EXPERIMENTS IN DATA ASSIMILATION USING THE ADJOINT MODEL TECHNIQUE

P. Courtier
Direction de la Météorologie
Paris, France

1. Introduction.

The general formalism of adjoint equations is presented in the article by Talagrand (this volume), which will be referred to as T85 in the following. The present article describes two series of data assimilation numerical experiments performed with the adjoint equations. The two series of experiments used models based respectively on the vorticity equation and on the shallow water equations. The data used were radiosonde observations of wind and geopotential.

Following the approach described in section 3 of T85, the adjoint equations are utilized in order to determine the gradient, with respect to the initial conditions, of a scalar function measuring the "distance" between a model solution and the available observations. The gradient thus computed is then introduced in a descent algorithm in order to determine the initial conditions which minimize the distance function.

2. The vorticity equation and its adjoint.

2.1 The continuous case

The vorticity equation at the surface of a rotating sphere reads

$$\frac{\partial \zeta}{\partial t} = J(\zeta+f,\Delta^{-1}\zeta) \qquad (2.1)$$

where ζ and f are the vorticities of the relative motion and basic rotation respectively, t is time and Δ^{-1} the inverse two-dimensional laplacian operator. J is the jacobian operator : $J(a,b) = \nabla a \times \nabla b = \nabla \times (a \nabla b)$ where ∇ denotes the first-order differential operator along the surface of the sphere.

For a given solution ζ of (2.1), the tangent linear equation is

$$\frac{\partial \delta\zeta}{\partial t} = J(\delta\zeta,\Delta^{-1}\zeta) + J(\zeta+f,\Delta^{-1}\delta\zeta) \qquad (2.2)$$

The determination of the adjoint of (2.2) requires the prior definition of an inner product on the space of all possible vorticity fields, i.e. the space of all regular functions on the sphere with mean equal to 0. The total kinetic energy K corresponding to a given vorticity field ζ is equal to

$$K = \frac{1}{2}\int_{\Sigma}(\nabla\Delta^{-1}\zeta.\nabla\Delta^{-1}\zeta)\,d\Sigma \qquad (2.3)$$

where the dot denotes scalar product of ordinary vectors in physical space, and Σ the surface of the sphere. Expression (2.3) defines a norm and it leads to the natural inner product of two vorticity fields ζ_1 and ζ_2

$$\langle \zeta_1 | \zeta_2 \rangle = \int_{\Sigma}(\nabla\Delta^{-1}\zeta_1 \cdot \nabla\Delta^{-1}\zeta_2)\,d\Sigma \qquad (2.4a)$$

which by using Green's formula can be written as

$$\langle \zeta_1 | \zeta_2 \rangle = -\int_{\Sigma}\zeta_1(\Delta^{-1}\zeta_2)\,d\Sigma = -\int_{\Sigma}(\Delta^{-1}\zeta_1)\zeta_2\,d\Sigma \qquad (2.4b)$$

It immediately results that the laplacian Δ and its inverse Δ^{-1} are self adjoint for this inner product, i.e. for any ζ_1 and ζ_2

$$\langle \Delta\zeta_1 | \zeta_2 \rangle = \langle \zeta_1 | \Delta\zeta_2 \rangle \qquad (2.5a)$$

$$\langle \Delta^{-1}\zeta_1 | \zeta_2 \rangle = \langle \zeta_1 | \Delta^{-1}\zeta_2 \rangle \qquad (2.5b)$$

The jacobian J verifies the property that for any three scalar fields a,b,c

$$\int_\Sigma J(a,b)c\, d\Sigma = \int_\Sigma aJ(b,c)d\Sigma \qquad (2.6)$$

as is seen from the following equalities

$$\int_\Sigma J(a,b)c\, d\Sigma = \int_\Sigma (\nabla\times(a\nabla b)c)d\Sigma = \int_\Sigma a(\nabla b\times\nabla c)d\Sigma = \int_\Sigma aJ(b,c)d\Sigma \qquad (2.7)$$

Three scalar fields α, $\delta\zeta$, $\delta\zeta'$ being given, use of (2.4b) and (2.6) yields

$$\begin{aligned} \langle J(\delta\zeta,\alpha)\,|\,\delta\zeta'\rangle &= -\int_\Sigma J(\delta\zeta,\alpha)\Delta^{-1}\delta\zeta' d\Sigma \\ &= -\int_\Sigma \delta\zeta J(\alpha,\Delta^{-1}\delta\zeta')d\Sigma \\ &= -\int_\Sigma \delta\zeta\Delta^{-1}(\,\Delta J(\alpha,\Delta^{-1}\delta\zeta')\,)d\Sigma \\ &= \langle\delta\zeta\,|\,\Delta J(\alpha,\Delta^{-1}\delta\zeta')\rangle \end{aligned}$$

which shows that for given α, the adjoint of the linear operator $\delta\zeta \to J(\delta\zeta,\alpha)$ is the operator $\delta\zeta' \to \Delta J(\alpha,\Delta^{-1}\delta\zeta')$. Using the fact that Δ^{-1} is self adjoint (2.5b) and the result that the adjoint of the product of two operators is the product of their adjoints taken in reverse order, we obtain for the adjoint equation of (2.2)

$$\frac{\partial\delta^*\zeta}{\partial t} = \Delta J(\Delta^{-1}\delta^*\zeta,\Delta^{-1}\zeta) + J(\zeta+f,\Delta^{-1}\delta^*\zeta) \qquad (2.8)$$

2.2 The discrete case

The experiments whose results will be presented below were performed with a pseudo-spectral model of the vorticity equation (2.1), built on the spherical harmonics Y_n^m, with triangular truncation T at degree N. A scalar field Ψ is then defined by its components Ψ_n^m along the spherical harmonics. The spherical harmonics expression for the laplacian operator

$$(\Delta\Psi)_n^m = -\frac{n(n+1)}{a^2}\Psi_n^m \quad \text{where a is the radius of the sphere}$$

leads to the following discretized expression for the inner product (2.4)

$$\langle \zeta_1 | \zeta_2 \rangle = \sum_T \frac{a^2}{n(n+1)} \zeta_{1n}^m \overline{\zeta_{2n}^m}$$

It is immediate that the operators Δ and Δ^{-1} remain self-adjoint with that expression for the inner product. As for the jacobian operator, the spatial derivatives which appear in its expression are computed in the model spectral space, and are therefore exact. Following the standard procedure for pseudo-spectral models (Jarraud, 1984), the quadratic terms are computed in physical space on a grid with enough resolution to ensure that the harmonics resolved in spectral space are free of aliasing errors. It is shown in Annex 1 that under these conditions the model adjoint equation is still given by (2.8) provided the continuous jacobian is replaced by its aliasing error free discretized form.

The time discretization of the model uses the leapfrog differentiating scheme, initialized with an Euler step. The corresponding adjoint turns out to be also a leapfrog integration of the adjoint equation as can be seen by taking the transpose of the matrix notation for the leapfrog scheme.

3. Numerical results with the vorticity equation.

The numerical model is truncated at degree N = 21. A vorticity field is then completely described by 483 independant real parameters (taking into account the fact that the m = n = 0 component is necessarily zero and the m = 0 components are real), among which 252 parameters are components along symmetric harmonics (n - m even) and 231 are components along the antisymmetric harmonics (n-m odd). However, all experiments were performed with antisymmetric vorticity fields, so that the effective number of parameters of our problem is 231.

The integration of the adjoint equation (2.8) required about twice as much computing time as the integration of the direct equation (2.1). This doubling is basically due to the fact that differentiation has produced two jacobians in (2.8).

3.1 Numerical validation of the adjoint model and the descent process.

The initial condition $\zeta = aY_1^0 + bY_n^m$, which defines a Haurwitz wave with wave number m, has been used for producing a complete space-time history $\zeta_{ob}(t)$ of the vorticity field over a 12 hour time interval (t_0,t_1). The distance-function to be minimized has been defined by

$$\mathcal{J}(\zeta(0)) = \sum_{t=0}^{12h} \langle \zeta(t) - \zeta_{ob}(t) | \zeta(t) - \zeta_{ob}(t) \rangle$$

where the summation extends over all time steps. The corresponding forcing term (see eq. (3.11), T85) which has to be added to the homogeneous adjoint equation (2.8) is readily seen to be equal to

$$- 2(\zeta(t) - \zeta_{ob}(t)).$$

The "observation" $\zeta_{ob}(t)$ having been produced by the model itself, the minimizing $\zeta(t)$ is of course $\zeta_{ob}(t)$ itself, and the corresponding value of $\mathcal{J}$ is 0. Starting the minimization from an atmosphere at rest, and using the conjugate gradient algorithm for the descent process, the value of $\mathcal{J}$ decreased by about two orders of magnitude at each descent step, and the vorticity field $\zeta_{ob}(t_0)$ was reconstructed to within an accuracy of 10^{-9} s^{-1} after five descent steps. The corresponding difference on the height field is of the order of a few centimetres.

3.2 Experiments with radiosonde data.

The basic correctness of the algorithm having been established as just described, a second series of experiments was performed on real observations. These were radiosondes observations of height and wind at the level 500 mb over the northern hemisphere and for the 24 hour period starting at 26 April 1984 0.00Z. The geographical distribution of the observations is shown on figure 1. A total of Nh = 1653 individual observations of height; Nv=1913 individual observations of each of the two components of the wind was used. These observations were of course mostly concentrated at the synoptic hours. The 500mb height for 26 april 1984 0.00Z as produced by the operational analysis of Météorologie Nationale, Paris is shown on figure 2.

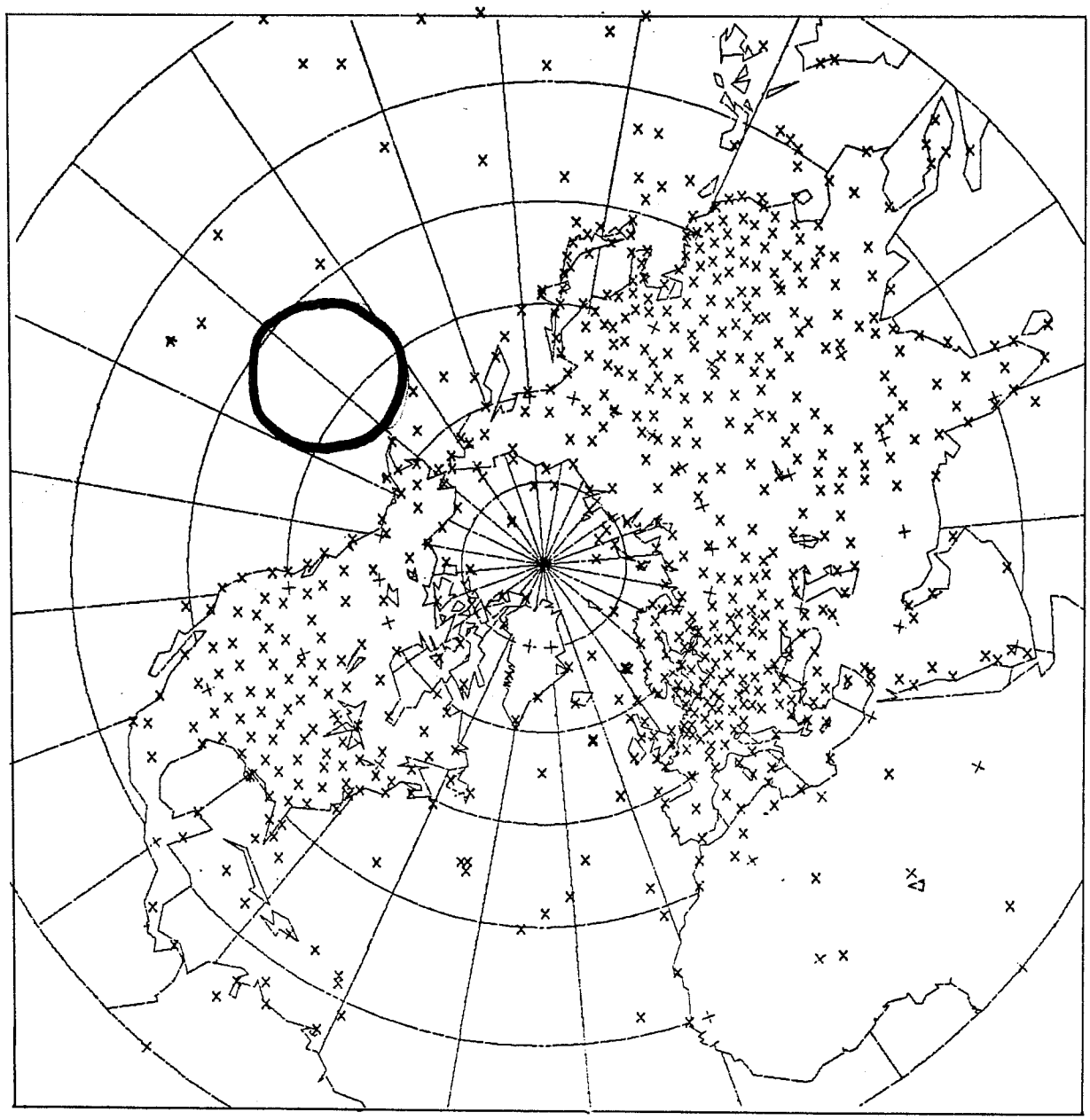

Figure 1. Distribution of radiosondes observations

26 April 1984

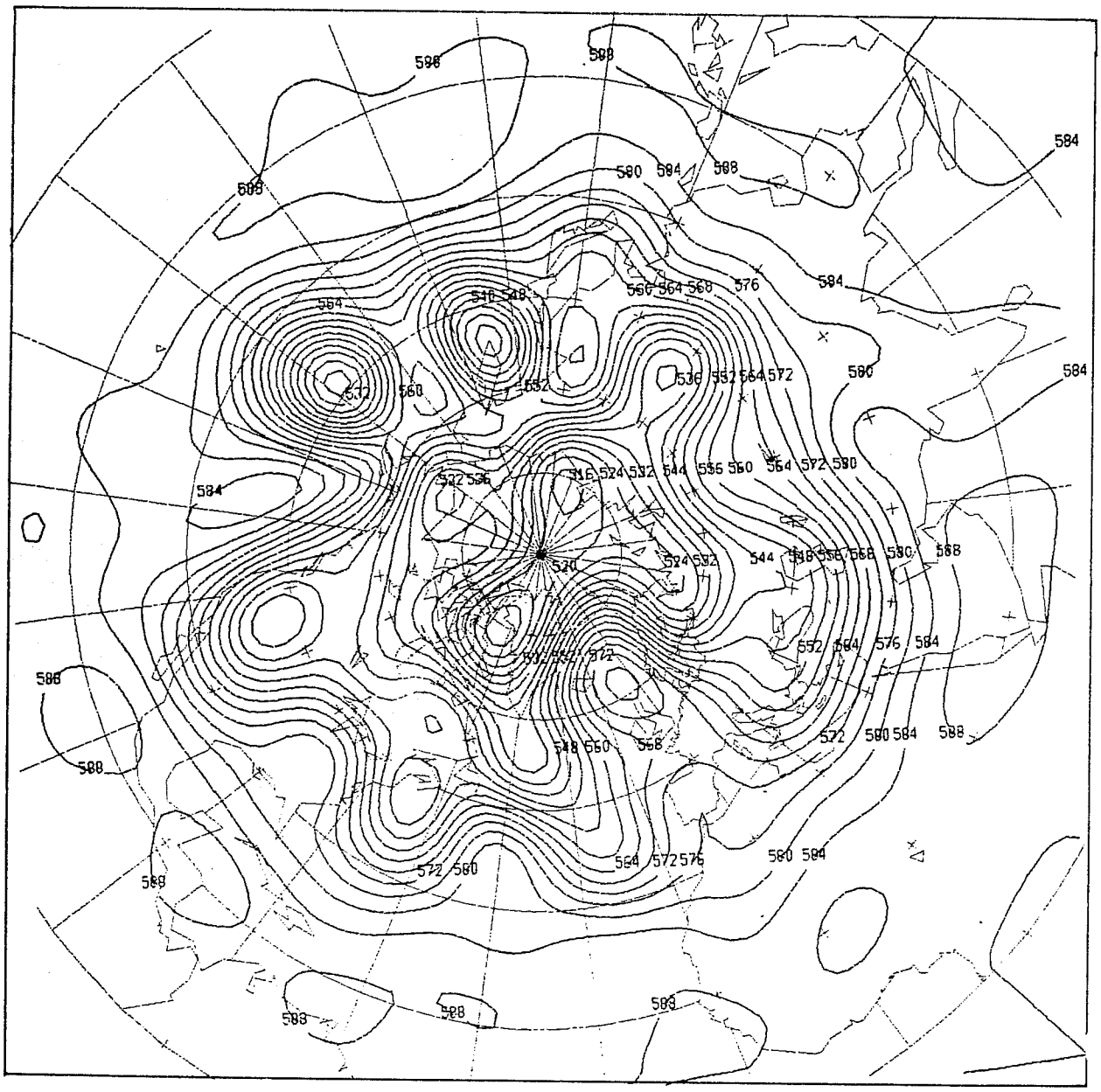

Figure 2. 500mb geopotential field of
Operational analysis, Paris.
26 April 1984 0.00Z

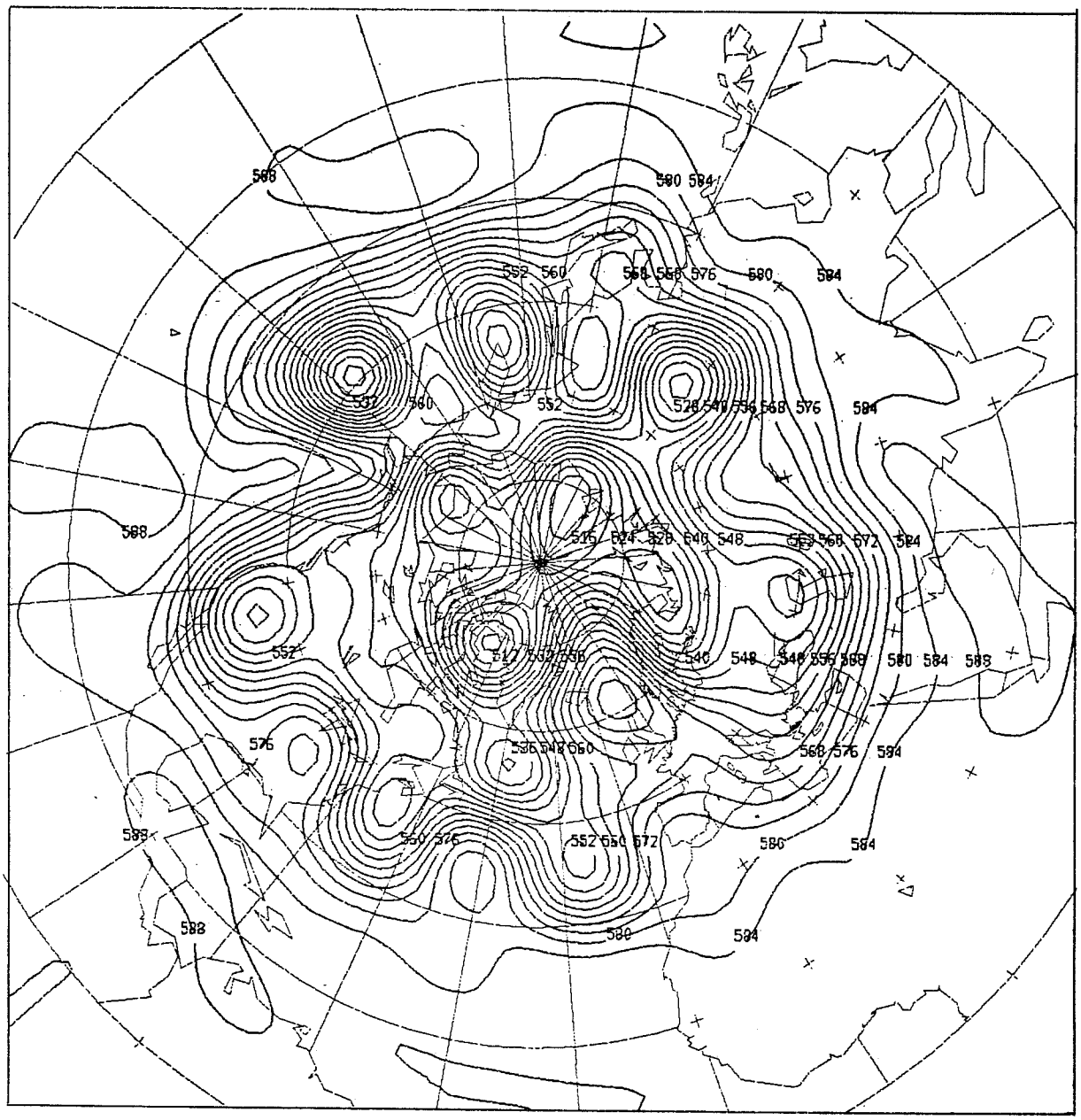

Figure 3. Same as figure 2 for

Variational analysis

26 April 1984 0.00Z

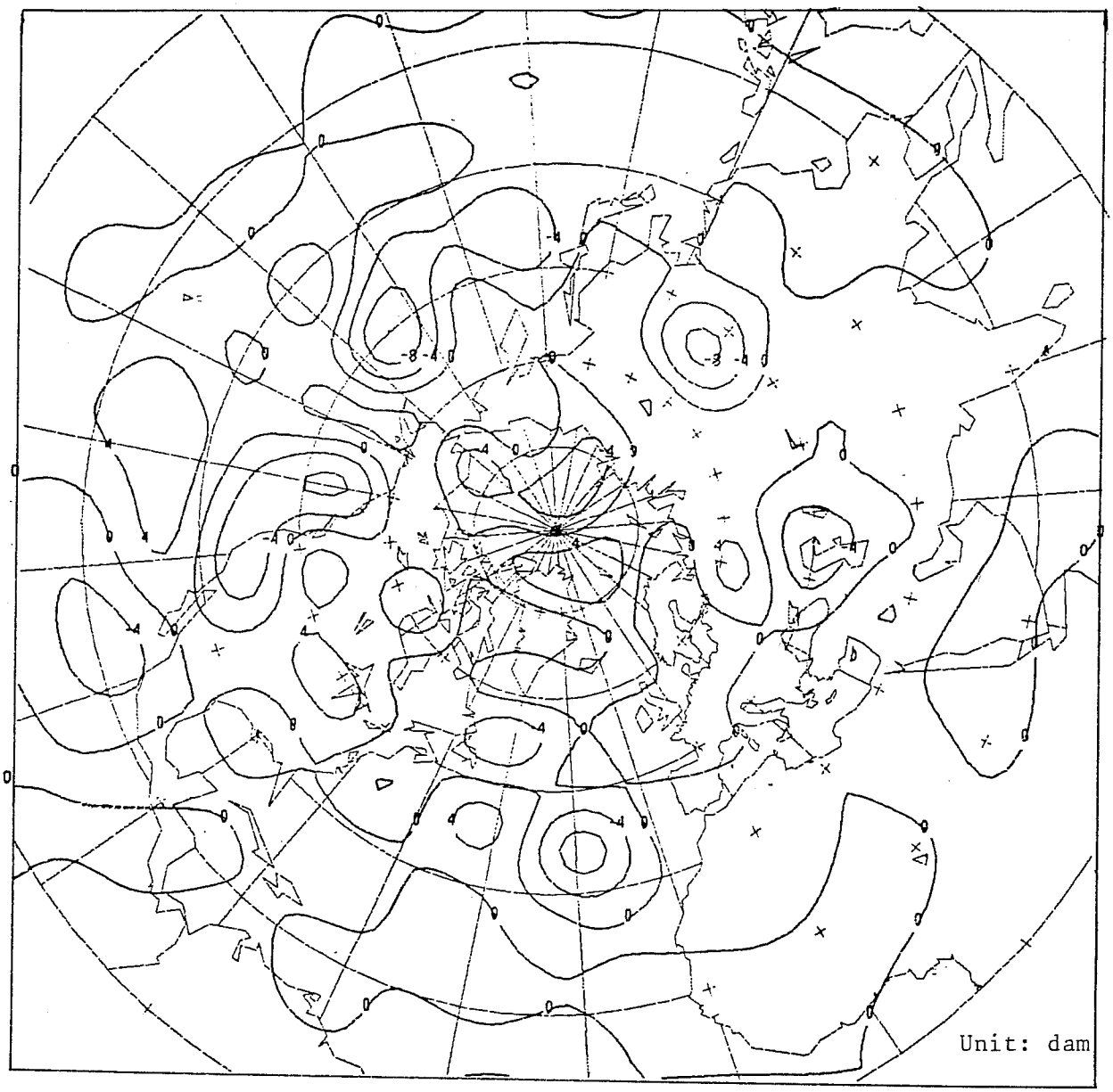

Figure 4. Differences between operational analysis and variational analysis. 26 April 1984 0.00Z

The functional to be minimized has been defined by (3.1) $\mathcal{J} = \alpha \mathcal{J}_\phi + \mathcal{J}_v$ where the contribution $\mathcal{J}_v$ of the wind observations (u_i, v_i) has been taken as

$$\mathcal{J}_v = \sum_{i=1}^{Nv} (\hat{u}_i - u_i)^2 + (\hat{v}_i - v_i)^2 \qquad (3.2)$$

where $\hat{u}_i$ and $\hat{v}_i$ are the values determined, as described in Annex 2, by solving the corresponding Poisson equation from the vorticity field ζ. The contribution $\mathcal{J}_\phi$ of the Nh height observations ϕ_i has been taken as :

$$\mathcal{J}_\phi = \sum_{i=1}^{Nh} \left(\hat{\phi}_i - \phi_i - \frac{1}{Nh} \sum_{j=1}^{Nh} (\hat{\phi}_j - \phi_j) \right)^2 \qquad (3.3)$$

where the values $\hat{\phi}_i$ of height at observation locations are obtained from ζ through a complete balance equation as described in Annex 2. The average $\sum_{j=1}^{Nh} (\hat{\phi}_j - \phi_j)$ has been introduced in order to make the cost function independant of the mean height which has to be arbitrarily choosen in the solution of the balance equation.

The integration of the inhomogeneous adjoint equation required at each observation time the explicit determination of the gradient of $\mathcal{J}_v$ and $\mathcal{J}_\phi$ with respect to the model variable, i.e. the spectral components of the vorticity field ζ. The mappings $\zeta \rightarrow (\hat{u}_i , \hat{v}_i)$ and $\phi \rightarrow \hat{\phi}_i$ are not analytically simple and the direct explicit determination of their gradients may not be easy. Accordingly, the adjoint of these mappings have been used, following again the general principle described in section 2 of T85, in order to determine these three gradients. The details of the corresponding computations are described in Annex 3.

A number of experiments have been performed, in which the spatial and temporal distribution of observations retained in the distance function $\mathcal{J}$ has been varied, together with a number of parameters, in particular the coefficient α of (3.1). Figure 3 shows the height field produced by a minimization performed with all available observations and the value $\alpha = .03\ m^{-2}s^2$. The minimization was started with an atmosphere at rest ($\zeta = 0$) and the descent process used a quasi-Newton algorithm.

Comparison of figure 2 and 3 shows that all major structures of the flow are reconstructed by the minimization process. It is particularly remarkable that the Aleutian depression is reconstructed. As can be seen from figure 1 where the edge of the depression has been delineated, no observations were available in the depression itself. Since that depression was no present in the state from which the assimilation was

started, it is necessarily through the time continuity implied by the evolution equation (2.1) and through the non-linear balance equation that the assimilation process has been able to "deduce" that a depression had to exist in the Aleutian area.

Figure 4 shows the difference field between figure 3 and 2. This difference field is not easy to interpret, especially because the operational analysis used observations performed at the analysis time or before it, while the variational analysis used observations performed at the analysis time and after it. The large differences over China, the Caspian Sea and the Atlantic Ocean are probably due to the differences between the two sets of observation. The differences over the Pacific Ocean are due to the lack of radiosonde reports there.

The root mean square difference corresponding to the final minimum of $\mathcal{J}$ was 28.5 m and 7.9 ms^{-1} per individual observation of height and wind vector respectively (against 185 m and 17.6 ms^{-1} at the start of the process). Theses values are two to three times as large as usual estimates of analysis error in four-dimensional assimilation procedures performed with operational forecasting models. Taking into account the simplicity of the evolution equation (2.1), the low truncation retained, and the fact that observations at only one level were used, they seem quite acceptable.

Study of the variation of the fields in the course of the minimization process shows that the first step of the process reconstructs the latitudinal gradient of geopotential. In the following ten steps, or so, structures are progressively built up over data rich areas and their vicinity, with no associated modification of the fields elsewhere. From that time, continuation of the minimization process modifies the fields only over data poor areas with no further significant decrease of the distance function. The modifications then introduced over data poor areas are only noise with no meteorological significance, as can be seen from figure 5 which shows the difference of the vorticity fields (which is much more noisy than geopotential fields because of the two differentiations) between the operational and variational analyses after twenty steps of descent process. Figure 5 shows noise over data poor areas (as in the Atlantic and Pacific Oceans and the Sahara) mostly concentrated in scales which turn out to be the smallest scales resolved by the model.

Figure 5. Differences between the vorticity fields produced by the operational and the variational analyses (fig 2 and 3 respectively)

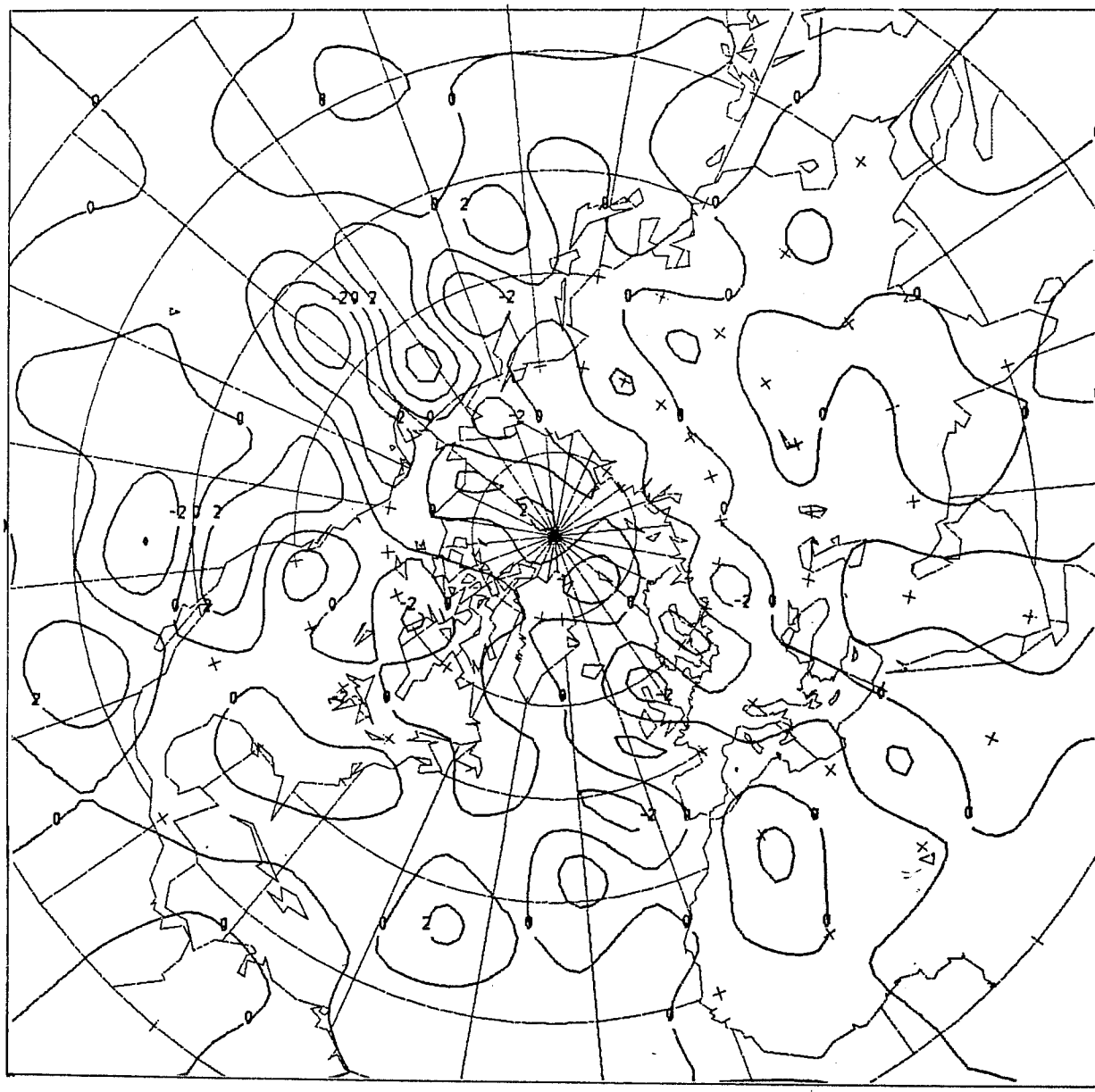

Figure 6. Same as fig.5, with a smoothing term in the functional of the variational analysis.

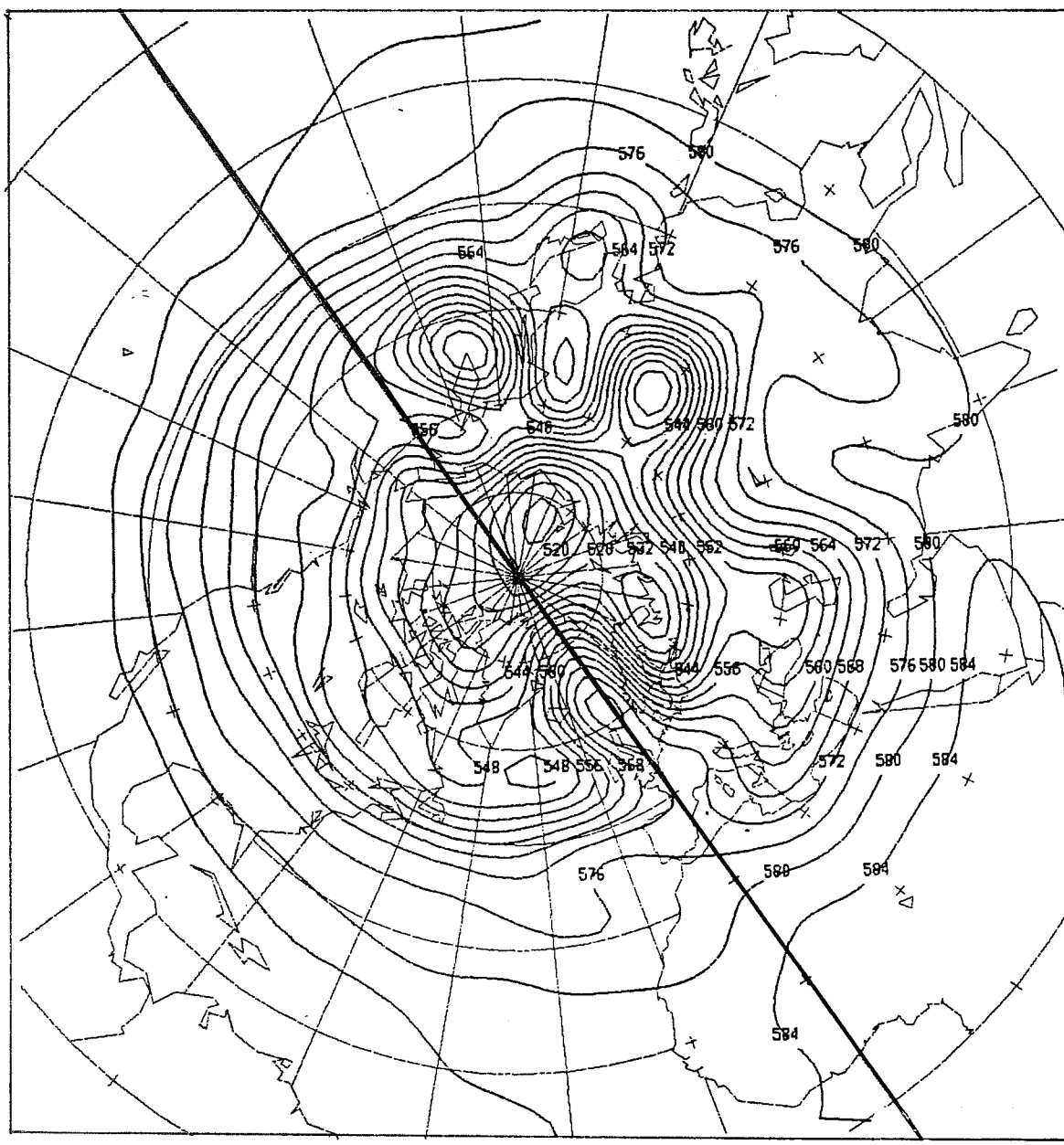

Figure 7. Geopotential field produced by a variational analysis without observations on the western hemisphere.

One possibility for avoiding the occurence of this noise could be to interrupt the minimization process before the noise appears. Another possibility is to add to the distance function a term measuring the amount of small scale noise in the field (see Wahba 1982). The minimization will then tend to produce fields in which the amount of small scale noise is reduced. Figure 6, similar to figure 5, is relative to an assimilation produced with a distance function containing the following term at the initial time

$$\mathcal{L} = A \sum_{n,m} a_n^2 \frac{1}{n(n+1)} |\zeta_n^m|^2 \qquad (3.4)$$

where A is equal to 5.10^{-4} s and a_n^2 is equal to n. It is seen that the noise, although still present, has been reduced. The corresponding geopotential field (not shown) has not been modified over data covered areas by the addition of the smoothing term. There is no particular reason to believe that the specific expression (3.4) provides the best way for avoiding unrealistic noise in the assimilation but, it is clear that the presence of a smoothing term can control the amount of noise.

One process through which adjustment of the fields to the observations is obtained, is advection by the flow. Since one model solution is adjusted globally to the data available over a period of time, there is not only downstream advection into the future, as in ordinary assimilation procedures but also upstream advection into the past. This is clearly visible on figure 7 which shows the height field produced at the initial time by an assimilation in which only observations <u>east</u> of 0 longitude (between 0 and 180 E) were used. The height field is reconstructed satisfactorily upstream of the observed area up to longitude 15 W. A similar situation (not shown) is observed at the final time. It is then the flow east of 180 W which is reconstructed by downstream advection.

4. The shallow water equations.

4.1 The continuous case.

The shallow water equations at the surface of a rotating sphere read

$$\left|\begin{array}{l} \dfrac{\partial V}{\partial t} = (\zeta+f)k\times V - \nabla(K+\phi) \\ \dfrac{\partial \phi}{\partial t} = -\nabla.\phi V \end{array}\right. \qquad (4.1a)$$

and they can be written in terms of only vorticity ζ, divergence η and geopotential ϕ

$$\left|\begin{array}{l} \dfrac{\partial \zeta}{\partial t} = J(\zeta+f,\Delta^{-1}\zeta) - \nabla.\left((\zeta+f)\nabla\Delta^{-1}\eta\right) \\ \dfrac{\partial \eta}{\partial t} = J(\zeta+f,\Delta^{-1}\eta) + \nabla.\left((\zeta+f)\nabla\Delta^{-1}\zeta\right) - \Delta\phi - \Delta K \\ \dfrac{\partial \phi}{\partial t} = J(\phi,\Delta^{-1}\zeta) - \nabla.(\phi\nabla\Delta^{-1}\eta) \\ \text{with, } K = \dfrac{1}{2}(\nabla\Delta^{-1}\zeta.\nabla\Delta^{-1}\zeta + \nabla\Delta^{-1}\eta.\nabla\Delta^{-1}\eta + 2J(\Delta^{-1}\zeta,\Delta^{-1}\eta)) \end{array}\right. \qquad (4.1b)$$

In these two sets of equations, the notations are the same as in section 2, with K denoting in addition the kinetic energy of the wind vector V.
The quadratic expression for energy

$$E = \int_{\Sigma} \left(\frac{1}{2}\phi^2 + \phi^{\circ}K\right) d\Sigma \qquad (4.2a)$$

is an invariant of the equations (3.1) linearized about a state of rest with mean geopotential ϕ°. Using the fact that the mean of a jacobian is 0 over the sphere, the expression (4.2a) becomes

$$E = \int_{\Sigma} \frac{1}{2}\phi^2 + \frac{1}{2}\phi^{\circ}(\nabla\Delta^{-1}\zeta\cdot\nabla\Delta^{-1}\zeta + \nabla\Delta^{-1}\eta\cdot\nabla\Delta^{-1}\eta) \qquad (4.2b)$$

For any two triplets (ζ_1,η_1,ϕ_1) and (ζ_2,η_2,ϕ_2) we define the corresponding scalar product

$$\langle \zeta_1,\eta_1,\phi_1 \mid \zeta_2,\eta_2,\phi_2 \rangle = \int_{\Sigma} \frac{\phi_1\phi_2}{2} + \phi^{\circ}(\nabla\Delta^{-1}\zeta_1\cdot\nabla\Delta^{-1}\zeta_2 + \nabla\Delta^{-1}\eta_1\cdot\nabla\Delta^{-1}\eta_2)$$

In addition to formula (2.6), we have for any three scalar fields a,b,c

$$\int_{\Sigma} \nabla\cdot(a\nabla b)c \, d\Sigma = -\int_{\Sigma} a\nabla b\cdot\nabla c \, d\Sigma \tag{4.4}$$

For a given solution (ζ,η,ϕ) of (4.1b), the tangent linear set of equations reads

$$\begin{aligned}
\frac{\partial\delta\zeta}{\partial t} &= J(\delta\zeta,\Delta^{-1}\zeta) + J(\zeta+f,\Delta^{-1}\delta\zeta) - \nabla.(\delta\zeta\nabla\Delta^{-1}\eta) \\
&\quad - \nabla.((\zeta+f)\nabla\Delta^{-1}\delta\eta) \\
\frac{\partial\delta\eta}{\partial t} &= J(\delta\zeta,\Delta^{-1}\eta) + J(\zeta+f,\Delta^{-1}\delta\eta) + \nabla.(\delta\zeta\nabla\Delta^{-1}\zeta) \\
&\quad + \nabla.((\zeta+f)\nabla\Delta^{-1}\delta\zeta) - \Delta\delta\phi \\
&\quad - \Delta(\nabla\Delta^{-1}\zeta.\nabla\Delta^{-1}\delta\zeta + \nabla\Delta^{-1}\eta.\nabla\Delta^{-1}\delta\eta) \\
&\quad - \Delta(J(\Delta^{-1}\delta\zeta,\Delta^{-1}\eta) + J(\Delta^{-1}\zeta,\Delta^{-1}\delta\zeta)) \\
\frac{\partial\delta\phi}{\partial t} &= J(\delta\phi,\Delta^{-1}\zeta) + J(\phi,\Delta^{-1}\delta\zeta) - \nabla.(\phi\nabla\Delta^{-1}\delta\eta) \\
&\quad - \nabla.(\delta\phi\nabla\Delta^{-1}\eta)
\end{aligned} \tag{4.5}$$

And, by using formulas (2.5a), (2.5b), (2.6) and (4.4), we find the adjoint of system (4.5)

$$\begin{aligned}
-\frac{\partial\delta\overset{*}{\zeta}}{\partial t} &= J(\Delta^{-1}\delta\overset{*}{\zeta},\zeta+f) + \Delta J(\Delta^{-1}\zeta,\Delta^{-1}\delta\overset{*}{\zeta}) + \Delta(\nabla\Delta^{-1}\eta.\nabla\Delta^{-1}\delta\overset{*}{\zeta}) \\
&\quad + \Delta J(\Delta^{-1}\eta,\Delta^{-1}\delta\overset{*}{\eta}) + \nabla.((\zeta+f)\nabla\Delta^{-1}\delta\overset{*}{\eta}) \\
&\quad - \Delta(\nabla\Delta^{-1}\zeta.\nabla\Delta^{-1}\delta\overset{*}{\eta}) + \nabla.(\delta\overset{*}{\eta}\nabla\Delta^{-1}\zeta) \\
&\quad + J(\delta\overset{*}{\eta},\Delta^{-1}\zeta) - \frac{1}{\phi}\circ J(\delta\overset{*}{\phi},\phi)
\end{aligned}$$

$$
\left|
\begin{aligned}
-\frac{\partial \delta\eta^*}{\partial t} &= J(\Delta^{-1}\delta\eta^*, \zeta+f) + \nabla.(\delta\eta^* \nabla\Delta^{-1}\eta) + J(\Delta^{-1}\zeta, \delta\eta^*) \\
&\quad - \nabla\cdot((\zeta+f)\nabla\Delta^{-1}\delta\zeta^*) + \frac{1}{\phi}\circ\nabla.(\phi\nabla\delta\phi^*) \\
-\frac{\partial \delta\phi^*}{\partial t} &= \phi\circ\delta\eta^* + J(\Delta^{-1}\zeta, \delta\phi^*) + \nabla\Delta^{-1}\eta.\nabla\delta\phi^*
\end{aligned}
\right. \qquad (4.6)
$$

4.2 The discretized case.

Numerical experiments have been performed with a pseudo-spectral model of equations (4.1b) with triangular truncation at order N = 21, this coresponds to 735 independent parameters. The collocation grid had enough resolution to prevent aliasing errors and, according to the result proved in Annex 1, it is sufficient for obtaining the adjoint of the discretized model to introduce in (4.6) the same discretized operators as in (4.1b).

The time integration was performed with a semi implicit leapfrog scheme and the corresponding adjoint was used for equation (4.6).

4.3 Numerical results.

In a first step, numerical experiments similar to those which have already been described in section 3.1 were performed on a Haurwitz wave for checkingthe numerical convergence of the descent process.

In a second step, assimilation experiments were performed with wind and geopotential radiosondes observations at 500mb distributed over a 24 hour period starting at 00Z 18 March 1985. When using the raw equations (4.1) and no smoothing term in the distance function, the assimilation produced an unacceptably large amount of gravity waves as can be seen in figure (8). This is not surprising since the model uses all its degrees of freedom, including gravity waves, for minimizing the distance function.

In order to circumvent the difficulty, non-linear normal mode initialization process was introduced at the beginning of the assimilation period. The adjoint of initialization process was also developed which allowed the minimization to be performed on the Rossby waves components only

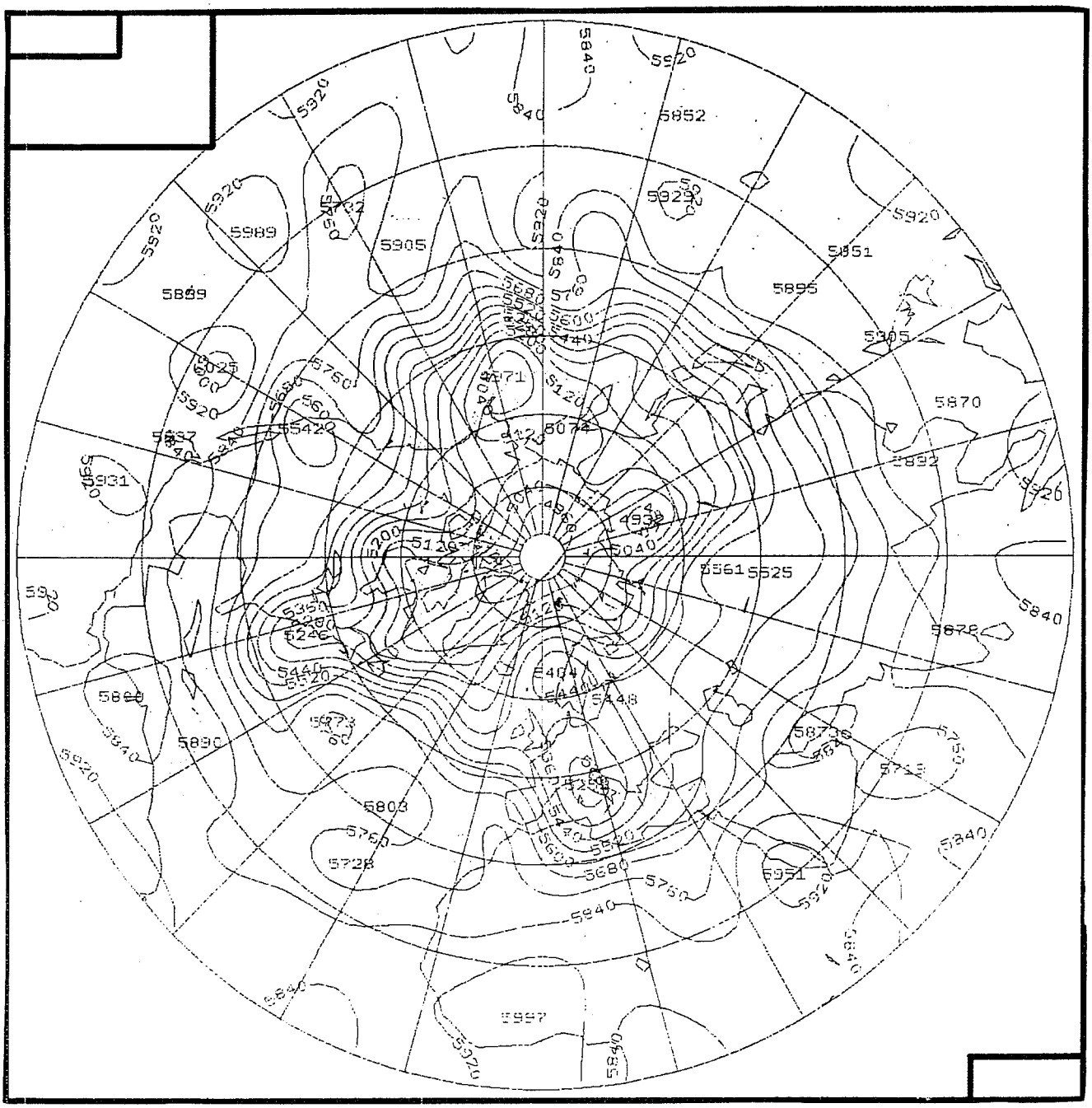

Figure 8. 500mb geopotential field of variational analysis performed with the shallow water equations. 18 March 1985 24.00Z

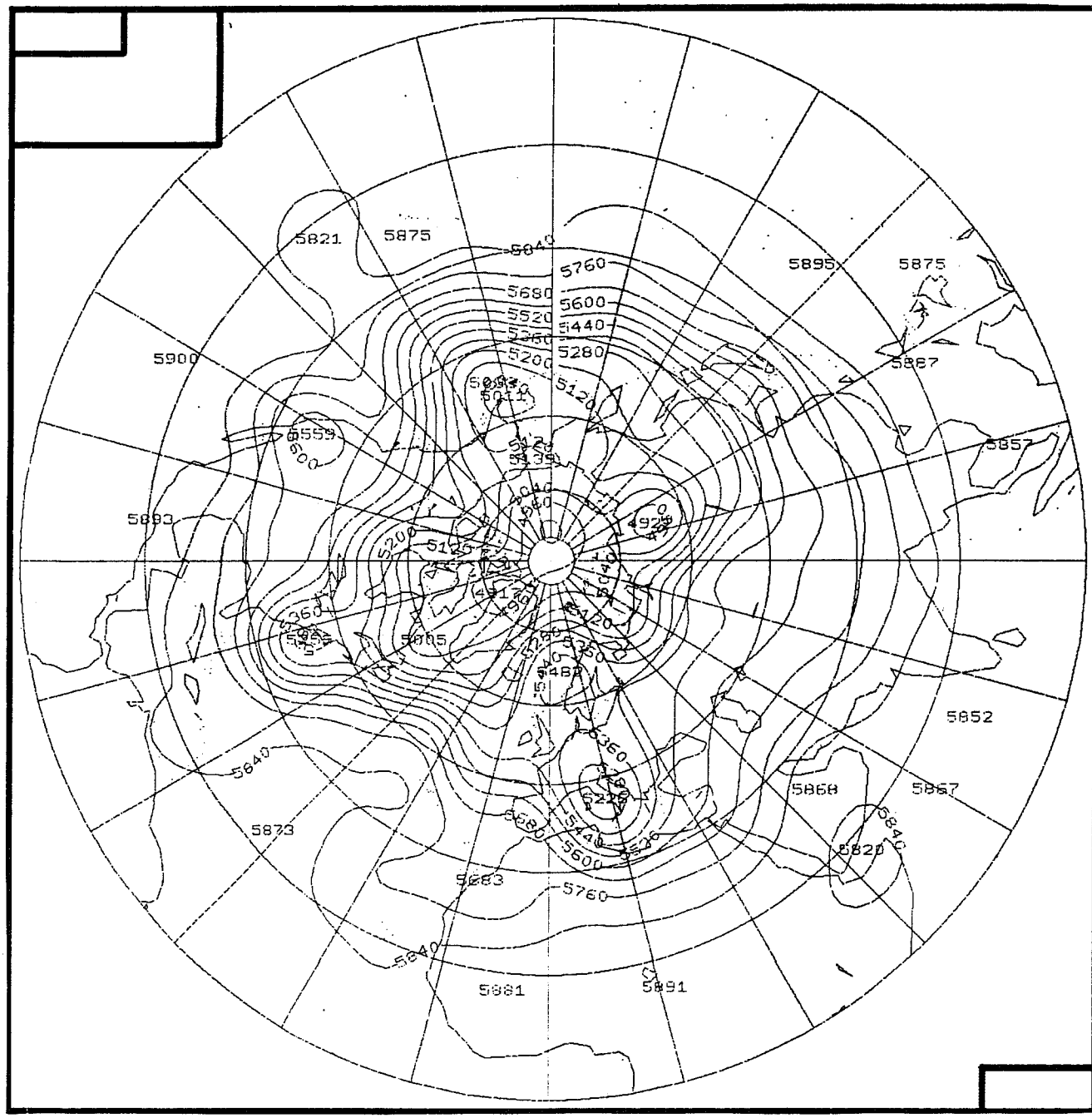

Figure 9. Same as figure (8) with the non-linear normal mode initialisation process introduced at the beginning of the assimilation process, 24 hours before.

and to avoid an unrealistic amount of gravity waves. It is important to stress that implementing the minimization on the Rossby waves and performing the non-linear normal mode initialization before the forecast, ensures that the final minimizing solution is property balanced, but this, of course, is achieved at the price of developing the adjoint of the initialisation process. Figure (9) shows the result of an assimilation performed with the Machenauer initialisation process; the amount of gravity waves has been greatly reduced and is five order of magnitude less than in the emeraude analysis. Another efficient way for reducing that noise is to add in the functional a term which measure the amount of gravity waves; we have tested the squared norm of the tendancies of the gravity waves and it contols satisfactorily the amount of gravity waves.

Conclusion.

The results presented in the article show that adjoint techniques can effectively be used for minimizing a scalar function measuring the distance between a model solution and a given set of observations. The fields obtained are meteorologically realistic.

Two major theorical advantages of adjoint techniques when applied to the problem of data assimilation are the following. First, they produce a sequence of analyzed fields which are exactly consistent with the dynamics of the model. Secondly, they are very versatile in the sense that there is no a priori limitation on the form of the function to be minimized (such as, for instance, being quadratic in the model's variable).

Much work remains of course to be done before the practical capabilities of adjoint techniques can be precisely assessed. It is clear however that their main disadvantages is their computational cost, at least in the context of operational use. However, in view of the continuous progress of computing power, we do not think that this disadvantage should remain in the long term.

Acknowledgments.

The author thanks Lemarechal who provided the minimization codes used in for the experiments described in the article, Talagrand who provided guidance for all this work. All the computation have been performed on the Cray 1.S. of the Centre de Calcul Vectoriel pour la Recherche.

References

Jarraud, M., and A.J. Simmons, 1984: The spectral technique. ECMWF Seminar on Numerical Methods for Weather Prediction, 5-9 September 1983, 1-59.

Talagrand, O., 1985: The adjoint model technique and meteorological applications ECMWF Workshop on High Resolution Analysis, 24-26 June 1985 (this volume).

Whaba, G., 1981: Spline interpolation and smoothing on a sphere. Siam.J.Sci.Stat.Comp., 2, No.1.

Annex 1. The adjoint of a discretized operator free of aliasing errors.

Let $\mathfrak{E}$ be the non discretized space of functions generated by the infinite set of spectral functions

$$Y_n \quad (n = 0, 1, \ldots)$$

$$\Psi = \sum_{n=0}^{\infty} \Psi_n Y_n$$

with scalar product $< \Psi \mid \chi > = \sum_{n=0}^{\infty} \Psi_n \bar{\chi}_n$ where the bar denotes the imaginary conjugation.

For given finite N, we denote by $\mathfrak{E}_N$ the space of functions whose spectral development is truncated at order N. We define on $\mathfrak{E}_N$ the scalar product $< \xi \mid \eta >_N = \sum_{n=0}^{N} \xi_n \bar{\eta}_n$

Let L be a continuous linear operator defined on $\mathfrak{E}$ and L_N a discretized analog of L on $\mathfrak{E}_N$ free of aliasing errors. This means that, for any ξ belonging to $\mathfrak{E}_N$, $L_N\xi$ is equal to the result obtained by performing successively the following operations :

i) extend the development of ξ by defining
$\xi_n = 0$ for $n > N$

ii) apply the non discretized operator L on the result of i

iii) truncate the result of ii) at order N

We will denote by Q_N operator i) which is an operator of $\mathfrak{E}_N$ into $\mathfrak{E}$ and by P_N operator iii which is an operator of $\mathfrak{E}$ into $\mathfrak{E}_N$. The condition that L_N is free of aliasing error therefore reads

$$L_N = P_N L Q_N \qquad (A1.1)$$

or by taking adjoints

$$L_N^* = Q_N^* L^* P_N^* \qquad (A1.2)$$

Now, for any ξ belonging to $\mathfrak{E}_N$ and any Ψ belonging to $\mathfrak{E}$

$<Q_N\xi \mid \Psi> = \sum_{n=0}^{\infty}(Q_N\xi)_n\bar{\Psi}_n = \sum_{n=0}^{N}\xi_n\bar{\Psi}_n = <\xi \mid P_N\Psi>_N$ which shows that

P_N and Q_N are adjoint of each other. (A1.2) then becomes

$$L_N^* = P_N L^* Q_N$$

which presents an obvious similarity with (A1.1) and shows that any discretisation of L^* which is free of aliasing errors is equal to L_N^*

This result can be extremely useful in practice for finding the adjoint of a given discretized operator.

Annex 2. Diagnostic of wind and height from vorticity field.

In order to compare the vorticity field with the radiosondes observations, it is necessary, starting from the vorticity field, to evaluate the two components of wind and the height at observation locations.

Diagnostic of wind.

The values $\hat{u}_i$ and $\hat{v}_i$ of (3.2) are determined as follow from the spectral components of the model vorticity field ζ . The two spatial components of the wind field

$$V = k \times \nabla\Delta^{-1}\zeta \qquad \text{(A2.1)}$$

are obtained at the points of the collocation Gaussian grid. These values are then interpolated bilinearly with respect to latitude and longitude in order to produce the components $\hat{u}_i$ and $\hat{v}_i$ at the observation location.

Diagnostic of height.

The local values $\hat{\phi}_i$ of (3.3) are obtained from the model vorticity ζ by solving in spectral components the nonlinear balance equation for the geopotential

$$\phi = \Delta^{-1}(\nabla.((\zeta+f)\nabla\Delta^{-1}\zeta)) - \nabla\Delta^{-1}\zeta.\nabla\Delta^{-1}\zeta/2 \qquad \text{(A2.2)}$$

The geopotential field is then computed on the collocation Gaussian grid and interpolated bilinearly, as for $\hat{u}_i$ and $\hat{v}_i$, at the observation location.

Annex 3. Adjoint of the diagnostic of wind and height.

The gradients of $\mathcal{J}_u$, $\mathcal{J}_v$ and $\mathcal{J}_\phi$ with respect to the Gaussian grid value of u, v and ϕ are analytically simple. So following the general principle described in section 2 of T85, the problem is to determine the adjoint of the operation which, starting from the spectral components of the voticity field, leads to the values of the fields u, v and ϕ at the points of the Gaussian grid.

Wind case.

The linear operator D_v $\zeta \rightarrow k\times\nabla\Delta^{-1}\zeta$ defines the continuous wind field from the vorticity field. The following equalities

$$\int_\Sigma k\times\nabla\Delta^{-1}\zeta\cdot V'd\Sigma = -\int_\Sigma \nabla\Delta^{-1}\zeta\cdot k\times V'd\Sigma = -\int_\Sigma \Delta^{-1}\zeta\ \nabla\cdot(k\times V')d\Sigma \quad (A3.1)$$

where ζ is a vorticity field and V' a wind field, shows that the adjoint of the operator D_v is the operator D_v^* $V' \rightarrow \nabla.(k\times V')$. We have chosen kinetic energy (L^2 norm) on to derive a scalar product on the space of all the possible wind field.

The operators $u \rightarrow u_g$ and $v \rightarrow v_g$ where u_g and v_g are the Gaussian grid fields are the product of the Legendre transform and the Fourier transform. Parseval's eqality means that Fourier transform is unitary, and the following equality

$$\sum_{k=1}^{N_k}(\Sigma\alpha_n^m P_n^m(\mu_k))\beta_g(\mu_k) = \sum_T \alpha_n^m \sum_{k=1}^{N_k} P_n^m(\mu_k)\beta_g(\mu_k) \quad (A3.2)$$

where P_n^m are the Legendre's polynomials, μ_k the Gaussian latitudes of the collocation grid, α_n^m the spectral components of a given scalar function and β_g a scalar field on the Gaussian grid, gives us the adjoint of Legendre Transform.

Geopotential case.

Equation (A2.2) is not linear, so we consider the tangent balance equation, which for a given field ζ reads

$$\zeta' \rightarrow \phi = \Delta^{-1}(\nabla\cdot((\zeta+f)\nabla\Delta^{-1}\zeta')+\nabla\cdot(\zeta'\nabla\Delta^{-1}\zeta)\ ) - \nabla\Delta^{-1}\zeta'\cdot\nabla\Delta^{-1}\zeta \quad (A3.3)$$

On the space of all possible height fields, we consider the potential energy which leads to be the L^2 norm

$$||\phi||^2 = \frac{1}{2} \int_{\Sigma} \phi^2 \, d\Sigma \qquad \text{(A3.4)}$$

and use of Green's formula leads to the adjoint of equation (A3.3)

$$\phi'^* \rightarrow -\nabla.((\zeta+f)\nabla\Delta^{-1}\phi'^*) + \Delta(\nabla\Delta^{-1}\phi'^*.\nabla\Delta^{-1}\zeta) - \nabla.(\phi'^*\nabla\Delta^{-1}\zeta) \qquad \text{(A3.5)}$$

The adjoint of the product of Legendre transform and Fourier transform is taken now as for u and v because we have also L^2 norm on the space of the continuous geopotential fields.