THE USE OF ADJOINT EQUATIONS IN NUMERICAL WEATHER PREDICTION

P Courtier and F Rabier
European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, UK

1. THE ADJOINT EQUATIONS

The adjoint equations have been introduced in Meteorology by *Marchuk* (1974) and their use is listed in *Courtier et al* (1993). Here we shall concentrate on the practical aspects of their applications to dynamical meteorology. Let G be an operator, differentiable almost everywhere, which computes the output parameters v from the input parameters u:

$$v = G(u) \tag{1}$$

G may be available in practice as a FORTRAN program like a numerical weather prediction model. For perturbations δu and δv of u and v, the Taylor formula provides the following equality, valid to first order:

$$\delta v = G'_u \, \delta u \tag{2}$$

where G'_u is the tangent-linear operator of G linearized in the vicinity of u. The matrix representing G'_u is the Jacobian matrix of G. In most meteorological applications, δu and δv are vectors of size 10^5 to 10^7 (the latter being valid for the T213L31 ECMWF operational model and the former for a T42L19 version). The Jacobian matrix cannot then be explicitly computed (and even stored). However, the following two results make it possible in practice to apply the operator G'_u to an input vector δu and its transpose to an input vector δv.

A It is possible to write a FORTRAN program which solves (2) once the FORTRAN program for (1) is available. Its cost is of similar order of magnitude as (1).

B It is possible to write a FORTRAN program of similar cost to (1) which computes an output vector δu for a given input vector δv:

$$\delta_u = G'^t_u \, \delta_v \tag{3}$$

where G'^t_u is the transpose of G'_u. This can be generalised to the adjoint G'^*_u of G'_u through appropriate metric (inner product) changes.

Result A is trivial in practice since the result of (2) could e.g. be obtained to first order by finite differences in the direction δu from two integrations of (1): this is the methodology generally followed in most sensitivity experiments. For an exact solution to (2), *Morgenstern* (1973) discusses the complexity of tangent-linear algorithms. However, result B is not trivial and relies on the Baur-Strassen theorem (*Baur* *and Strassen*, 1983; *Morgenstern*, 1985). In the following we shall describe the practical applications of results A and B.

2. SENSITIVITY ANALYSIS

Let us consider $J(v)$ a function of the output parameters v. We assume J "simple" in that an analytic expression of the gradient of J with respect to v, $\nabla_v J$ is available (note that this is not a restriction, it is always possible to consider J as a simple function of something, e.g. $J=J$).

By definition of $\nabla_v J$; one has for any perturbation δv (and to first order)

$$\Delta J = J(v+\delta v) - J(v) = \langle \nabla_v J, \delta v \rangle.$$

δv may be related to δu using (2)

$$\Delta J = \langle \nabla_v J, G'_u \delta u \rangle$$

by definition of the adjoint operator

$$\Delta J = \langle G'^*_u \nabla_v J, \delta u \rangle$$

which implies that

$$\nabla_u J = G'^*_u \nabla_v J \tag{4}$$

Result B then implies that it is possible to compute the sensitivity of J with respect to the input parameters u: the gradient of J with respect to u is obtained by applying the adjoint of the tangent-linear operator G'^*_u to the gradient of J with respect to v.

2.1 Sensitivity to initial conditions

Here we consider u as the initial conditions of a numerical weather prediction model and J as one output. *Courtier* (1987), using a shallow-water model, identified a tidal wave problem in nonlinear normal mode initialization. *Errico and Vukicevic* (1992) presented a case of lee cyclogenesis. Fig 1 (their Fig 2) presents the initial conditions of surface pressure and 500 hPa height while Fig 2 (their Fig 1) presents a 36h forecast produced by the PSU/NCAR MM4 model. Fig 3 presents the sensitivity to the initial conditions of the predicted surface pressure at point P (centre of the low). One notices that the sensitivity is localized. A more intense synoptic wave (stronger ridge over the Atlantic and deeper over France) would lead to a deeper lee cyclogenesis.

Rabier et al (1992) studied the sensitivity of the baroclinic instability of *Simmons and Hoskins* (1978) to the initial conditions 24 hours before, during the 24 hours of most intense cyclogenesis. They showed that it is easy to eliminate the gravity waves signal present in the sensitivity pattern using the adjoint of nonlinear normal mode initialization as can be seen comparing Fig 4 (their Fig 3) and Fig 5 (their Fig 5). The vertical

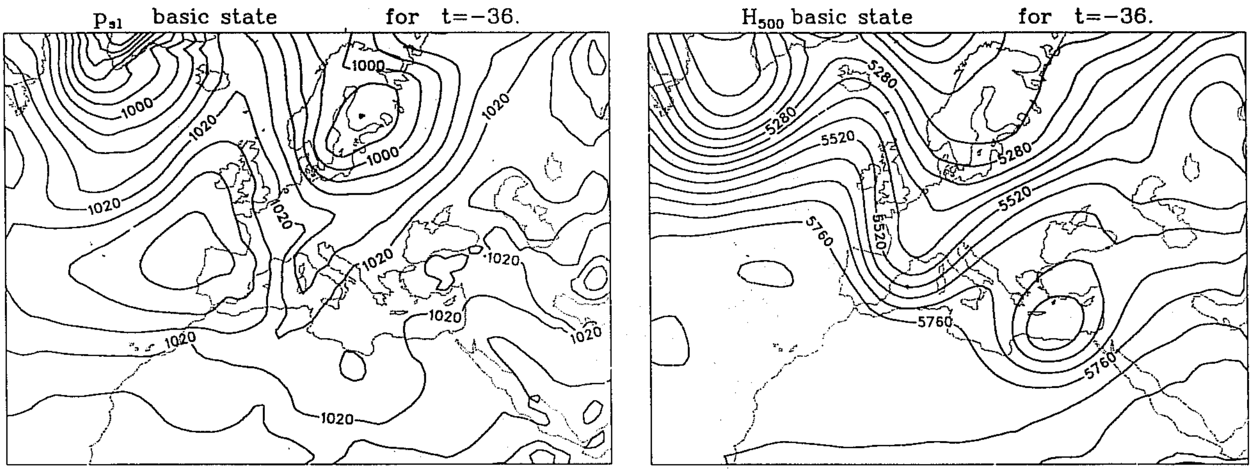

Fig 1 Initial conditions of the forecast. Left: surface pressure, right: 500 hPa height. 4 March 1982 18h00. (From *Errico and Vukicevic*, 1992).

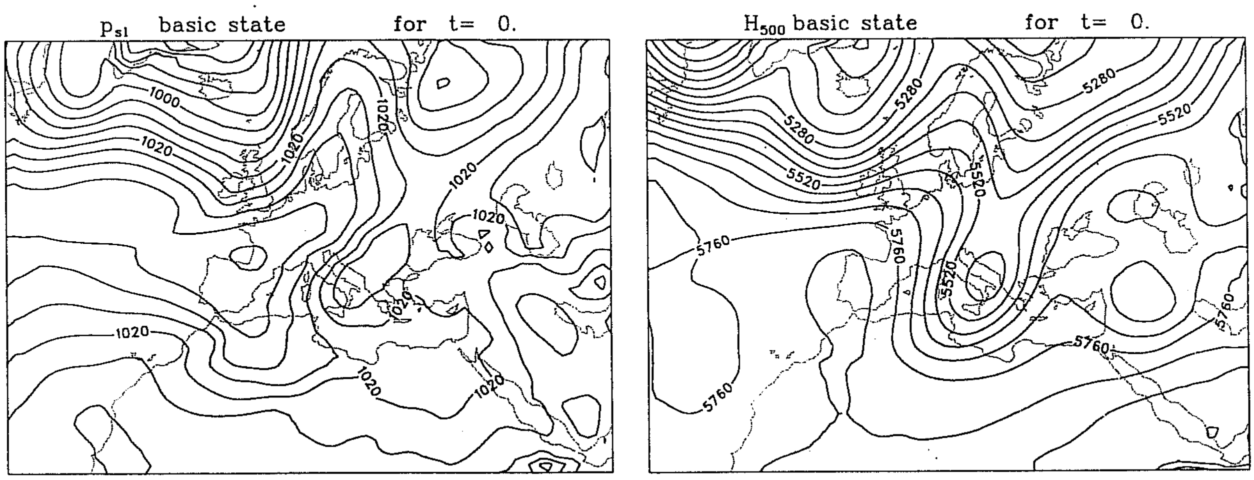

Fig 2 Same as Fig 1 but 36 hr forecast valid for 6 March 1982 6h00. (From *Errico and Vukicevic*, 1992).

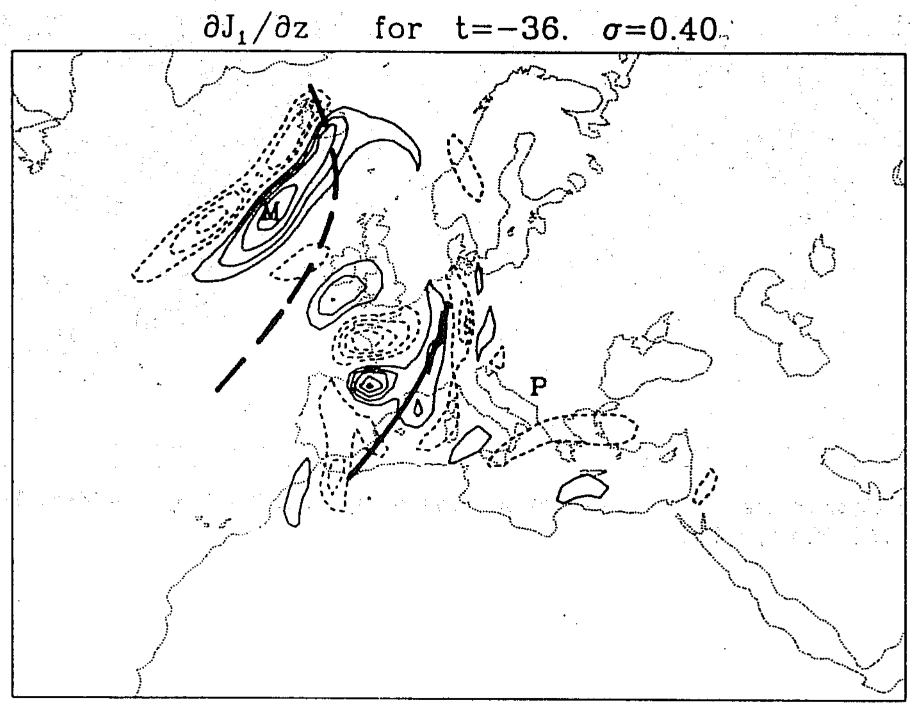

Fig 3 Sensitivity of the surface pressure forecast at point P with respect to the height field at level σ = .4 (~ 400 hPa) initial conditions. (From *Errico and Vukicevic*, 1992).

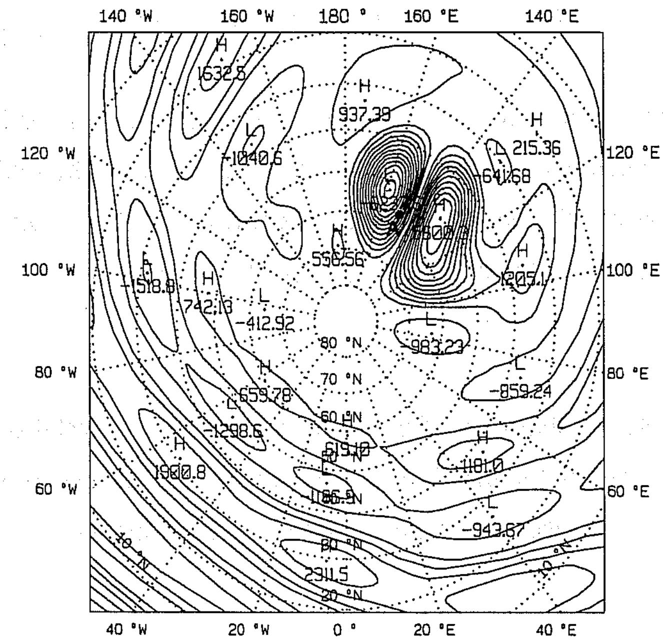

Fig 4 Sensitivity of the logarithm of the surface pressure at point A (middle of the surface pressure low) to the 500 hPa meridional wind 24 hours earlier. (From *Rabier et al*, 1992).

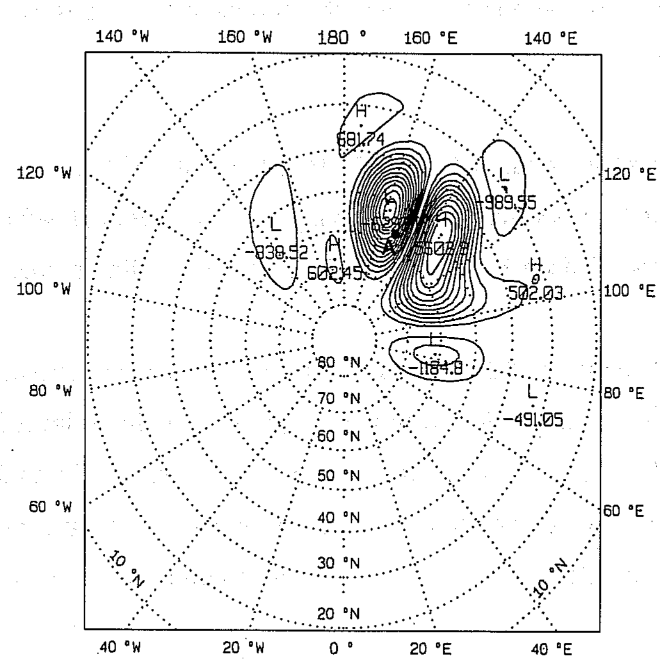

Fig 5 Same as Fig 4 but with the adjoint of non linear normal mode initialization included. (From *Rabier et al*, 1992).

structure of the sensitivity showing a maximum in the low troposphere is depicted in Fig 6 (their Fig 10) which is consistent with other baroclinic instability studies. This gravity wave signal was not visible at the 36-hour range in the Errico and Vukicevic work reported in the previous paragraph since it had already been dissipated at the boundaries (however, it was visible at very short range).

Rabier et al (1993) performed a feasibility study of the use of the adjoint equations for routinely monitoring the sensitivity of the short-range (48 hour) forecast errors with respect to the initial conditions. The norm used is the quadratic invariant of the linearized primitive equations in the vicinity of a state of rest and the cost function is the square norm of the difference between the operational 48h forecast and the verifying analysis.

The experiments used a T63L31 adiabatic model (with horizontal and vertical diffusion and a surface drag). The adjoint integration is performed in the vicinity of an adiabatic trajectory originated from the ECMWF analysis valid 48 hour before. At the end of the adjoint integration, the adjoint of nonlinear normal mode initialization is performed.

Figs 7, 8 and 9 present the rms over the month of January 1993 of the gradient of the cost function with respect to the initial conditions of vorticity at model level 11 (~250 hPa), level 18 (~500 hPa) and level 26 (~850 hPa). In agreement with *Rabier et al* (1992), the gradient is stronger with respect to mid troposphere vorticity. It is stronger over the oceans than over the continents which is consistent with better data coverage over the continents.

Fig 10 shows the average over the month of January 1993 of the analysed 500 hPa geopotential height. The maxima in the Atlantic of Fig 9 are located in the cyclone track going from Newfoundland to Norway. The maximum located north of the Hudson Bay corresponds to the descending branch of the arctic jet.

Fig 11 presents the temporal variation of the cost function (forecast errors square norm), of the square norm of the gradient and of the rms of the gradient with respect to the vorticity at level 26. All three curves are normalized and of average 1. It is clear that there is few day to day variability in the forecast error norms. However, the sensitivity to the initial conditions can vary a lot and by as much as a factor of 4. There is no apparent correlation with the day to day variability of the medium range scores (Fig 12).

There is significant vertical variation of the rms of the gradient with respect to vorticity as can be seen from Fig 13. Furthermore the day to day variability is significant (±25%) with, in addition, a few exceptional cases. The average over the 30 cases is presented in Fig 14. Fig 15 presents the vertical variation of the contribution of vorticity to the cost function (rotational part of the kinetic energy of the error multiplied by the layer depth) for the individual cases of January 1993 and Fig 16 the average. The maximum of error is at the jet level whereas, as we have seen, the maximum of sensitivity is located in the low troposphere.

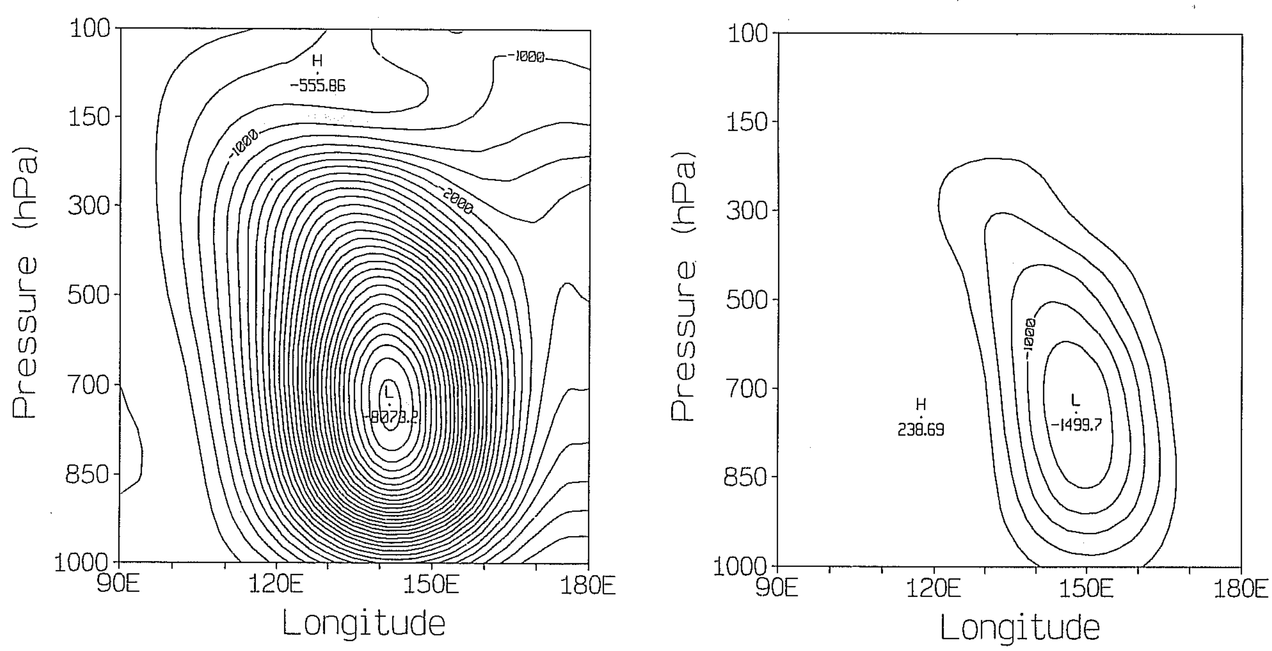

Fig 6 Cross section of the gradient of the average of the logarithm of the surface pressure over the low with respect to the initial conditions 24 hours earlier. Left: vorticity, right: temperature. (From *Rabier et al*, 1992).

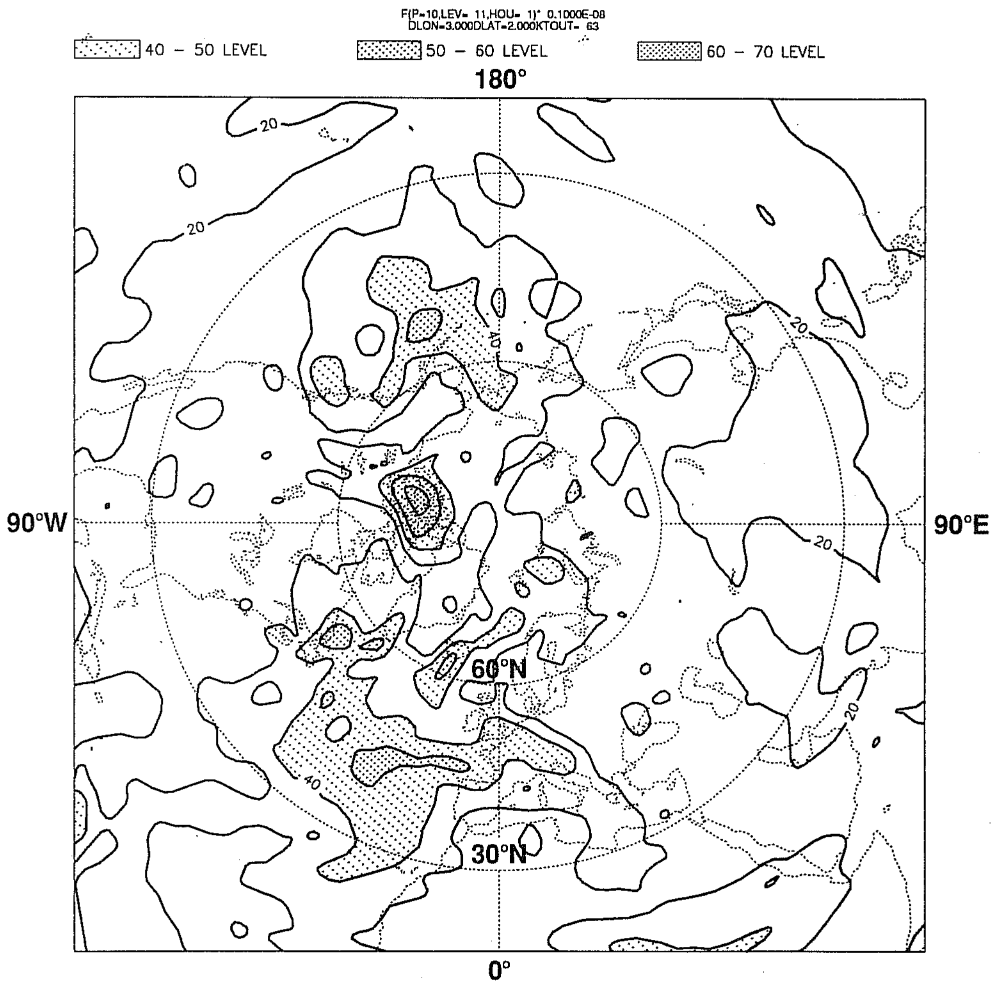

Fig 7 RMS computed over the month of January of the gradient of the 48 hour forecast errors with respect to the vorticity at level 11 (~ 250 hPa).

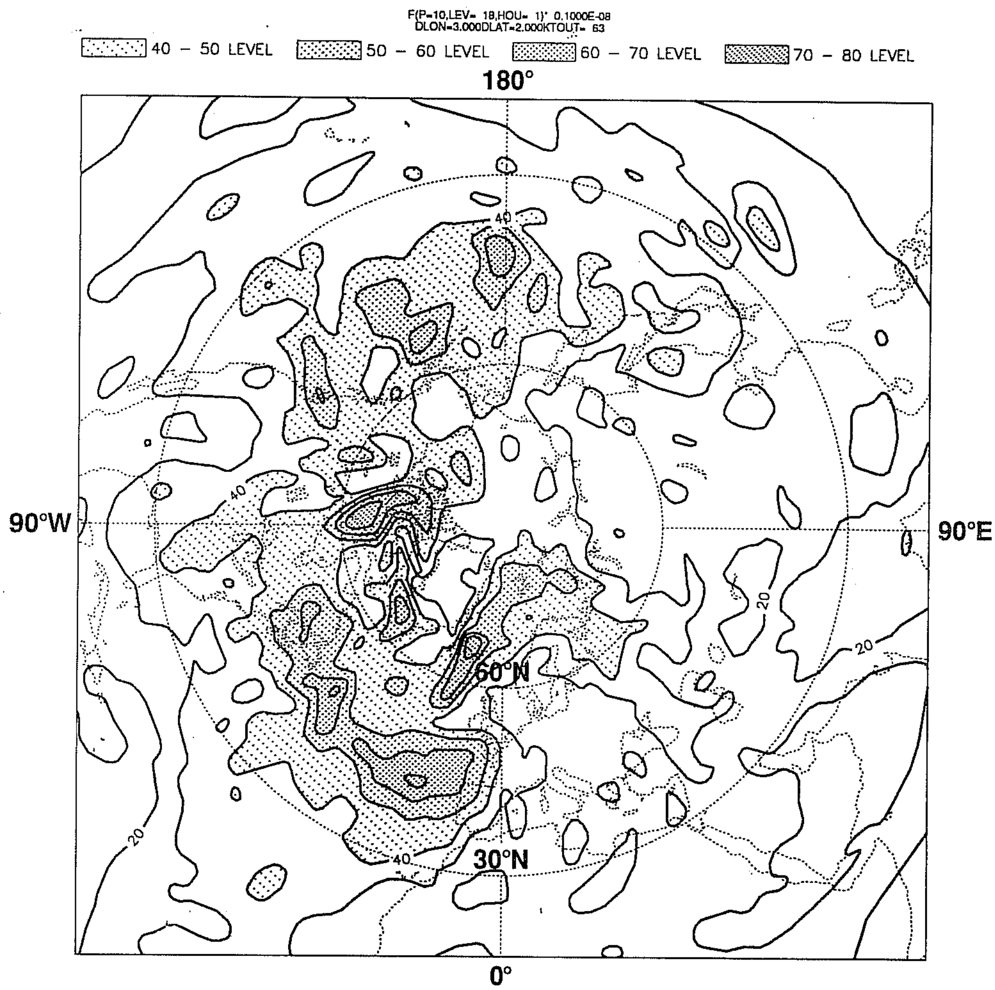

Fig 8 Same as Fig 7 for level 18 (~ 500 hPa).

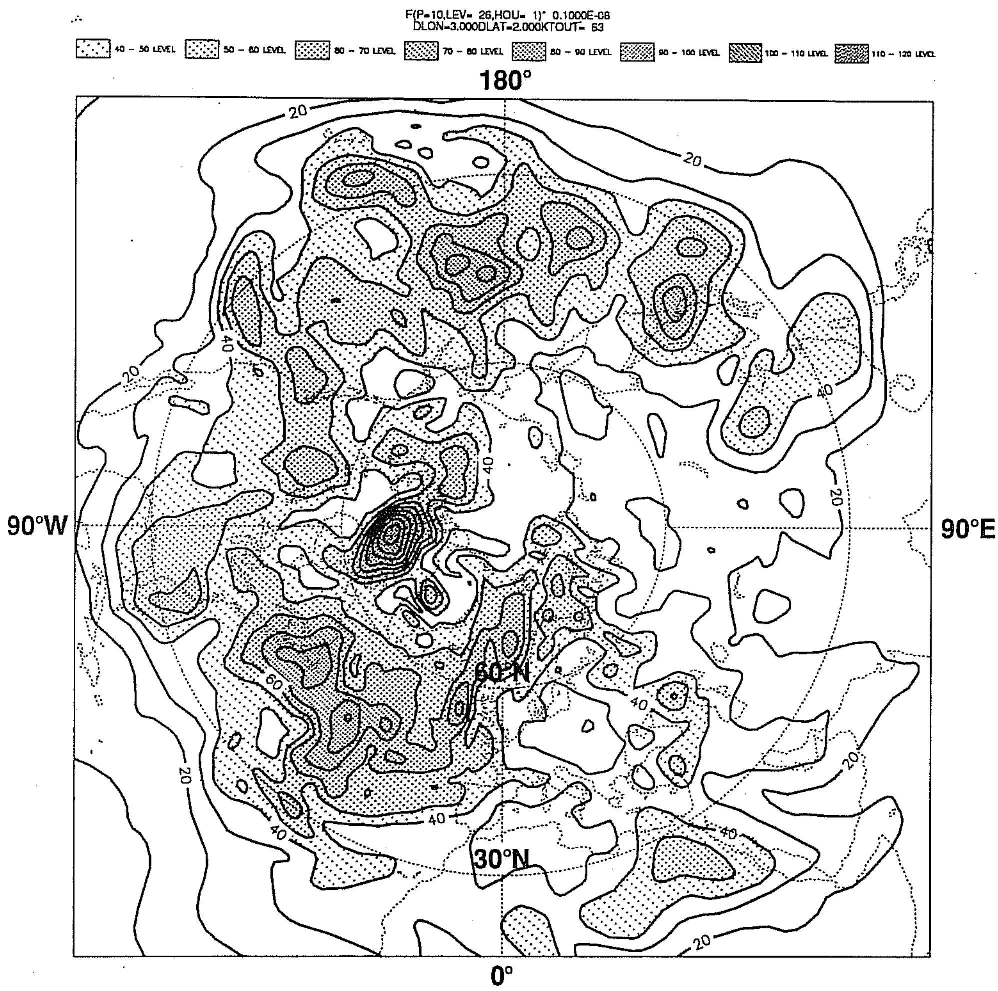

Fig 9 Same as Fig 7 for level 26 (~ 850 hPa).

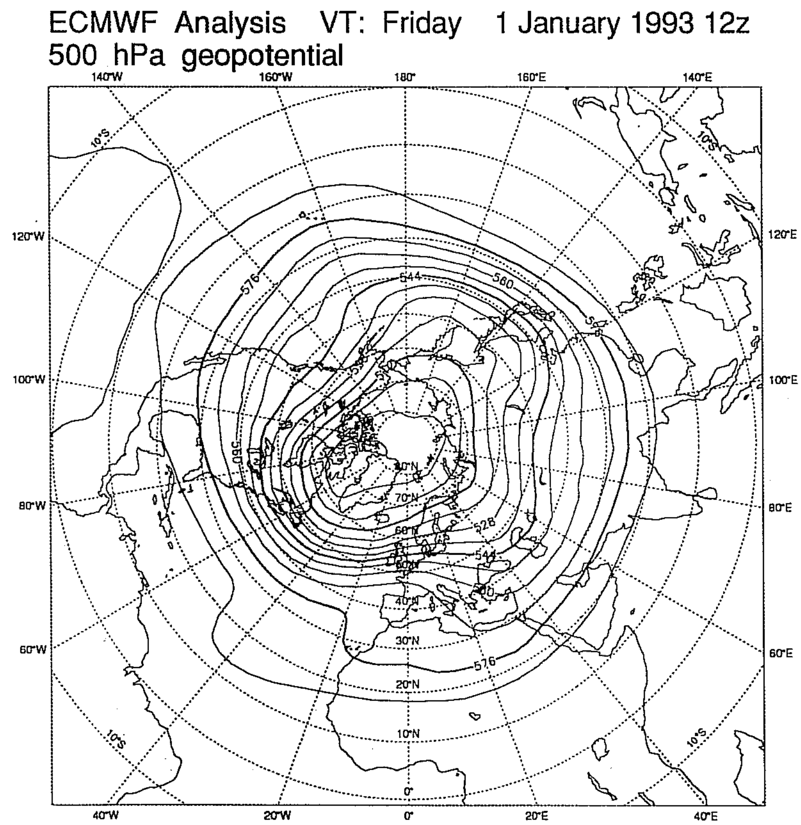

Fig 10 500 hPa geopotential height averaged over the month of January 1993.

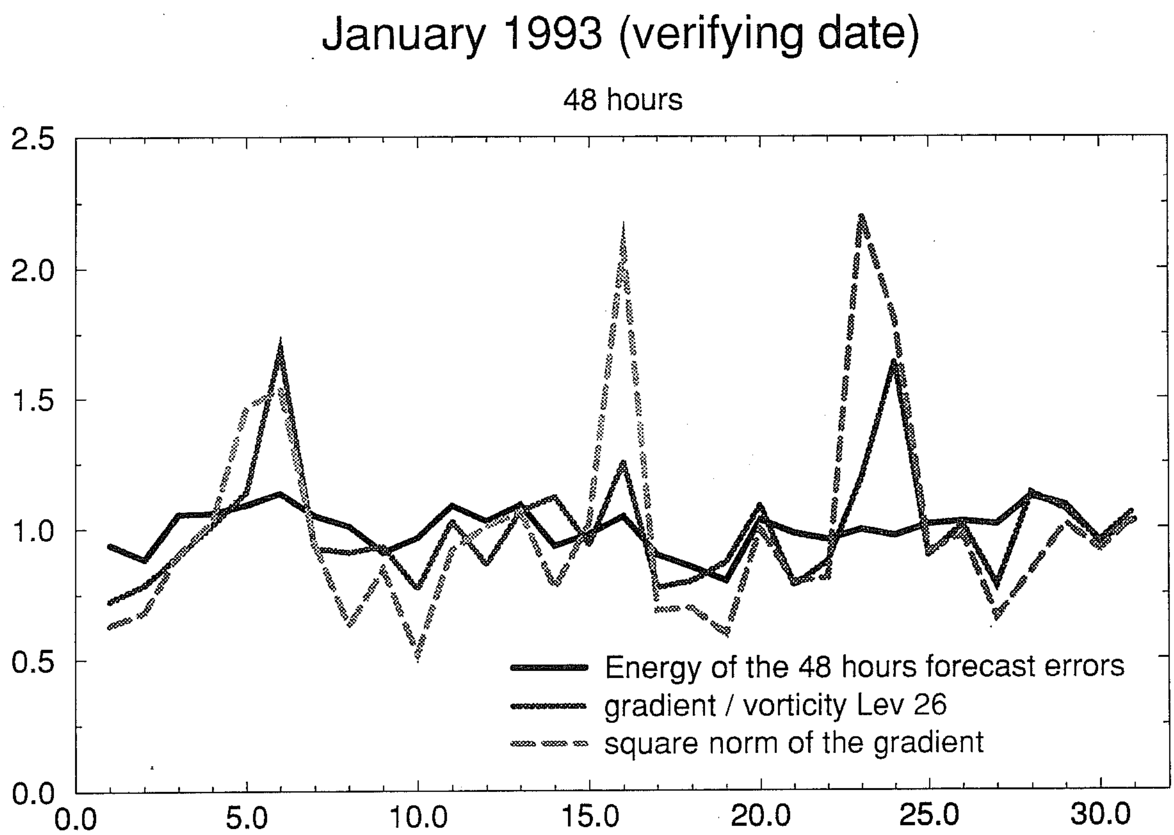

Fig 11 Temporal variation over January 1993 of the cost function (forecast errors square norm), of the square norm of the gradient and of the RMS of the gradient with respect to vorticity at level 26. All curves are normalized and of mean unity.

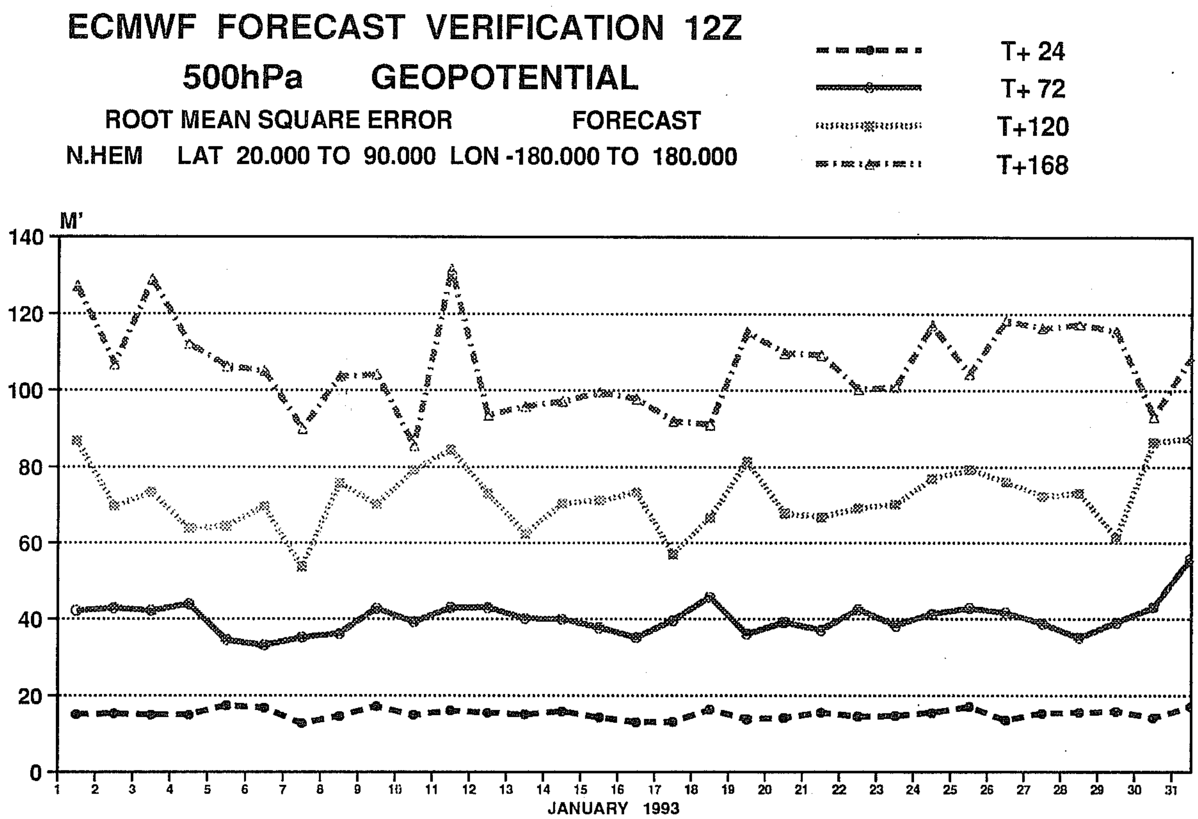

Fig 12 ECMWF operational forecast scores for the month of January 1993.

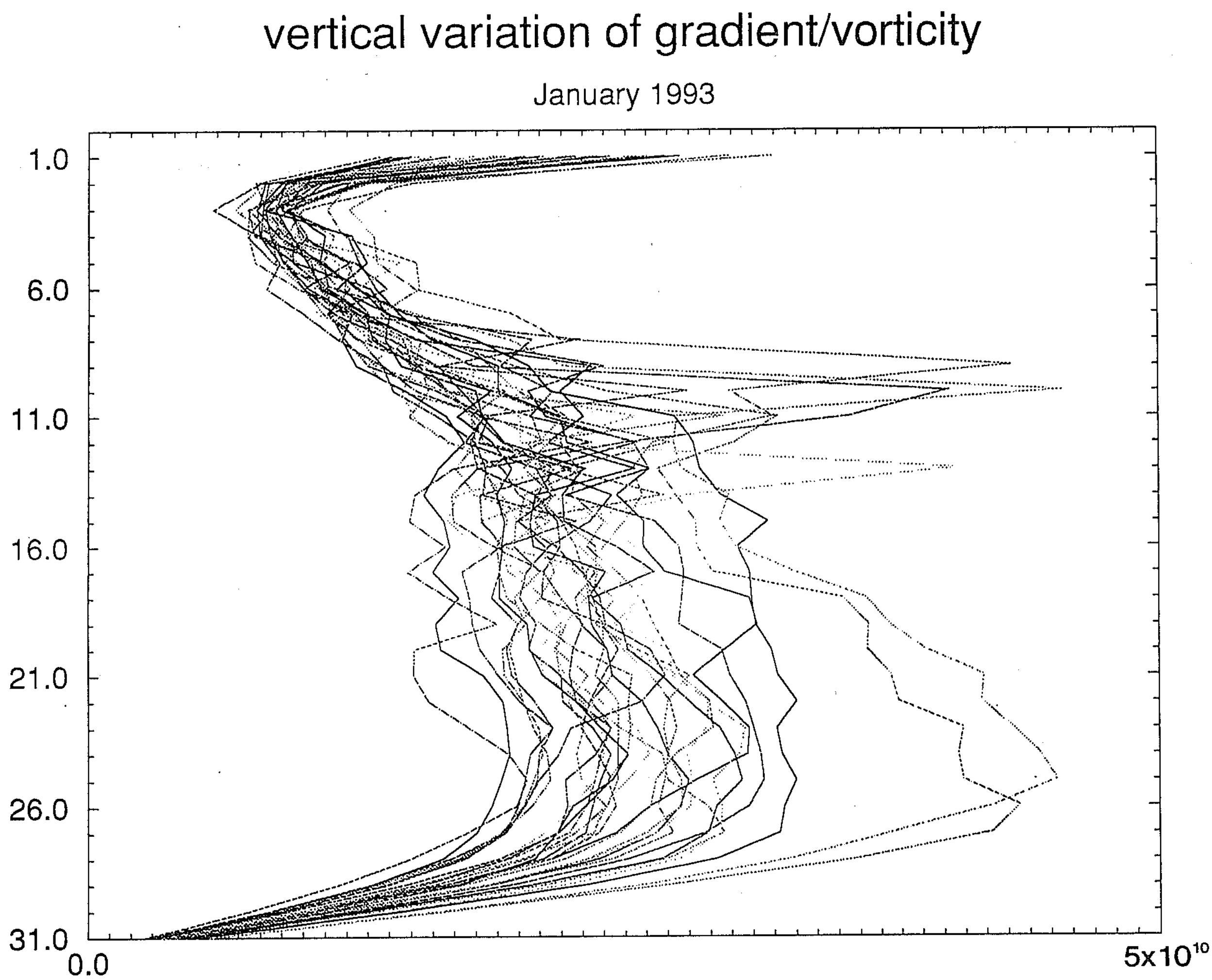

Fig 13 Vertical variation of the RMS of the gradient of the 48 hour forecast errors with respect to vorticity. All individual cases for the month of January 1993. Vertical scale: model level.

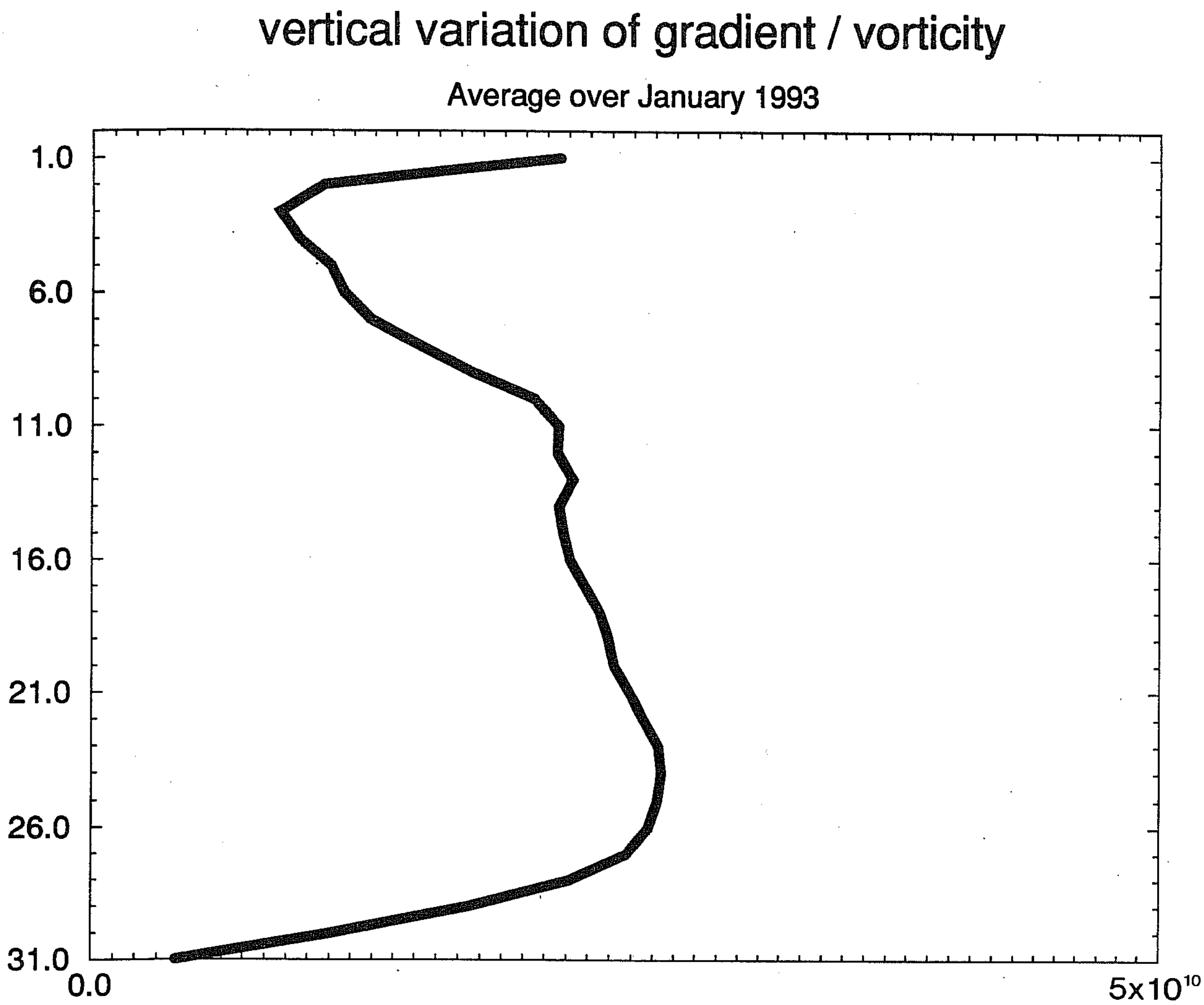

Fig 14 Same as Fig 13 but for the average.

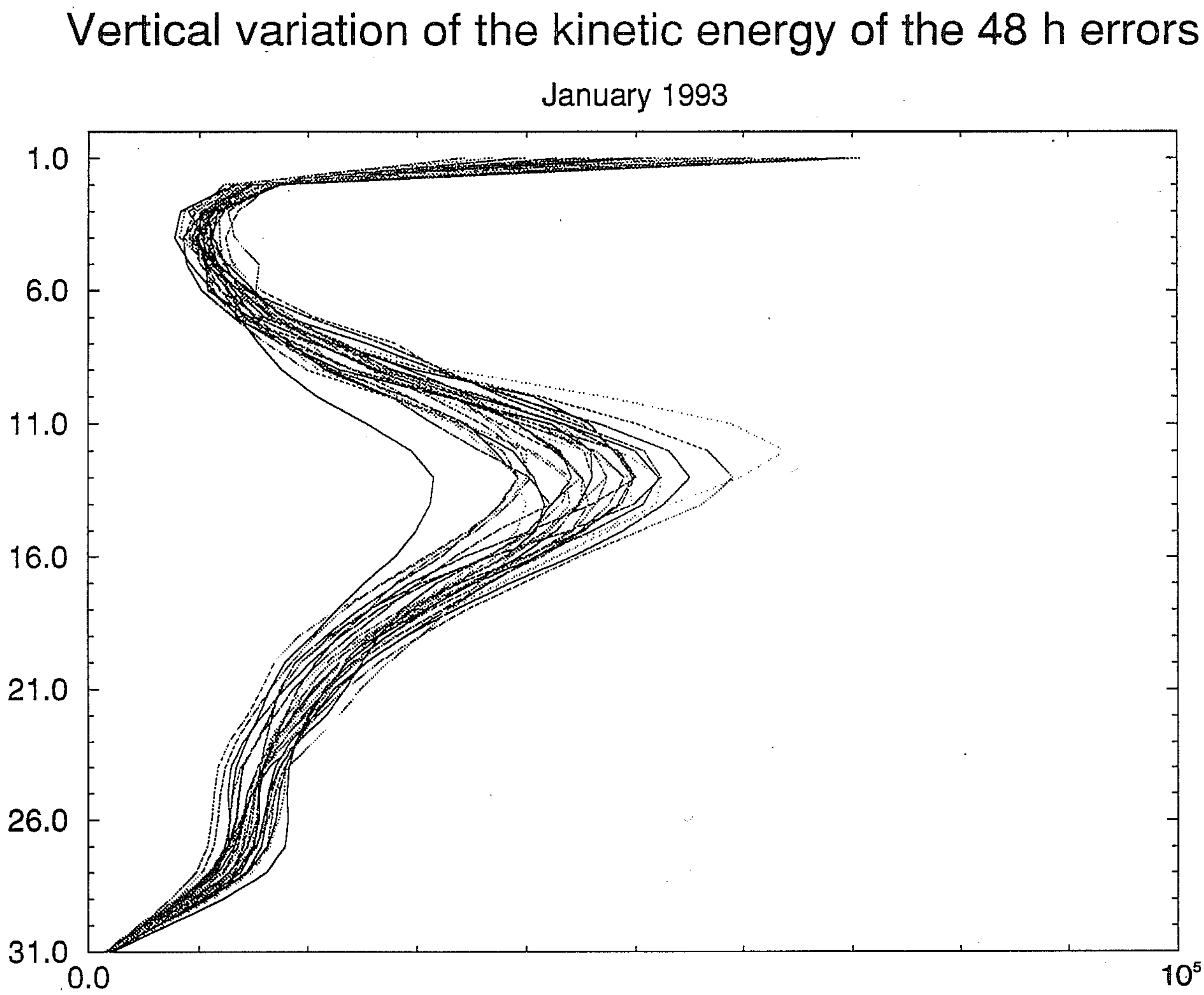

Fig 15 Vertical variation of the kinetic energy of the 48 hour forecast error. All individual cases of January 1993. Vertical scale: model level.

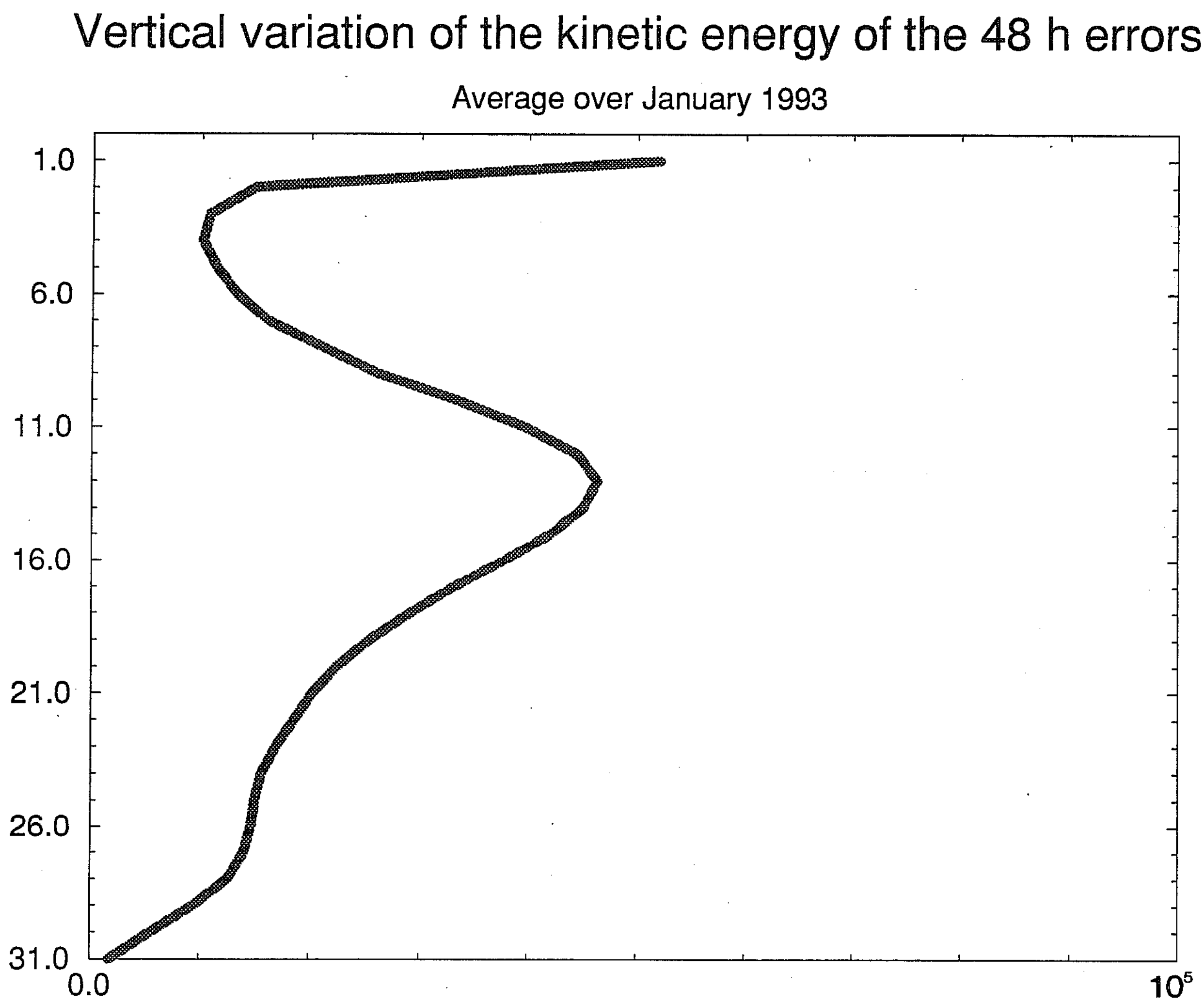

Fig 16 Same as Fig 15 for the average.

This is consistent with the results obtained by *Rabier et al* (1992). This feasibility study led to an operational implementation of the sensitivity computations in February 1994. The results of this routine monitoring and case studies are described in *Rabier et al* (1994).

2.2 Sensitivity to parameters

The input parameters u are not necessarily the initial conditions of a numerical weather prediction model but may rather be some parameters like the orography for an atmospheric model or the atmospheric forcing for an ocean circulation model. *Hall et al* (1982) applied the approach to compute the sensitivity of a radiative convective model to 312 parameters (including basic physical constants). *Courtier* (1987) computed the sensitivity of short range forecast errors to orography in a shallow-water model.

3. ESTIMATION

The adjoint method allows computation of the gradient of a cost function with respect to some parameters in an efficient way. It is then possible to minimize the cost function using descent algorithms like conjugate gradient or quasi-Newton (*Le Dimet and Talagrand,* 1986). This may be applied to several problems like:

3.1 Variational assimilation

Pailleux (1990) proposed a 3D variational analysis scheme (3D-Var) as an alternative to the Optimal Interpolation (OI) scheme. A cost function measuring the departure between the state to be estimated and the various sources of information (background, observations, slow-manifold) is minimized. The adjoint of the operators involved in the distance to the background, to the observations or to the slow manifold allows computation of the gradient of the cost function. 3D-Var is under pre-operational evaluation at ECMWF. It is described in detail in *Courtier et al* (1993). Similar ideas have been implemented operationally at NMC (*Parrish and Derber,* 1992).

The main a priori advantages of 3D-Var as compared to the OI are, first, more flexibility for the description of the spatial structure of the short-range forecast error: both the NMC and ECMWF implementation use spatial correlations which are a non-separable (between the vertical and horizontal direction) function of the spatial distance. The second advantage lies in the easy use of weakly nonlinear observation operators, which is the case for most of the satellite data.

Another strength of the variational formulation is that it readily extends to the time dimension: 4D-Var. One then seeks a model trajectory which best fits the available information (minimize a cost function measuring the distance of the model trajectory to the available information). The gradient of the cost function is evaluated integrating the adjoint of the forecast model. *Thépaut et al* (1993) and *Rabier et al* (1993) demonstrated the ability of 4D-Var to generate flow dependent structure functions. Being iterative, 4D-Var is expensive since it requires several integrations of the model and its adjoint. Ideas to reduce the cost by introducing simplifications in the 4D-Var formulation are discussed in *Courtier et al* (1994): the incremental approach approximates the full sized minimization problem by a quadratic problem involving a lower resolution model. The cost reduction is sufficient to foresee an operational implementation within a few years.

3.2 Parameter estimation - inverse problems

Variational assimilation can be seen as a particular inverse problem (*Tarantola*, 1987) where one estimates some geophysical parameters from observations.

Thépaut and Moll (1990) developed the adjoint of a fast radiative transfer model for inverting TOVS radiances. *Smedstad and O'Brien* (1991), together with data assimilation, retrieved parameters of their ocean model. *Marais and Musson-Genon* (1992) estimated soil parameters fitting a vertical column model to screen level observations.

3.3 Nonlinear equilibration

Vautard and Legras (1988) were looking for "weather regimes" defined as the large scale patterns of the flow which are on average stationary (a statistical equilibration occurs between self interaction and feedback from the small scales). They solve a nonlinear optimization problem in which the cost function is the statistical average of the large scale tendencies. With the statistical averaging defined using ergodicity hypothesis through a long time interval averaging, the authors faced a difficulty: while the time interval was becoming longer and longer, the cost function was converging to an asymptotic value, but not its gradient.

A trivial (and non meteorological) example is the following:

Let $f(t,x) = x \cos tx$ where x is the phase space variable and t time. We have $\frac{1}{T}\int_0^T f(t,x)dt = \frac{1}{T}\sin Tx$. Since $\sin Tx$ is bounded, we have $\lim_{T\to\infty}\frac{1}{T}\sin Tx = 0$ whereas its derivative with respect to x: $\cos Tx$ has no limit when T tends to infinity.

Vautard and Legras solved the problem using an ensemble mean. This points to a difficulty in using the adjoint technique for climatic application: the adjoint is still useful for computing the gradient of a cost function with respect to its input parameters. However, care has to be taken in the definition of the cost function for its gradient to keep a physical meaning.

4. KALMAN FILTERING

Assuming that the forecast error's evolution is governed by the tangent-linear model, the evolution from time t to time $t+T$ of the covariances of forecast errors B reads

$$B(t+T) = R(t,t+T)\ B(t)\ R^{t}(t,t+T) \qquad (5)$$

where $R(t,t+T)$ is the resolvent of the tangent-linear model between times t and $t+T$ (see e.g. *Jaszwinksi*, 1970). Here we have assumed the model to be perfect, no source term is present (the reader will easily generalise to an imperfect model in the following discussion).

This equation is often solved as

$$B(t+T) = R(t,t+T)\ (R(t,t+T)\ B(t))^t$$

by doing product of matrices. This generally requires less CPU time than the original formulation since R has to be evaluated only once. However, it requires the storage of $R(t,t+T)\ B(t)$.

Let us now consider the vector e_i which has 0's everywhere except at the i st position where it is a 1. Then $B(t+T)e_i$ provides the covariances of forecast errors between e_i and all the other model variables: if e_i represents, for example, the surface pressure at a given location, then $B(t+T)e_i$ provides the covariances of the forecast errors of the surface pressure at this particular location with the surface pressure and all other variables at all grid points of the model. Thus we also obtain the variance of error of the surface pressure at this point. $R(t,t+T)$, $B(t)$ and $R(t,t+T)$ being available as operators, $B(t+T)e_i$ is computed efficiently. This relies on the use of the adjoint model R^t.

Bouttier (1993) applied this approach to a global vorticity equation model. Fig 17 (his Fig 4) presents the autocorrelation function of a point with its neighbour at the initial time and for 6 and 24 hour forecasts. One should notice the significant deformation induced by the flow. In order to go from covariances to correlations, he had to compute the diagonal of the covariance matrix (and then the full matrix).

He also computed the 24 hour prediction of the variances of errors induced by idealised flows and by a real situation. Fig 18 (his Fig 20) shows a case of difluence. He relates the maximum of error to the position of the maximum of the jet through barotropic instability which is then advected (and not to difluence in itself). Fig 19 (his Fig 21) presents the forecast of variances of error in the case of a large-scale wave. The maximum of variance is located in the eastern part of the trough. It is consistent with a result obtained by *Barkmeijer* (1992). As in *Veyre* (1991), the approach has been applied to a real situation. The maximum amplification of error is located in the areas of barotropic instability.

5. SINGULAR VECTORS

Let $R(t,t+T)$ be the resolvent of the tangent-linear model over time t, $t+T$. Let $<,>$ be a norm which measures the forecast errors, for example the quadratic invariant of the primitive equations linearized in the vicinity of a state of rest. Let $\delta u(t)$ be the initial errors and $\delta u(t+T)$ the errors at time $t+T$. We have:

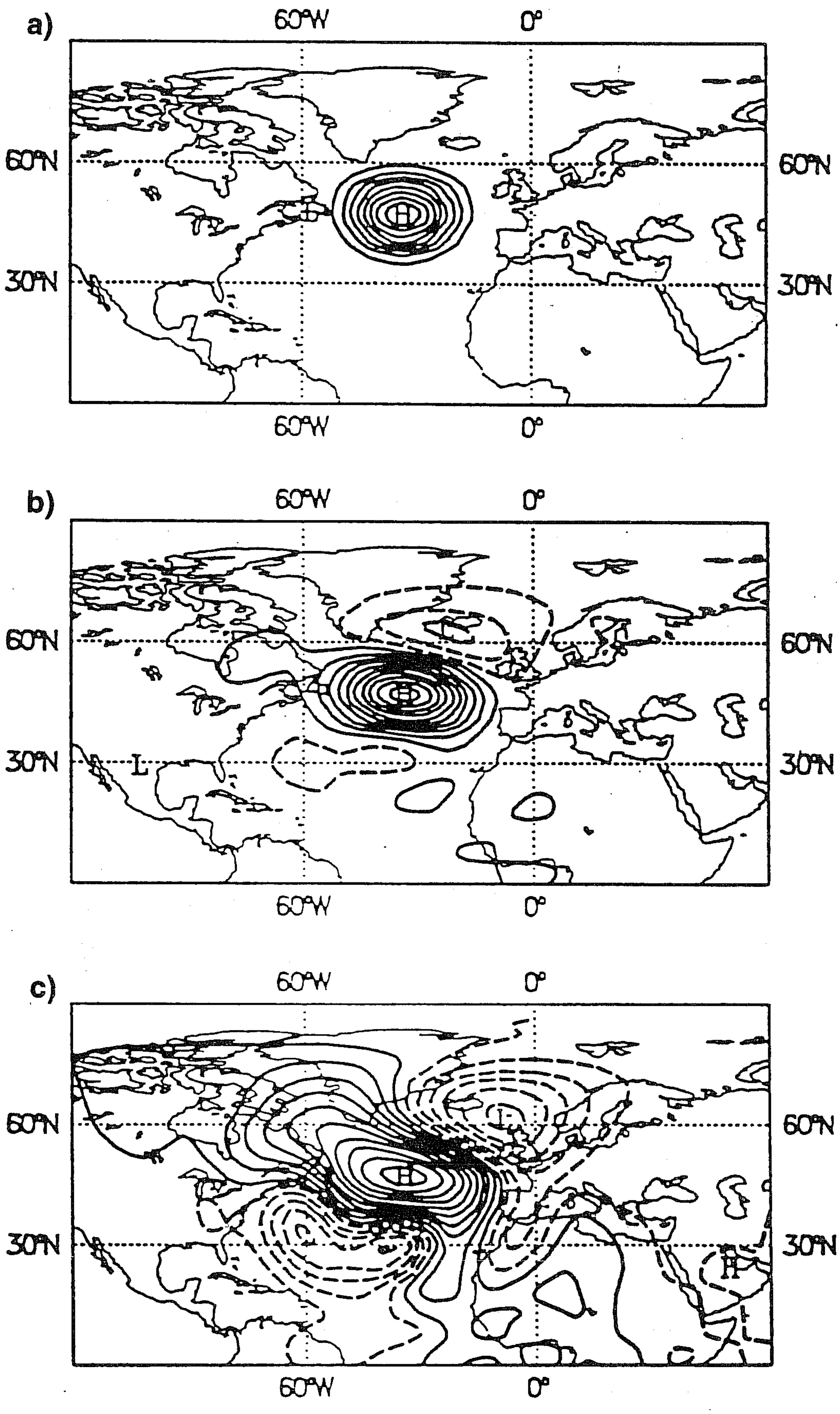

Fig 17 Evolution of the autocorrelation field relative to point (45N, 35 W) in a barotropic vorticity equation model (from *Bouttier*, 1993).

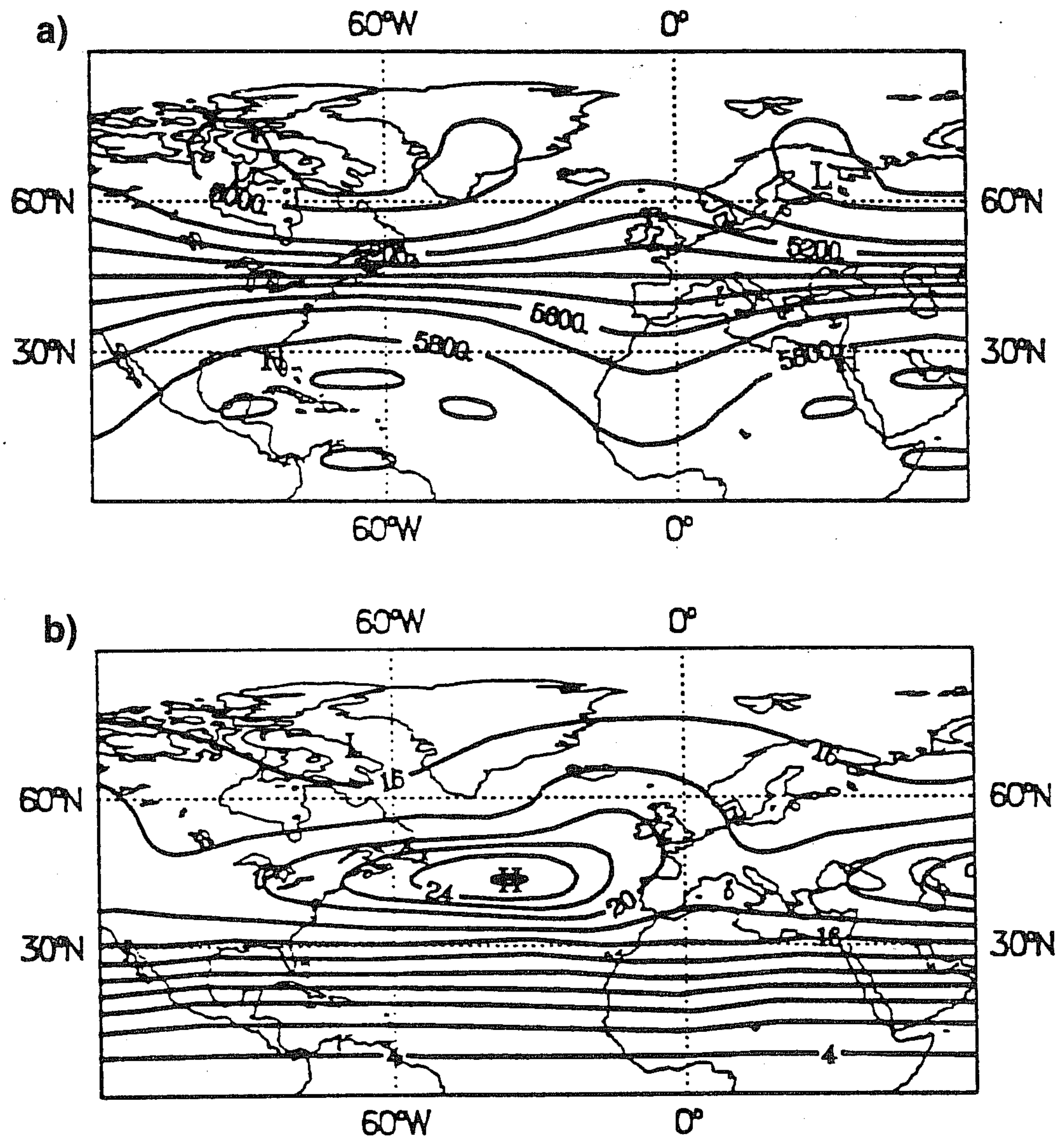

Fig 18 Forecast of height error (6) in the vicinity of the idealised meteorological situation (a) (from *Bouttier*, 1993).

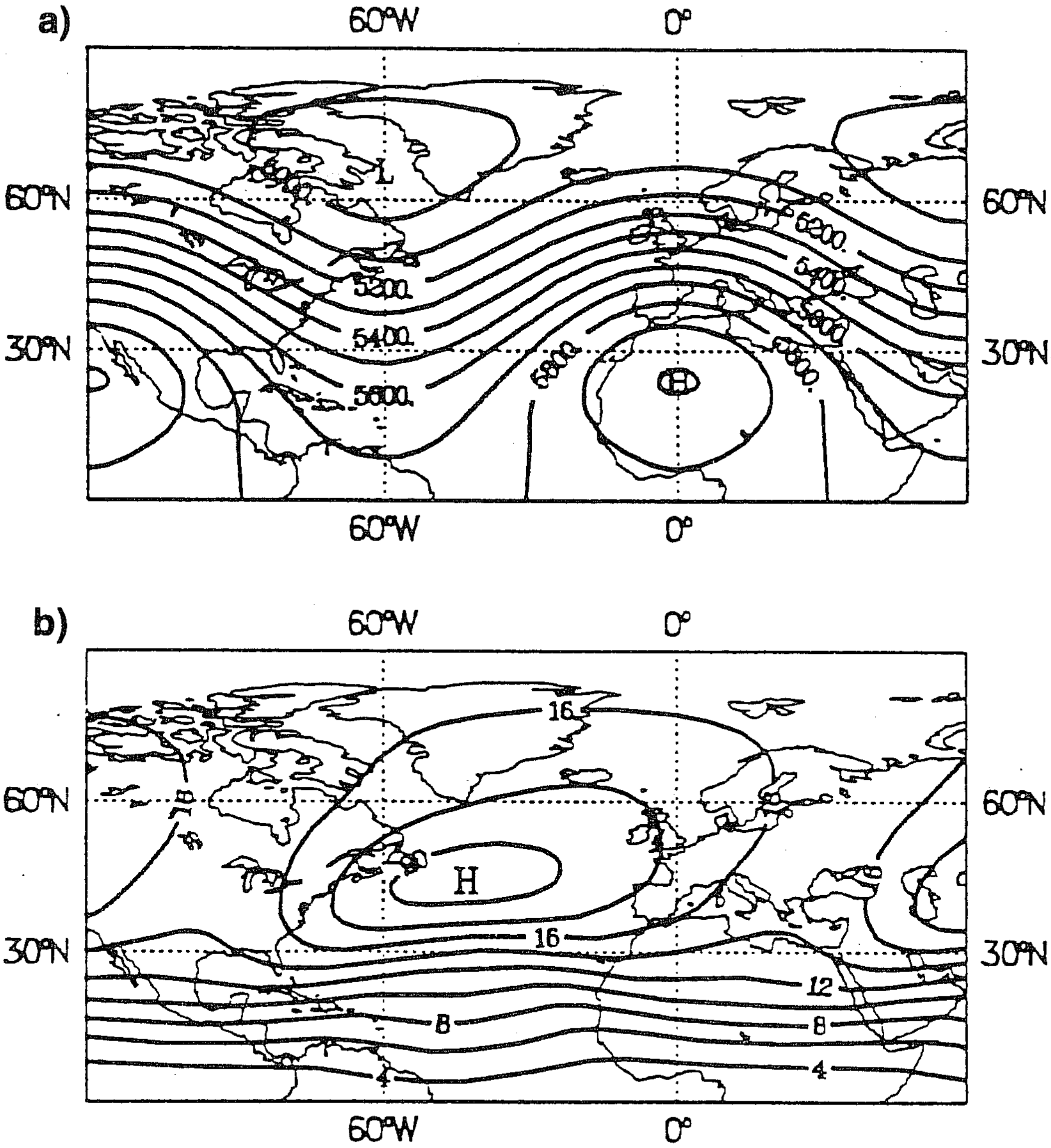

Fig 19 Same as 18 for large scale wave.

$$<\delta u(t+T),\ \delta u(t+T)> \ = \ <R(t,t+T)\delta u(t),\ R(t,t+T)\delta u(t)>$$

$$= \ <R^*(t,t+T)\ R(t,t+T)\ \delta u(t),\ \delta u(t)>$$

which is extreme for the considered norm when $R^*(t,t+T)\ R(t,t+T)\delta u(t)$ is proportional to $\delta u(t)$.

The initial perturbations $\delta u(t)$ which will lead to the largest error according to the chosen norm are then given by the dominant eigenvectors of $R^*(t,t+T)\ R(t,t+T)$. The eigenvectors/eigenvalues of $R^*(t,t+T)\ R(t,t+T)$ are called the singular vectors/values of R in linear algebra. Finding the most unstable perturbations is thus reduced to an eigenvalue problem which may be solved using a Lanczos iterative algorithm. This requires only to be able to compute $R^*(t,t+T)\ R(t,t+T)$ applied to some vectors which makes the algorithm tractable, even for a large-scale problem.

The important point to note is that the eigenvalues of R^*R can be very different from the square of those of R. Let us consider the simple and purely dissipative example illustrated in Fig 20. The trajectory starts from a circle of equi-energy. As there is a strong dissipation along the horizontal axis, the trajectory reaches the axis of weak dissipation without moving significantly along that direction. Then it converges toward the state of rest. This illustrates that, even in a purely dissipative system, it is possible to have growth of energy for a finite time. This happens when two eigenvectors of R are close to parallel but associated to significantly different eigenvalues.

This is a common feature of meteorology and *Farrel* (1989) used this concept for understanding baroclinic instability. *Molteni and Palmer* (1993) computed the singular vector of the tangent-linear version of a 3-level quasi-geostrophic model. The singular vectors computed depend significantly on the time interval chosen: in Fig 21 (their Fig 5), 12 hours, two days and height days. This subsequent evolution is also clearly different, with a strong impact of being optimal for a given range.

Buizza (1993) computed the singular vectors of a T21L19 primitive equation model. He found it necessary to include a simple vertical diffusion and surface drag in order to prevent spurious structures at low level. Fig 22 (his Fig 5) presents the spectrum of the eigenvalues detained for a different time interval. When the time interval is increased, the separation between the first few singular vector increases. The meteorological structures are stable from a one day time interval up to three days. 36 hours is a good compromise between the cost of the method and the meteorological significance of the singular vectors.

Fig 23 (his Fig 6) presents the 6 dominant eigenvectors for a 36 hour time interval at time t_o and at time t_o + 36 hours. They all grow very fast and propagate eastward. The vertical structure depicted in Fig 24

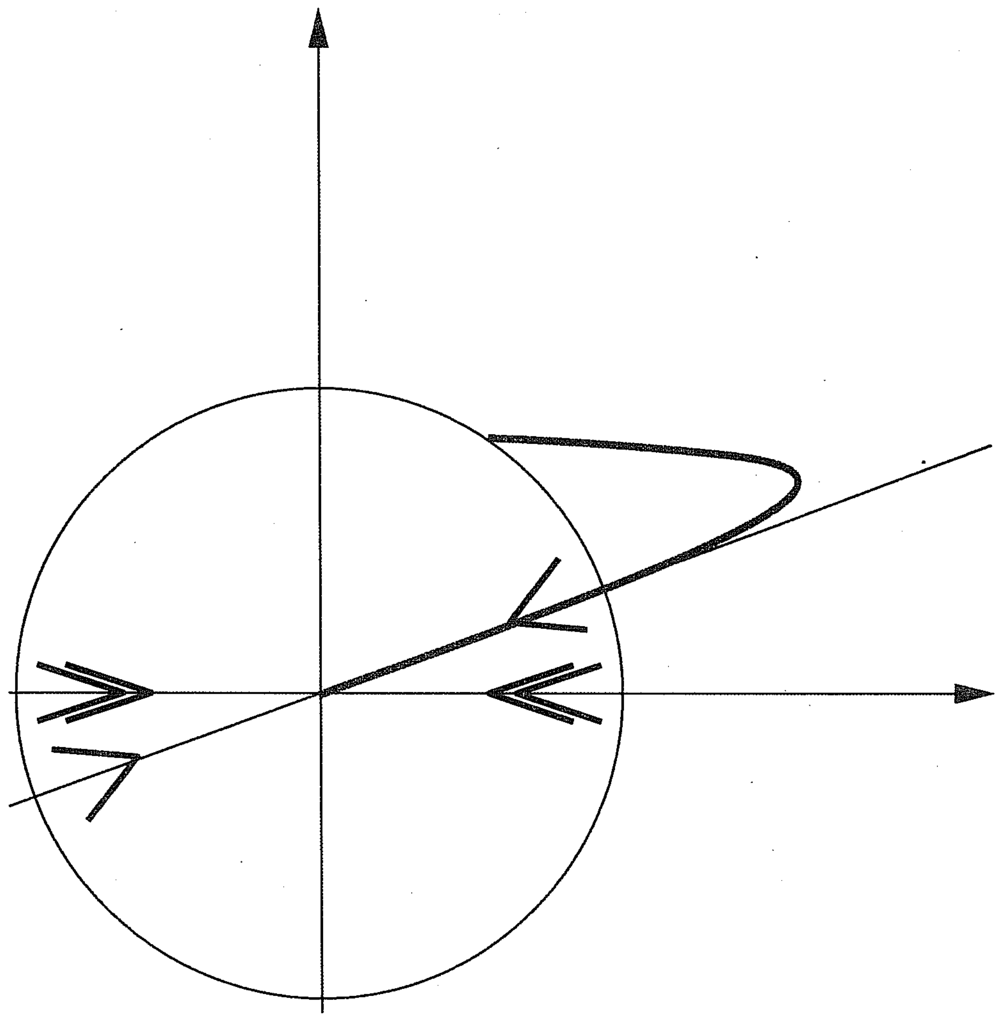

Fig 20 Schematic phase space diagram of a purely dissipative 2D dynamical system. The >> << denotes a strong dissipation in this direction whereas the > < denotes a weak dissipation. The trajectory starting from a circle of equi-energy starts with an energy increase and then converges toward the state of rest.

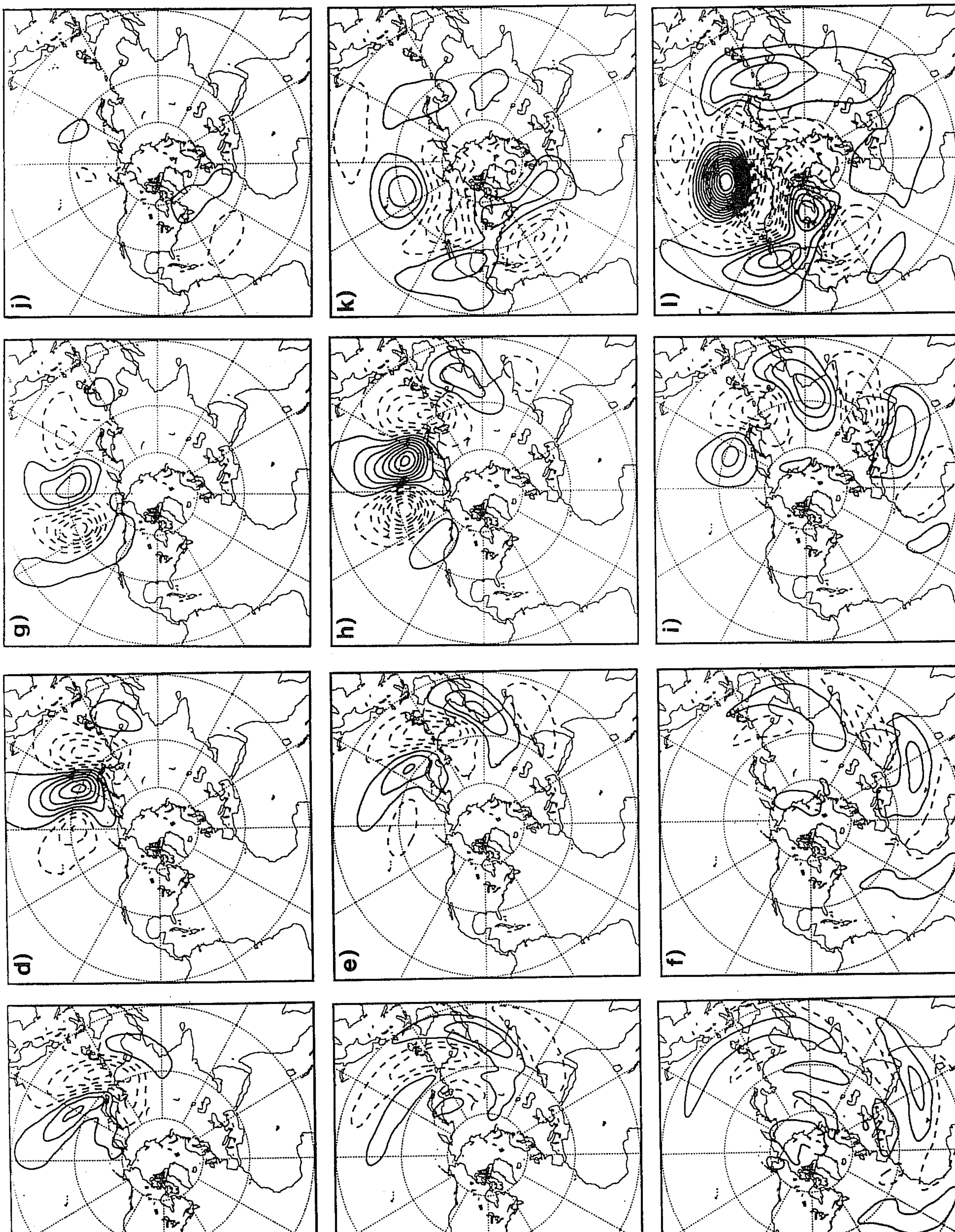

Fig 21 Singular vectors and their subsequent linear evolution computed over a time interval of 12 hours, 2 days and 8 days. (From *Molteni and Palmer*, 1993).

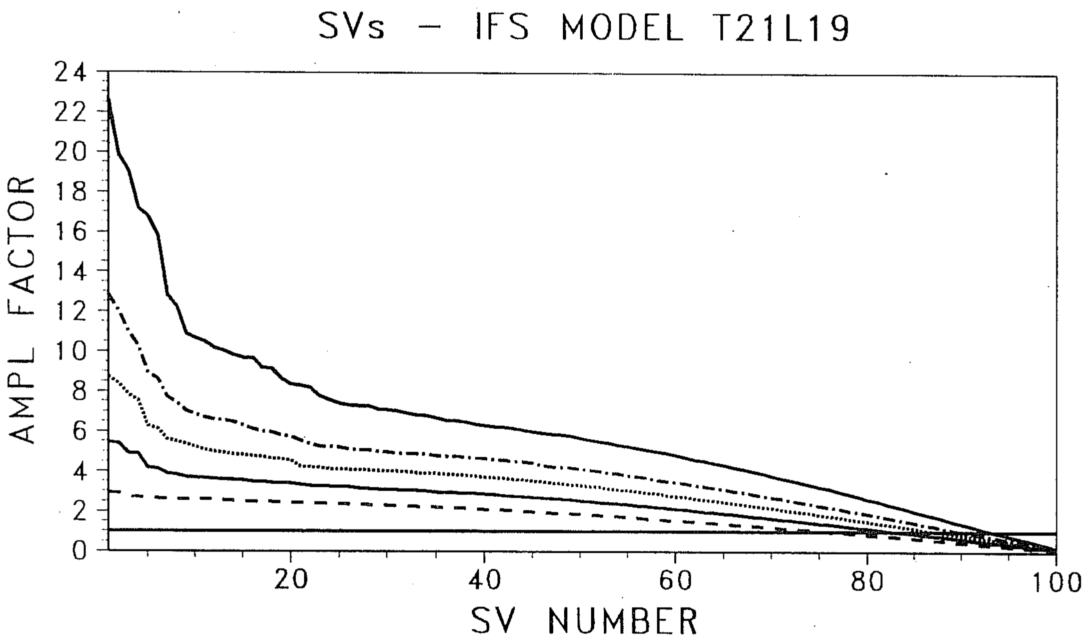

Fig 22 Spectrum of the singular values for different time intervals. (From *Buizza*, 1993).

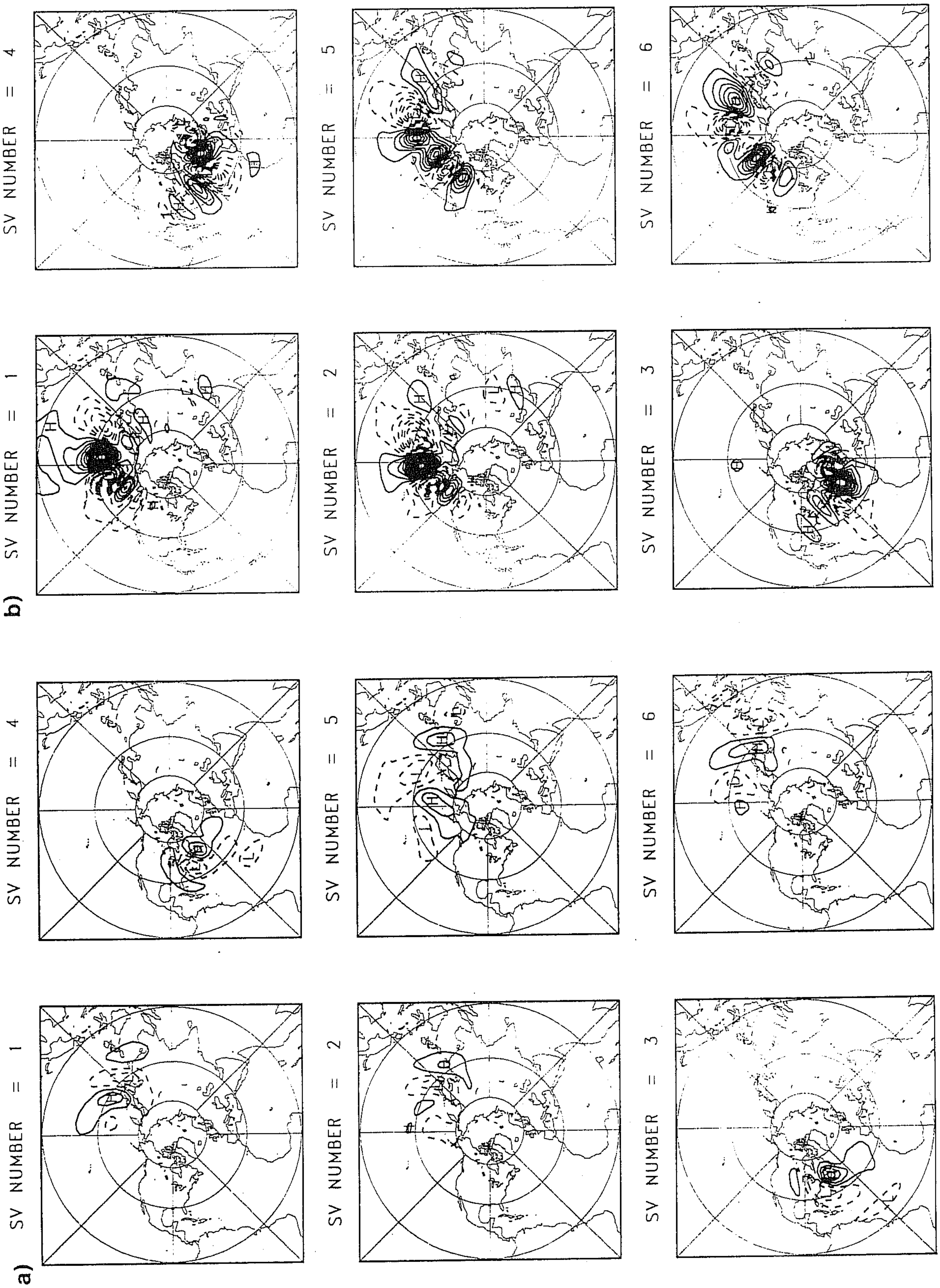

Fig 23 Dominant singular vectors computed over 36 hours a: at time t_o b: at time t_o + 36 hours (From *Buizza*, 1993)

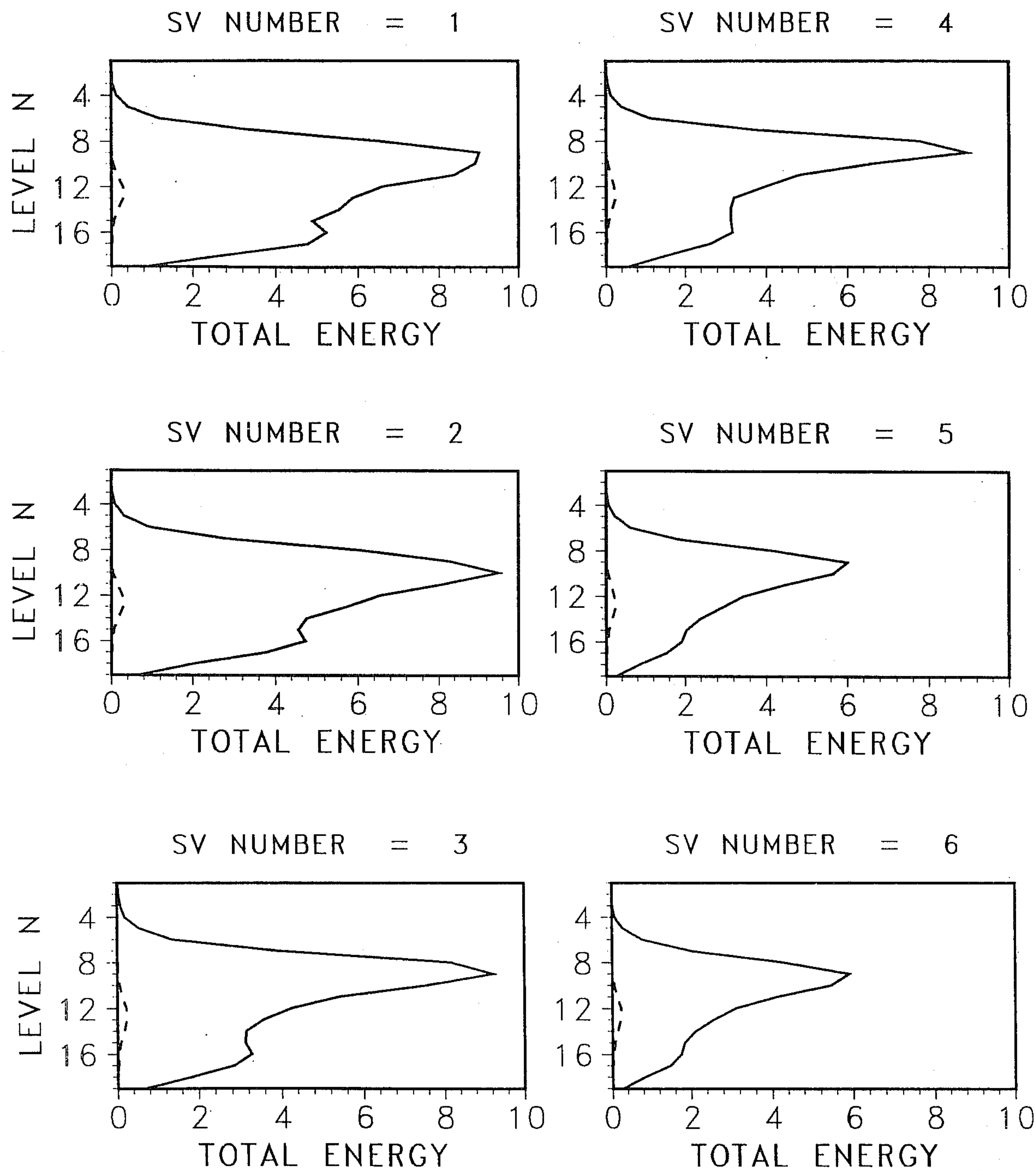

Fig 24 Vertical structure of the dominant singular vectors. (From *Buizza*, 1993).

(his Fig 7) shows a maximum at around 600 hPa at time t_o which is then located at 300 hPa at time t_o + 36 hours. This is consistent with baroclinic instability theories and with the sensitivity experiments reported in section 2.

6. CONCLUSION

The concepts presented in this paper were introduced in the early days of numerical weather prediction. *Sasaki* (1958) may be seen as a precursor to 3D-Var, *Thomson* (1969) introduces the ideas behind 4D-Var, *Jones* (1965) introduces Kalman filtering while *Lorenz* (1965) uses the singular vectors. The adjoint models bring the feasibility of the practical implementation for large-scale problems. Other applications may emerge in the future if applying the transpose of a matrix to a vector is required for an efficient algorithm. At this stage, we should mention the second order adjoint (*Wang et al*, 1992) which allows the computation of the Hessian of a cost function applied to a vector in a minimization problem.

The use of the adjoint model has a major limitation: it requires the tangent-linear model to be meteorologically realistic for a finite amplitude perturbation. 4D-Var results and, quite convincingly, the recent sensitivity study of *Rabier et al* (1994) indicate that this is indeed the case for the adiabatic evolution of perturbations of order of magnitude comparable to the analysis error, over a 2-day time interval. The stiffness of the adiabatic primitive equations does not seem to be a critical problem.

However, the introduction of physical parametrizations seems to be more critical. Significant progress is being made by *Vukicevic and Errico* (1993), *Zou et al* (1993) and *Zupanski* (1993), but it is probably fair to say that much work remains to be done in this area. This might even impact on the physical parametrization design with regularisation widely introduced in the formulation.

REFERENCES

Barkmeijer, J, 1992: Local error growth in a barotropic model. Tellus, 44A, 314-323.

Baur, W and V Strassen, 1983: The complexity of partial derivatives. Theoretical Comp Sci, 22, 317-330.

Bouttier, F, 1993: The dynamics of error covariances in a barotropic model. Tellus, 45A, 408-423.

Buizza, R, 1993: Impact of a simple vertical diffusion scheme and of the optimization time interval on optimal unstable structures. ECMWF Tech Memo No 192 (available from ECMWF).

Courtier, P, 1987: Application du contrôle optimal à la prévision numérique en météorologie. Thèse de Doctorat de l'Université Pierre et Marie Curie, Paris, France.

Courtier, P, E Andersson, W Heckley, G Kelly, J Pailleux, F Rabier, J-N Thépaut, P Undén, D Vasiljevic, C Cardinali, J Eyre, M Hamrud, J Haseler, A Hollingsworth, A McNally and A Stoffelen, 1993: Variational assimilation at ECMWF. ECMWF Tech Memo 194 (available from ECMWF).

Courtier, P, J Derber, R Errico, J-F Louis and T Vukicevic, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. Tellus, 45A, 342-357.

Courtier, P, J-N Thépaut and A Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. Q J R Meteorol Soc, 120, 1367-1387.

Errico, R and T Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model. Mon Wea Rev, 120, 1644-1660.

Farrel, B F, 1989: Optimal excitation of baroclinic waves. J Atmos Sci, 46, 1193-1206.

Hall, M C G, D G Cacuci and M E Schlesinger, 1982: Sensitivity analysis of a radiative-convective model by the adjoint method. J Atmos Sci, 39, 2038-2050.

Jaszwinksi, A H, 1970: Stochastic processes and filtering theory. Academic Press, New York, 376 pp.

Jones, R H, 1965: Optimal estimation of initial conditions for numerical prediction. J Atmos Sci-, 22, 658-663.

Le Dimet, F-X and O Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations and theoretical aspects. Tellus, 38A, 97-110.

Lorenz, E N, 1965: A study of the predictability of a 28-variable atmosphere model. Tellus, 17, 321-333.

Marais, L and L Musson-Genon, 1992: Forecasting the surface weather elements with a local dynamical adaptation method using a variational technique. Mon Wea Rev, 120, 1035-1049.

Marchuk, G I, 1974 (Russian version 1967). The numerical solution of problems of atmospheric and oceanic dynamics, 387 pp. Gidrometeoizdat, Leningrad, USSR. (English translation, Rainbow systems, Alexandria, va.).

Molteni, F and T N Palmer, 1993: Predictability and finite-time instability of the Northern Hemisphere circulation. Q J R Meteorol Soc, 119, 269-298.

Morgenstern, J, 1985: How to compute a function and all its derivatives fast, a variation on the theorem of Baur-Strassen. Sigact News, 16, 60-62.

Morgenstern, J, 1973: Algorithmes linéaires tangents et complexité. Compte rendu de l'Academie des Sciences t.277:367.

Pailleux, J, 1990. A global variational assimilation scheme and its application for using TOVS radiances. Proceedings World Meteorological Organisation, International Symposium on Assimilation of Observations in Meteorology and Oceanography, Clermont-Ferrand, France, 9-13 July 1990, 325-328.

Parrish, D F and J C Derber, 1992: The National Meteorological Center's Spectral Statistical Interpolation analysis system. Mon Wea Rev, 110, 1747-1766.

Rabier, F, P Courtier and O Talagrand, 1992: An application of adjoint models to sensitivity analysis. Beitr Phys Atmosph, 65, 177-192.

Rabier, F, P Courtier, J Pailleux, O Talagrand and D Vasiljevic, 1993: A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. Q J R Meteorol Soc, 119, 845-880.

Rabier, F, P Courtier, M Herveou, B Strauss and A Persson, 1993: Sensitivity of forecast error to initial conditions using the adjoint model. ECMWF Tech Memo 197, available from ECMWF.

Rabier, F, E Klinker, P Courtier and A Hollingsworth, 1994: Sensitivity of two-day forecast errors over the Northern Hemisphere to initial conditions. Submitted to Q J R Meteorol Soc.

Sasaki, Y, 1958: An objective analysis based on the variational method. J Meteor Soc, Japan, 36, 77-88.

Simmons, A J and B Hoskins, 1978: The life cycles of some nonlinear baroclinic waves. J Atmos Sci, 35, 414-432.

Smedstad, O M and J J O'Brien, 1991: Variational data assimilation and parameter estimation in the equatorial Pacific Ocean. Prog Oceanog, 26, 179-241.

Tarantola, A, 1987: Inverse problem theory, method for data fitting and model parameter estimation. Elsevier, Amsterdam, 613 pp.

Thépaut, J-N and P Moll, 1990: Variational inversion of simulated TOVS radiances using the adjoint technique. Q J R Meteorol Soc, 116, 1425-1448.

Thépaut, J-N, R Hoffman and P Courtier, 1993: Interaction of dynamics and observations in a four-dimensional variational assimilation. Mon Wea Rev, 121, 3393-3414.

Thomson, P N, 1969: Reduction of analysis error through constraints of dynamical consistency. J Appl Meteor, 8, 738-742.

Vautard, R and B Legras, 1987: On the source of midlatitude low-frequency variability. Part II: nonlinear equilibration of weather regimes. J Atmos Sci, 45, 2845-2867.

Veyre, P, 1991: Direct prediction of error variances by the tangent-linear model: a way to forecast uncertainty in the short range. ECMWF workshop on new developments in predictability, 13-15 November 1991, pp 65-86.

Vukicevic, T and R M Errico, 1993: Linearization and adjoint of parametrized most diabatic processes. Tellus, 45A, 493-510.

Wang, Z, I-M Navon, F-X Le Dimet and X Zou, 1992: The second order adjoint analysis: theory and application. Meteorology and Atmospheric Physics, 50, 3-20.

Zou, X, I M Navon and J G Sela, 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. Tellus, 45A, 370-387.

Zupanski, D, 1993: The effects of discontinuities in the Betts-Miller cumulus convective scheme on four-dimensional variational data assimilation. Tellus, 45A, 511-524.