown how to compute the projection matrices  $\mathbf{S}$  and  $\mathcal{P}$  for linear and cubic finite elements.

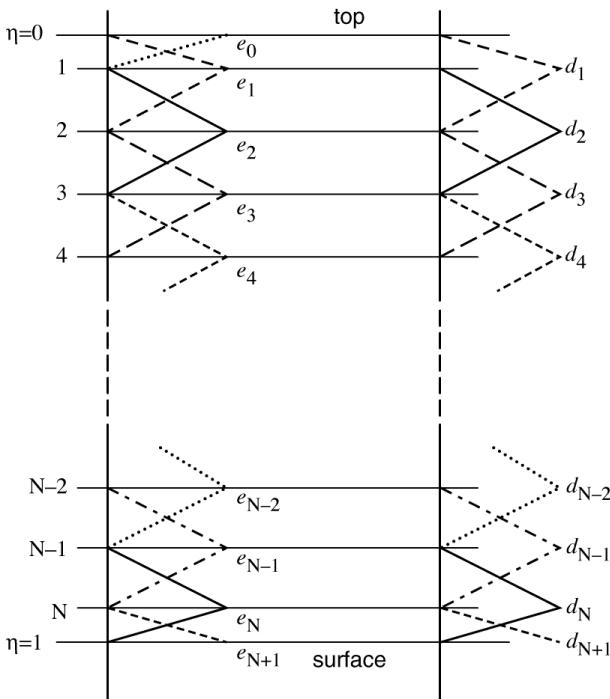


Figure 1. Basis functions for the linear finite-element scheme: the set of hat functions on the left are used as basis functions for expanding the integrand, and the set of functions on the right are the basis functions used for expanding the integral and as test functions. The horizontal lines numbered from 1 to N are the full model levels, while the lines at  $\eta = 0$  and  $\eta = 1$  indicate the top and the bottom of the atmosphere, respectively.

The matrix  $\mathbf{I} \equiv \mathcal{P} \mathcal{A}^{-1} \mathcal{B} \mathcal{S}^{-1}$  is the integration operator in FE formulation which, applied to a function given at full model levels, yields the integrals of this function from the top of the atmosphere to each individual full model level and to the surface. As the integration operator is a matrix acting on the model-level values of the function to be integrated, the scheme can be considered as a high-order FD scheme even if the machinery used in the development is the FE machinery.

### (c) Choice of basis functions

The FE method outlined above has been implemented in two versions, one based on linear splines (hat functions; Fig. 1) and the other based on cubic B-splines (Prenter 1975) (Fig. 2). Spline functions are the most regular polynomial functions for a given order. Although in Figs. 1 and 2 the basis functions are shown for a regular spacing of the model levels, both versions have been implemented for arbitrary spacing of levels.

A hat function centred on level  $\eta_i$  is simply given by:

$$e_i(\eta) = \begin{cases} (\eta - \eta_{i-1})/(\eta_i - \eta_{i-1}) & \text{for } \eta_{i-1} \leq \eta \leq \eta_i \\ (\eta_{i+1} - \eta)/(\eta_{i+1} - \eta_i) & \text{for } \eta_i \leq \eta \leq \eta_{i+1} \\ 0 & \text{otherwise.} \end{cases} \quad (3.8)$$

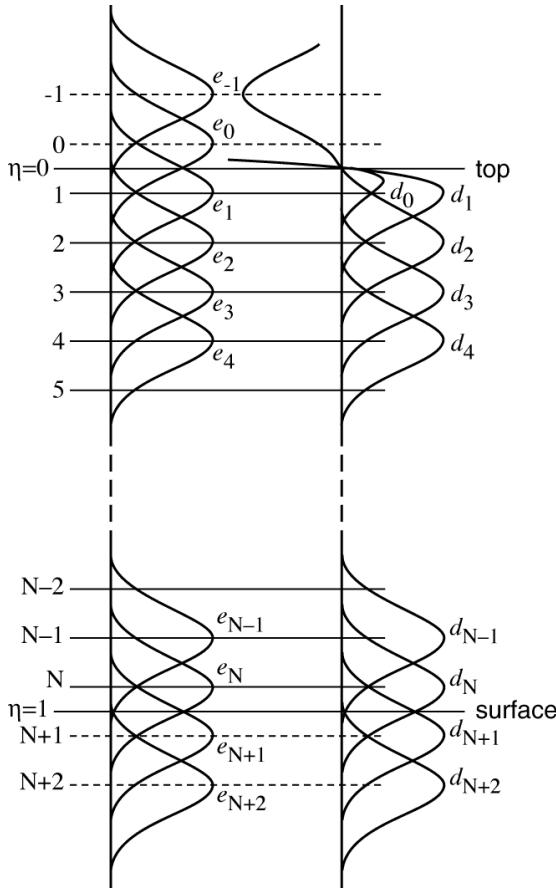


Figure 2. Basis functions for the cubic finite-element scheme. The layout of this figure is the same as Fig. 1. The additional horizontal lines (dashed) above the top of the atmosphere and below the surface represent the levels at which the additional cubic B-splines, needed for a full coverage of the domain  $\eta \in [0, 1]$ , are located.

The equivalent cubic B-spline covers four intervals around  $\eta_i$ , and an analytic expression for a regular spacing of levels is given by (Prenter 1975):

$$B_i(\eta) = \frac{1}{4h^3} \times \begin{cases} (\eta - \eta_{i-2})^3 & \text{for } \eta_{i-2} \leq \eta \leq \eta_{i-1} \\ h^3 + 3h^2(\eta - \eta_{i-1}) + 3h(\eta - \eta_{i-1})^2 - 3(\eta - \eta_{i-1})^3 & \text{for } \eta_{i-1} \leq \eta \leq \eta_i \\ h^3 + 3h^2(\eta_{i+1} - \eta) + 3h(\eta_{i+1} - \eta)^2 - 3(\eta_{i+1} - \eta)^3 & \text{for } \eta_i \leq \eta \leq \eta_{i+1} \\ (\eta_{i+2} - \eta)^3 & \text{for } \eta_{i+1} \leq \eta \leq \eta_{i+2} \\ 0 & \text{otherwise.} \end{cases} \quad (3.9)$$

Here  $h$  denotes the distance between levels for regular spacing.

For the expression of a cubic B-spline for non-regular spacing of levels see the appendix.

In order to cover the whole atmosphere with hat functions,  $N + 2$  linearly independent functions are needed, two more than the number of full model levels (one at

each end of the atmosphere). This is necessary because the top and bottom full model levels do not coincide in our model with the top and the bottom of the atmosphere. With cubic B-splines four more functions than model levels are required, two at each end of the atmosphere because each B-spline covers four intervals. These additional basis functions are chosen, in the linear case, to peak at the top ( $\eta = 0$ ) and the bottom ( $\eta = 1$ ) of the atmosphere, thus having only about half the width of the other linear elements. In the cubic case the option was taken of a more uniform spacing of the additional nodes (levels) required above the first and below the last model level, and the additional basis functions were centred at  $-3\eta_1$ ,  $-\eta_1$ ,  $1 + (1 - \eta_N)$  and at  $1 + 3(1 - \eta_N)$ , i.e. outside of the atmosphere.

In order to have sufficient conditions to determine the cubic spline interpolate of function  $f$  (i.e. the coefficients  $c_j$  in spline space), of which the values are given at the full model levels  $f_i$ ,  $1 \leq i \leq N$ , one needs also to specify values for  $f_{-1}$ ,  $f_0$ ,  $f_{N+1}$  and  $f_{N+2}$ , and for the first or the second derivatives of  $f$  at level  $-1$  and at level  $N + 2$ . The option chosen was to specify  $f_{-1} = f_0 = f_1$  and  $f_{N+1} = f_{N+2} = f_N$ , and also the first derivatives of  $f$  as  $f'_{-1} = 0$  and  $f'_{N+2} = 0$ . This choice is the one consistent with the boundary conditions used in the previous scheme, which was to keep the functions to be integrated constant above the uppermost full level and below the lowermost full level. The impact of different choices for the boundary conditions will be investigated and reported upon in the near future. In the linear version it is sufficient to specify  $f_0$  and  $f_{N+1}$ , and again the option of keeping the function constant at the ends of the domain ( $f_0 = f_1$  and  $f_{N+1} = f_N$ ) was chosen. The coefficients  $c_j$  can now be determined, in the cubic case, from the linear system of equations

$$\left\{ \begin{array}{l} \sum_{i=-1}^{N+2} c_i e'_i(\eta_{-1}) = f'_{-1} \\ \sum_{i=-1}^{N+2} c_i e_i(\eta_j) = f_j \quad \text{for } -1 \leq j \leq N + 2, \\ \sum_{i=-1}^{N+2} c_i e'_i(\eta_{N+2}) = f'_{N+2} \end{array} \right. \quad (3.10)$$

which simplifies in the linear case to

$$\sum_{i=0}^{N+1} c_i e_i(\eta_j) = f_j \quad 0 \leq j \leq N + 1. \quad (3.11)$$

The solution written in vector form is  $\mathbf{c} = \mathbf{S}^{-1}\mathbf{f}$ . Matrix  $\mathbf{S}^{-1}$  is the projection operator from physical space to FE space. Its existence is guaranteed for spline functions, because the matrix  $\mathbf{S}$  of system (3.10) is diagonally dominant. In the linear version this projection operator is the unity matrix, since the hat functions are chosen to have the maximum value of unity at the node to which they belong ( $e_i(\eta_i) = 1$  for all  $0 \leq i \leq N + 1$ ) and they are zero at all other nodes ( $e_i(\eta_j) = 0$  for  $i \neq j$ ), i.e. the basis is ‘interpolatory’ unlike in the cubic case.

In both versions, linear and cubic, the set of test functions  $\{t_j\}$  is chosen to be identical to the basis set for the integral  $\{d_i\}$ .

The basis functions in set  $\{d_i\}$  are identical to those in set  $\{e_i\}$  except for the functions extending into the first half layer (from the top of the atmosphere to the

TABLE 1. RELATIVE INTEGRATION ERROR USING DIFFERENT INTEGRATION SCHEMES FOR THE FUNCTION  $\sin(6\pi x)$ ,  $x \in [0, 1]$

Resolution: number of nodes	FD scheme	Linear FE	Cubic FE	Cubic collocation
60	0.82	$0.14 \cdot 10^{-2}$	$0.90 \cdot 10^{-8}$	$0.14 \cdot 10^{-2}$
90	0.37	$0.27 \cdot 10^{-3}$	$0.32 \cdot 10^{-9}$	$0.27 \cdot 10^{-3}$
120	0.21	$0.85 \cdot 10^{-4}$	$0.31 \cdot 10^{-10}$	$0.85 \cdot 10^{-4}$
150	0.13	$0.35 \cdot 10^{-4}$	$0.55 \cdot 10^{-11}$	$0.35 \cdot 10^{-4}$
Estimated order of the scheme	2	4	8	4

See text for details of integration schemes.

first full model level). These basis functions have been modified to be zero at the top of the atmosphere in order to ensure that the value of the integral at the top of the atmosphere, where the integration starts, is zero ( $F(\eta = 0) \equiv 0$ ). For the linear elements it is straightforward to modify the topmost hat function to fulfil this condition. For the cubic elements three new, linearly independent, functions were constructed by linearly combining the B-spline located at node  $-1$  (function  $e_{-1}$ ) with those at nodes  $0, 1$ , and  $2$ , respectively, (i.e.  $e_0, e_1$ , and  $e_2$ ) such that the new basis functions  $d_0, d_1$ , and  $d_2$  are zero at the top of the atmosphere (notice that  $e_2$  only becomes zero above node  $0$ , not at the top of the atmosphere, where  $\eta = 0$ , which is not a node). Outside the atmosphere the new basis functions are negative, but this does not matter. Due to this ‘boundary’ condition for the integral, the basis set  $\{d_i\}$  has one function less than the set  $\{e_i\}$ .

The projection matrix  $\mathcal{P}$ , which is required in the computation of matrix  $\mathcal{I}$ , is obtained in the same way as matrix  $\mathcal{S}$  but using the basis set  $\{d_i\}$  instead of the set  $\{e_i\}$ .

#### (d) Accuracy of the finite-element integration

In order to check the accuracy of the derived linear and cubic FE integral operators and compare it to that of the FD scheme, all three schemes were applied to the integration of several smooth trial functions for which the analytical integral is known.

Table 1 shows as an example the results obtained for the function  $\sin(6\pi x)$ ,  $x \in [0, 1]$  at the different resolutions  $N = 60, 90, 120$  and  $150$ , where  $N$  is the number of nodes, positioned in the integration domain  $[0, 1]$  at  $(n - 1/2)/N$ ,  $n = 1, 2, \dots, N$  to correspond to the positions of the full levels in our model for uniform spacing.

The relative error (%), with respect to the analytical solution, of the integration over each inter-nodal interval, is compared for the FD scheme, for the linear and the cubic FE schemes, and for the analytic integration of the cubic spline interpolate of the integrand (which will be referred to as the cubic collocation scheme). The values given in Table 1 are averaged values over all inter-nodal intervals located within one wavelength of the integrand (the full wave located in the middle of the integration domain  $[0, 1]$ ).

As expected, the linear FE scheme is seen to be significantly more accurate than the FD scheme, and the cubic FE scheme is again much more accurate than the linear FE scheme. From the values in Table 1, the order of convergence for the different schemes can be estimated; the result is given in the last row of the table. The FD scheme is second order accurate (as was to be expected for uniform spacing of the nodes), and for the linear and cubic FE schemes roughly order four and order eight, respectively, is obtained. The cubic collocation scheme (which uses the same cubic basis functions to fit the integrand with a cubic spline as the cubic FE scheme) is only fourth order accurate like the linear FE scheme.

Applying the Galerkin method in addition to a cubic spline fit of the integrand doubles the order of accuracy of the integration scheme (from four for the cubic collocation to eight for the cubic FE). The same factor of two is obtained by going from linear collocation (which is identical to the FD scheme and therefore of second order) to the linear FE scheme. This suggests the following expression for the order of the FE integration scheme as a function of the degree  $k$  of the basis functions used for the scheme:  $O\{2(k + 1)\}$ . This expression is in good agreement with error estimates derived in text books for FE approximations to the solution of differential equations (with sufficiently smooth right-hand sides) valid at the nodes (see, for example, Oden and Reddy (1976)). In between the nodes the order is only  $k + 1$  according to the theory on finite elements (Oden and Reddy 1976; Becker *et al.* 1981), the same as for the collocation schemes. This was not tested for the schemes presented in this paper, because the integration operators return by design only the integrals at the nodes. The accelerated convergence at the nodes with FE schemes is referred to in the literature as super convergence, and is one reason for the superiority of the FE schemes over the collocation schemes.

#### 4. IMPACT OF THE FINITE-ELEMENT DISCRETIZATION ON THE FORECAST MODEL

##### (a) *Impact on the linear vertical modes*

A first impression of how changes to the vertical discretization affect the vertical structure of the model can be obtained by looking at the vertical modes of the linearized model. In the ECMWF model, the linear vertical modes are used in the semi-implicit time discretization, i.e. the implicit treatment of the linearized adjustment terms of the momentum, the thermodynamic, and the continuity equation. They are computed in the following way.

The linearized form of the forecast equations (2.1) to (2.5) (linearized about an isothermal reference state at rest, of constant temperature  $T_r$  and constant surface pressure  $(p_s)_r$ ) can be written as (Simmons and Burridge 1981; Simmons and Temperton 1997):

$$\frac{\partial D}{\partial t} = -\nabla_H^2 \{\gamma_r T + R_d T_r (\ln p_s)\}, \quad (4.1)$$

$$\frac{\partial T}{\partial t} = -\tau_r D, \quad (4.2)$$

$$\frac{\partial}{\partial t} (\ln p_s) = -v_r D, \quad (4.3)$$

where  $D$  is the horizontal divergence and Greek letters indicate vertical operators, each depending on the reference state ( $r$ ). These operators are defined as:

$$(\gamma_r T)_\eta = - \int_1^\eta R_d T \frac{d}{d\eta'} (\ln p_r) d\eta', \quad (4.4)$$

$$(\tau_r D)_\eta = \kappa T_r \left( \frac{1}{p_r} \right)_\eta \int_0^\eta D \frac{dp_r}{d\eta'} d\eta', \quad (4.5)$$

$$v_r D = \frac{1}{(p_s)_r} \int_0^1 D \frac{dp_r}{d\eta} d\eta. \quad (4.6)$$

Here  $p_r$  is the function defined by (3.1) for the finite-difference scheme or by the vertical integral of (3.4) for the FE scheme, using as surface pressure the reference

surface pressure ( $p_s$ )<sub>r</sub>. In this study the same reference state is used ( $T_r = 350$  K and ( $p_s$ )<sub>r</sub> = 1000 hPa) as in the current operational semi-implicit semi-Lagrangian two-time-level scheme at ECMWF.

Equations (4.1) to (4.3) can easily be combined into the following equation for divergence

$$\frac{\partial^2 D}{\partial t^2} - \nabla_H^2 \Gamma D = 0,$$

where  $\Gamma$  is the vertical structure operator given by  $\Gamma \equiv \gamma_r \tau_r + R_d T_r \nu_r$ .

By expanding  $D$  in terms of eigenfunctions of the Laplace operator  $\nabla_H^2$ , i.e. the spherical harmonics  $Y_n^m$ , one obtains for each spectral component  $(n, m)$

$$\frac{\partial^2}{\partial t^2} D_n^m(\eta, t) + \frac{n(n+1)}{a^2} \Gamma D_n^m(\eta, t) = 0, \quad (4.7)$$

where  $a$  is the radius of the earth. In discretized form with  $N$  levels in the vertical, (4.7) represents a set of  $N$  equations coupled in the vertical by a matrix  $\Gamma$ , the discrete representation of the operator  $\Gamma$ . By projecting onto the eigenspace of  $\Gamma$  one can decouple these equations. The eigenvectors of  $\Gamma$  are the linear vertical (gravity) eigenmodes and the corresponding eigenvalues are proportional to the square of the frequencies of the wave solutions of (4.7).

Matrix  $\Gamma$  combines all vertical integral types that have to be evaluated in the model, so changes to the numerical integration scheme should translate into differences in its eigenmodes, rendering them good test objects for assessing different vertical integration schemes.

The eigenmodes of  $\Gamma$  obtained numerically with the FE integration schemes (linear and cubic) are compared with those obtained with the Simmons and Burridge FD scheme in Lorenz arrangement of variables (FD-Lorenz scheme) and also with the eigenmodes computed with FD in Charney–Phillips arrangement (Charney and Phillips 1953; henceforth the FD-Charney–Phillips scheme), where the horizontal winds and the temperatures are held at different sets of levels (the winds at full levels and the temperatures at half levels).

Figure 3 shows the eigenvalues of matrix  $\Gamma$  as a function of mode number obtained with these four integration schemes at two different vertical resolutions using 50 and 100 levels. The spacing of the levels is equidistant in  $\ln(p)$ , and the top full level is located at 0.1 hPa. As a reference, the eigenvalues computed with 1000 levels are used. At this high resolution the four integration schemes converged well over the range of eigenmodes plotted in Fig. 3, and the four curves are virtually indistinguishable.

For coarser vertical resolutions, the eigenvalues obtained with the FD-Lorenz scheme are smaller than the reference values, i.e. the waves have phase speeds that are too slow and lag behind the reference solution. The error increases rapidly towards the end of the spectrum. In Charney–Phillips arrangement, the eigenvalues are larger than the reference values, i.e. the waves have phase speeds that are too fast, but the error is smaller than with the Lorenz staggering and does not show accelerated growth at the end of the spectrum (see also Beland *et al.* 1983).

The eigenvalues obtained with the two FE schemes are smaller than the reference values and the curves have a similar shape to the FD-Lorenz curve, but the errors are very much smaller. For the cubic FE scheme, the first 70 modes out of 100 (in the 100-level resolution) have an error of less than 1%. For the linear FE scheme the first 47 modes have this level of error, while this is true only for the first 13 modes for the FD-Lorenz scheme and for 17 modes for the FD scheme in Charney–Phillips arrangement. In the 50-level resolution scheme the numbers of eigenvalues with a relative error of less than 1%

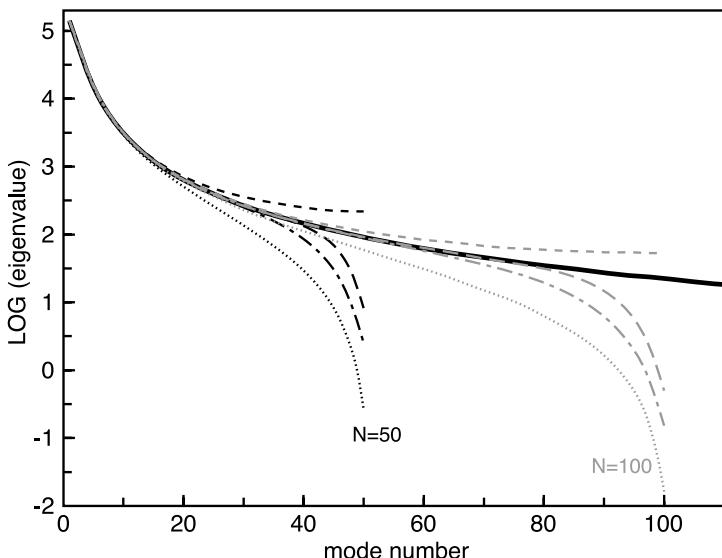


Figure 3. Eigenvalues of the vertical structure matrix for the linearized divergence equation as a function of eigenvector number, using a uniform distribution of levels in terms of  $\ln(p)$ . The full line is obtained with 1000 levels and is used as the reference. Results are shown for: the FD scheme in Lorenz staggering (dotted lines); the FD scheme in the Charney–Phillips staggering (short-dashed line); with the linear FE scheme (dash-dot line); and with the cubic FE scheme (long-dashed line). Lines ending at mode number 50 (100) are for a resolution of 50 (100) levels. See text for further details.

are 35 for the cubic and 24 for the linear FE schemes, only 7 for the FD-Lorenz scheme and 8 for the FD-Charney–Phillips scheme, giving very similar percentage values (i.e.  $100N_{<1\%}/N$ , where  $N$  is the total number of modes) as for the 100-level resolution.

Several resolutions between 50 and 200 levels were tested, and it was found that the percentage of modes with errors less than 1% is remarkably independent of vertical resolution for all four schemes. The average values are: 70% for the cubic, 47% for the linear FE, but only 13% for the FD-Lorenz, and 17% for the FD-Charney–Phillips scheme. The improvement of 34 percentage points brought about by the linear FE scheme over the FD scheme is impressive, and is due to the application of the Galerkin method in constructing the FE integral operator. Using cubic instead of linear elements leads to a further improvement by 23 percentage points.

Figure 3 also shows that an impressive 43 out of 50 modes computed with the cubic FE scheme at the 50-level resolution have smaller errors than the corresponding modes obtained at twice this resolution with the FD-Lorenz scheme. For the linear FE scheme this is true for the first 34 modes. Compared with the FD-Charney–Phillips scheme at 100-level resolution, the first 41 out of 50 modes have smaller error for the cubic and 29 for the linear FE scheme at only half the resolution.

At the tail of the spectrum, the error with the FE schemes is quite large (larger than with the FD-Charney–Phillips scheme for the last three modes of the cubic FE scheme at  $N = 50$  and seven at  $N = 100$ ) but significantly smaller than with the FD-Lorenz scheme. The wavelengths of these modes are, nevertheless, at the limit of the resolution and therefore they cannot be expected to be treated accurately by any discretization scheme.

The spacing of levels in the vertical resolution currently operational at ECMWF for the analysis–forecast system (which we will call here L60) is not equidistant in

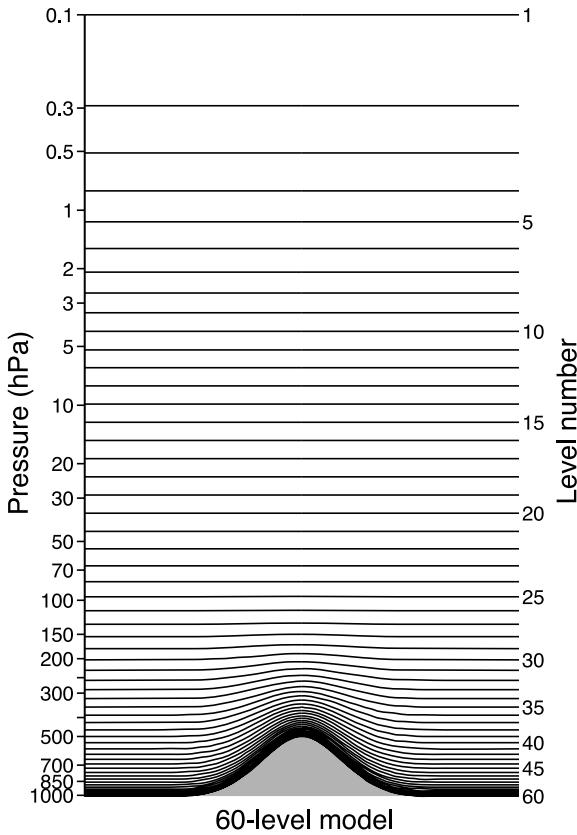


Figure 4. Distribution of the hybrid levels in the currently operational 60-level version of the ECMWF forecast model.

$\ln(p)$  as in these tests. Near the surface the layer thickness is as small as 20 m, and it increases gradually throughout the troposphere and lower stratosphere to reach 1.5 km near 50 hPa. It is kept constant at this value up to about 5 hPa and then grows again towards the top of the model. The distribution of these 60 levels is shown in Fig. 4. (If not specifically stated otherwise, the numerical experiments discussed in this paper were run at this vertical resolution.)

Figure 5 shows the eigenvalues of the matrix  $\Gamma$  as a function of mode number for this L60 resolution, obtained with the four numerical schemes discussed above. Also included is the curve of eigenvalues computed with the cubic collocation scheme; it is almost identical to the curve for the linear FE scheme. This is not surprising since the integration errors found with the convergence tests described in subsection (d) are very similar for the two schemes (see Table 1). With this non-uniform spacing of levels in the L60 distribution, the shape of the eigenvalue curve obtained with the FD-Charney–Phillips scheme differs substantially for large mode numbers from that obtained with regular spacing in  $\ln(p)$ , and is now more similar to the shape of the FD-Lorenz curve. The eigenvalues at the end of the spectrum decrease rapidly, as is the case also for the FD-Lorenz scheme, but stay more than an order of magnitude larger than those of the FD-Charney–Phillips scheme. The kink in the FD-Charney–Phillips curve at around mode 32 is linked to the equidistant spacing of levels in height between 50 and 5 hPa in the L60 distribution. The eigenvalues computed with the two FE schemes are again significantly

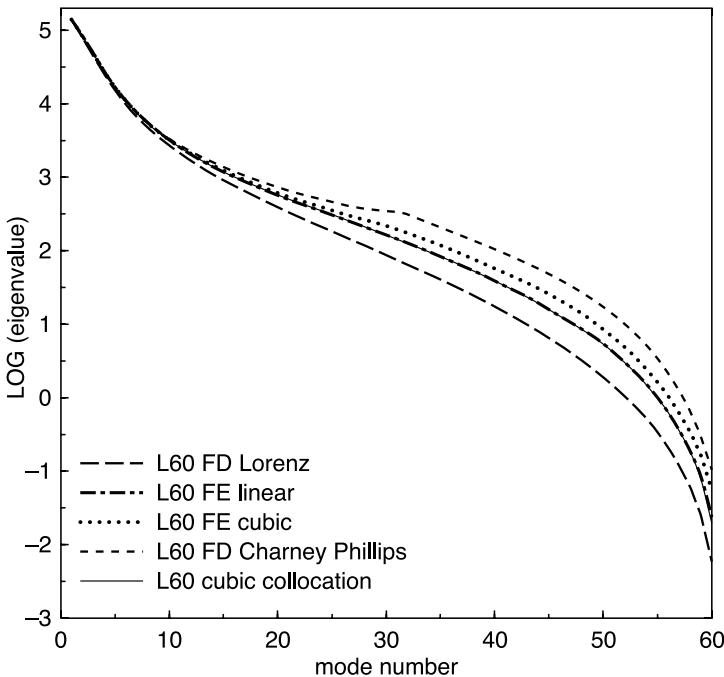


Figure 5. Eigenvalues of the vertical structure matrix for the linearized divergence equation as a function of eigenvector number, using the 60 levels of Fig. 4. Results are shown for the cubic FE scheme (dotted line); the linear FE scheme (dash-dotted line); the cubic collocation method (thin full line (on top of the dash-dotted line)); the FD scheme in the Lorenz staggering (long-dashed line); and with the FD scheme in the Charney–Phillips staggering (short-dashed line). See text for details.

larger than the FD-Lorenz eigenvalues. The smallest eigenvalue for the cubic FE scheme is roughly an order of magnitude larger than the smallest FD-Lorenz eigenvalue. For the case of the linear FE scheme it is larger by about a factor of four. This increase in the smallest eigenvalues with the cubic FE scheme, together with the decrease in amplitude of the grid-wave oscillations in the eigenfunctions of  $\Gamma$  discussed below, proved to be important in making data assimilation experiments work when run with the ECMWF four-dimensional variational (4D-Var) system using potential vorticity as a control variable (where the operator  $\Gamma$  needs to be inverted; Cullen (2002)).

The most striking difference in the eigenvectors computed with the different integration schemes at L60 occurs in the upper part of the model (stratosphere) for eigenvectors corresponding to medium and small eigenvalues (i.e. for mode numbers from 35 to about 55). The eigenvectors computed with the FD scheme in the Lorenz arrangement have large amplitude oscillations of a  $2\Delta\eta$  character, while the corresponding eigenvectors obtained with the FD scheme in the Charney–Phillips arrangement are almost zero at these levels, and have significant amplitude only in the lower part of the model (where the numbers of levels are greatest). The eigenvectors computed with the FE schemes have  $2\Delta\eta$  oscillations of significantly smaller amplitude than the FD-Lorenz eigenvectors. The cubic FE scheme reduces these oscillations more than the linear scheme.

Figure 6 illustrates this for eigenvector number 40 and Fig. 7 for eigenvector number 50.

These large amplitude  $2\Delta\eta$  oscillations in the eigenvectors of the FD-Lorenz scheme are probably linked to the existence of a computational grid-wave mode with the Lorenz

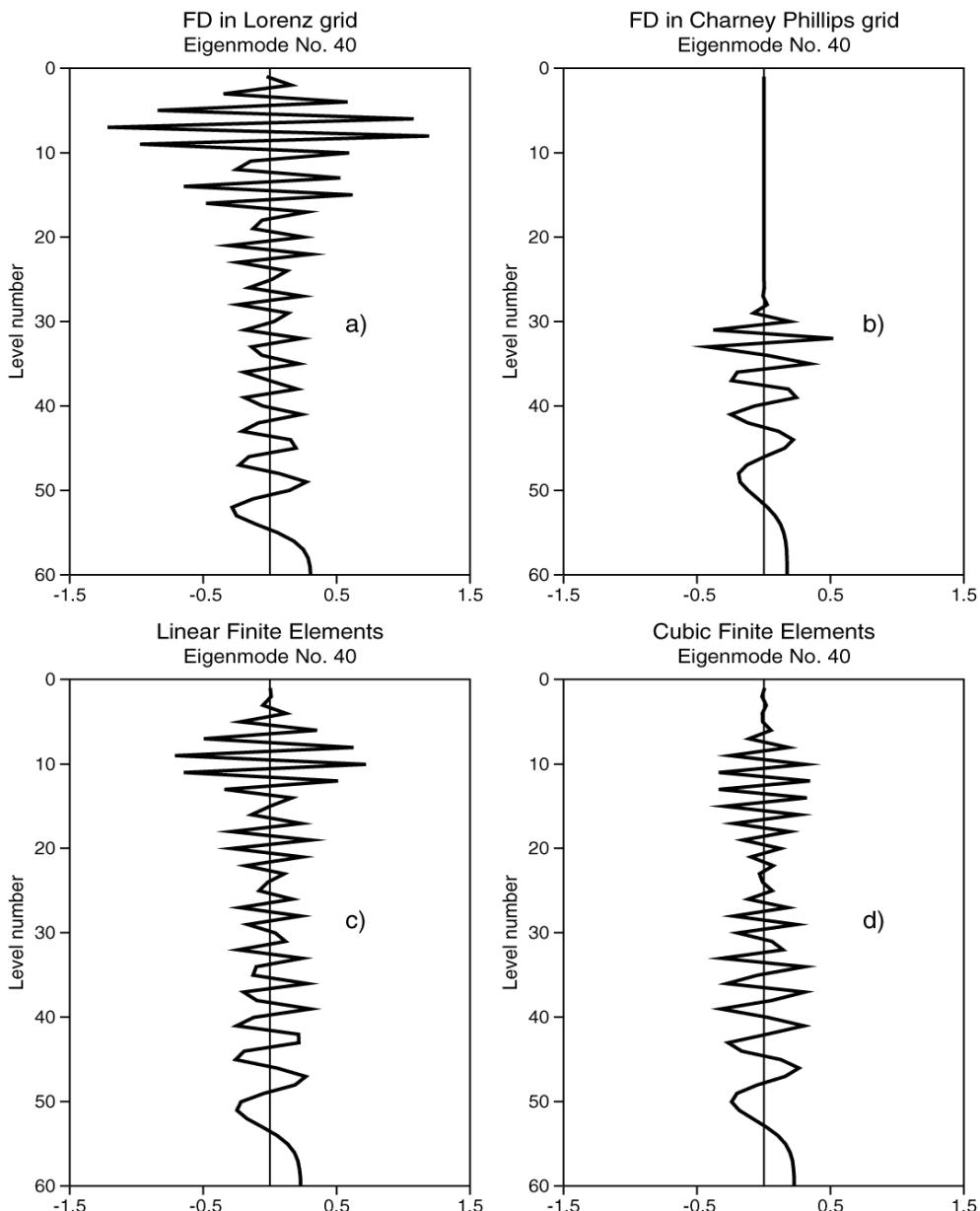


Figure 6. Structure of eigenmode number 40 of the vertical structure matrix  $\Gamma$ , for the 60-level model computed with FD schemes in: (a) the Lorenz, and (b) the Charney–Phillips arrangements; and with (c) the linear, and (d) the cubic FE schemes. See text for details.

staggering (Hollingsworth 1995; Arakawa and Konor 1996) which is absent with the Charney–Phillips staggering. Certainly they are an undesirable feature, since they are at the limit of the vertical resolution and can only be regarded as vertical noise. Therefore, a numerical scheme that reduces the amplitude of these oscillations is to be preferred.

Judging by the results obtained with the different integration schemes for the linearized model, the nonlinear model integrated with the FE schemes is expected

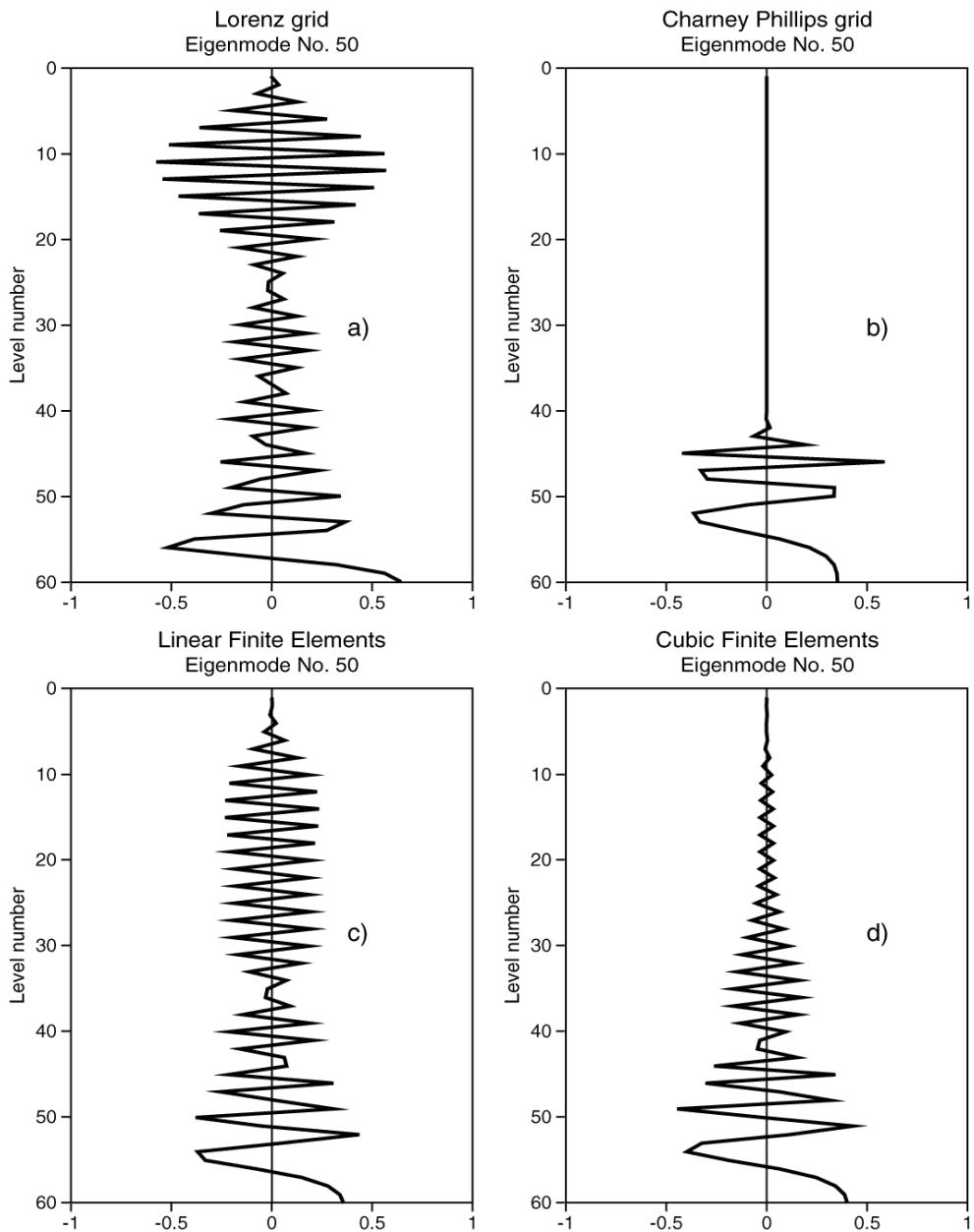


Figure 7. As Fig. 6 but for eigenvector number 50.

to be less susceptible to the excitation of large amplitude  $2\Delta\eta$  waves than the FD-Lorenz scheme. To test this, a simple numerical experiment was performed with the nonlinear dynamical core of the ECMWF model, where a vertical grid-wave structure in the temperature field was excited by diabatic heating at only one level, and the response of the model with the FD-Lorenz scheme, and with the linear and the cubic FE schemes considered. Unfortunately, a nonlinear model with the Charney-Phillips

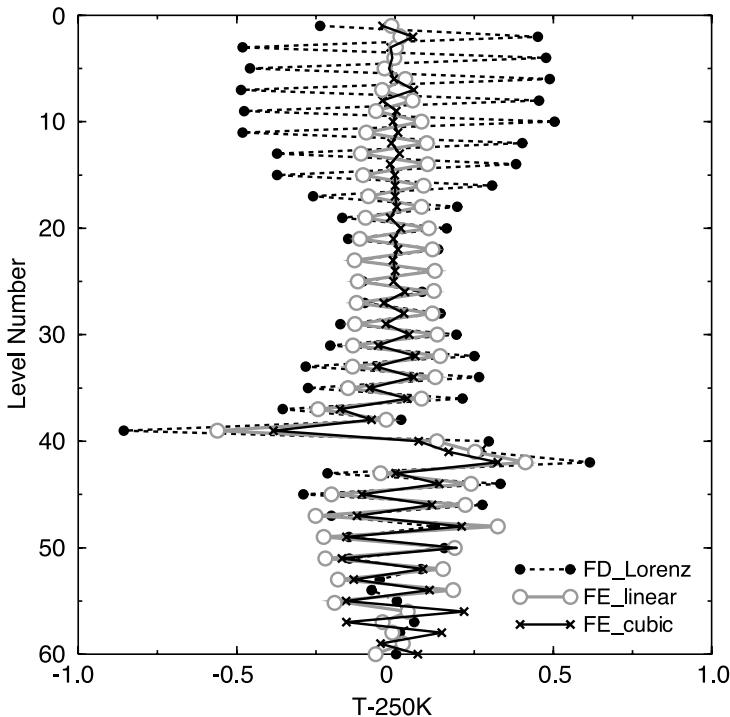


Figure 8. Vertical profiles of temperature at  $0^{\circ}\text{N}$  and  $90^{\circ}\text{E}$  obtained in a dynamical core experiment with heating at only one level to excite the computational mode of the Lorenz grid, with: the FD scheme in the Lorenz staggering (dotted line, full dots); the linear FE scheme (grey line, open dots); and with the cubic FE scheme (full line, crosses). See text for more details.

staggering of variables was not available for this test. This experiment is similar in design and set-up to one of the experiments used by Arakawa and Konor (1996) to demonstrate the existence of the computational mode in the Lorenz grid. The experiment was started from an isothermal atmosphere of 250 K at rest, and the dynamical core (no tendencies from physical parametrizations) of the nonlinear model was integrated at a resolution  $\text{T}_L95\text{L}60^*$  with flat orography using semi-Lagrangian time discretization with a time step of 1 h. At level 40 ( $\sim 540$  hPa) a temperature tendency of  $+0.01$  K  $(24\text{h})^{-1} \sin \lambda \cos \theta$  (where  $\lambda$  is longitude and  $\theta$  latitude) was added at each time step.

Figure 8 shows vertical profiles of temperature at  $0^{\circ}\text{N}$ ,  $90^{\circ}\text{E}$  obtained with the three discretization schemes at the end of the 600-day run. In the upper part of the model, the temperature profiles for the different schemes resemble, to some degree, the vertical structure of their eigenvectors (see Fig. 7)). Correspondingly, the vertical grid oscillations are greatly reduced with the two FE schemes, mainly with the cubic version. This is an indication that the spurious mode in Lorenz arrangement can be damped with FE schemes (or maybe even eliminated if high order elements are used). In other equatorial locations the profile of temperature is similar. The oscillations in the temperature profile outside the tropical region are much less marked, and therefore the influence of the FE scheme is much less noticeable.

\* Horizontal spectral triangularly truncated, representation up to total wave number 95 and using the linear Gaussian grid for the grid point representation.

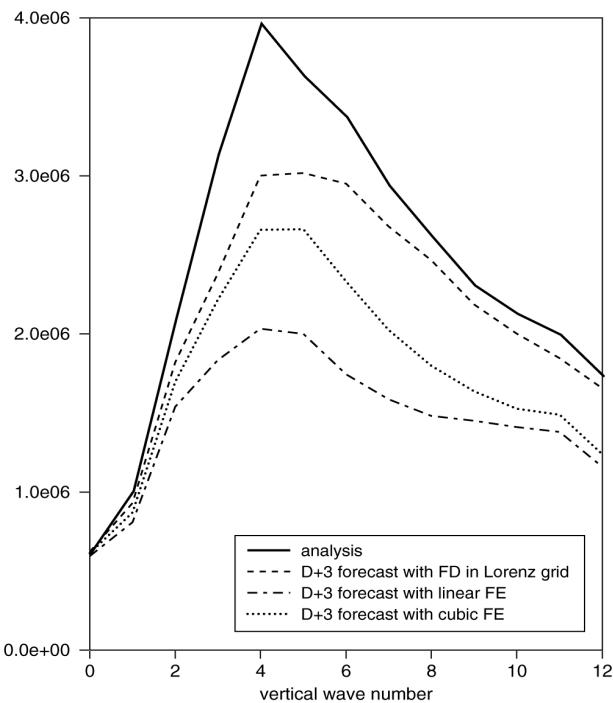
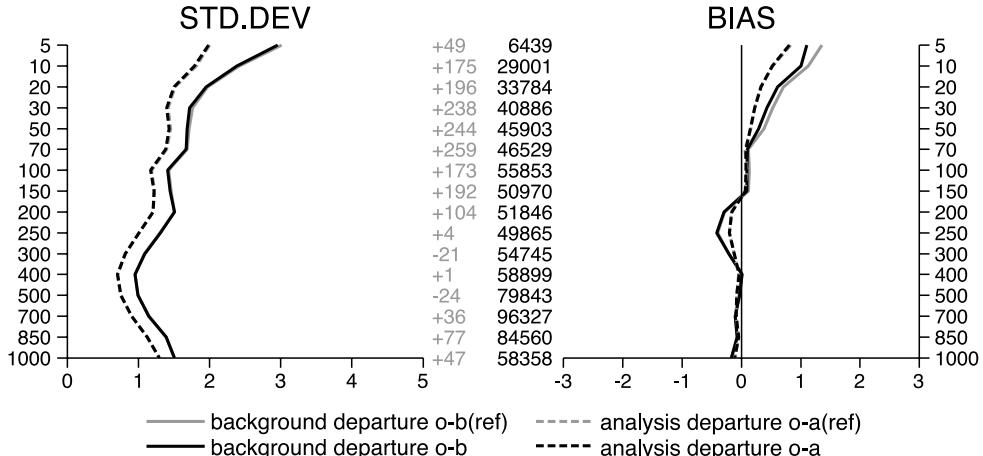


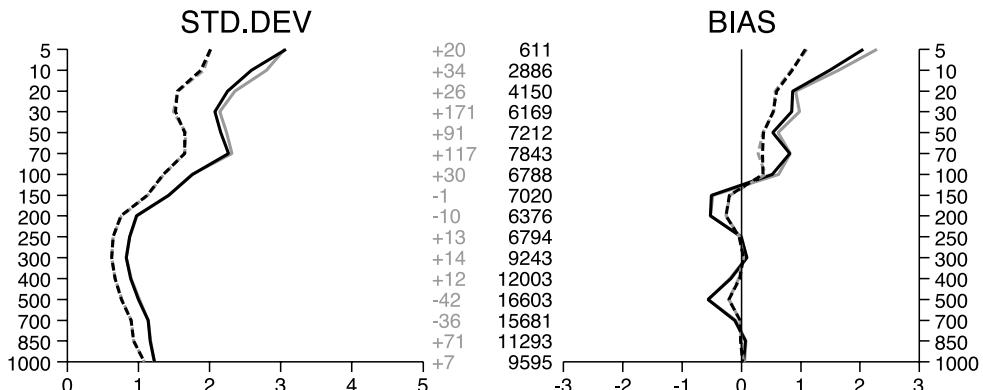
Figure 9. Fourier decomposition of vertical profiles of divergence (as a function of level number) over the top 24 levels (i.e. above 100 hPa) averaged over the deep tropics (from  $10^{\circ}\text{N}$  to  $10^{\circ}\text{S}$ ). The solid line is the spectrum of the analysis from which forecasts with the FD scheme in the Lorenz arrangement, with the linear FE scheme and with the cubic FE scheme were started. The broken lines show the spectra at day 3 of the forecast as indicated in the legend.

The two FE schemes also reduce the amplitude of the  $2\Delta\eta$  oscillations, often seen in vertical profiles of divergence in the tropical stratosphere, in forecasts run with the full model (with physical parametrizations switched on) using the FD-Lorenz scheme. In order to quantify this visual impression somewhat, the Fourier decomposition of vertical profiles of divergence was computed (as a function of level number) over the top 24 levels of L60 (i.e. above 100 hPa), and these Fourier coefficients were averaged over the deep tropics (from  $10^{\circ}\text{N}$  to  $10^{\circ}\text{S}$ ). Figure 9 shows ‘spectra’ computed in this way for 3-day forecasts, run with the FD-Lorenz scheme, and with the linear and cubic FE schemes, together with the spectrum of the analysis from which these forecasts were started. For the analysis the FD-Lorenz scheme was used. After 3 days, the cubic FE scheme has reduced the amplitude of the  $2\Delta\eta$  structure (wave number 12) and other poorly resolved vertical structures by almost a third, while there is only a very small decrease with the FD-Lorenz scheme. The similarity in the level of vertical noise (i.e. poorly resolved vertical structures) in the stratospheric analysis and forecasts suggests that this is a feature of the model numerics and is not forced in by the assimilated data. Assimilations using the FE schemes have a reduced level of vertical noise in the stratosphere (not shown). The linear FE scheme reduces the poorly resolved structures slightly more than the cubic FE scheme, but damps the better resolved structures even more than the cubic.

exp:cubFE, ref:FD L-grid, 2001010100-2001013112, TEMP-T N.Hemis, used T



exp:cubFE, ref:FD L-grid, 2001010100-2001013112, TEMP-T Tropics, used T



exp:cubFE, ref:FD L-grid, 2001010100-2001013112, TEMP-T S.Hemis, used T

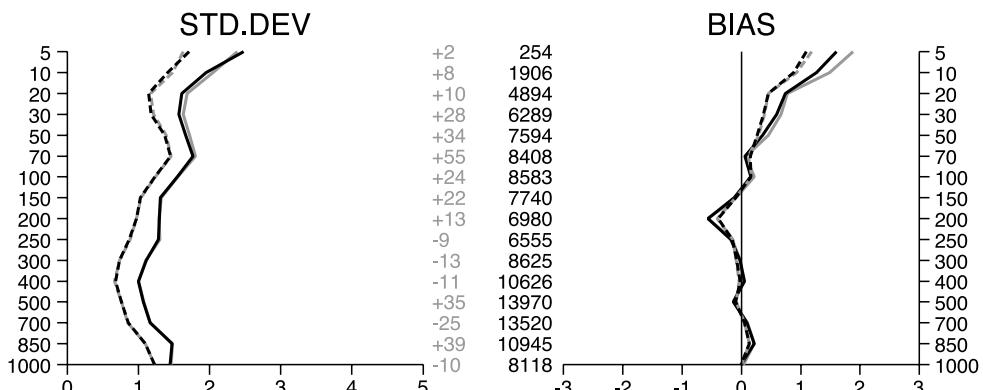


Figure 10. Fit of upper-air temperature observations from radiosonde soundings to the first guess (full line) and to the analysis (broken line), averaged over 62 assimilation cycles for an assimilation experiment using the cubic FE scheme (black lines) compared with the control experiment (grey) which uses FDs in the Lorenz grid. The numbers in black show the total number of observations used, while those in grey give the difference with respect to the control experiment (positive means more observations have been used in the FE experiment than in the FD control run). See text for details.

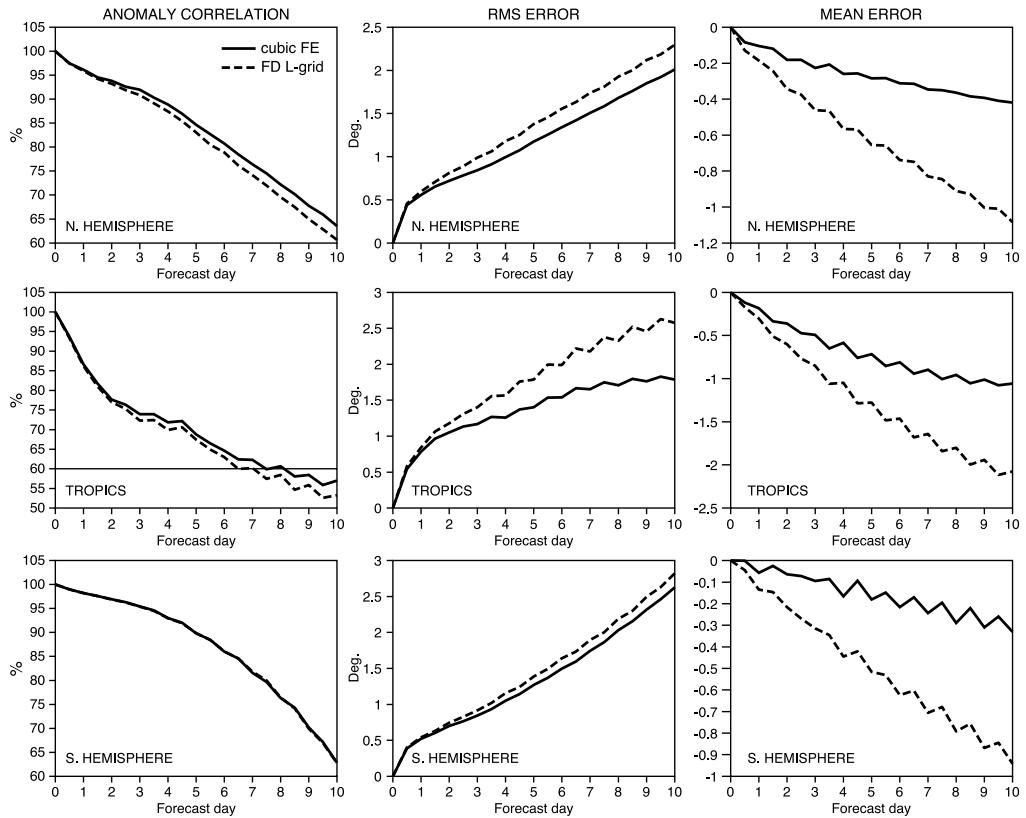


Figure 11. Mean temperature verification scores at 50 hPa for 98 10-day forecasts (67 summer and 31 winter cases) showing the anomaly correlation, root-mean-square error and bias (mean error) for the northern hemisphere, the tropics and the southern hemisphere. The cubic FE scheme (solid line) is compared with the control experiments using FDs in the Lorenz grid (FD L-grid, dashed line). See text for details.

### (b) Impact of the cubic finite-element scheme in assimilation–forecast experiments

Extended periods of data assimilation with the 4D-Var method have been run using the cubic FE scheme in the forecast model, as well as in the tangent linear and adjoint models. From the analyses corresponding to 12 UTC, 10-day forecasts were run. The results are compared to control experiments using the (then operational) FD scheme in the Lorenz grid. Figure 10 compares the fit of first guess and analysis to reported temperature observations, averaged over 62 assimilation cycles, from the FE and the control FD experiments. The number of observations accepted into the analysis is larger in the FE experiment at most levels, with the largest differences (of 0.5–3%) in the tropical stratosphere. Also, the bias and the standard deviation of the difference between the first guess and the observations are markedly reduced in the stratosphere in the FE experiment compared to the control, in spite of more observations having been accepted into the analysis (first guess and analysis are compared only to the accepted observations and not to all available observations). Figure 10 shows results for a winter period (January 2001); very similar results have been obtained for the summer period (not shown). The improved fit of the first guess to observations at stratospheric levels in the FE experiments is partly due to the reduction of vertical noise in the stratosphere

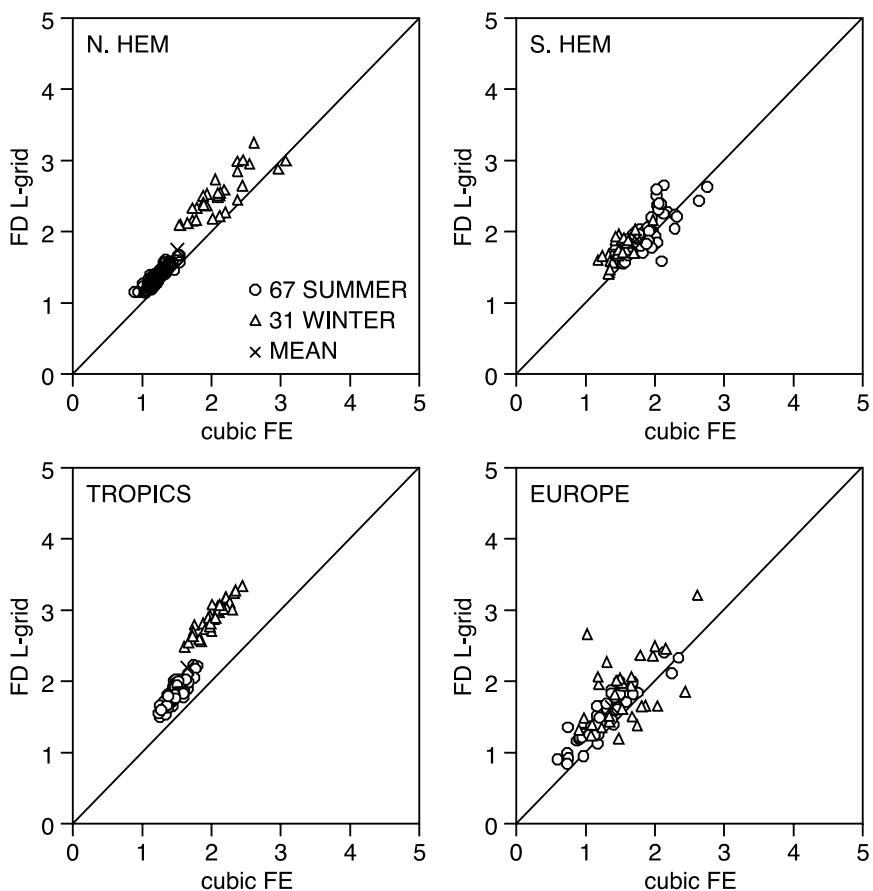


Figure 12. Scatter diagrams comparing root-mean-square errors of temperature at 50 hPa for the same 98 forecasts as in Fig. 11, for different verification areas. The difference between results using the cubic FE and the FD L grid is statistically significant at the 0.1% level. See text for details.

with the FE scheme, and partly to a reduction in the cold bias in the first guess of the lower stratosphere, present with the FD scheme.

The main impact of the FE discretization on the 10-day forecast is also found in the stratosphere. A substantial reduction of root-mean-square error (RMSE) for temperature is found in the lower stratosphere at all forecast ranges and over all verification areas, with the largest improvement occurring in the tropics. The tropospheric scores show a neutral impact on all the parameters and all verification areas (not shown).

Figure 11 compares mean temperature verification scores (anomaly correlation and RMSE) at 50 hPa in the northern hemisphere (NH), tropics and southern hemisphere (SH) from 10-day forecasts run with the cubic FE scheme and with the FD scheme (control). The average is taken over 98 forecasts comprising 67 summer and 31 winter cases. All forecasts were started from analyses run with the same vertical discretization as the forecasts. They are also verified against their corresponding analyses. The improvements in temperature RMSE at this level are very consistent (with high statistical significance) as shown in the corresponding scatter diagrams of the RMS error at day 7 into the forecast presented in Fig. 12. Notice that for the tropics *all* the forecasts score better in terms of RMSE of temperature at 50 hPa using the FE scheme than using the

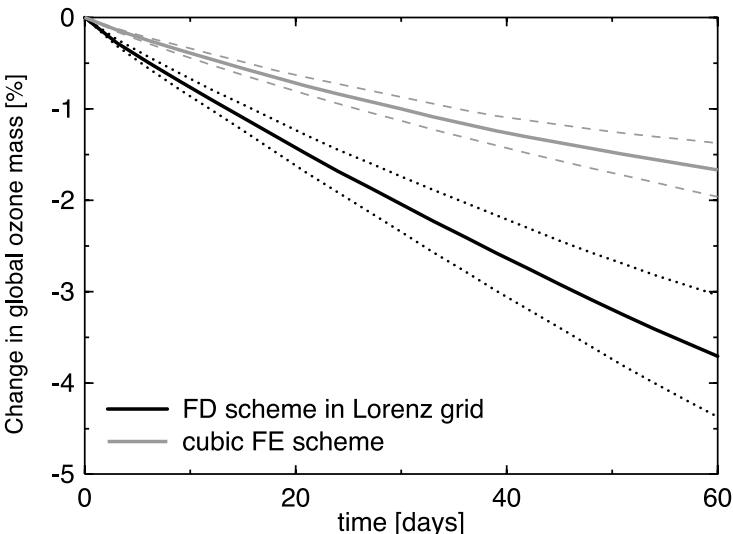


Figure 13. Conservation of ozone in the semi-Lagrangian model, showing the change in % in global ozone mass during 60-day forecasts, using the FD scheme in the Lorenz arrangement (black line) and the cubic FE scheme (grey line). The full lines are the means over 24 forecasts (two for each month) and the dotted (FD) and dashed (FE) lines indicate the spreads, i.e.  $\pm$  one standard deviation. See text for details.

FD scheme, and only very few of them score better using the FD scheme than using the FE scheme in the other areas. The large reduction in RMSE with the FE scheme is mostly due to a marked reduction in the cold bias present with the FD scheme in the lower stratosphere. The evolution of the mean error in temperature during 10-day forecasts at 50 hPa is shown in the right-hand column of Fig. 11. In the NH and in the tropics the cold bias of the FD scheme is roughly halved by the cubic FE scheme, and in the SH it is reduced to about a third. Wind scores in the stratosphere are also improved (not shown). In the troposphere the impact of the FE scheme on the verification scores is neutral (not shown), except in the upper troposphere in the tropics (at 200 and 100 hPa) where the existing positive temperature bias with the FD scheme is made slightly worse by the FE scheme.

The improvements in the stratosphere with the FE scheme translate into a better conservation of ozone, as shown in Fig. 13. The semi-Lagrangian advection scheme used in the ECMWF model does not conserve tracers by construction, and about 4% of the global mass of ozone is lost, on average, over 60 days of integration (with the ozone chemistry parametrization switched off) when using the FD scheme.

The two main sources for this ozone loss are errors in the advective wind field (mainly in the vertical wind) and errors in the interpolation of the ozone field at the departure point of the semi-Lagrangian trajectory (again mainly in the vertical interpolation). With the FE scheme the ozone loss is reduced by about 50%. This is due to a better vertical transport in the stratosphere (mainly to a slower upward movement in the tropics), resulting from improved computation of the vertical integral in the expression for the vertical velocity  $\dot{\eta}$ , (2.11). The remaining ozone loss can be substantially further reduced by using cubic spline interpolation for the vertical interpolation of ozone at the departure point, instead of four-point Lagrange interpolation used here (not shown).

The lack of conservation of mass in the model is not significantly changed by the FE scheme, because the influence of the stratosphere on this quantity is small.

## 5. SUMMARY AND DISCUSSION

A FE discretization scheme for the vertical dimension of the ECMWF model has been presented. This discretization is designed to be used in conjunction with the semi-Lagrangian advection scheme. Only non-local vertical operations are performed in FE space, products of dependent variables are evaluated in physical space ('transform method'). Since the only non-local vertical operations in the semi-Lagrangian version of the ECMWF model are vertical integrations, the FE discretization involved only the computation of an integration operator in the FE representation. The Galerkin method was used, and B-splines were employed as basis functions with compact support. Two versions of the integration operator were computed, a linear version based on piecewise linear functions and a cubic version based on cubic B-splines. In both versions, the integration operator is a full matrix and operates directly on the integrand, given on the model levels, and returns the integral from the top of the model to each model level and to the surface. The transformations from physical space to FE space and back are incorporated into the integration matrix. They are not performed explicitly, as is the case with the spectral transform method used for the horizontal.

Since the integration matrix is a full matrix for both the cubic as well as the linear FE scheme, the computational cost of applying these two schemes is basically the same (there is a small extra cost for the cubic scheme in the initialization phase), i.e. it is practically independent of the degree of the B-splines used as basis functions. For a 10-day forecast with the full model the cost of the FE scheme is 3% greater than with the FD scheme.

No staggering of dependent variables is used in this vertical discretization, to avoid the difficulties staggering causes for semi-Lagrangian advection, in that it requires the winds and all the advected variables to be given at the same levels. However, keeping the geopotential and the temperature at the same levels (Lorenz staggering) can lead to a large amplitude vertical 'zigzag' mode in the temperature field ('computational mode'; Arakawa and Konor 1996). The FE schemes described here reduce the amplitude of this mode in temperature considerably (the cubic scheme much more than the linear) as shown in subsection 4(a). By using Charney–Phillips staggering in the vertical, this computational mode can be suppressed completely (Arakawa and Konor 1996), but at the price of having to interpolate the temperature field to the levels where the winds are held (or vice versa) for semi-Lagrangian advection. This interpolation can introduce errors if not done carefully and adds to the computational cost. The cubic FE scheme makes a very good compromise between the convenience of the Lorenz grid with semi-Lagrangian advection and the desirable property of the Charney–Phillips grid of being free of the computational mode in temperature, present with the Lorenz grid.

The linear FE integration is fourth order accurate on the nodes, away from the upper and lower boundaries, and the cubic is of order eight when used for integrating smooth functions. This high accuracy in evaluating integrals translates into a much improved accuracy in the phase speeds of the linear gravity waves obtained with the FE schemes compared with the FD schemes.

Due to the much higher accuracy at almost no additional cost of the cubic FE scheme compared to the linear scheme, the cubic version was chosen for implementation into the operational version of the ECMWF model, and was tested extensively in assimilation–forecast experiments. The main benefits from the cubic FE scheme are a reduction of the persistent cold bias in the lower stratosphere (by about 50% at 50 hPa) and a marked reduction in vertical noise in the stratosphere. Both these improvements lead to a better fit of the first-guess fields to observations in the lower stratosphere (mainly for

radiosonde temperature data) and to more observations being accepted into the analysis. They also lead to a significant improvement in the lower-stratospheric verification scores, with the largest improvement found in temperature in the tropics; wind scores are also improved. Tropospheric scores are neutral except at 200 and 100 hPa in the tropics, where the FE scheme slightly worsens an existing warm bias.

The FE scheme also improves the vertical transport in the lower stratosphere and around the tropopause, leading to better conservation of ozone. The loss in global ozone mass of almost 4% in 60-day forecasts with the FD scheme at T<sub>L</sub>159L60 is halved by the cubic FE scheme.

Based on these positive results, the cubic FE scheme became operational at ECMWF on 22 January 2002 in all the configurations (deterministic forecast mode, variational assimilation mode, ensemble forecast mode) which are run using the semi-Lagrangian advection scheme. It was also used in the 40-year re-analysis project (ERA40, Gibson *et al.* 1999).

#### ACKNOWLEDGEMENTS

The authors are grateful to A. Hollingsworth, M. Miller and A. Staniforth for helpful comments and suggestions to improve the manuscript.

#### APPENDIX

##### *Cubic B-splines for non-uniform node spacing*

For a non-regular spacing of levels, the cubic B-spline belonging to the  $i$ th node,  $B_i(\eta)$ , can be constructed, as in the case of regular spacing, from four piecewise cubic polynomials with the conditions that it is twice continuously differentiable on the entire real line and that  $B_i(\eta) \equiv 0$  for  $\eta \leq \eta_{i-2}$  and  $\eta \geq \eta_{i+2}$ . These continuity conditions provide 15 linear equations for the 16 coefficients of the four cubic polynomials that make up a cubic B-spline. An additional normalization condition is needed to complement this set of 15 equations to fully determine a cubic B-spline.

Let us consider the B-spline belonging to node  $i$ . The  $j$ th polynomial ( $j = 1, 4$ ), which forms part of that spline can be written as:

$$f_j(\eta) = a_j \eta^3 + b_j \eta^2 + c_j \eta + d_j, \quad (\text{A.1})$$

where we have dropped the node reference ‘ $i$ ’ for simplicity. The continuity conditions at  $\eta_{i-2}$ , namely  $B_i(\eta_{i-2}) = 0$ ,  $B'_i(\eta_{i-2}) = 0$ ,  $B''_i(\eta_{i-2}) = 0$ , imply that the first polynomial (defined in the interval  $[\eta_{i-2}, \eta_{i-1}]$ ) can be written as

$$f_1(\eta) = a_1(\eta - \eta_{i-2})^3, \quad (\text{A.2})$$

and similarly the continuity conditions at node  $\eta_{i+2}$  imply that the fourth polynomial (defined in the interval  $[\eta_{i+1}, \eta_{i+2}]$ ) can be written as

$$f_4(\eta) = a_4(\eta_{i+2} - \eta)^3. \quad (\text{A.3})$$

The coefficients  $a_1$  and  $a_4$  will be determined, together with the four coefficients of polynomial  $f_2$  and those of  $f_3$ , from the nine continuity conditions at nodes  $\eta_{i-1}$ ,  $\eta_i$  and  $\eta_{i+1}$  and a normalization condition to be specified later.

From the continuity conditions at node  $\eta_{i-1}$ , namely  $f_1(\eta_{i-1}) = f_2(\eta_{i-1})$ ,  $f'_1(\eta_{i-1}) = f'_2(\eta_{i-1})$ ,  $f''_1(\eta_{i-1}) = f''_2(\eta_{i-1})$ , we obtain:

$$\begin{aligned} b_2 &= 3(\eta_{i-1} - \eta_{i-2})a_1 - 3\eta_{i-1}a_2 \\ c_2 &= -3(\eta_{i-1}^2 - \eta_{i-2}^2)a_1 + 3\eta_{i-1}^2a_2 \\ d_2 &= (\eta_{i-1}^3 - \eta_{i-2}^3)a_1 - \eta_{i-1}^3a_2, \end{aligned} \quad (\text{A.4})$$

and similarly, from the continuity conditions at node  $\eta_{i+1}$ :

$$\begin{aligned} b_3 &= -3\eta_{i+1}a_3 + 3(\eta_{i+2} - \eta_{i+1})a_4 \\ c_3 &= 3\eta_{i+1}^2a_3 - 3(\eta_{i+2}^2 - \eta_{i+1}^2)a_4 \\ d_3 &= -\eta_{i+1}^3a_3 + (\eta_{i+2}^3 - \eta_{i+1}^3)a_4. \end{aligned} \quad (\text{A.5})$$

Finally, the continuity conditions at node  $\eta_i$ , together with (A.4) and (A.5), give:

$$\begin{aligned} a_1 &= a_4 \frac{\Delta_+^3 - \Delta_i^3 + (\Delta_+^2 - \Delta_i^2)(\Delta_{i-1} - \Delta_i) - \Delta_{i-1}\Delta_i\Delta_{i+1}}{\nabla} \\ a_2 &= -a_4 \frac{\Delta_{i+1}\Delta_+[\Delta_-^3 - \Delta_{i-1}^3 + (\Delta_-^2 - \Delta_{i-1}^2)(\Delta_i + \Delta_+) + \Delta_{i-2}\Delta_i\Delta_+]}{\Delta_{i-1}(\Delta_{i-1} + \Delta_i) \cdot \nabla} \\ a_3 &= a_4 \frac{\Delta_{i-2}\Delta_-[\Delta_+^3 - \Delta_i^3 + (\Delta_+^2 - \Delta_i^2)(\Delta_{i-1} + \Delta_-) + \Delta_{i-1}\Delta_{i+1}\Delta_-]}{\Delta_i(\Delta_{i-1} + \Delta_i) \cdot \nabla}, \end{aligned} \quad (\text{A.6})$$

where the  $\Delta_j$  are the inter-nodal distances ( $\Delta_{i-2} \equiv \eta_{i-1} - \eta_{i-2}$ ,  $\Delta_{i-1} \equiv \eta_i - \eta_{i-1}$ ,  $\Delta_i \equiv \eta_{i+1} - \eta_i$  and  $\Delta_{i+1} \equiv \eta_{i+2} - \eta_{i+1}$ ),  $\Delta_- \equiv \Delta_{i-2} + \Delta_{i-1}$ ,  $\Delta_+ \equiv \Delta_i + \Delta_{i+1}$ , and  $\nabla \equiv \Delta_-^3 - \Delta_{i-1}^3 + (\Delta_-^2 - \Delta_{i-1}^2)(\Delta_i - \Delta_{i-1}) - \Delta_{i-2}\Delta_{i-1}\Delta_i$ . Since all  $\Delta$ s are positive and non-zero, the denominators in (A.6) can be shown to be non-zero, i.e. a solution exists for any arbitrary choice of node distribution.

Inserting (A.6) back into (A.4) and (A.5) yields analytic expressions for  $b_j$ ,  $c_j$  and  $d_j$ , for  $j = 2, 3$ . Because these expressions are rather lengthy we do not write them out explicitly. Nevertheless, it is easily seen that all coefficients contain  $a_4$ —the only unknown left to be determined—as a factor that can be fixed by imposing a normalization condition. We have chosen to set the value of the B-spline at its central node equal to one, i.e.

$$a_2\eta_i^3 + b_2\eta_i^2 + c_2\eta_i + d_2 = 1. \quad (\text{A.7})$$

In the case of uniform node spacing this condition is equivalent to setting the maximum of the B-spline to one, because in this case the B-spline is symmetric.

## REFERENCES

- |   |      |  |
|---|------|--|
| Arakawa, A. and Konor, C. S.                    | 1996 | Vertical differencing of the primitive equations based on the Charney–Phillips grid in hybrid vertical coordinates. <i>Mon. Weather Rev.</i> , <b>124</b> , 511–528                          |
| Becker, E. B., Carey, G. F. and Oden, J. T.     | 1981 | <i>Finite elements: An introduction</i> . Volume I. Prentice–Hall Inc., Englewood Cliffs, New Jersey, USA  |
| Beland, M., Cote, J. and Staniforth, A.         | 1983 | The accuracy of a finite-element vertical discretization scheme for primitive equation models: Comparison with a finite-difference scheme. <i>Mon. Weather Rev.</i> , <b>111</b> , 2298–2318 |
| Burridge, D. M., Steppeler, J. and Strüfing, R. | 1986 | ‘Finite element schemes for the vertical discretization of the ECMWF forecast model using linear elements’. Technical Report No. 54. ECMWF, Shinfield Park, Reading, UK                      |

- Charney, J. G. and Phillips, N. A. 1953 Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. *J. Meteorol.*, **10**, 71–99
- Cullen, M. J. P. 2002 ‘Use of potential vorticity as a control variable within a 4D variational data assimilation system’. Technical Memorandum 358. ECMWF, Shinfield Park, Reading, UK
- Francis, P. E. 1972 The possible use of Laguerre polynomials for representing the vertical structure of numerical models of the atmosphere. *Q. J. R. Meteorol. Soc.*, **98**, 662–667
- Gibson, J. K., Fiorino, M., Hernandez, A., Källberg, P., Li, X., Onogi, K., Saarinen, S. and Uppala, S. 1999 ‘The ECMWF 40-Year Re-Analysis (ERA-40) Project—Plans and current status’. Pp. 369–372 of Proceedings of the tenth symposium on global change. American Meteorological Society, Boston, USA
- Hollingsworth, A. 1995 ‘A spurious mode in the ‘Lorenz’ arrangement of  $\phi$  and  $T$  which does not exist in the ‘Charney–Phillips’ arrangement’. Technical Memorandum 211. ECMWF, Shinfield Park, Reading, UK
- Hoskins, B. 1973 Comments on ‘The possible use of Laguerre polynomials for representing the vertical structure of numerical models of the atmosphere’ by P. E. Francis. *Q. J. R. Meteorol. Soc.*, **99**, 571–572
- Lorenz, E. N. 1960 Energy and numerical weather prediction. *Tellus*, **12**, 364–373
- Machenhauer, B. and Daley, R. 1972 ‘A baroclinic primitive equation model with a spectral representation in three dimensions’. Report 4, Institut for Teoretisk Meteorologi, University of Copenhagen, Denmark
- Oden, J. T. and Reddy, J. N. 1976 *An introduction to the mathematical theory of finite elements*. Springer-Verlag, Berlin, Germany
- Prenter, P. M. 1975 *Splines and variational methods*. John Wiley and Sons. New York, USA
- Ritchie, H., Temperton, C., Simmons, A. J., Hortal, M., Davies, T., Dent, D. and Hamrud, M. 1995 Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. *Mon. Weather Rev.*, **123**, 489–514
- Simmons, A. J. and Burridge, D. M. 1981 An energy and angular momentum conserving vertical finite difference scheme and hybrid vertical coordinates. *Mon. Weather Rev.*, **109**, 758–766
- Simmons, A. J. and Strüfing, R. 1983 Numerical forecasts of stratospheric warming events using a model with hybrid vertical coordinates. *Q. J. R. Meteorol. Soc.*, **109**, 81–111
- Simmons, A. J. and Temperton, C. 1997 Stability of a two-time-level semi-implicit integration scheme for gravity wave motion. *Mon. Weather Rev.*, **125**, 600–615
- Staniforth, A. N. and Daley, R. W. 1977 A finite-element formulation for the vertical discretization of sigma-coordinate primitive equation models. *Mon. Weather Rev.*, **105**, 1108–1118
- Steppeler, J. 1986 ‘Finite elements scheme for the vertical discretization of the ECMWF forecast model using quadratic and cubic elements’. Technical Report No. 55. ECMWF, Shinfield Park, Reading, UK
- 1987 Quadratic Galerkin finite element schemes for the vertical discretization of numerical models. *Mon. Weather Rev.*, **115**, 1575–1588
- Tokioka, T. 1978 Some considerations on vertical differencing. *J. Meteorol. Soc. Jpn.*, **56**, 89–111