COMPARISON OF FOUR-DIMENSIONAL VARIATIONAL ASSIMILATION WITH SIMPLIFIED SEQUENTIAL ASSIMILATION

Florence Rabier [1] Philippe Courtier [1] Jean Pailleux [2] Olivier Talagrand [3]
Jean-Noël Thépaut [4] Drasko Vasiljevic [4]

1 Météo-France, CNRM, France, currently at ECMWF.

2 ECMWF, currently at Météo-France, CNRM, France.

3 Laboratoire de Météorologie Dynamique, Paris, France.

4 ECMWF.

Abstract

The main part of this study is devoted to perform a strict comparison between two assimilation algorithms, sequential and four-dimensional variational, on a 24-hour period extracted from a baroclinic instability situation representative of mid-latitude dynamics. In the case of linear dynamics and under the hypothesis of a perfect model, these two four-dimensional algorithms are known to lead to the same optimal estimate of the atmosphere at the end of the assimilation period, and both methods can be generalised in the nonlinear case. The full sequential algorithm being too resource-demanding to be implemented as such, we will test in the sequel the four-dimensional variational method (4D-VAR) and a simplified sequential method based on three-dimensional variational analysis (3D-VAR). We deliberately do not exceed the range of validity of the tangent-linear model in the experiments. 4D-VAR is then expected to be almost equivalent to the generalisation of the sequential Kalman filter in the nonlinear case, i.e. the Extended Kalman Filter. As for the simplified sequential algorithm, it can be seen as an approximation of this full Extended Kalman Filter, for which the forecast error matrices are crudely evaluated before each analysis instead of being explicitly computed from the complete dynamical equations. In the 4D variational scheme, the consistency of the propagation of information with the dynamics is illustrated in an experiment assimilating some localized AIREP data. The large impact of these additional observations over a large geographical area appears to be very beneficial for the quality of the analysis. Comparing the results of both methods in various configurations, 4D-VAR is systematically found to behave substantially better than the simplified sequential algorithm, with a more accurate analysis at the end of the assimilation period and a much smaller error growth rate in subsequent forecasts. On the one hand, extremely bad specifications of initial forecast errors are found to be detrimental to both algorithms. On the other hand, the 4D variational algorithm proves to be more robust to the way gravity waves control is implemented.

In the second part of this study, a similar comparison between 4D-VAR and a simplified operational assimilation cycle is carried out using real data on an extreme event, the October '87 Storm over France and Britain. The model used for both the assimilation and the subsequent forecast is a T63L19 model, adiabatic or with very simplified physics. Although the 36-hour forecasts started from the 4D-VAR analysis and the Optimal interpolation analysis are different, none of them is really successful in predicting the storm. It appears that, for this explosive cyclogenesis, the model used in the experiments is too crude to extract a clear signal from the results obtained with different assimilation methods.

1 Introduction.

Numerical weather prediction is an initial-value problem, and high-quality analyses of the atmosphere are necessary for the production of reliable forecasts. The basis of four-dimensional assimilation is to use all the available information in order to produce the best possible initial state for the prediction. The available information consists of the observations, distributed in time, and of the physical laws which govern the evolution of the flow. The latter are in practice available under the form of the numerical prediction model to be used for the forecast. Detailed reviews of data assimilation are given in Ghil and Malanotte-Rizzoli (1991) and Daley (1991) and we will just here briefly introduce the two basic algorithms used to solve the problem, which are shown to lead to equivalent results under certain assumptions (Lorenc, 1986). The more common approach is sequential and comes from estimation theory (Ghil et al., 1981) where one seeks a minimum error variance estimate. Optimally, in the case of linear dynamics, it takes the form of the resource-demanding Kalman filter, and is naturally extended to the weakly nonlinear case under the form of the Extended Kalman Filter (EKF). In most operational centres, a degraded form of this algorithm is implemented, consisting of a cycle of analyses and 6-hour forecasts. The main approximation is that the forecast error covariance matrices are prescribed instead of being explicitly computed from the relevant dynamical equations. An alternative to sequential estimation is the variational approach, coming from control theory (for the underlying theory see Lions, 1971 and for a general description of its application to meteorology, see Le Dimet and Talagrand, 1986). The goal can be described as optimising an objective criterion which quantifies the distance between model states and the available information (data and prior knowledge). Assuming the model to be perfect, one seeks a model trajectory over the assimilation period and the optimisation problem depends solely upon the initial state at the beginning of the period. Using an adjoint model, the minimizing solution can be found at a cost which is not prohibitive, although still high for operational implementation.

Current operational systems are performing globally well (see Hollingsworth et al., 1986) but are nevertheless sub-optimal (Daley, 1991; Daley, 1992). They do not extract properly all the information contained in the available data. One of the most striking examples of their weaknesses is the prediction of the 15-16 October 1987 Storm over England and France, for which various authors have attributed the failure of most models to properly forecast the intensity of the storm in the short range to data assimilation problems (see Lorenc et al., 1988 and Jarraud et al., 1989). Indeed, cyclogenesis conditions are such that the approximation of separability between horizontal and vertical correlations of bakground errors commonly used in the analysis step is expected to break down. Therefore, this kind of meteorological situation seems appropriate to perform a comparison of new methods such as

4D-VAR with a more classical assimilation cycle. A preliminary study aiming to understand the behaviour of 4D-VAR in the case of an idealised baroclinic instability situation was carried out in Rabier and Courtier (1992), hereafter denoted RC92. The major result was that, in most cases, 4D-VAR is able to reconstruct the entire atmospheric state from a time series of observations of only part of the spectrum, because of the explicit use of the dynamics in the assimilation process. In particular, when initialised by an atmospheric state far away from the desired result, the method leads to impressive results, taking advantage of the dynamical coupling between various components of the flow. Such a benefit of the consistent use of the dynamics had also been found in the case of a hydraulic jump by Lorenc (1988a). When starting the algorithm from a reasonable prior estimate of the atmospheric flow, which will be the case in the following, 4D-VAR is also expected to improve upon classical cycles. In particular, it has been shown by Dee (1991) for a 1-D linear model that even a simplified Kalman Filter leads to substantial improvements over optimal interpolation. Recently, Gauthier et al. (1992) have illustrated the benefit of using the full EKF over simplified methods.

The aim of this paper is first to estimate the benefit of using an elaborate assimilation scheme in a realistic primitive-equation model, and a representative dynamical problem. The same baroclinic instability situation as in RC92 was chosen to compare 4D-VAR with a sequential assimilation cycle relying on a three-dimensional variational analysis (3D-VAR). The 3D-VAR scheme is the one being developed at the European Centre for Medium-range Weather Forecasts (Pailleux et al., 1991), but other centres such as the National Meteorological Center are already using a three-dimensional variational approach as an alternative to optimal interpolation (Parrish and Derber, 1992). The 4D-VAR scheme is also the one included in the ARPEGE/IFS model, developed as a cooperative project between Météo-France and ECMWF (see Courtier et al., 1991 for a comprehensive description). It has recently been compared with a sequential estimation in the case of real data, although without using a background term, by Thépaut et al.(1992). The methodology followed here is that of identical twin experiment where synthetic observations are extracted from a reference run of the assimilating model thus implicitly assuming a perfect model.

In a second time, we shall compare four-dimensional variational assimilation with a simplified operational sequential assimilation for the prediction of the October 87' Storm.

Firstly, in section 2 we state the basic equivalence between 4D variational and sequential methods when they are used optimally, and describe the particular set-up of the experiments and in particular the approximations made on the forecast error covariance matrices in the case of sequential assimilation. The most significant results on the academic baroclinic instability situation are presented and commented in section 3 in the case of perfect data, in section 4 in the case of noisy data and in section

5 using a later version of the variational schemes currently still under development. Then, real data experiments on the October' 87 Storm are presented in section 6. Finally, a discussion concludes the paper in section 7.

2 Framework of the experiments

2.1 Theoretical equivalence between the two algorithms.

All assimilation algorithms existing at present can be described as a more or less approximate form of statistical linear estimation, which can be summed up as follows.

We assume that observations are available, making up a vector y, and that we want to estimate quantities making up a vector x. We also assume that y can be related to x by the relation

$$y = Hx + \epsilon \tag{1}$$

where H is a known matrix of coefficients defining the components (or combination of components) of x which have been observed, and ϵ is an observational error. This error is of course unknown, but it is supposed to be unbiased, and to have a known covariance matrix Σ. Among all unbiased linear estimates of x, the one which minimizes the error variance is

$$\hat{x} = (H^T\Sigma^{-1}H)^{-1}H^T\Sigma^{-1}y \tag{2}$$

and the corresponding estimation error

$$\hat{x} - x = (H^T\Sigma^{-1}H)^{-1}H^T\Sigma^{-1}\epsilon \tag{3}$$

is proportional to the observational noise. It possesses a variance-covariance matrix equal to

$$(H^T\Sigma^{-1}H)^{-1} \tag{4}$$

The estimate $\hat{x}$ is called the Best Linear Unbiased Estimator (BLUE) of x from y. It can also be found by minimizing a distance function

$$J(\xi) = \frac{1}{2}(y - H\xi)^T\Sigma^{-1}(y - H\xi) \tag{5}$$

with respect to ξ (Jazwinski, 1970). In the particular case when the error ϵ is Gaussian, the distance function is the argument of the exponential in the conditional probability function $p(x/y)$, and the

value of ξ minimimizing J is the least-square estimate of x, either linear or nonlinear. It is also the maximum likelihood estimate of x (Lorenc, 1988a).

In the case of time evolving systems when one wants to reconstruct the state of the system at, say, the final time of an observing period, the BLUE $\hat{x}$ can be determined either by explicit minimization of a functional J similar to Eq. 5, which measures the misfit between a model trajectory and the available data, or by implementation of sequential estimation over the observing period. The first method is called variational assimilation (Le Dimet and Talagrand, 1986). The second leads to Kalman filtering, which is a classical tool in estimation theory (Ghil et al., 1981).

In the case of an exact linear model, both algorithms lead to exactly the same result at the end of the assimilation period. If the model is linear, but not exact, there is still identity between the results produced by both algorithms, provided the model equations are introduced in the variational assimilation as weak constraints, i.e. as additional terms in the distance function to be minimized. However, in that case, 4D-VAR becomes as expensive as the full Kalman filter, since the control variable now contains all the successive atmospheric states over the assimilation period.

Both methods can be implemented in nonlinear cases. Sequential estimation requires the specification of an appropriate law for the time evolution of the estimation error. This can be done by using the tangent-linear equation which leads to Extended Kalman Filtering.

Using these methods remains justified from a theoretical point of view, at least as long as the local tangent-linear hypothesis is valid.

Let us denote $x = (x_0, x_1, ..., x_n)$ the four-dimensional vector of the time-series of atmospheric states over the period one wishes to analyse $[t_0, t_n]$. It is assumed that the numerical model M which gives the time evolution of the atmospheric state

$$x_0 \longrightarrow x_i = M(t_i, t_0)(x_0) \tag{6}$$

is exact. We further assume that the tangent-linear approximation is valid, i.e. the deviation from the true state remains small enough so that the expansion of M in a Taylor series can be approximated by its first order term only

$$M(t_i, t_0)(x_0 + \delta x_0) = M(t_i, t_0)(x_0) + M'(t_i, t_0).\delta x_0 \tag{7}$$

for $i = 1, n$, where $M'(t_i, t_0)$ is equal to the integration of the tangent-linear model from t_0 to t_i, the linearization being performed in the vicinity of the trajectory starting from x_0.

Let us write the observations taken over the time period $[t_0, t_n]$ as a four-dimensional vector

$y = (y_0, y_1, ..., y_n)$. In this notation, some of the $y_i, (i = 0, n)$ might be empty if no data is available at that time. For the operator H_i, generally nonlinear, giving y_i for a given x_i: $y_i = H_i(x_i)$, we will make the same tangent-linear hypothesis as for the numerical model

$$H_i(x_i + \delta x_i) = H_i(x_i) + H_i'.\delta x_i \tag{8}$$

One further assumption is that we have a prior knowledge about the atmospheric state, given by a background field x_b at time t_0, coming generally from a previous forecast.

Some information about the statistics involved in this estimation problem are also available. It is assumed that the $(y_i)_{i=0,n}$ and x_b are unbiased variables, for which the knowledge of the error variance matrices $(O_i)_{i=0,n}$ and P_b is available. It is also assumed that the observation errors are uncorrelated in time.

Firstly, in the case of variational estimation, one minimizes

$$J(x_0) = \frac{1}{2}\sum_{i=0}^{i=n}(y_i - H_i(M_i(x_0)))^T O_i^{-1}(y_i - H_i(M_i(x_0))) + \frac{1}{2}(x_b - x_0)^T P_b^{-1}(x_b - x_0) \tag{9}$$

In most previous experiments (Rabier and Courtier, 92; Thépaut and Courtier, 91), no account was taken of the background term which is a valid approximation as long as the assimilation period is sufficiently long to contain enough observations to determine x_0 without other prior information (Lorenc, 1986).

Minimization of J requires the knowledge of its gradient given by

$$J'_{x_0} = \sum_{i=0}^{i=n} M'(t_i, t_0)^T H_i'^t O_i^{-1}(y_i - H_i(M_i(x_0))) + P_b^{-1}(x_b - x_0) \tag{10}$$

Assuming that the tangent-linear approximations are valid on $[x_b, x_a]$ where x_a is the desired analysed state, the development of J in a Taylor series is quadratic

$$J(x_a) = J(x_b) + J'_{x_b}.(x_a - x_b) + \frac{1}{2}(x_a - x_b)^T J''_{x_b}.(x_a - x_b) \tag{11}$$

where J'_{x_b} is the gradient (Jacobian matrix) of J evaluated at point x_b and J''_{x_b} is the Hessian matrix evaluated at the same point. If the Hessian matrix was exactly known, we could then get the minimum of J in one single iteration of the Newton method

$$x_a = x_b - J''^{-1}_{x_b}.J'_{x_b} \tag{12}$$

However, in most cases, this Newton step is too expensive, the storage of the Hessian matrix is impossible or nonlinearities are present and several iterations of a Quasi-Newton algorithm are necessary (see Lorenc, 1988b, for a simplified implementation of this algorithm).

As for the sequential estimation, it takes the form of the Extended Kalman Filter (EKF) (for a complete derivation, see Ghil and Malanotte-Rizzoli, 1991). Starting from a background $x_{0,g} = x_b$, with error covariance matrix $P_{0,g} = P_b$, the following steps have to be performed between times t_i and t_{i+1}, for (i=0,n-1):

(i) analysis step to update the background field (current estimate) $x_{i,g}$

$$x_i = x_{i,g} + K_i(y_i - H_i(x_{i,g})) \tag{13}$$

where the so-called gain matrix K_i is given by

$$K_i = P_{i,g}H_i'^T(H_i'P_{i,g}H_i'^T + O_i)^{-1} \tag{14}$$

and the corresponding estimation error covariance matrix P_i

$$P_i = (I - K_iH_i')P_{i,g} \tag{15}$$

(ii) forecast step to carry the analysis vector x_i and its error variance matrix P_i to time t_{i+1}

$$x_{i+1,g} = M(t_{i+1}, t_i)x_i \tag{16}$$

$$P_{i+1,g} = M'(t_{i+1}, t_i)P_i(M'(t_{i+1}, t_i))^T + Q_i \tag{17}$$

where Q_i represents the model-generated error covariance. In the following we will always take $Q_i = 0$ implicitly assuming that the model-generated errors are negligible.
The analysis step (i) can be performed either directly as in classical optimal interpolation, or equivalently by minimizing

$$J(x_i) = \frac{1}{2}(y_i - H_i(x_i))^T O_i^{-1}(y_i - H_i(x_i)) + \frac{1}{2}(x_{i,g} - x_i)^T P_{i,g}^{-1}(x_{i,g} - x_i) \tag{18}$$

which is what is being done in 3D-VAR (see Pailleux at al., 1991). One of the main advantages of three-dimensional methods upon optimal interpolation is that no geographical selection of observations is needed.

What should be kept in mind is that, in the linear context and if the model is perfect, 4D-VAR and the Kalman Filter lead to the same result which is the best estimate of the atmospheric state at the end of the assimilation period. In the nonlinear case and if the model is still perfect, although no general result exists for the strict equivalence, if the tangent-linear hypothesis is valid, results are expected to be very close (by the way, it would be interesting to quantify the difference that exists between them, even if it is small).

The question is now to quantify the impact of the approximations of the EKF which have to be done in the operational practice on the quality of the analysis. One of the prevailing weaknesses of current assimilation systems being the approximate specification of the forecast error covariance matrix ($P_{i,g}$ in our notations), we will simulate a degraded EKF by running an assimilation cycle composed of 3D-VAR analyses for which the forecast error covariance matrices are prescribed instead of being explicitly computed according to Equation 17. Comparing the result of this simplified sequential algorithm and of 4D-VAR will then give an indication on the benefit of using a more elaborate scheme in an identical twin-experiment context.

The main limitation of this approach is that we do not take into account the model-generated error. Since the perfect is perfect in this case, and that the four-dimensional variational scheme assumes that the model is perfect, 4D-VAR is certainly given an advantage. However, it would not have been straightforward to perform a similar study taking into account the model error for various reasons. The first one is that our knowledge of such model deficiencies is rather limited. The second one is that, although the full Kalman Filter contains a model error covariance term (Equation 17), no such obvious way of introducing this information exists in 4D-VAR. Different approaches have been tried by several authors. Following the simple remark that, in the case of assimilation, one wants to make a forecast from the final time of the assimilation period, Courtier and Talagrand (1990) gave a larger weight to more recent observations in the definition of the cost function. Derber (1989) showed how one could use the variational context to exhibit the model error which, under certain assumptions, would lead to the best fit of a model trajectory to the available data over a given time period. More recently, Zupanski (1992) extended this study by looking for both the initial conditions and the model error field which would minimize an appropriate cost function. The inclusion of model-error information in 4D-VAR is thus very much at a research stage and it was not our purpose to investigate this area. In any case, one has to remember that the comparison performed in this study is valid only under the assumption of a perfect model.

2.2 Set-up of the experiment.

The baroclinic instability situation which will be used in order to compare the two assimilation methods is the same as in RC92. It was built in the same way as in Simmons and Hoskins (1978) or in Thorncroft and Hoskins (1990). The basic zonal flow symmetric about the equator is baroclinically unstable and the linearly most unstable mode at zonal wavenumber six is selected. It provides the initial conditions of the simulation of the life cycle of a baroclinic wave. The primitive equation model used in this study is the ARPEGE/IFS model. The vertical discretization consists of the 19 levels of the hybrid pressure / σ vertical coordinate (Simmons and Burridge, 1981) of the 1990 operational ECMWF model, and a T21 triangular truncation of the spherical harmonics representation of the fields is used. No diabatic effects are considered in this study. The results are consistent with those of Simmons and Hoskins (1978): the wave grows for nine days and then decays because of the reduction in baroclinicity of the flow and barotropic effects. The most intense cyclogenesis occurs between days six and eight with a maximum drop in surface pressure of 8 hPa per day. We will concentrate on the characteristics of the flow at this time, with a 24-hour assimilation period over the interval (day 6,day 7), and a subsequent forecast period from day 7 to day 8. The surface pressure fields at days 5, 6, 7 and 8 in the Northern Hemisphere are illustrated in Fig. 1. It was shown in RC92 that the tangent linear approximation is valid in this particular meteorological situation for ranges up to one or two days, even for substantial amplitudes of initial perturbations (20 hPa for the surface pressure at midlatitudes). This confirms results previously obtained by other authors (in particular Lacarra and Talagrand, 1988 or Vukicevic, 1991) and is of significance in our study to guarantee a good approximate equivalence between 4D-VAR and the EKF.

The observational pattern is intended to represent a schematic daily situation, where only conventional observations are considered. Data are gathered at "synoptic" hours (0, 6, 12, 18 and 24 UTC), with a constant distribution. A simple representation of the heterogeneous space coverage (alternation ocean-land) is given by the simulated radiosonde network illustrated in Fig. 2 which consists of 464 stations. Each report consists of wind and geopotential observations at the standard 16 pressure levels ranging from 1000 hPa to 10 hPa. It should be noted that no direct use is made of temperatures, as in ECMWF operational analysis. The error statistics (matrix O_i) are those used operationally at ECMWF, with in particular a vertical correlation for geopotentials (ECMWF, 1992). In terms of number of individual scalar quantities, we have 111360 data over the assimilation period (for the observations alone), compared with an equivalent 28072 spectral components necessary to define an atmospheric state at truncation T21.

The background field used either for the sequential algorithm at the beginning of the assimila-

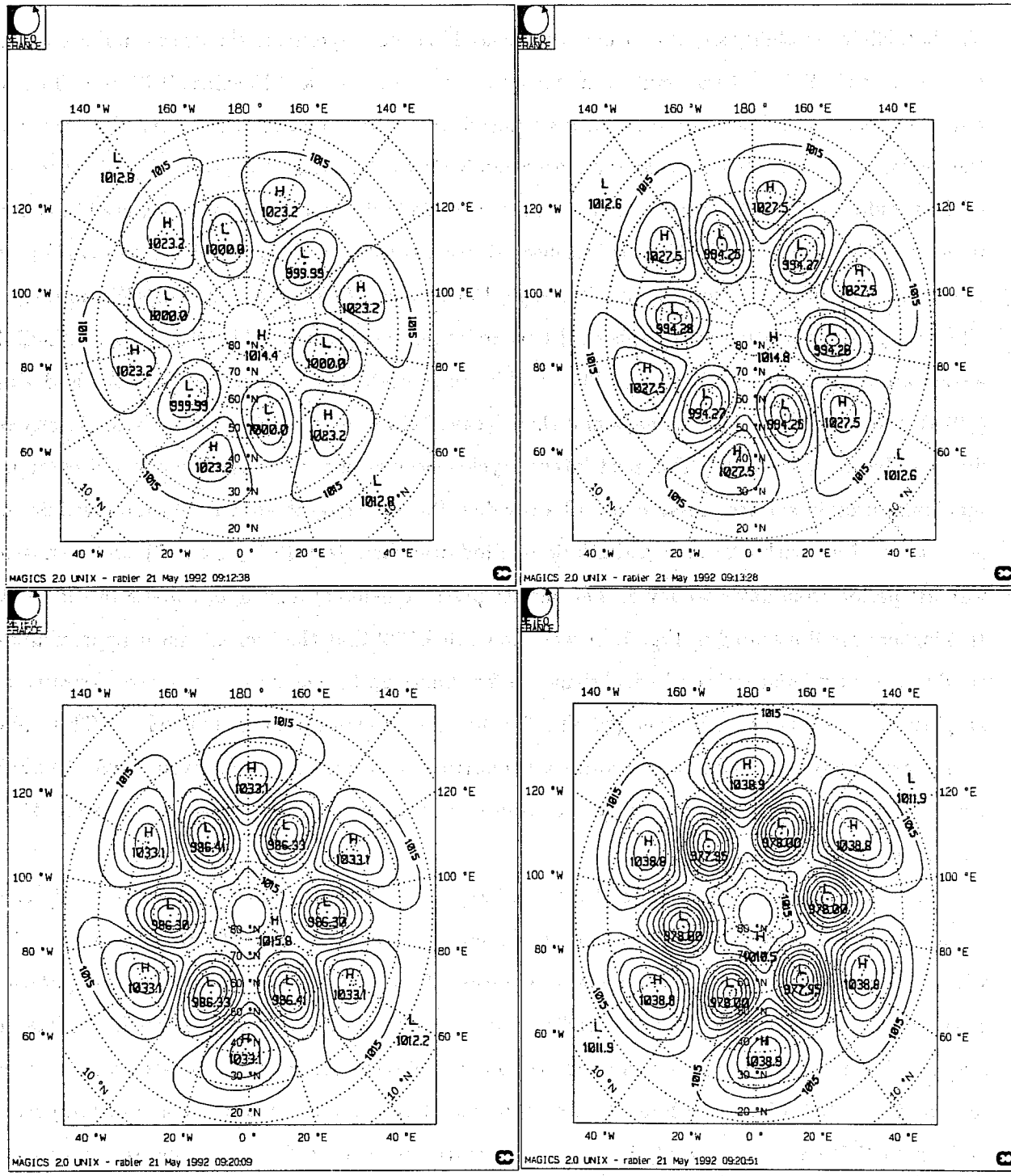


Fig. 1 *Surface pressure fields in the Northern hemisphere at days 5, 6, 7 and 8 of the life cycle of the baroclinic wave, (respectively panels (a), (b), (c) and (d)). Isolines are drawn every 5 hPa.*

tion period or for 4D-VAR is derived from the reference run, but with some time-lag from the true atmospheric state (-24 hours) to simulate a forecast error in which the baroclinic system is slightly out of phase and too weak. This background field is used in two ways: it is included in the definition of the cost function and is also the starting point of the minimization (first-guess). The subsequent background fields used at 6, 12, 18 and 24 hours for the sequential scheme will of course be the result of a 6-hour forecast from the previous analysis. The statistical forecast errors will be approximated by the differences existing between the effective background fields and the truth (the reference run). In particular the error standard-deviations will be taken as the rms of error at model levels. In most of the experiments, for each observed variable (wind components u and v, geopotential and surface pressure), we will compute the rms over latitude rows and we will take 4 different values for each row, following the data coverage which vary every 90 degrees in longitude. Moreover, we will impose a lower bound on the standard deviations equal to one tenth of the maximum value, in order to avoid numerical instability during the spectral transforms. This choice of standard-deviations cannot be considered as optimal, as it has no statistical basis. As for correlations, the choice was to take them as simple as possible. In most of the experiments, we have no vertical correlations, no cross-correlations, and we use the same horizontal correlation for all variables. It is a simple exponential auto-correlation function $\exp(-\frac{1}{2}\frac{x^2}{a^2})$, where the extension a is equal to 600 kms. Being very simple, these approximations are also very crude and much less realistic than what is currently used in operations (Mitchell et al., 1990; Hollingsworth and Lönnberg, 1986; Lönnberg and Hollingsworth, 1986). The absence of vertical correlations may not be of much consequence in the case of complete radiosonde soundings (Mitchell et al., 1990), but the absence of geostrophic coupling may be more detrimental. In Section 5, a geostrophic coupling, vertical correlations, and an updated version of horizontal correlations will be included in the definition of the forecast error covariances, to confirm the results in a framework closer to operational practice. In any case, since a background constraint is included in both 4D variational and sequential schemes at the beginning of the assimilating period, we are sure that both algorithms are penalised by the mis-specifications present in this background term. Furthermore, whatever sophistications are included in correlation definition in realistic systems, those will probably be far from being optimal unless computed by a Kalman filter (see Cohn and Parrish, 1991). One of the major weaknesses of present correlation models are the basic assumptions of separability between the horizontal and the vertical and of time-invariance, and the impact of such weaknesses is precisely what we want to investigate.

As we are looking for a balanced atmospheric state, we will include Nonlinear Normal Mode Initialization during the course of the minimization in most of the experiments. As in Courtier and

Talagrand (1990) or Thépaut and Courtier (1991), one can use a Nonlinear Normal Mode Initialization scheme (NNMI) and its adjoint to perform the minimization mostly in Rossby space. Three iterations of the Machenhauer algorithm are inserted both in the sequential scheme (at each analysis) and in 4D-VAR, and five vertical modes are retained (Machenhauer, 1977). As explained in Pailleux et al. (1991), it amounts to looking for the quasi-balanced state which fits the observations. If one writes $NMI(x)$ the result of the operation consisting in applying Normal Mode Initialization to the vector x, 3D-VAR consists in minimizing the cost-function, at each analysis step

$$J(x_i) = J_G(x_i) + J_O(x_i) \tag{19}$$

where

$$J_G(x_i) = +\frac{1}{2}(x_{i,g} - NMI(x_i))^T P_{i,g}^{-1}(x_{i,g} - NMI(x_i)) \tag{20}$$

$$J_O(x_i) = \frac{1}{2}(y_i - H_i(NMI(x_i)))^T O_i^{-1}(y_i - H_i(NMI(x_i))) \tag{21}$$

which leads to the analysed vector $NMI(x_i)$, and 4D-VAR consists in minimizing the global cost-function

$$J(x_0) = J_G(x_0) + J_O(x_0) \tag{22}$$

where

$$J_G(x_0) = \frac{1}{2}(x_b - NMI(x_0))^T P_b^{-1}(x_b - NMI(x_0)) \tag{23}$$

$$J_O(x_0) = \frac{1}{2}\sum_{i=0}^{i=n}(y_i - H_i(M_i(NMI(x_0))))^T O_i^{-1}(y_i - H_i(M_i(NMI(x_0)))) \tag{24}$$

which leads to the minimizing solution equal to the model integration starting at $NMI(x_0)$. Theoretically, if NNMI can be considered as a strict projection on the slow manifold, the effect of this inclusion of NNMI is to make changes in Rossby waves only in order not to get more gravity waves than those included in the starting point of the minimization. However, as explained in Thépaut and Courtier (1991), NNMI is not strictly a projection onto the slow manifold and it can partly be inverted. But, in any case, in the experiments performed in this study, the analyses were found to possess a small amount of unbalanced gravity waves.

The background fields are always used as the starting points of the minimization: they initialize the minimization process. The minimization algorithm is stopped when either it has performed 50

simulations, i.e. 50 computations of the cost function and of its gradient, or when the norm of the gradient has been divided by a factor 100. It is of a limited-memory Quasi-Newton type and has been provided by Institut National de Recherche en Informatique et Automatique (Gilbert and Lemaréchal, 1989).

In terms of computer cost, at truncation T21 with a 2400s timestep, a 4D-VAR over 24 hours extends over 39 timesteps (3 iteration of NMI, 36 forecast steps), and each 3D-VAR is equivalent to 3 timesteps corresponding to the NNMI stage. On the whole, the sequential cycle, composed of 5 analyses is 2.6 times less expensive than 4D-VAR. It should be noted that this is a particular context. In a more refined three-dimensional analysis, we would probably not need to use NNMI which would then reduce drastically the cost of the sequential scheme, consequently increasing the cost-ratio between the two algorithms.

3 Assimilation with perfect data.

The investigation of the behaviour of both algorithms has first been carried out in the context of perfect data and we present in the following the most significant results. The first question we are asking is: which algorithm is the more efficient to correct for a previous exceptional forecast error by inserting new observations?

3.1 Assimilation experiment, starting from an unexpectedly high forecast error.

In the case of rapidly evolving situations, it is known that current operational systems sometimes have problems correcting a wrong forecast with the available data to produce a coherent analysis. The impact of using a more elaborate assimilation scheme such as 4D-VAR is then likely to be significant in this kind of situations.

To initiate the 24-hour assimilation, we take a relatively poor background field and we introduce statistics of errors representative of a higher quality backgound field. The background is chosen as day 5 of the life cycle (Fig. 1,panel (a)), which is the atmospheric situation one day before the actual time (see the surface pressure field for day 6, 0 UTC in Fig. 1,panel (b)). The rms errors over a latitude row, at the latitude where they are maximum, are 3.84 ms^{-1} for the zonal wind, 8.20 ms^{-1} for the meridional wind, 4.31 K for the temperature and 6.8 hPa for the surface pressure. Instead of those, the standard-deviations inserted in the definition of the cost-function are taken as the rms of errors relative to the difference (day5, 12 UTC; day6, 0 UTC). The maximum values over a latitude row (to be compared with the values given above) are about twice as small: 2.09 ms^{-1} for the zonal wind, 4.63 ms^{-1} for the meridional wind, 2.21 K for the temperature and 3.7 hPa for the surface pressure.

error rms	at the end assim. period		after 24-hour fcst	
	surf. pres.	500 hPa height	surf. pres.	500 hPa height
Exp. 1	1.89 hPa	15.1 m	2.63 hPa	20.6 m
Exp. 2	1.48 hPa	11.6 m	1.66 hPa	12.1 m
Exp. 2+AIREP	1.28 hPa	11.3 m	1.53 hPa	11.4 m
Exp. 3	1.33 hPa	12.3 m	1.71 hPa	14.4 m
Exp. 4	0.92 hPa	7.2 m	1.09 hPa	7.8 m
Exp. 5	1.61 hPa	14.2 m	2.07 hPa	16.9 m
Exp. 6	1.04 hPa	7.9 m	1.23 hPa	8.3 m
Exp. 7	2.21 hPa	19.0 m	5.95 hPa	51.5 m
Exp. 8	1.14 hPa	8.9 m	1.28 hPa	9.0 m
Exp. 9	2.37 hPa	21.4 m		
Exp. 10	2.08 hPa	17.4 m		

Table 1 : *Global rms of errors for surface pressure (in hPa) and height at 500 hPa (in m), at the end of the assimilation period and after a 24-hour forecast for various experiments.*

- *Exp. 1: Sequential algorithm. Forecast error matrices are kept constant and initial error variances are underestimated. Experiment is performed without any gravity-wave control.*
- *Exp. 2: Variational algorithm. Background error variances are underestimated. Experiment is performed without any gravity-wave control.*
- *Exp. 2+AIREP: Variational algorithm. Background error variances are underestimated. Experiment is performed without any gravity-wave control. Additional AIREP reports.*
- *Exp. 3: Sequential algorithm. Updating of forecast error matrices, and initial error variances are correctly estimated. Experiment is performed with gravity-wave control.*
- *Exp. 4: Variational algorithm. Background error variances are correctly estimated. Experiment is performed with gravity-wave control.*
- *Exp. 5: Sequential algorithm. Updating of forecast error matrices, and initial error variances are correctly estimated. Experiment is performed with gravity-wave control. Standard observational error is included in the data.*
- *Exp. 6: Variational algorithm. Background error variances are correctly estimated. Experiment is performed with gravity-wave control. Standard observational error is included in the data.*
- *Exp. 7: Sequential algorithm. Updating of forecast error matrices, and initial error variances are correctly estimated. Experiment is performed with gravity-wave control after each analysis as a separate step. Standard observational error is included in the data.*
- *Exp. 8: Variational algorithm. Background error variances are correctly estimated. Experiment is performed without gravity-wave control. Standard observational error is included in the data.*
- *Exp. 9: Sequential algorithm. Updating of forecast error matrices, and initial error variances are correctly estimated. Experiment is performed with gravity-wave control. Twice the standard observational error is included in the data.*
- *Exp. 10: Variational algorithm. Background error variances are correctly estimated. Experiment is performed with gravity-wave control. Twice the standard observational error is included in the data.*

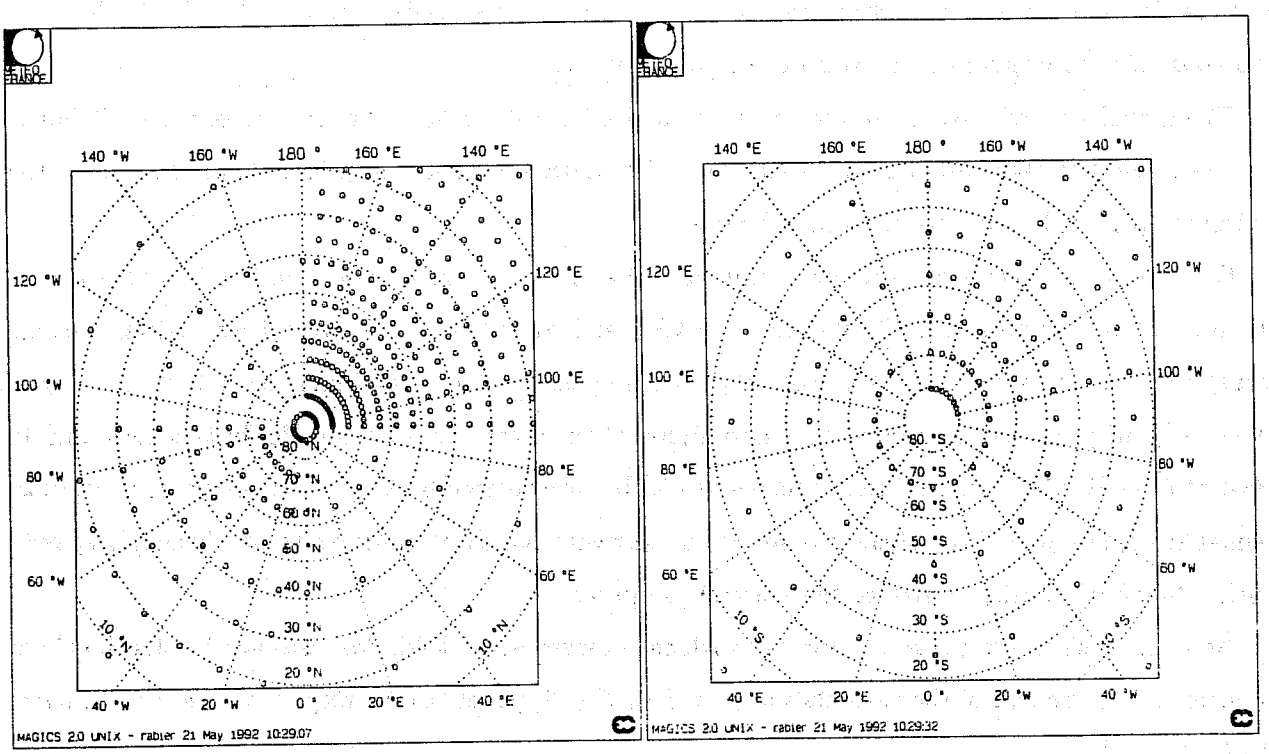


Fig. 2 *Radiosonde network as simulated in the assimilation experiments. Each radiosonde is indicated by a circle. Panel (a) represents the Northern hemisphere and Panel (b) the Southern hemisphere.*

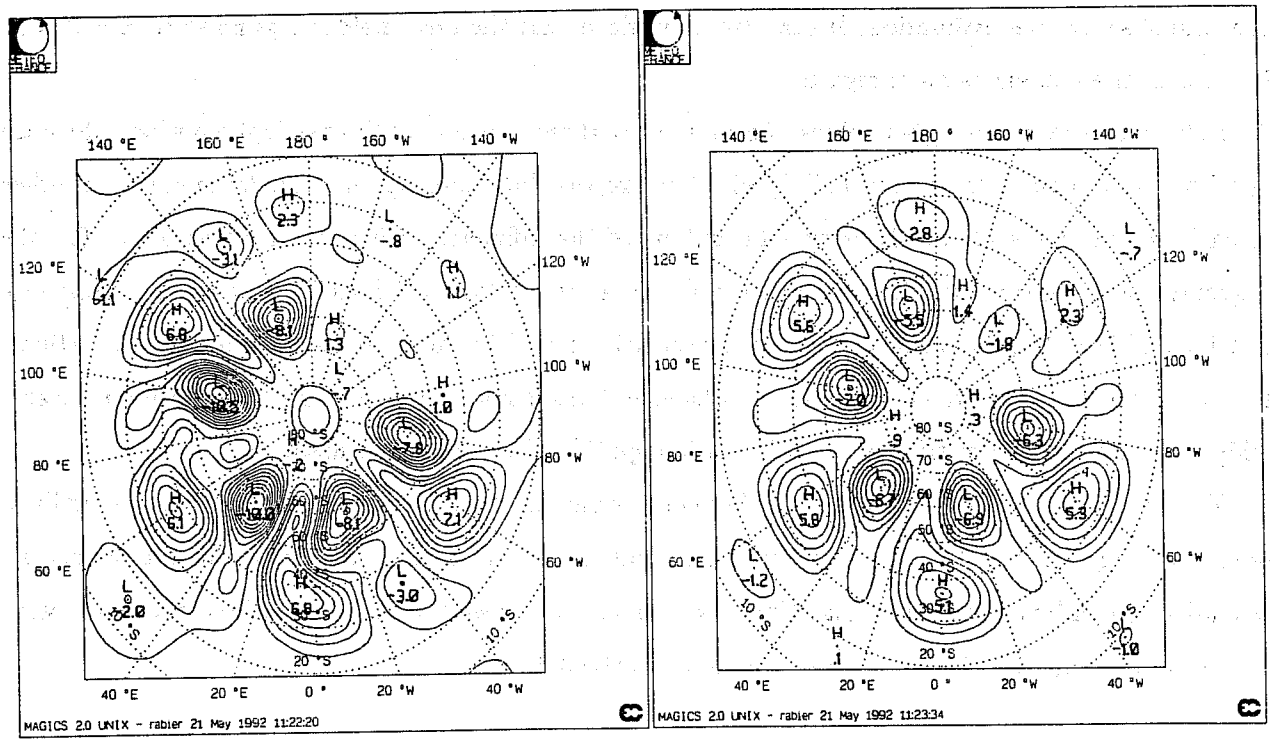


Fig. 3 *Surface pressure error fields in the Southern hemisphere at the end of the assimilation period, for the sequential algorithm (Exp. 1) and the variational algorithm (Exp. 2), in respectively panels (a) and (b). Forecast error matrices are kept constant and initial error variances are underestimated. Experiments are performed without any gravity-wave control. Isolines are drawn every 1 hPa.*

The matrix defining the forecast errors is kept constant over the assimilation cycle. No use is made of a constraint to impose an analysis free of gravity waves.

The global rms of error of surface pressure and of height at 500 hPa are presented in Table 1, for the sequential algorithm (Exp. 1) and the 4D variational algorithm (Exp. 2) at the end of the assimilation period, and after a 24-hour forecast.

For the surface pressure, the error rms at the end of the assimilation period is 1.89 hPa for the sequential scheme and 1.48 hPa for 4D-VAR , and after 24-hour forecast these values become respectively 2.63 hPa and 1.66 hPa. Not only is the sequential algorithm found to produce an analysis of worse quality but the error growth is also higher than in the case of 4D-VAR. In this case, and in most of the following experiments, conclusions can be drawn from looking at the error fields in both hemispheres, but we will illustrate the results by showing the charts relative to one hemisphere only (either North or South), to reduce the amount of figures.

Here the results are presented for the surface pressure error fields for instance in the Southern hemisphere, at the end of the assimilation period in Fig. 3, panel (a) for Exp. 1 and in Fig. 3, panel (b) for Exp. 2.

One can instantly notice the weaker errors in the case of 4D-VAR, in agreement with the better global rms of errors. The maximum absolute error value is 7.04 hPa to be compared with 10.5 hPa in the case of sequential estimation. It can also be noticed that the error field is less noisy in Exp.2 than in Exp. 1: the pattern is more regular.

In both cases we find higher values of errors in the three-quarters of the hemisphere where the data density is low than in the region (90W,180W) where the data coverage is more favourable. Another remark can be made about the horizontal extent of the influence of isolated observations. For the sequential scheme, the error field exhibits low values at about 45S, which clearly shows up for instance at 20W and 70E. Such areas of weak errors are clearly related to the presence of isolated observations, for which they represent the local radius of influence. This is not so clear for the 4D variational scheme which then appears to distribute the information spatially in a smoother manner.

This first comparison clearly shows the benefit of using the 4D variational approach over a simplified sequential algorithm. We will see in the following how some of the weaknesses of the sequential assimilation can be alleviated by inserting a control of gravity waves, but we will first in the following section concentrate on the use of the dynamics in 4D-VAR.

3.2 Propagation of information in 4D variational assimilation.

Let us now consider the same experiment as Exp. 2, and see how the introduction of AIREP reports will be used in 4D-VAR to help coming closer to reality. In addition to the previous simulated radiosonde network, some information about the horizontal wind at the 250 hPa level is now inseerted, at latitudes 58N and 58S, every 80 minutes. For each of these two latitudes, 64 "airplane reports" are equally distributed in longitude. During the 24-hour assimilation period we thus have, at 18 different times, reports on the upper-level wind at 128 locations. It should be noted that this additional information is quite negligible in terms of number of observations, compared with the whole radiosonde network.

The impact of the AIREP reports is illustrated in Table 1 where the rms of errors of (Exp. 2+ AIREP) should be compared with those relative to Exp. 2. One can see that this impact is globally positive, with in particular an improvement of about 15 percent on the quality of the surface pressure at the end of the assimilation period, which is far from being negligible. The sensitivity of surface pressure to upper-level winds was found in another context on the same baroclinic instability case by Rabier et al. (1992), where the adjoint method was used for sensitivity problems.

In Fig. 4 are presented the differences between pressure fields at the end of the assimilation period in the Northern hemisphere for instance: panel (a) shows the difference between the truth and the experiment without AIREPs, and panel (b) shows the difference between the experiment without AIREPs and the experiment with AIREPs (multiplied by 10). These two difference fields are of opposite sign, which shows that there has really been an improvement due to the new information. One can also notice that, as anticipated from linear estimation theory, the changes are greater in the regions where the radiosonde coverage is not very dense, that is where the estimation errors were relatively high (from 0E to 90E, and from 180W to 90W). The overall agreement between the structures of the difference fields over the whole hemisphere, between 70N and 20N is remarkable: the alternation of negative and positive values is quite similar. Of course the impact of AIREP data is bigger in the immediate vicinity of the latitude 58N where the data is inserted, but it is also significant further South. The changes brought by the introduction of new data amount to an important portion of the initial errors, i.e. the errors in the experiment Exp. 2. For instance, the correction is of about 1 to 1.5 hPa in regions where the initial errors were ranging from 2 to 6 hPa.

Another illustration of the propagation of information can be found in Fig. 5, for the temperature fields at the begining of the assimilation period. Panel (a) shows the temperature differences between the truth and the experiment without AIREPs, and panel (b) the temperature differences between the experiment without AIREPs and the experiment with AIREPs, both of them being multiplied by 10. First of all, one notices that the maximum impact on the temperature field is not located at

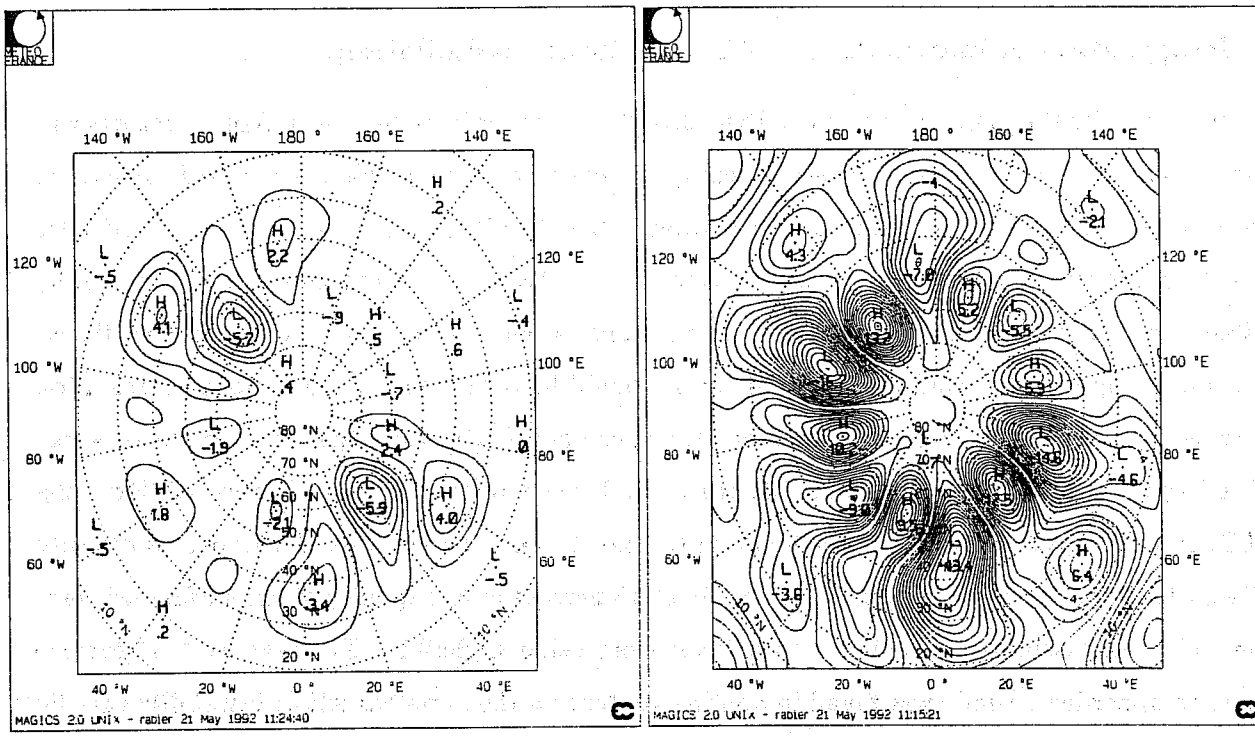


Fig. 4 *Surface pressure difference fields in the Northern hemisphere at the end of the assimilation period for 4D-VAR. Panel (a): difference between the truth and the experiment without AIREPs (Exp. 2). Panel (b): difference between the experiment without AIREPs (Exp. 2) and the experiment with AIREPs (Exp. 2+AIREP), multiplied by 10. Isolines are drawn every 1 hPa.*

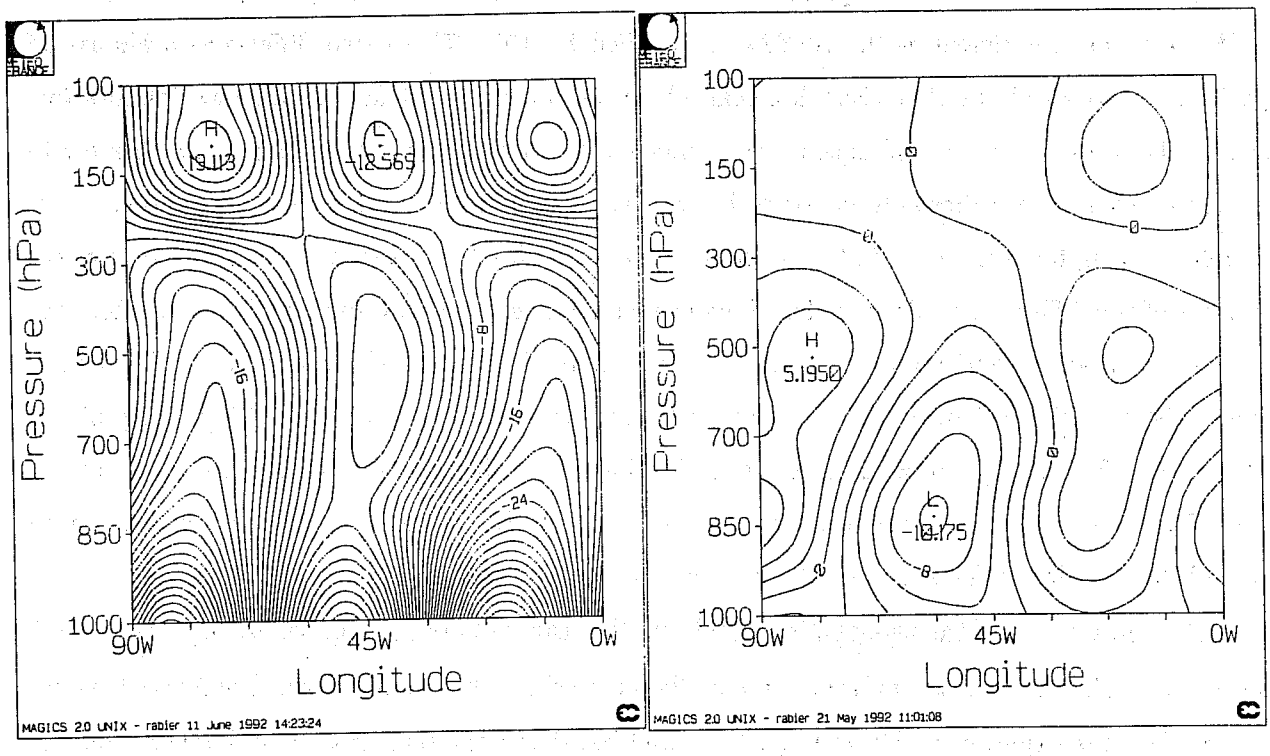


Fig. 5 *Cross-sections longitude-height of temperature error fields along the latitude 58S at the beginning of the assimilation period for 4D-VAR. Panel (a): temperature differences between the truth and the experiment without AIREPs (Exp. 2), multiplied by 10. Panel (b): temperature differences between the experiment without AIREPs (Exp. 2) and the experiment with AIREPs (Exp. 2+AIREP), multiplied by 10. Isolines are drawn every 2 K.*

the same pressure level as the observations, i.e. at the 250 hPa level. It is located in the mid and low-troposphere, between 500 hPa and 850 hPa. Secondly, the temperature differences exhibit well correlated patterns: in particular, the horizontal wavelength and the vertical tilt are similar. Using a Kalman filter in the context of a two-layer shallow-water model, Todling (1992) found that the presence of shear enhances exchange of information between the layers with consequent reduction of the analysis errors, which is perfectly consistent with the vertical propagation of information described in our experiment. Our results are also a confirmation of those obtained in a similar experiment using real observations by Thépaut et al.(1992).

The major conclusion which can be drawn from this experiment is that four-dimensional variational assimilation is truly propagating information in an efficient manner, since new isolated observations have a large impact on the quality of the analysis over a large geographical area and for various meteorological parameters, even if they are not directly observed.

3.3 Influence of gravity-wave control and of a better specification of forecast error variances.

We have seen that, in the case of badly specified forecast errors and without any constraint on the degree of balance of the analysed state, 4D-VAR behaved substantially better than the sequential algorithm, and we have illustrated in the previous section how the dynamics inserted in the 4D variational scheme helped propagating observational information. The next step is to perform the comparison in a more favourable context for the sequential algorithm, and two major differences from the previous experiments are described.

The first improvement is related to the forecast error variances. These are now updated before each analysis in the sequential algorithm, and are taken as the geographical error variances (background minus truth) over each quarter of each latitude row, at every horizontal level. The forecast errors are then no longer underestimated and the fact that the quality of the forecast improves more in data-rich areas is taken into account. The second improvement is linked with the gravity-wave problem: as we are looking for a atmospheric state, Nonlinear Normal Mode Initialization and its adjoint will be used to perform the minimization mostly in Rossby space.

Intermediate experiments were carried out (not shown in Table 1), and we will sum up the results before giving an illustration on the final experiment, with both modifications.

The introduction of NNMI is not found to have a great impact on the quality of the analyses, this impact being even weaker for the 4D variational scheme than for the sequential one, but it changes drastically the error growth rate in the subsequent forecast in the case of the sequential algorithm.

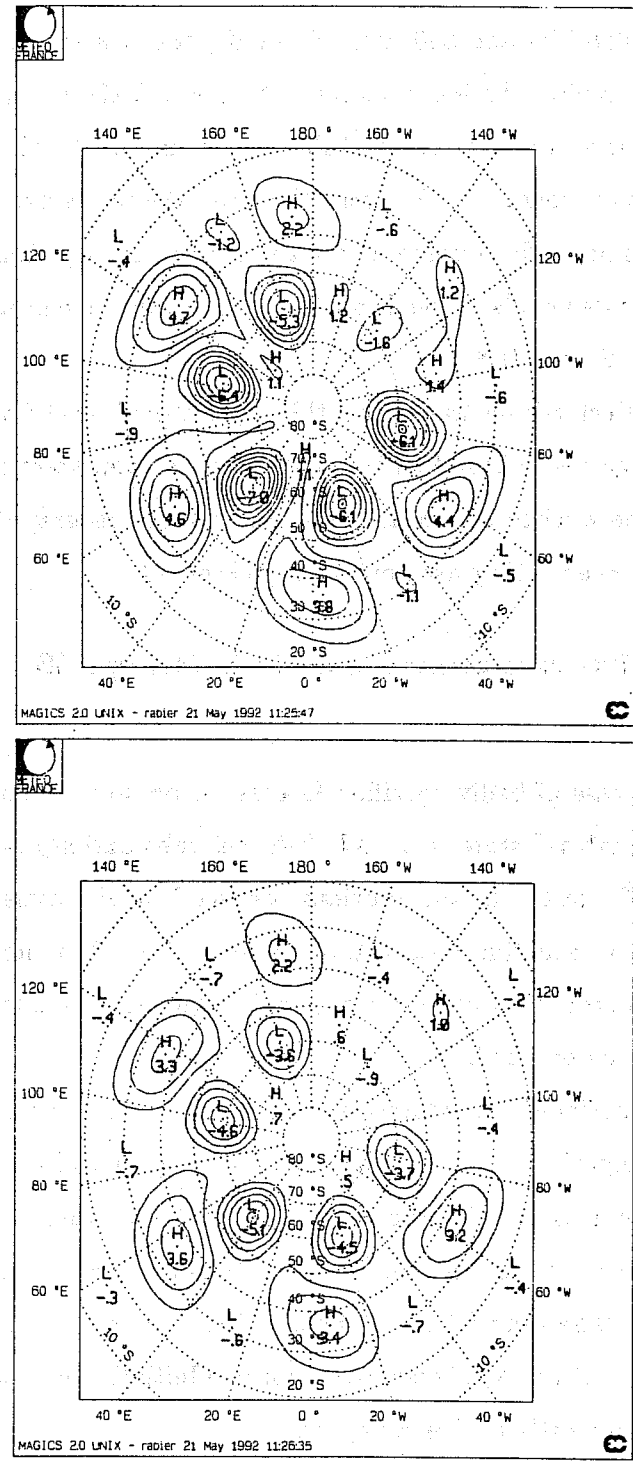


Fig. 6 *Surface pressure error fields in the Southern hemisphere at the end of the assimilation period, for the sequential algorithm (Exp. 3) and the variational algorithm (Exp. 4), in respectively panels (a) and (b). Forecast error matrices are updated in the sequential scheme and initial error variances are correctly estimated. Experiments are performed with gravity-wave control. Isolines are drawn every 1 hPa.*

The error in Exp. 1 was increased by 40 percent in 24 hours, whereas it does not exceed 30 percent in the experiments with gravity wave control.

The most important changes in the quality of the analysed state at the end of the assimilation period come from the better specification of forecast error variances. These improvements are about 4 times bigger than the ones brought by NNMI. A further remark is that the differential updating of the variances on a latitude row leads to a more homogeneous error field in the sequential scheme, its spatial homogeneity becoming comparable to that of the results of the 4D variational scheme.

The results of the experiments are presented in Table 1 and Fig. 6, where panel (a) and panel (b) show the analysed situation at the end of the assimilation period in the Southern hemisphere for respectively the sequential and the 4D variational algorithms (Exp. 3 and Exp. 4).

In terms of rms, it is clear from the comparison between Exp. 3 and Exp. 1 for the sequential scheme, and of Exp. 4 and Exp. 2 for the 4D variational scheme in Table 1, that those two changes have been highly beneficial to the quality of the estimations in both cases. The error is also increasing more moderately during the subsequent forecast for Exp. 3 than for Exp. 1. Comparing Exp. 4 and Exp. 3, the gain of using 4D-VAR is about 44 percent at the end of the assimilation period. Moreover, the error is approximately multiplied by a factor of 1.29 for the sequential algorithm, and of 1.18 for the 4D variational one, over the 24-hour forecast.

Fig. 6 can be compared with Fig. 3 to see the improvements in the Southern hemisphere (they are very similar in the Northern hemisphere).

Firstly, the overall better quality of the analysis due to the better specification of forecast error variances is immediately noticeable for both algorithms. It is clear that giving a more appropriate weight to the background results in a better agreement with the truth. One could have expected 4D-VAR to be rather insensitive to the specification of forecast errors, as these are only inserted in the cost function at the initial time. However, if we establish a parallel between 4D-VAR and the EKF, it has been explained in Daley (1991) that it takes a long integration of the EKF to get the correct asymptotic forecast error matrix. In the present experiment 4D-VAR is not implemented on a sufficiently long period of time to thoroughly correct for the extreme mis-specification of the forecast error covariance matrix given at the beginning of the assimilation period in Exp. 2. As for the more homogeneous analysis errors in the sequential scheme in Exp. 3 than in Exp. 1, it is linked with the differential updating of the variances on a latitude row, and is easily seen by comparing Fig. 6, panel (a) with Fig. 3, panel (a). An indication being given to the algorithm that the forecast is less accurate over the oceans (data-void areas), more weight is given to the observations in these regions, and more improvement due to the observation is allowed for. The opposite happens over the continents

(data-rich areas) where the analysis tends to stick to the accurate background. The fact that 4D-VAR produces errors which are also rather homogeneous spatially, can be thought of as coming from the implicit update of the forecast error variances, equivalent to what is done in the Extended Kalman Filter.

The major conclusion which can be drawn from this first series of experiments with perfect data is that the dynamical information included in the 4D variational assimilation is clearly beneficial, as confirmed by the superiority of the results over the sequential algorithm. The model dynamics is able to make an efficient use of observations. The approximate similarity of 4D-VAR with the EKF also illustrates the implicit update of the error variances according to the data density, which has to be explicitly simulated in the assimilation cycle to lead to similar results. We now proceed with the systematic comparison of the two algorithms in the context of noisy data.

4 Noisy data.

The previous experiments illustrate the prevailing differences between the algorithms (simplified sequential and 4D variational) used to assimilate meteorological observations. It is well-known that observations are not perfect, and data assimilation algorithms also have to be compared in the case of noisy data. In a practical situation, data assimilation methods have to act as filters to retain only the useful information provided by each observation. We will first run what we call a "basic experiment", similar to the previous experiment (control of gravity waves, updating of the forecast error covariance matrix before each analysis), for which some random noise has been added to the observations. The observational noise is built in such a way as to satisfy the statistics underlying the definition of observational error matrix introduced in the definition of the cost-function.

4.1 Basic experiment.

The basic experiment is intended to simulate a realistic situation, where the background field is relatively accurate and the data rather noisy (the observation error is taken from current operational values at ECMWF and, as an example, at the 1000 hPa level, the error standard-deviations for the height and for the wind are respectively 4.3 m and 2.0 ms^{-1}). The two assimilation algorithms are given the same handicap, which is an approximate definition of the forecast error covariance matrix, in particular in the definition of the correlations. As previously, for the sequential algorithm, the forecast error covariance matrix is updated taking approximately into account the spatial variability of observations. A control of gravity waves is inserted through the use of a NNMI scheme.

In Table 1 the two experiments (sequential and 4D variational) are referenced as Exp. 5 and Exp. 6. There is a clear deterioration in quality from the previous comparison (Exp. 3 and Exp. 4) due to the introduction of noise in the observations. At the end of the assimilation period, the surface pressure rms are 1.61 hPa for the the sequential algorithm (Exp. 5) and 1.04 hPa and for the 4D variational algorithm (Exp. 6); after 24-hour forecast these values become respectively 2.07 hPa and 1.23 hPa; and after 5-day forecast they become 2.57 hPa and 1.57 hPa (these 5-day figures are not indicated in the table).

As in the perfect data case, the error is approximately multiplied by a factor of 1.29 for the sequential algorithm, and of 1.18 for the 4D variational one, over the first 24-hour forecast. Comparing 4D-VAR with a sequential algorithm relying on optimal interpolation , Thépaut et al.(1992) also found a slower growth of error in the case of the 4D variational scheme.

The difference in quality between the results of the two algorithms saturates after a few days, with a global difference in rms of errors equals to 1 hPa at the 5-day range. The saturation of the difference between the two algorithms comes from the meteorological situation in which the maximum instability of the flow occurs during the assimilation period and then decreases afterwards.

For the 500 hPa height, we have respectively 14.2 meters for Exp. 5 and 7.9 meters for Exp. 6 at the end of the assimilation period, which corresponds to approximately 44 percent improvement.

Results are shown in Fig. 7 for the surface pressure errors in the Northern hemisphere (panels (a) and (b)) and the Southern hemisphere (panels (c) and (d)). Comparing the error fields in the Southern hemisphere for the perfect data case (Fig. 6, panels (a) and (b)) and for the noisy data case (Fig. 7, panels (c) and (d)), one can see that the error structures are very similar, although an overall degradation of the quality of the analyses is present. In the Northern hemisphere, the data-void areas clearly show up for both algorithms (Fig. 7, panels (a) and (b)), but the maximum errors are much worse in the sequential case (panel (a)). These important differences over the oceans are also visible on the one-day surface pressure forecasts, displayed in Fig. 8, panels (a) and (b). When comparing those with the truth (Fig. 1, panel (d)), we see that the errors on the ocean lows are of 3 hPa for the four-dimensional case, whereas they reach more than 5 hPa in the sequential case. Moreover, the corresponding weather systems are better located in 4D-VAR.

Of course, the figure of about 40 percent improvement at the end of the assimilation period when using 4D-VAR is by no means intrinsic, as it corresponds to this particular assimilation experiment. This experiment was built in such a way as to be rather representative of meteorological reality, nevertheless it cannot have the same representativeness as extensive statistics which are normally used to compare different assimilation schemes. But what can be argued from this result is that,

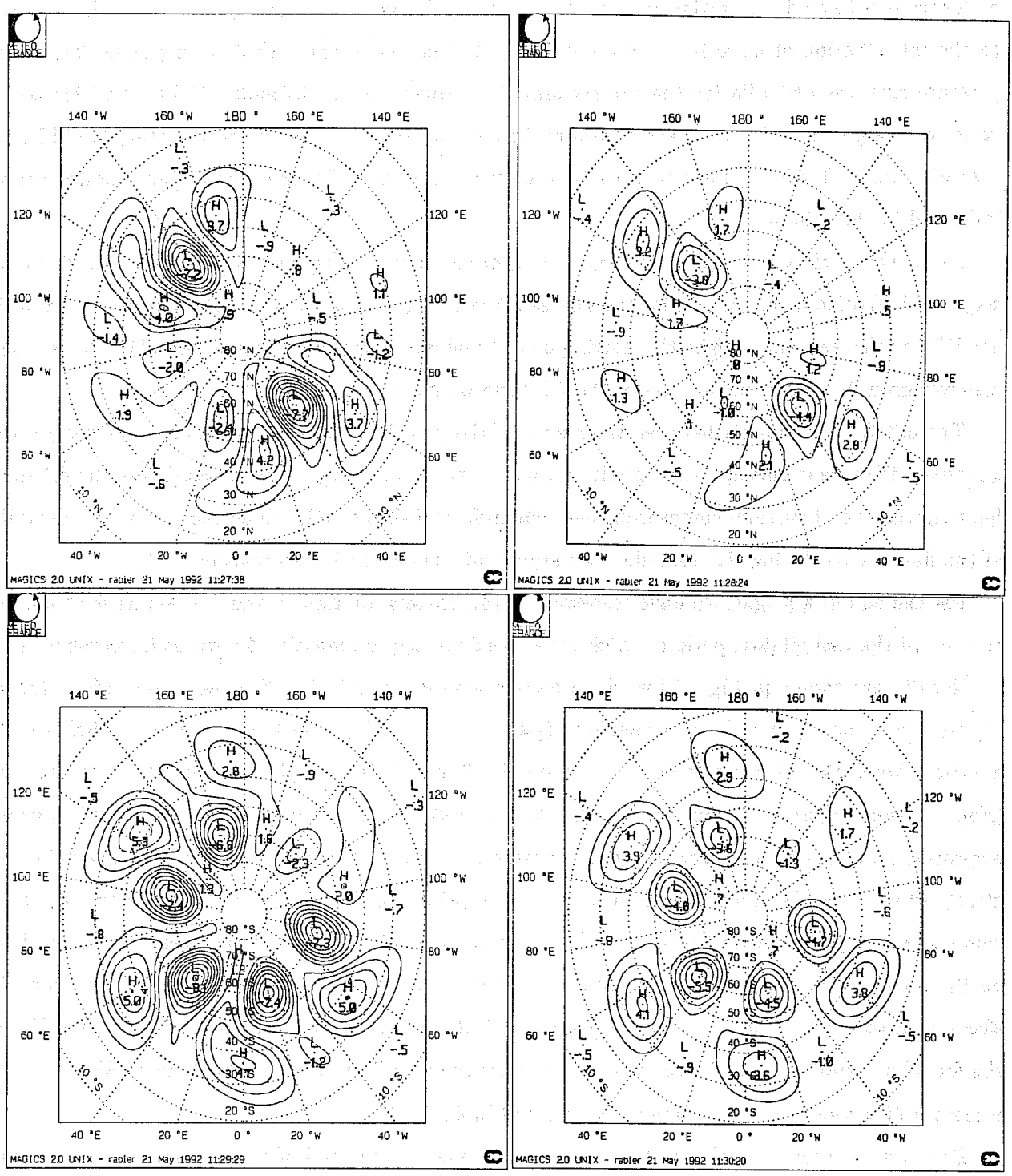


Fig. 7 *Surface pressure error fields at the end of the assimilation period for the sequential algorithm (Exp. 5) and the variational algorithm (Exp. 6) are presented in respectively panels (a) and (b) for the Northern hemisphere and (c) and (d) for the Southern hemisphere. Forecast error matrices are updated in the sequential scheme and initial error variances are correctly estimated. Experiments are performed with gravity-wave control. Standard noise is inserted in the observations. Isolines are drawn every 1 hPa.*

indeed, the fully four-dimensional scheme is potentially extremely beneficial to the quality of the analysis.

Worth noting is that the same experiment was run with forecast errors and observation errors divided by a factor of 10, to be absolutely sure to be in the range of validity of the tangent linear hypothesis. The same relative difference in the quality of the analyses is found between the two algorithms. It proves that we are really in the context where the beneficial impact of 4D-VAR is not related to its better treatment of nonlinearities but to its more coherent use of the dynamical information, linear or not.

At this stage, it is interesting to investigate the robustness of the methods, by testing them in slightly different conditions. In particular, we will study the impact of relaxing the constraint on the balance of the flow and of a larger observational noise on the two algorithms.

4.2 Sensitivity of the results to various parameters.

Practically, optimality is not within our reach, as we will never have enough information to derive the exact error statistics necessary for the estimation process. And this is even more true for the forecast error statistics than for the observational errors, because since the former are horizontally correlated they require the knowledge of too many coefficients (Dee, 1991). We thus have to implement sub-optimal filters, and as suggested by Gelb (1974), these should not be implemented without performing a set of experiments in order to test the sensitivity of the algorithms to the various assumptions used. Here, our point is not to extensively study such sensitivities as neither of the two algorithms will be used as such in practice (in particular, the forecast error matrix will be more sophisticated). But we are interested in testing the robustness of the methods to the specification of the slow manifold constraint, which can be treated in many ways (see Courtier and Talagrand, 1990; Navon and Zou, 1991; Zou and Navon, 1991). We are looking for a state, and the question we now ask is whether a constraint on the gravity-waves should be included in the analysis step, or whether is it sufficient to perform a NNMI as a separate step after the analysis. One should remember that our forecast error covariance matrix is univariate, not including any information on geostrophic coupling for instance. The result of the analysis step is then very likely to possess a fair amount of gravity waves, but these might be filtered out by the initialisation step. We perform an experiment similar to the basic one, except that no NNMI is included in 4D-VAR and that a separate initialisation step is performed at the end of each analysis in the assimilation cycle. Results show a drastic decrease in the quality of the analysis at the end of the assimilation cycle for the sequential algorithm (compare in Table 1 the results for current Exp. 7 and Exp. 5), the rms of surface pressure errors going from 1.61 hPa up to

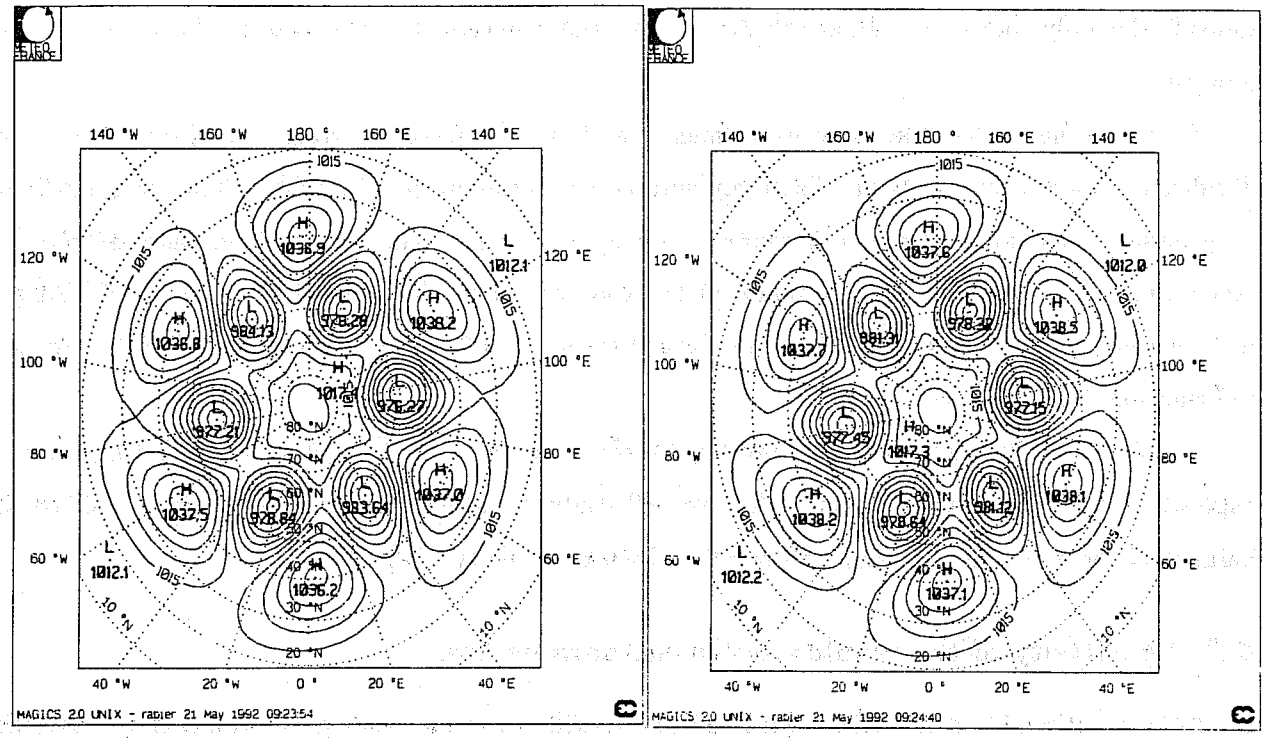


Fig. 8 *One-day forecast surface pressure fields for the Northern hemisphere for the sequential algorithm (Exp. 5) and the variational algorithm (Exp. 6) are presented in respectively panels (a) and (b). Forecast error matrices are updated in the sequential scheme and initial error variances are correctly estimated. Experiments are performed with gravity-wave control. Standard noise is inserted in the observations. Isolines are drawn every 5 hPa.*

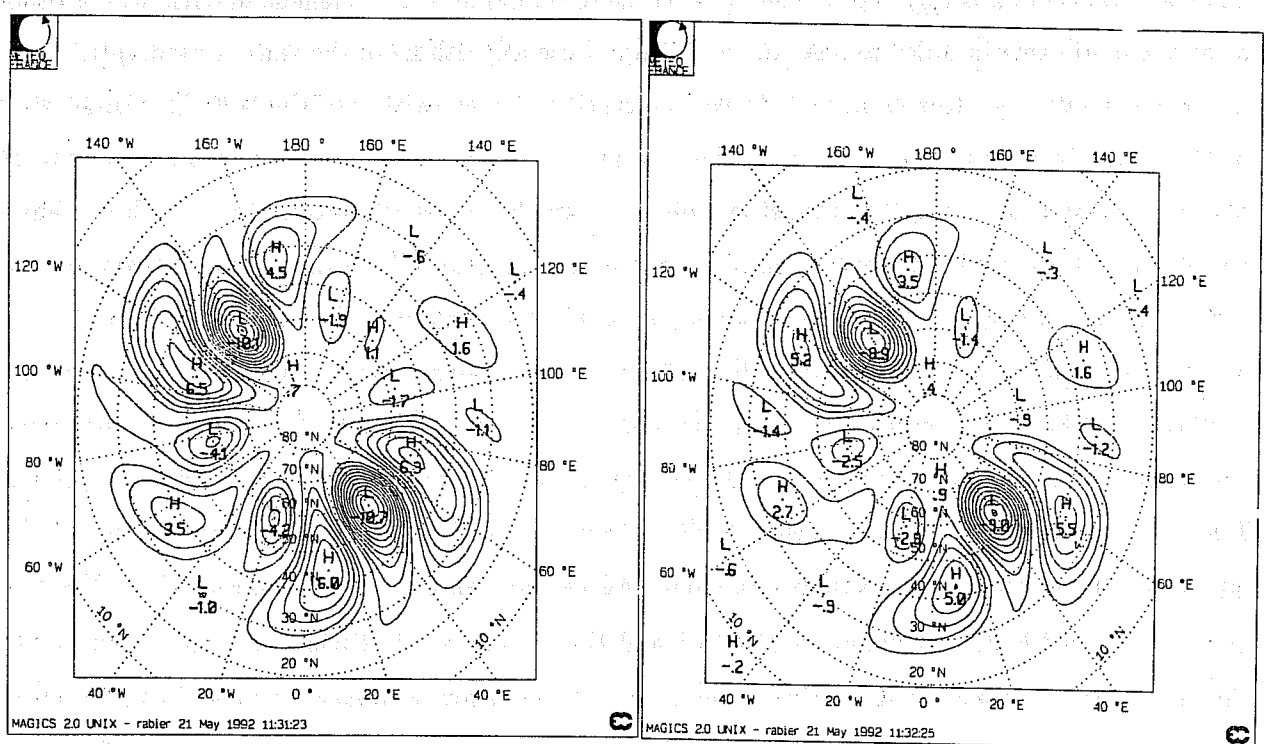


Fig. 9 *Surface pressure error fields for the Northern hemisphere at the end of the assimilation period for the sequential algorithm (Exp. 9) and the variational algorithm (Exp. 10) are presented in respectively panels (a) and (b). Forecast error matrices are updated in the sequential scheme and initial error variances are correctly estimated. Experiments are performed with gravity-wave control. Twice the standard noise is inserted in the observations. Isolines are drawn every 1 hPa.*

2.21 hPa. No such phenomenon is seen to happen for the 4D variational scheme, with only a slight increase in the rms from 1.04 hPa for Exp. 6 to 1.14 hPa for current Exp. 8. Of course, this is not a thorough investigation of the problem, but it points out the remarkable robustness of 4D-VAR, as far as the slow manifold constraint is concerned. In the case of 3D-VAR much more care must be taken in the definition of the forecast error covariance matrix and/or the way the balance is directly imposed.

The other sensitivity problem we study is not related to approximations made during the estimation process, but to an external parameter, namely the variability in the quality of the observations. What happens if the observations are noisier? Again, an experiment similar to the basic one is performed, the only difference being that the noise is doubled. The sequential and 4D variational experiments are respectively called Exp. 9 and Exp. 10. As can be seen in Table 1, the four-dimensional scheme (Exp. 10) leads to an approximate global doubling in errors, with in particular the surface pressure rms jumping from 1.04 hPa in Exp. 6 to 2.08 hPa. It is perfectly consistent with the tangent-linear hypothesis, as can be seen from Eq. 3 relating the estimation error to the observation error. This result suggests that 4D-VAR acts roughly as a linear estimation of the atmosphere from the observations, with a negligible impact of the background field on the quality of the estimate in this particular experiment. On the other hand, the estimation error of the sequential scheme is multiplied by a factor smaller than 2 (the rms of 1.61 hPa in Exp. 5 becomes 2.37 hPa in Exp. 9). The quality of the result of the sequential algorithm does not depend solely upon the observational noise. It appears that the successive weights given to the background fields in the assimilation process are certainly over-estimated, although the forecast error standard deviations had been approximately updated. This excessive weight is then likely to come from the bad specification of the forecast error correlations, and in particular the lack of a geostrophic coupling.

Results are also shown in Fig. 9 for the surface pressure errors in the Northern hemisphere (panels (a) and (b) for respectively the sequential and 4D variational algorithms) to be compared with the standard observational noise (Fig. 7, panels (a) and (b)). In the case of 4D-VAR, the degradation is rather uniform in space, athough for the sequential scheme, comparing Fig. 9, panel (a) with Fig. 7, panel (a), it can be seen that the degradation is the most important in relatively data-dense areas. In data-void areas, too much weight is certainly given to the background terms, and as a consequence, the contribution of the background to the estimation error is important.

Therefore, in this particular experiment, the quality of 4D-VAR is found to be mainly governed by the quality of the observations, whereas the quality of the sequential scheme also depends drastically upon the accuracy of the specification of forecast error variances.

These results highlight the importance of a very accurate definition of the various parameters

included in the implementation of the assimilation algorithms, and in particular in the description of the forecast error covariance matrices, or of the slow manifold constraint for the sequential scheme. They also raise a subject which has not been treated here, namely the observation quality control. Both algorithms are expected to deal with this problem, and in particular Lorenc and Hammon (1988) have shown how information about gross errors in data can be inserted in the variational context.

5 Results with a different formulation of the J_G term.

So far, the approximations in the forecast error correlations have been exaggerated, compared with those performed in the operational practice. We will now present the results of an experiment similar to the basic one (section 4.1), but with a more elaborate version of the J_G term included in the cost-functions to be minimized by the variational schemes.

5.1 Formulation of the J_G term.

As previously, the forecast error covariance matrix to be used in the J_G term is separated into its standard deviations component and its correlations component. The main difference comes from a better specification of the correlations, with the inclusion of both a vertical and a geostrophic coupling (and an updated version of the horizontal correlations).

The vertical correlations are those used operationally at ECMWF. The geostrophic coupling is introduced in a way comparable to the one used in the Spectral Statistical Interpolation scheme, at the National Meteorological Center (Parrish and Derber, 1992). The spectral fields of both the control variable and of the background field are separated into their Rossby and gravity components. Then, two contributions are included in the definition of the J_G term: one measures the misfit between the Rossby parts of the vorticity fields, the other measures the misfit between the unbalanced parts of the height fields and the misfit between the divergence fields (we accept a fraction of 10 percent of the flow to be out of the Rossby modes). More details can be found in Heckley et al. (same volume).

As for the forecast error standard deviations, they will be taken constant on the horizontal at each model level.

5.2 Results.

We will now present the results equivalent to experiments 5 and 6 described in section 4.1 with this new formulation of the J_G term. As a first experiment, we performed one 4D-VAR and one sequential assimilation using as forecast error standard deviations the global rms of error over each model level.

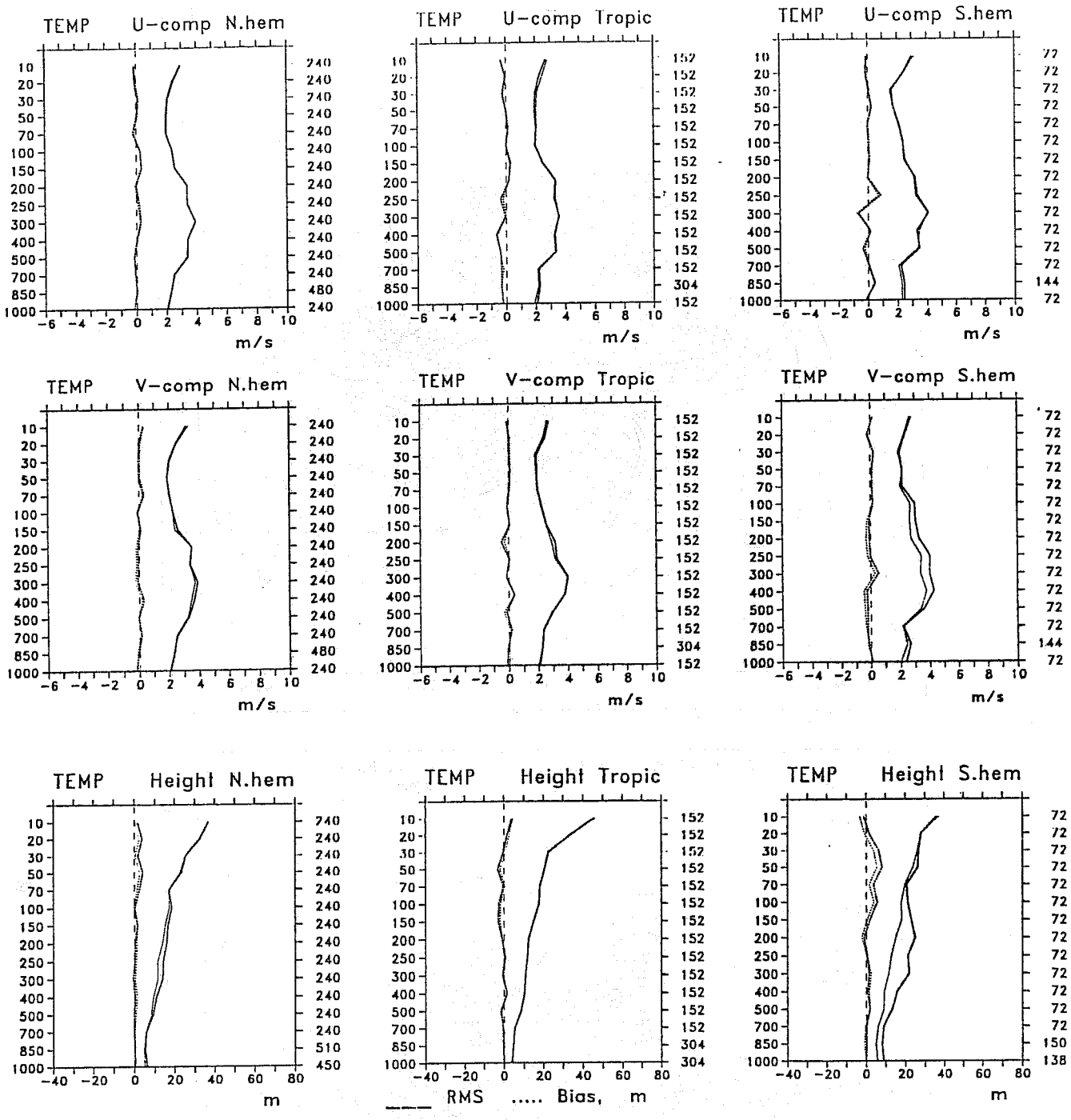
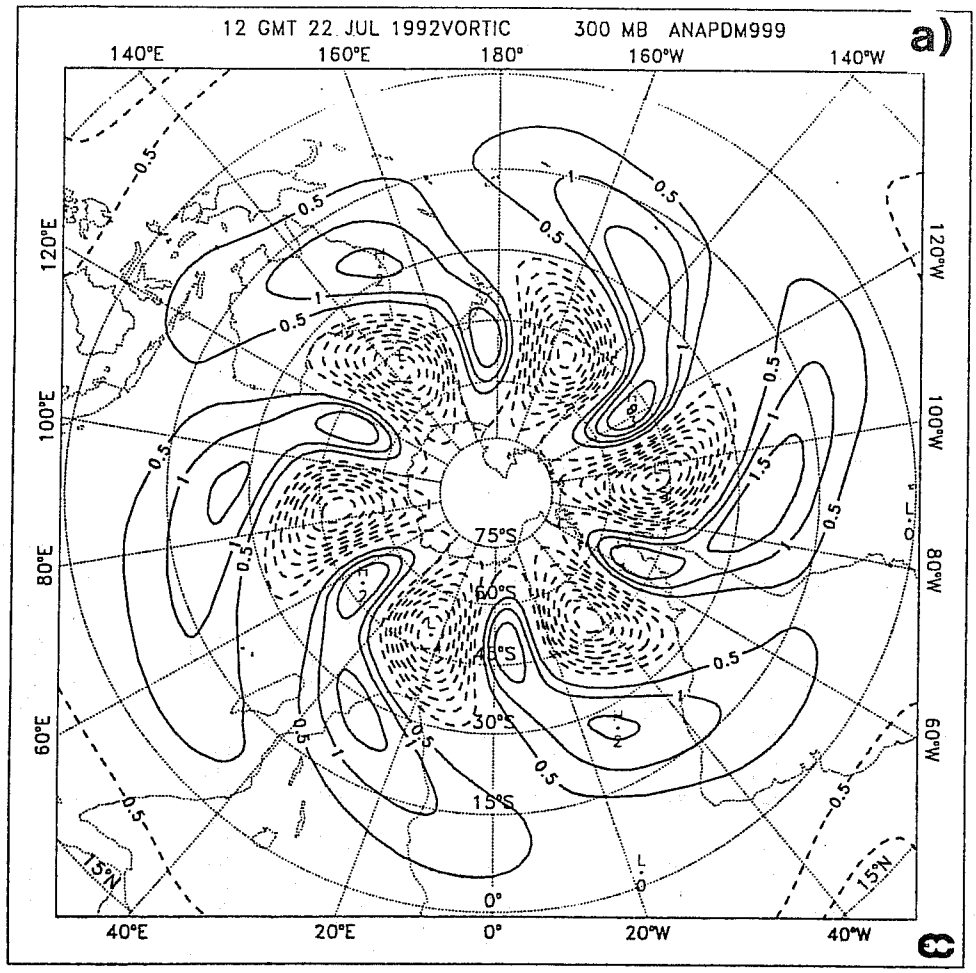
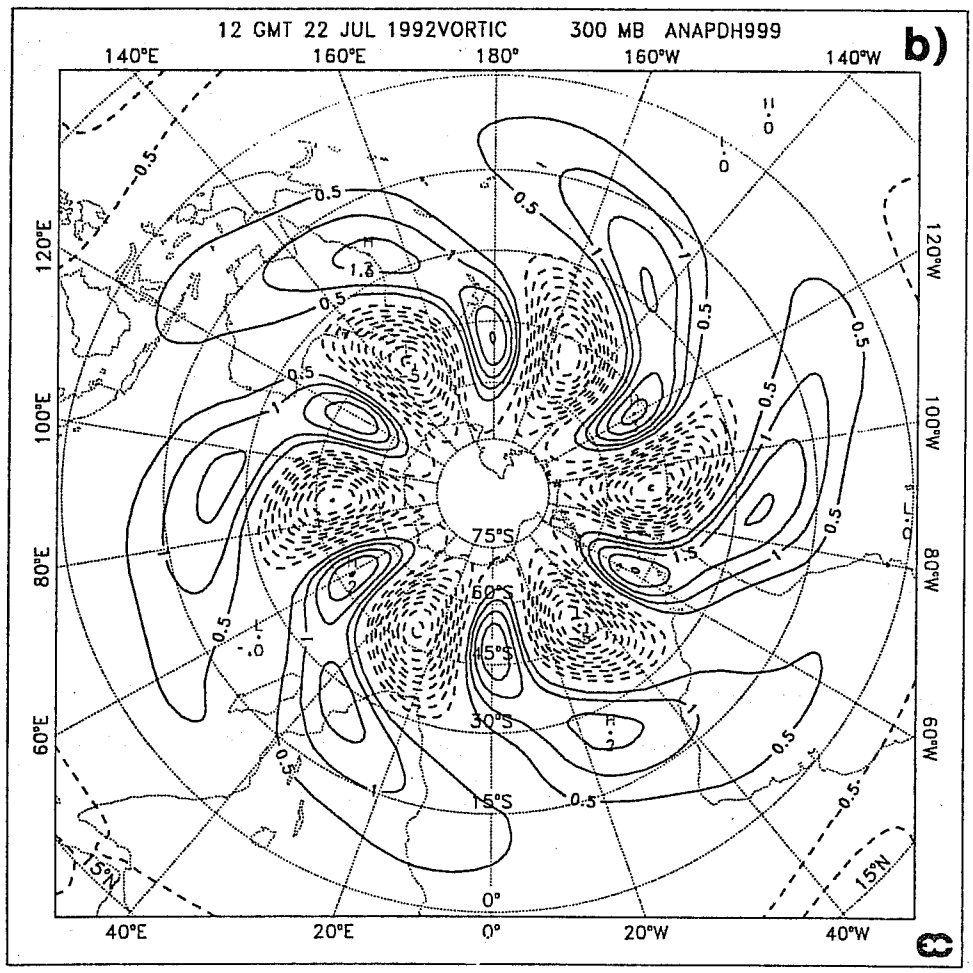


Fig. 10 *Fits to the data at the end of the assimilation period for the sequential scheme (thick curves) and the 4D variational scheme (thin curves). The full lines represent the rms of the differences (analysis-observations), the dotted lines represent the bias. The observations are zonal and meridional components of the wind and heights. The values are computed separately for three different areas: Northern Hemisphere, Tropics and Southern Hemisphere. The number of data is indicated on the right-hand side of the plots, the pressure in hPa on the left-hand side.*


12 GMT 22 JUL 1992VORTIC 300 MB ANAPDM999
a)
140°E
160°E
180°
160°W
140°W
120°E
100°E
80°E
60°E
120°W
100°W
80°W
60°W
75°S
60°S
45°S
30°S
15°S
0°
15°N
40°E
20°E
0°
20°W
40°W


12 GMT 22 JUL 1992VORTIC 300 MB ANAPDH999
b)
140°E
160°E
180°
160°W
140°W
120°E
100°E
80°E
60°E
120°W
100°W
80°W
60°W
75°S
60°S
45°S
30°S
15°S
0°
15°N
40°E
20°E
0°
20°W
40°W

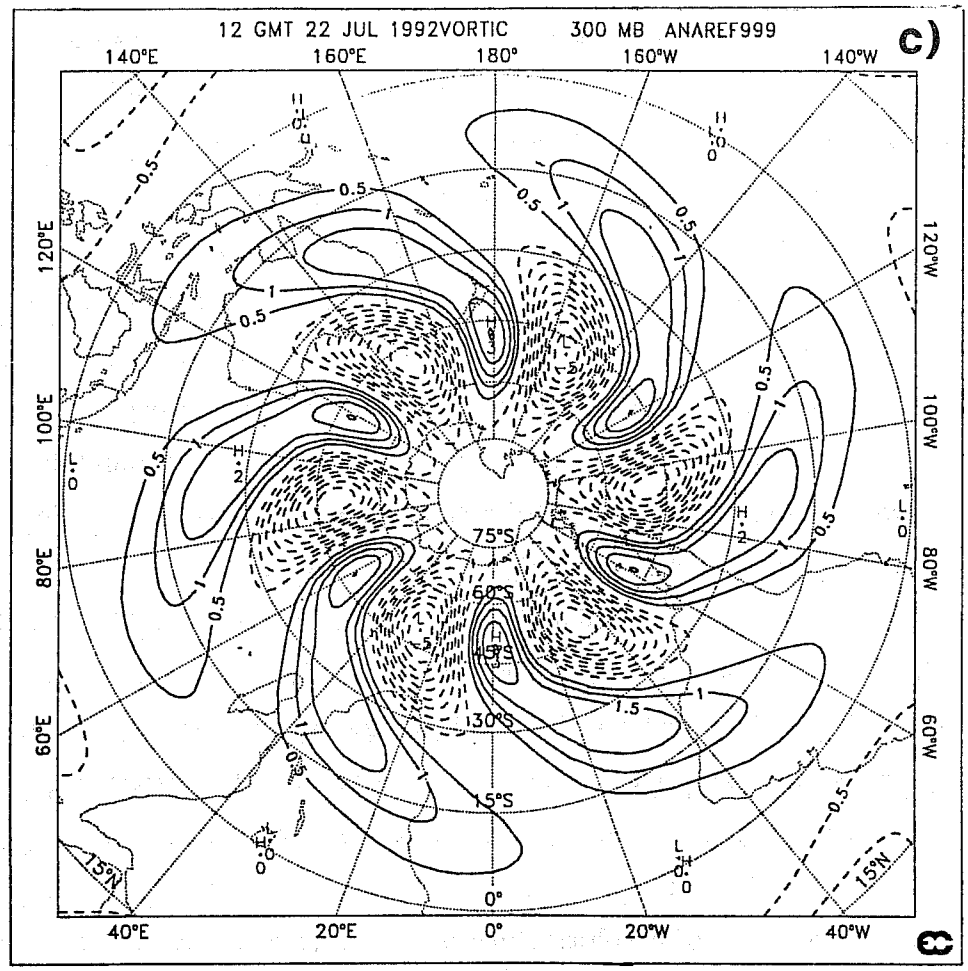


Fig. 11 *One day forecast 300 hPa vorticity fields for the Southern hemisphere. Panel a) represents the forecast from the analysis produced by the sequential scheme, panel b) the forecast from the analysis produced by the 4D variational scheme and panel c) the reference field. The units are* $10^{-5}s^{-1}$.

A representative value for the initial backgound field is 3.5 ms^{-1} for the wind at 1000 hPa. The analysis error rms at the end of the assimilation period for the height at 500 hPa is 19.5 m for the sequential algorithm and 12.4 m for the 4D variational algorithm. These rms are higher than those computed for the basic experiment (respectively 14.2 m for Exp. 5 and 7.9 m for Exp. 6). This comes from the fact that in the basic experiment, although the forecast error correlations were crude, the forecast error standard deviations were computed in a rather sophisticated way, taking into account horizontal variations. It is clear that, for this academic situation, most of the forecast error is located at mid-latitudes, with almost no error over the tropical area. A global averaged standard deviation is then by no means representative of the values of the forecast errors at mid-latitudes. A better approximation would be about twice this global rms, and we have redone the previous experiments multiplying by 2 the computed background global standard deviations.

As a remark, the forecast error standard deviation for the initial background field (at the beginning of the assimilation period) is now taken as 7 ms^{-1} for the wind at 1000 hPa, to be compared with a typical value of 3.5 ms^{-1} in operations. The initial background is then less accurate by a factor 2 than a typical operational background. In the sequential scheme, the forecast standard deviations are updated before each analysis, and the value obtained before the last analysis at the end of the 24-hour period is 3.4 ms^{-1}. The final background field is now as accurate as a typical background. On the other hand, the accuracy of the observations is taken everywhere equal to the accuracy of the American radiosonde network, that is 2.0 ms^{-1} for the wind at 1000 hPa. On the whole, we are then giving relatively more weight to the data and less weight to the background than in the operational practice.

As far as the analysis error rms at the end of the assimilation period for the height at 500 hPa are concerned, they now drop to 10.9 m for the sequential algorithm and 5.7 m for the 4D variational algorithm. Firstly, one can notice that the quality of the results has been drastically improved by the better specification of the forecast error standard deviations, for both algorithms. Secondly, the quality of the analyses is now better than that produced by the basic experiment. Thirdly, the difference in quality between the analysis produced by 4D-VAR and the one produced by the sequential scheme is still very significant.

In Fig. 10, one can see the fit of the final analyses to the available data at the end of the assimilation period (these data correspond to the data used in the last analysis of the sequential scheme). The full curves represent the rms and the dotted curves the biases. The thick and light curves show the fit of the analysis produced by the sequential and 4D variational schemes, respectively, to the available observations. In the tropics, for all types of observations (zonal and meridional winds, heights), the

curves are almost the same for the results of the two assimilation algorithms. This reflects the fact that, as the backgound was already of excellent quality in this area, both analyses also are of excellent quality and the fits to the data just show the noise present in the observations.

In the Northern hemisphere, due to the large number of data inserted in the course of the assimilation, the resulting fits to the data are also quite similar. In contrast, the results are very different in the Southern hemisphere. This is particularly noticeable for the fits to the height data and the meridional wind in the troposphere. On top of the observational noise, the error of the result of the sequential scheme clearly seems to be large around 300 hPa.

A 24-hour forecast was run from the results of the two assimilation algorithms. In terms of global error rms the values for the height at 500 hPa are now 12.8 m for the sequential scheme and 5.8 m for the 4D variational scheme. The 300 hPa vorticity forecast charts for the Southern Hemisphere are displayed in Fig. 11 where panel a) represents the reference field, panel b) the forecast from the result of the sequential scheme and panel c) the forecast from the result of the 4D variational scheme.

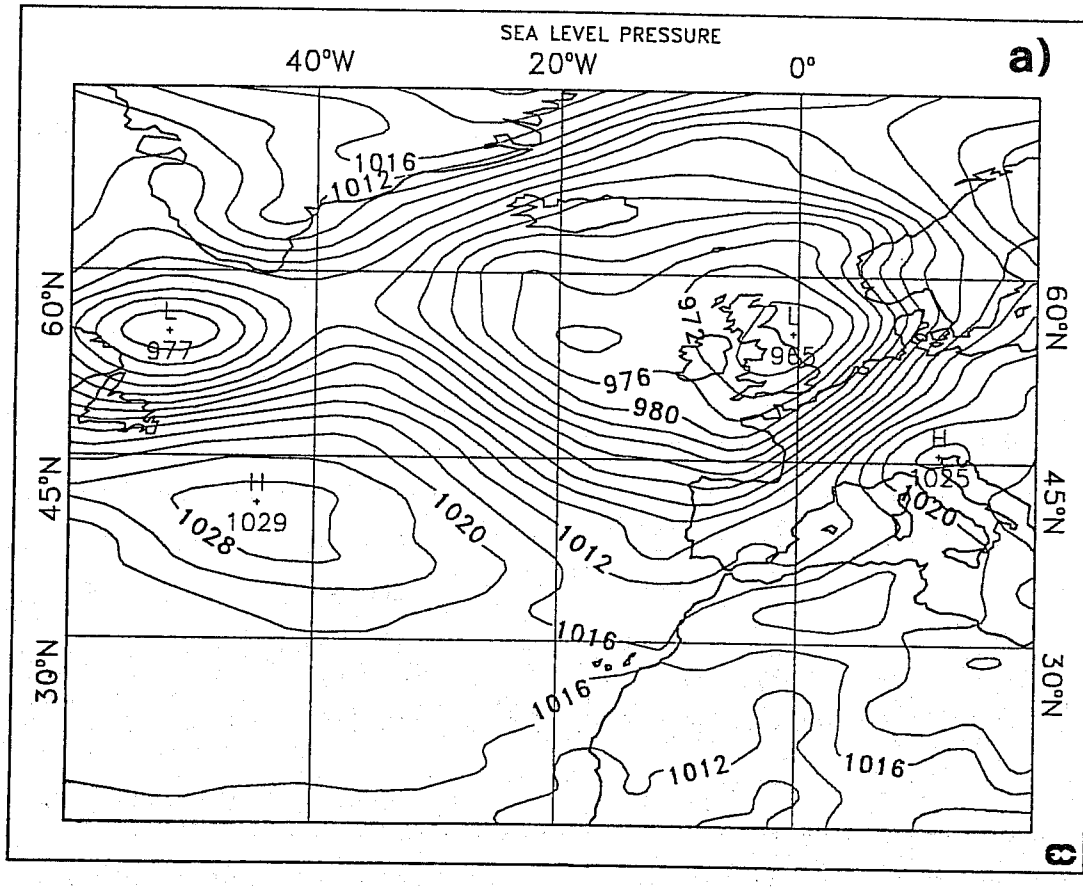
In the area between longitudes 90°W and 180°W, where the data coverage is reasonable, the differences between the forecasts and the reference are hardly noticeable. In the rest of the hemisphere, the better quality of the forecast displayed in panel c) clearly shows up. In particular, the locations of the maxima/minima are always more accurate in panel c) than in panel b) (one can for instance check it for the maximum located at 60 ° E).

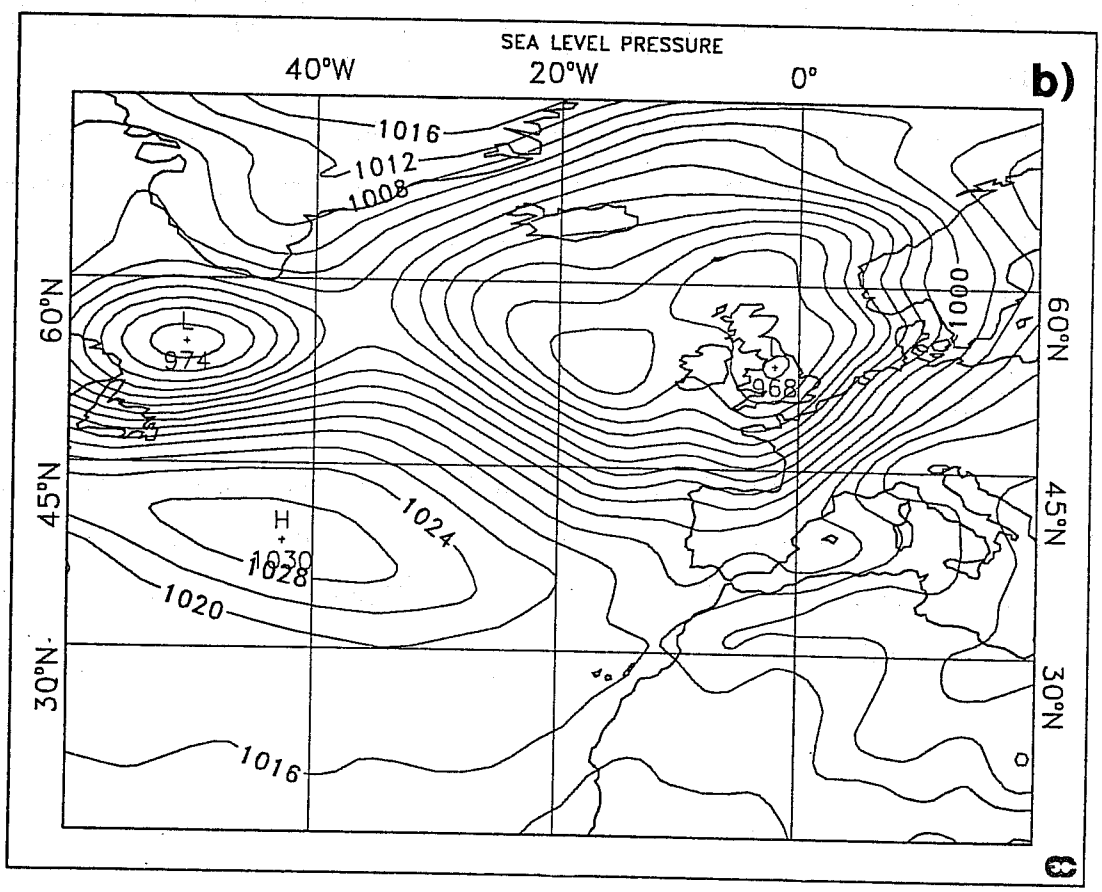
It seems that the sequential scheme has not been able to correct thoroughly for the phase-error of the initial background.

These additional results are particularly encouraging in that they confirm the former comparisons performed in this study, with a multivariate 3D formulation of the J_G term in the variational schemes. They also illustrate the importance of a correct specification of the forecast error standard deviations, for both algorithms.

6 OCTOBER 87' STORM

Now that we have seen how beneficial is the use of four-dimensional variational assimilation upon a simplified sequential scheme in the case of academic baroclinic instability situation, let us extend the comparison over a real meteorological situation. The October 87' Storm was selected as it is a typical example of an explosive cyclogenesis.


SEA LEVEL PRESSURE
40°W
20°W
0°
a)
60°N
45°N
30°N
1016
1012
977
976
980
1029
1028
1020
1025
1016
1012
1016
1012
1016


SEA LEVEL PRESSURE
40°W
20°W
0°
b)
60°N
45°N
30°N
1016
1012
1008
974
1000
968
1024
1030
1028
1020
1016

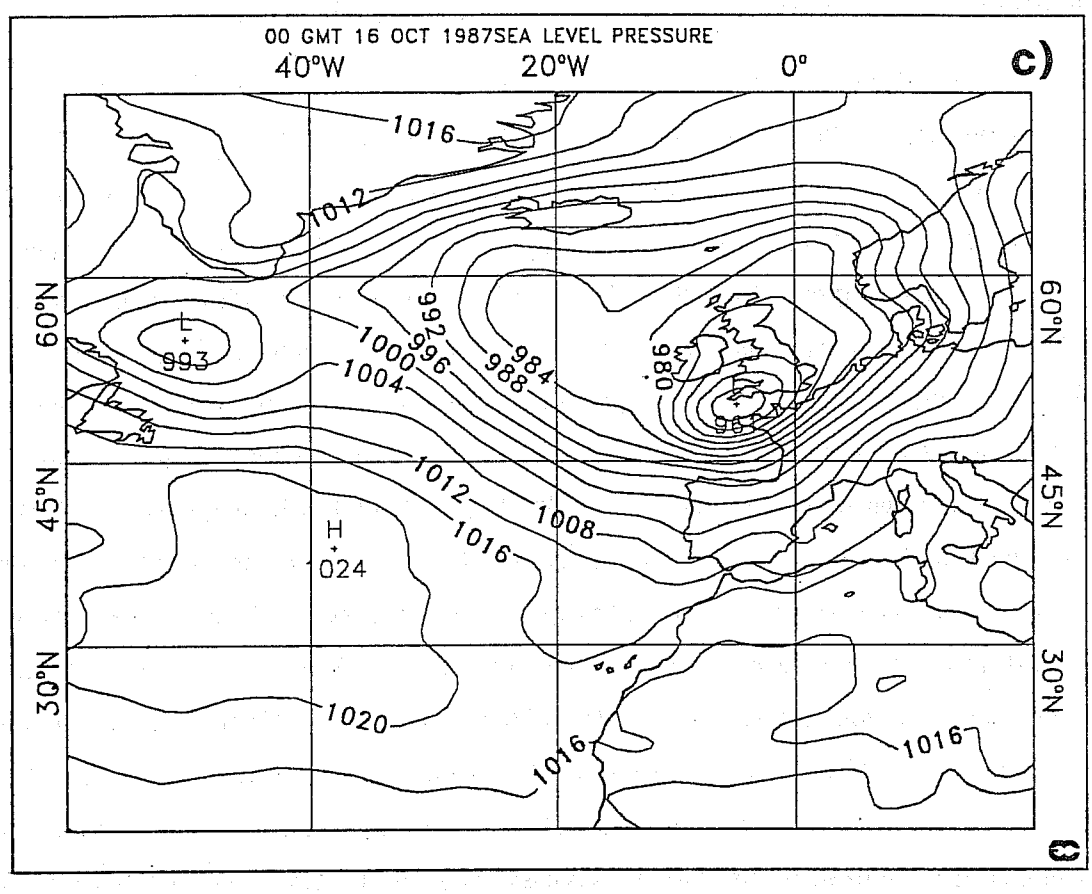


Fig. 12 *36-hour forecasts for the sea-level pressure in the purely adiabatic case (no vertical diffusion), valid for 16/10/87,0 UTC, and corresponding adiabatic T63 analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis and panel c) the corresponding analysis.*

6.1 Meteorological situation and set-up of the experiments

The meteorological situation has been documented in a number of papers (Morris and Gadd, 1988; Hoskins and Berrisford, 1988; Lorenc et al., 1988; Jarraud et al., 1989;...), and we will not go into details. In short, between Thursday and Friday 15-16 October 1987, a violent storm hit France (Brittany and Normandy) and the South of England. A surface low rapidly developed over the sea, with its pressure dropping from 985 hPa to 951 hpa in 24 hours, wind gusts of 50 ms^{-1} were recorded and the damage was considerable on both sides of the Channel. From a forecast point of view, the model outputs available for 16/10/87,0 UTC showed some inconsistencies. Global models of Météo-France and UKMO were forecasting a low near the tip of Brittany, at the 24-hour range, but underestimated its intensity (giving respectively 970 hPa and 965 hPa). In contrast, both the UKMO fine-scale model at the 24-hour range and the ECMWF T106 spectral model at the 36-hour range failed to predict the rapid evolution of the atmosphere over this region. As pointed out by Jarraud et al. (1989), these short-range inconsistencies seem to indicate problems in the data assimilation part of the numerical prediction systems. As a matter of fact, Lorenc et al. (1988) succeeded in simulating the storm with the UKMO fine-scale model, starting from a different analysis for the 15/10/87, 0 UTC, increasing both the the cut-off period and the weights given to the observations. Simmons (pers. comm.) also obtained better 24-hour forecasts with the ECMWF model using a later version of the analysis scheme.

From an analysis point of view, there are two prevailing features in the atmospheric circulation 24 and 36 hours before the storm. These are an upper-level jet at about 250 hPa moving rapidly Eastwards, and an intense low-level baroclinic zone, located over the Atlantic Ocean. When the jet gets in phase with the low-level baroclinic zone, between 15/10/87, 0 UTC and 15/10/87, 12 UTC, the secondary surface low starts developing explosively.

Our purpose is to test 4D-VAR on this situation, and compare a 36-hour forecast valid for 16/10/87,0 UTC coming from a 4D-VAR analysis with a 36-hour forecast coming from an analysis of an operational type, produced in similar conditions. The experiments will be run at resolution T63L19 with the ARPEGE/IFS model, with no physics or simplified physics.

The first experiment is adiabatic. We have run an adiabatic T63L19 assimilation cycle with the ECMWF model and Optimal Interpolation analysis (OI) for four days between 12/10/87,12 UTC and 16/10/87,12 UTC. A 4D-VAR is performed over the period (13/10/87,15 UTC; 14/10/87,15 UTC). The result of the minimizing solution of this 4D-VAR will be taken at 14/10/87,12 UTC, to be compared with the OI valid for the same time. Due to 4D-VAR assimilation period which is extending until 3 hours later, we have then used the same amount of data for these two analyses (4D-VAR and

OI) at this date.

4D-VAR is minimizing a function of the type $J = J_O + J_G + J_C$. J_O represents the misfit between the model trajectory and the observations. J_G quantifies the discrepancy between the initial point of the 4D-VAR trajectory and a first-guess. The first-guess is in this case a 9-hour forecast from the OI at 13/10/87,6 UTC. The standard deviations of this first-guess are taken as those computed by the assimilation system for the 6-hour forecast valid for the 14/10/87,12 UTC, averaged at each horizontal level. We are then giving a disadvantage to 4D-VAR by over-estimating the accuracy of the first-guess (6-hour forecast instead of 9-hour forecast errors), and more particularly over data-void areas due to the horizontal averaging. The J_G formulation is the same as the one used in section 5: a balance is included in the formulation of the forecast error correlations. As in Courtier and Talagrand (1990), a penalty term (J_C) on the tendency of the gravity modes of the analysed state is included in the cost-function. We also used two iterations of Normal mode Initialization. The minimization algorithm was stopped after 30 iterations. We have then performed two 36-hour forecasts from the two different analyses valid at 14/10/87,12 UTC to be compared with the OI coming from the adiabatic assimilation cycle valid for the 16/10/87,0 UTC.

In Fig. 12 are presented the results of the sea-level pressure forecasts and the corresponding T63L19 adiabatic analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis, and panel c) the OI analysis. Both forecasts are underestimating the intensity of the low by a few hPa, although the location of the low is slightly better in the case of 4D-VAR. At this stage, one has to be very careful in interpreting the results because of the lack of physical processes. As a matter of fact, if one compares the intensities of the low at 50°W and of the high pressure at 40°W, it is noticeable that both forecasts deepen the low and strenghten the high exaggerately compared with the corresponding OI. Furthermore, winds at 1000 hPa are too strong (not shown). It was then decided to rerun a similar experiment including vertical diffusion and surface drag in the assimilation cycle, 4D-VAR and the forecasts. In 4D-VAR, a simplified vertical diffusion scheme and its adjoint were introduced. These were originally developed for computing realistic most unstable modes for the model (Buizza, pers. comm.). It is then possible to use surface observations such as winds from SYNOPs and DRIBUs in 4D-VAR, which makes it even more comparable with OI, in terms of observations used. For the forecasts, the vertical diffusion used is the one operational at Météo-France for the ARPEGE/IFS model.

Firstly, let us compare the analyses produced by 4D-VAR and the assimilation cycle for the 14/10/87,12 UTC. In Fig. 13 are presented the analysed winds at 250 hPa for OI (panel a)) and 4D-VAR (panel b)). Both analyses exhibit an area of high winds noticeable at (45°N;40°W) corre-

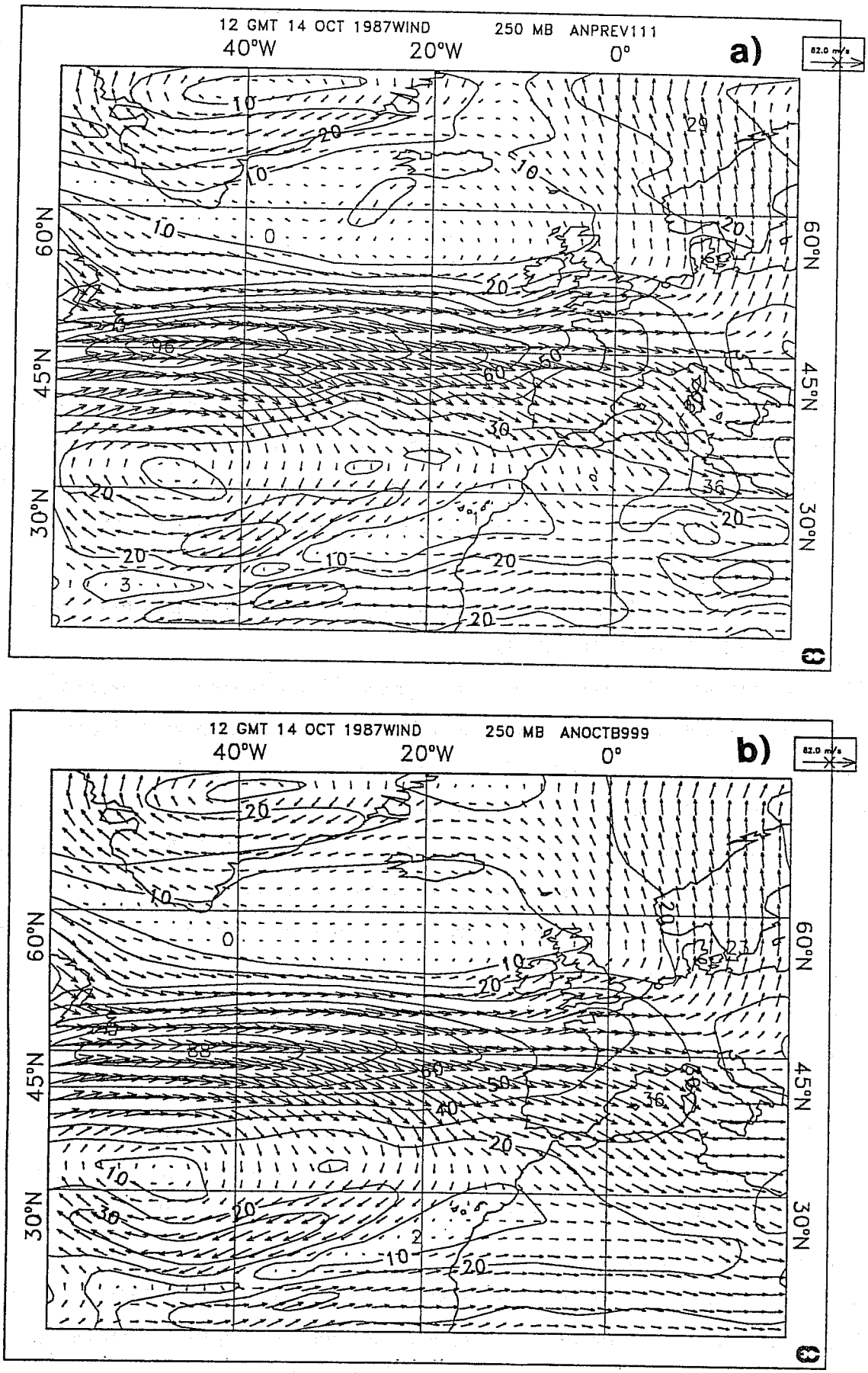
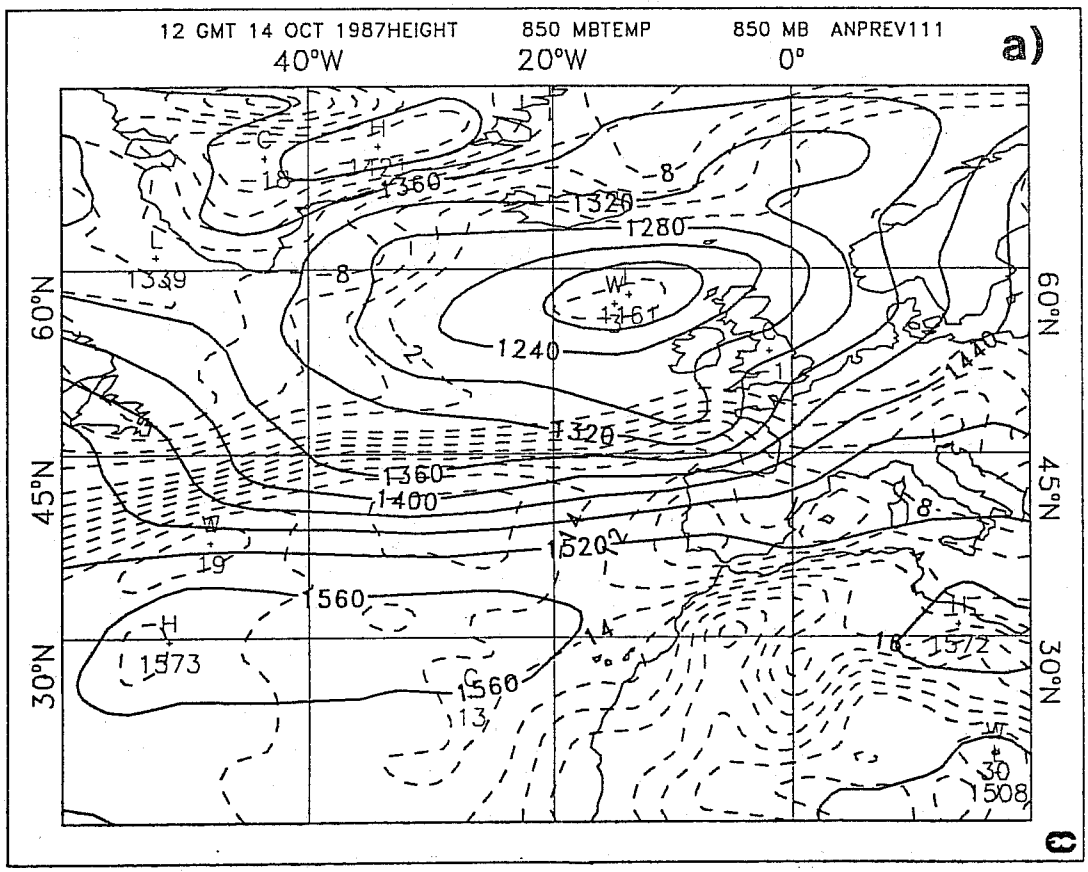


Fig. 13 *Analyses for the wind at 250 hPa, valid for 14/10/87,12 UTC. Panel a) represents the OI T63 initialized analysis, panel b) the 4D-VAR analysis.*



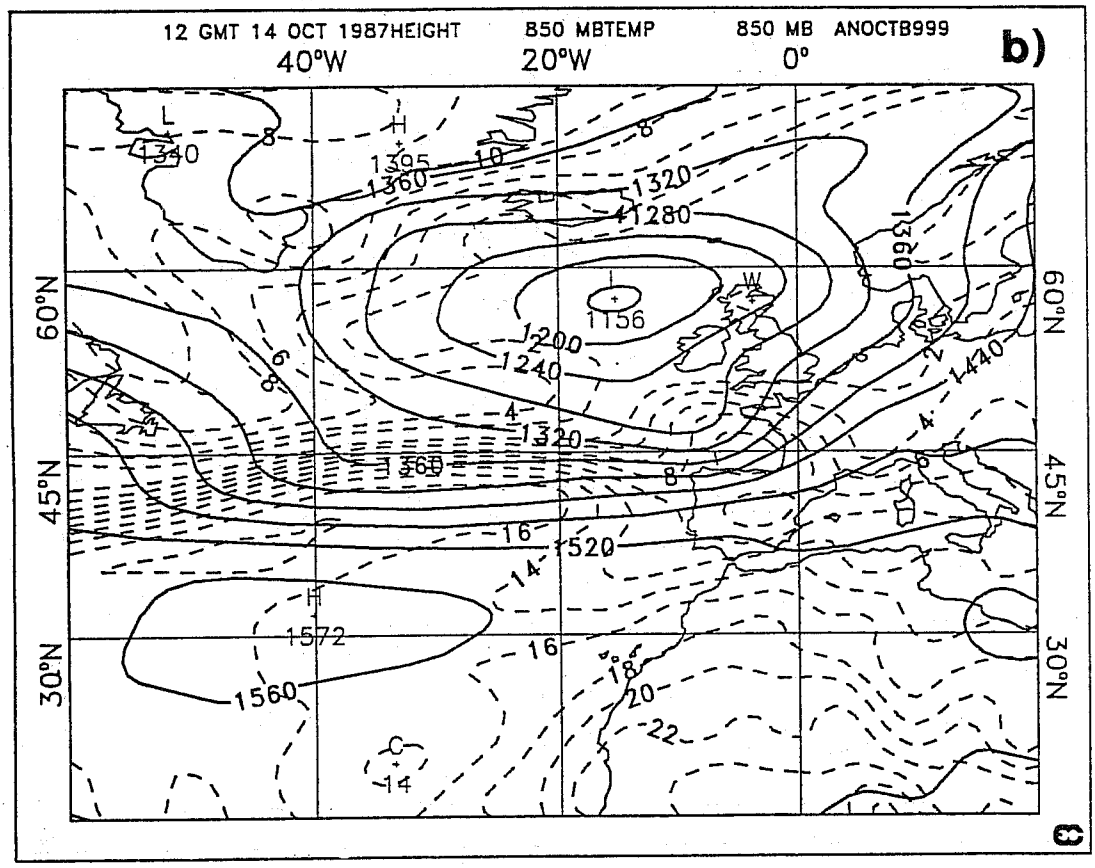


Fig. 14 *Analyses for the height and temperature at 850 hPa, valid for 14/10/87,12 UTC. Panel a) represents the OI T63 initialized analysis, panel b) the 4D-VAR analysis.*

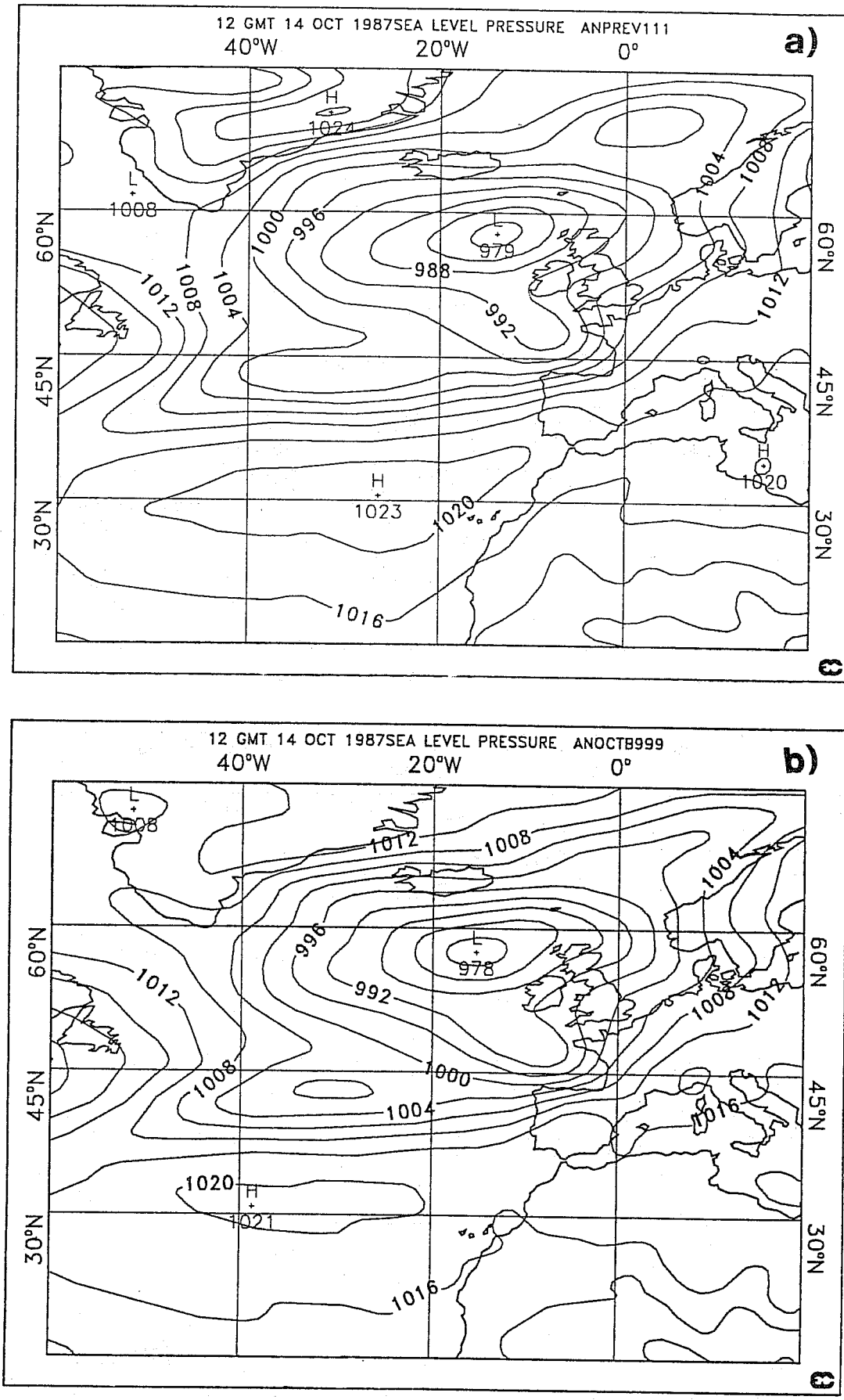


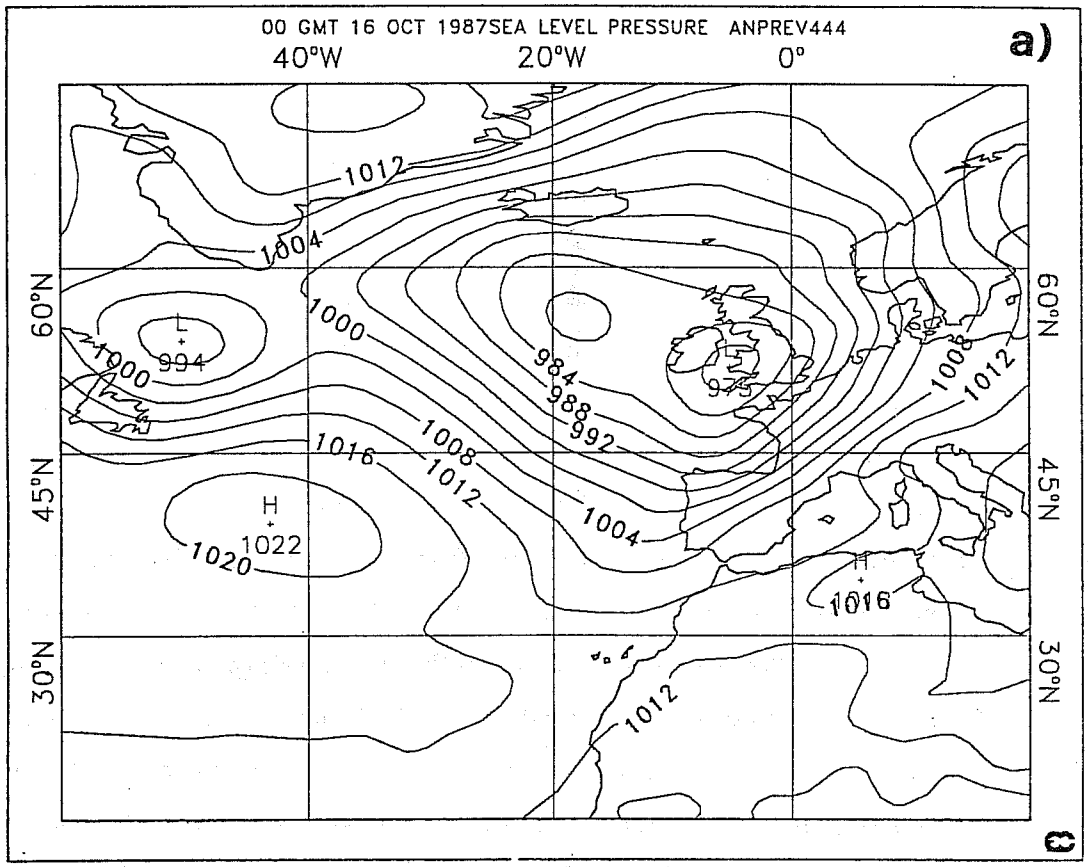
Fig. 15 *Analyses for the sea-level pressure, valid for 14/10/87,12 UTC. Panel a) represents the OI T63 initialized analysis, panel b) the 4D-VAR analysis.*

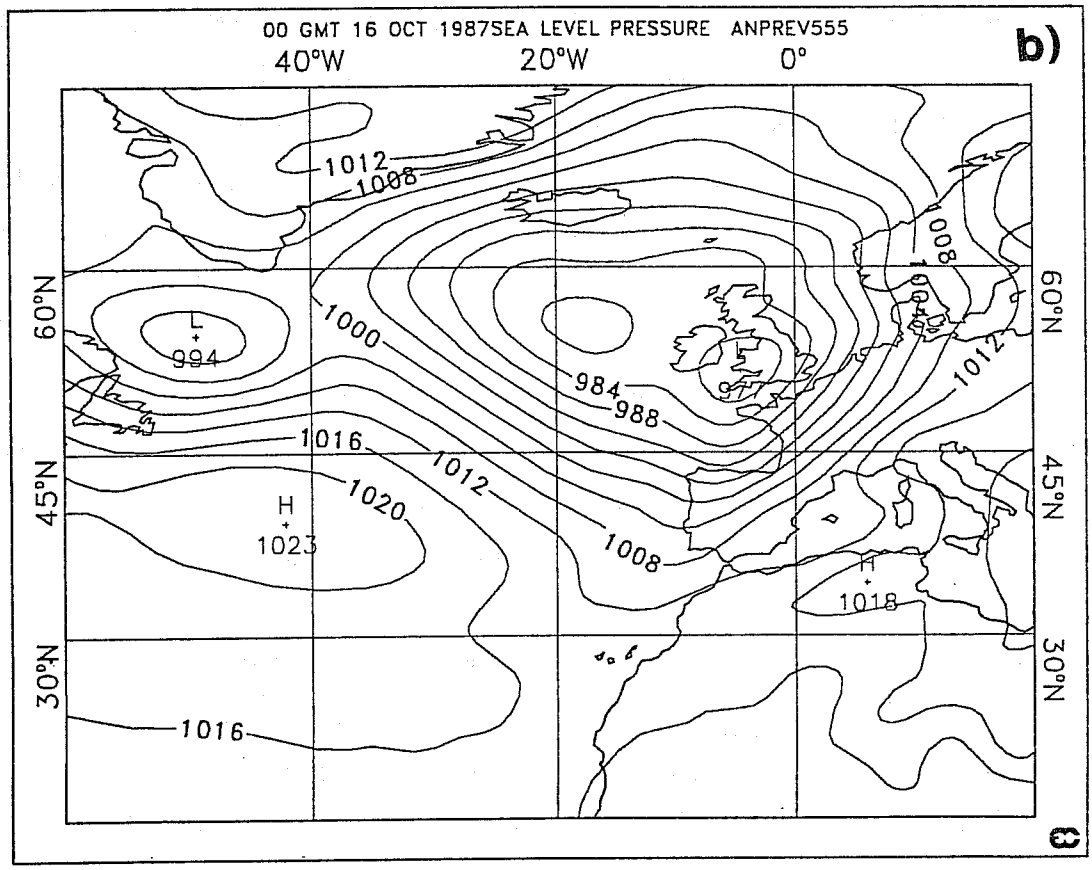
sponding to the upper-level jet, with a maximum of 96 ms^{-1} for OI and 88 ms^{-1} for 4D-VAR. The lower-level baroclinic zone is also present in both analyses, as can be seen from the strong temperature gradient around (45°N;30°W) in Fig. 14. As far as the sea-level pressure is concerned, the analyses are displayed in Fig. 15. The main low is located at (55°N;15°W), with a value of 979 hPa for the OI analysis and 978 hPa for the 4D-VAR analysis. This low pressure area is extended by two troughs, one towards the tip of Brittany, and one towards the mid-Atlantic. The later will be the one responsible for the storm. The corresponding pressure minimum is approximately located at (42°N;30°W) with a typical value of 1000 hPa. What can be said about these two analyses is that, although slightly different, they both exhibit the basic features of the atmospheric flow.

Now, let us compare the 36-hour forecasts started from these analyses with the corresponding analysis coming from the assimilation cycle, valid for the 16/10/87,0 UTC. The sea-level pressure fields are displayed in Fig. 16. The main low (963 hpa) in the OI analysis (panel c)) is located slightly to the South of Cornwall, with a strong pressure gradient over Brittany. In contrast, the lows produced by the 36-hour forecasts are not deep enough and located too far to the North. On the one hand, the value of 975 hPa is slightly better in the case of the forecast started from the OI analysis (panel a)), compared with 977 hPa in panel b). On the other hand, the location is slightly better in the case of 4D-VAR. At this stage, one could think that the introduction of vertical diffusion has been detrimental, but looking at other features (the low and the high in the Western Atlantic), these are now in much better agreement with the analysis than in the purely adiabatic case. The problem is then very probably coming from either the lack of other physical processes or the lack of resolution, which are important for this explosive case, and may not be so important for more standard situations. (Simmons (pers. comm.) improved further the 24-hour forecast of the storm performed with the ECMWF model and a later version of the analysis scheme, from going to a T106L19 resolution to a T213L31 resolution).

As far as the low level height and temperature fields are concerned, Fig. 17 shows the analysed situation (panel c)) and the forecasts. The low in the height field centred at the tip of Cornwall is not very well reproduced by the forecasts, with a slight advantage for the forecast produced from the OI analysis. As for temperature, although the very warm air over Brittany is not forecast properly, the temperature gradient is definitely stronger in the case of 4D-VAR. Looking now at the vorticity fields at 500 hPa displayed in Fig. 18, it is clear that both the intensity and the location of the vorticity maximum are not correct in the forecasts. However, the structure of the field is better in 4D-VAR with two maxima, one West of Brittany and the other one West of Portugal.

It should be pointed out that both the better temperature gradient at 850 hPa and the better


00 GMT 16 OCT 1987SEA LEVEL PRESSURE ANPREV444
a)
40°W
20°W
0°
60°N
45°N
30°N
1012
1004
1000
994
1000
984
988
992
1008
1012
1016
1008
1012
1004
H
1022
1020
H
1016
1012


00 GMT 16 OCT 1987SEA LEVEL PRESSURE ANPREV555
b)
40°W
20°W
0°
60°N
45°N
30°N
1012
1008
1000
994
984
988
1012
1016
1012
1020
H
1023
1008
H
1018
1016

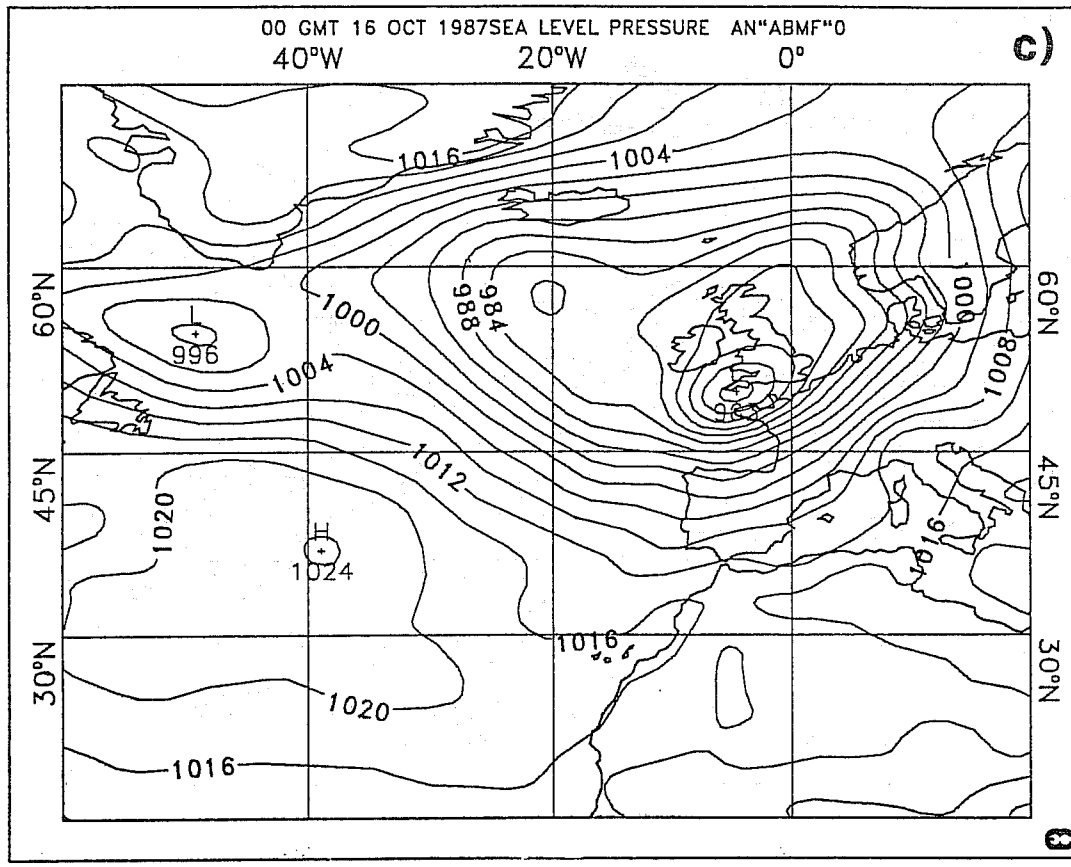
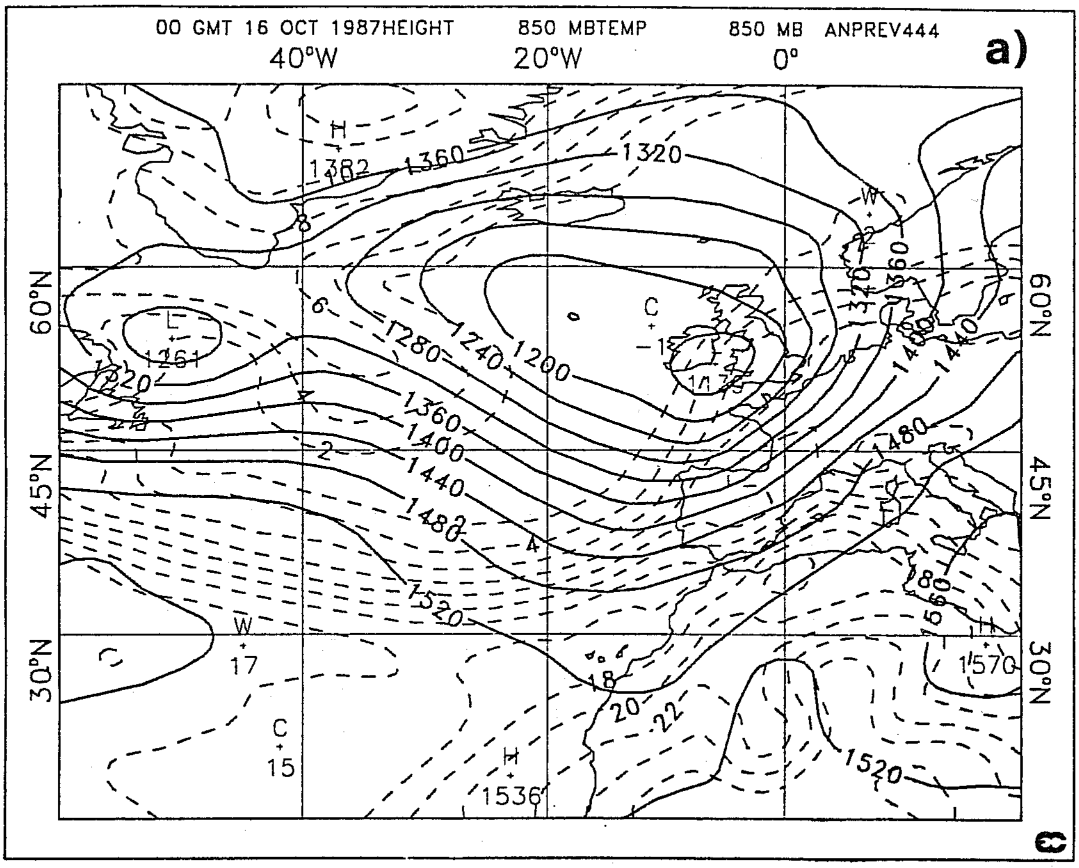
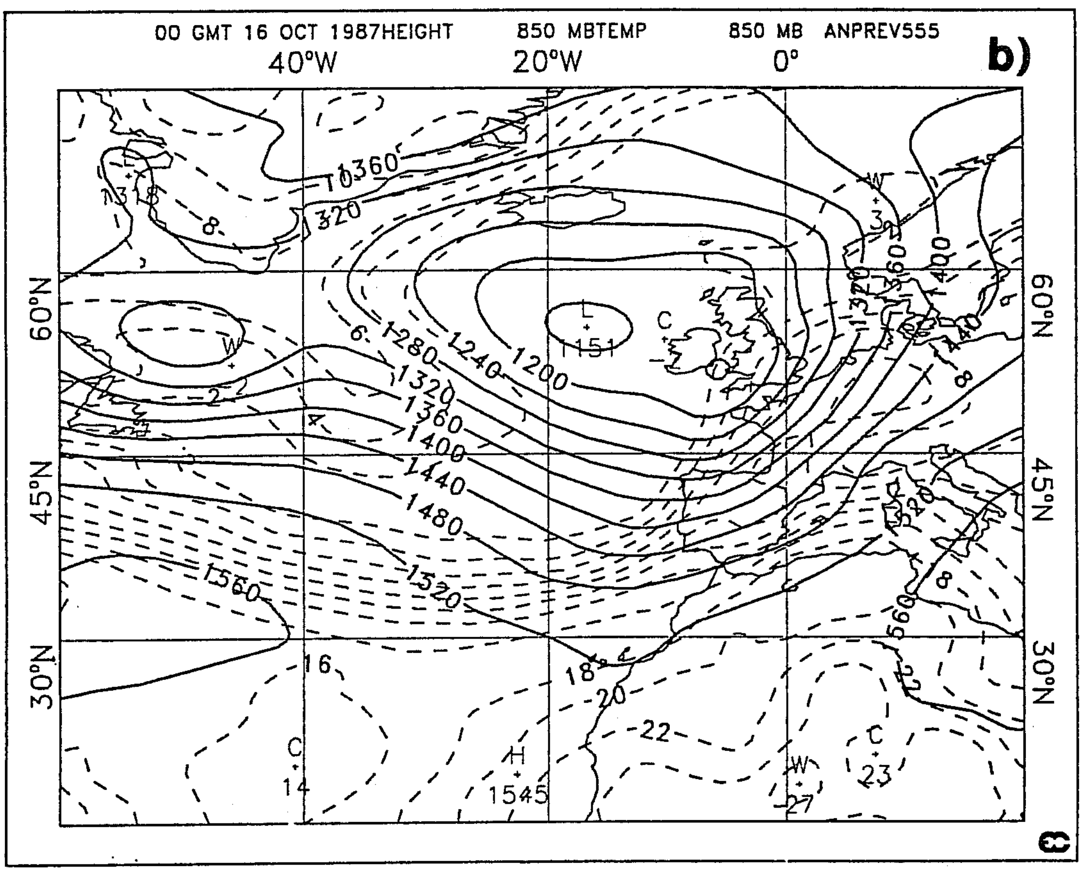


Fig. 16 *36-hour forecasts for the sea-level pressure, valid for 16/10/87,0 UTC, and corresponding T63 analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis and panel c) the corresponding analysis.*


00 GMT 16 OCT 1987HEIGHT 850 MBTEMP 850 MB ANPREV444
40°W
20°W
0°
a)
60°N
45°N
30°N


00 GMT 16 OCT 1987HEIGHT 850 MBTEMP 850 MB ANPREV555
40°W
20°W
0°
b)
60°N
45°N
30°N

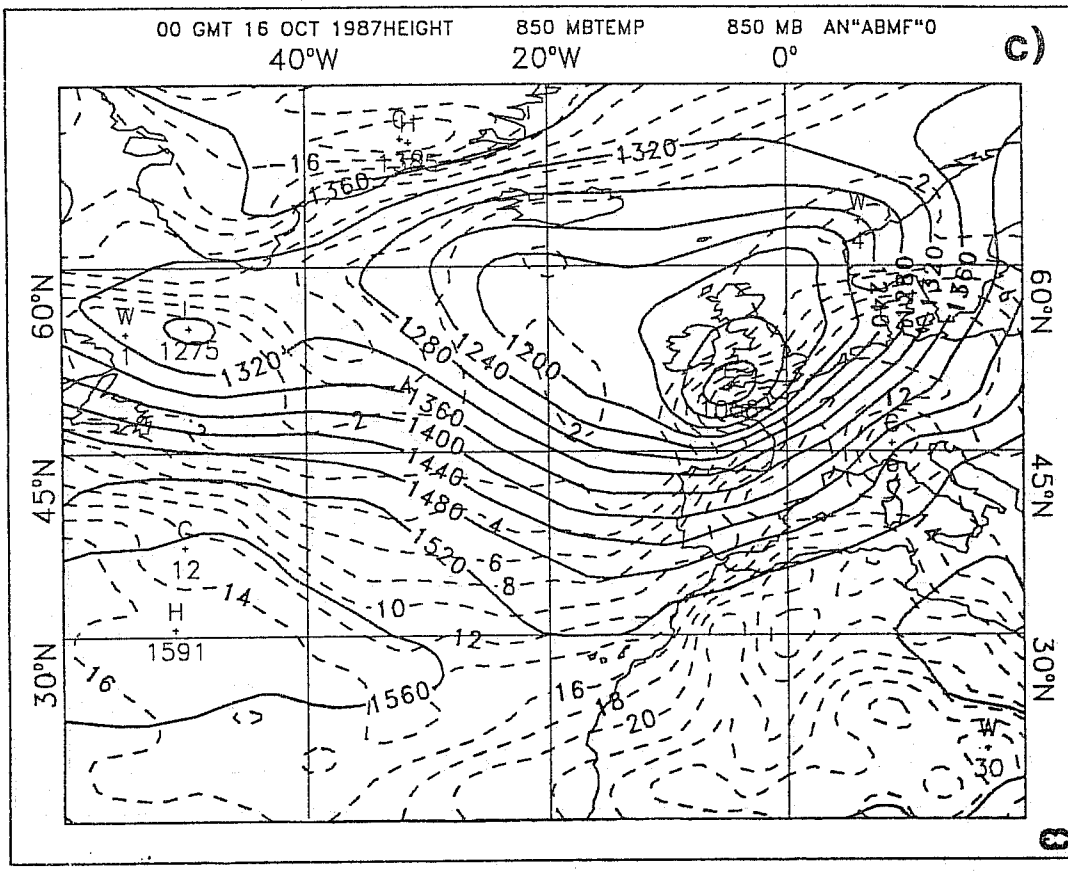
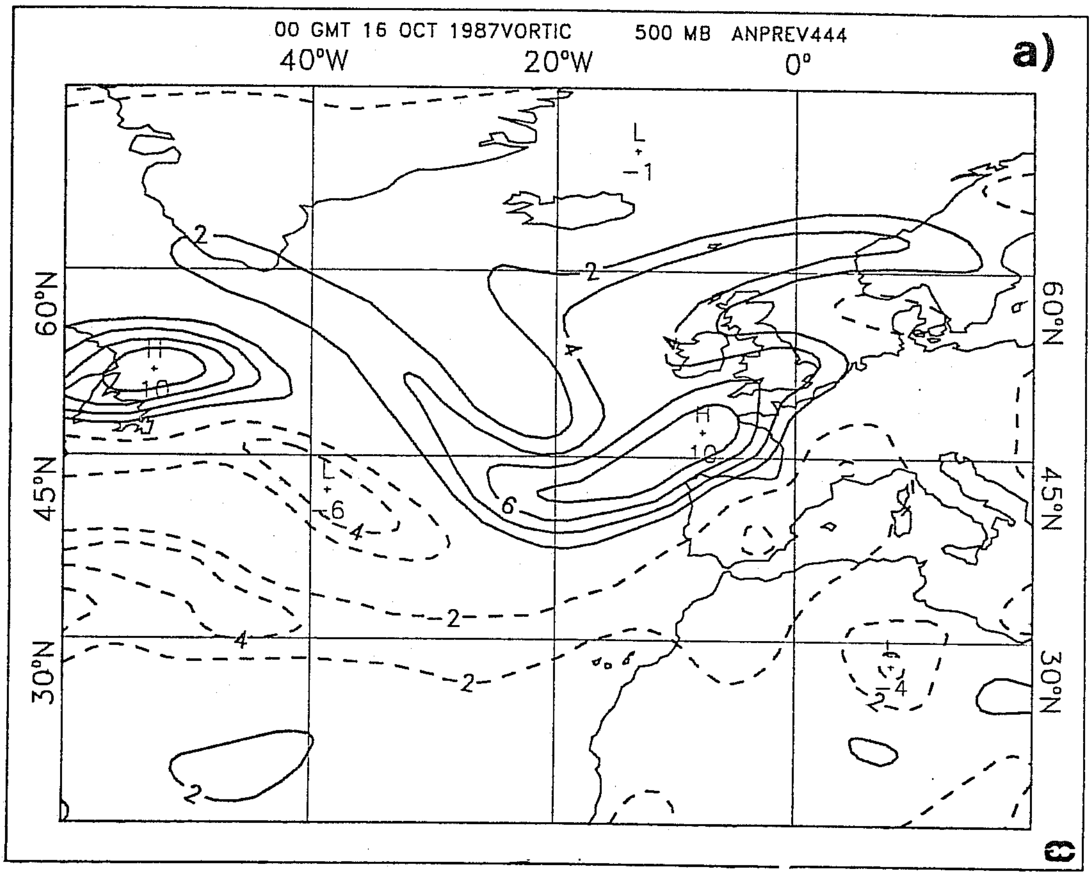
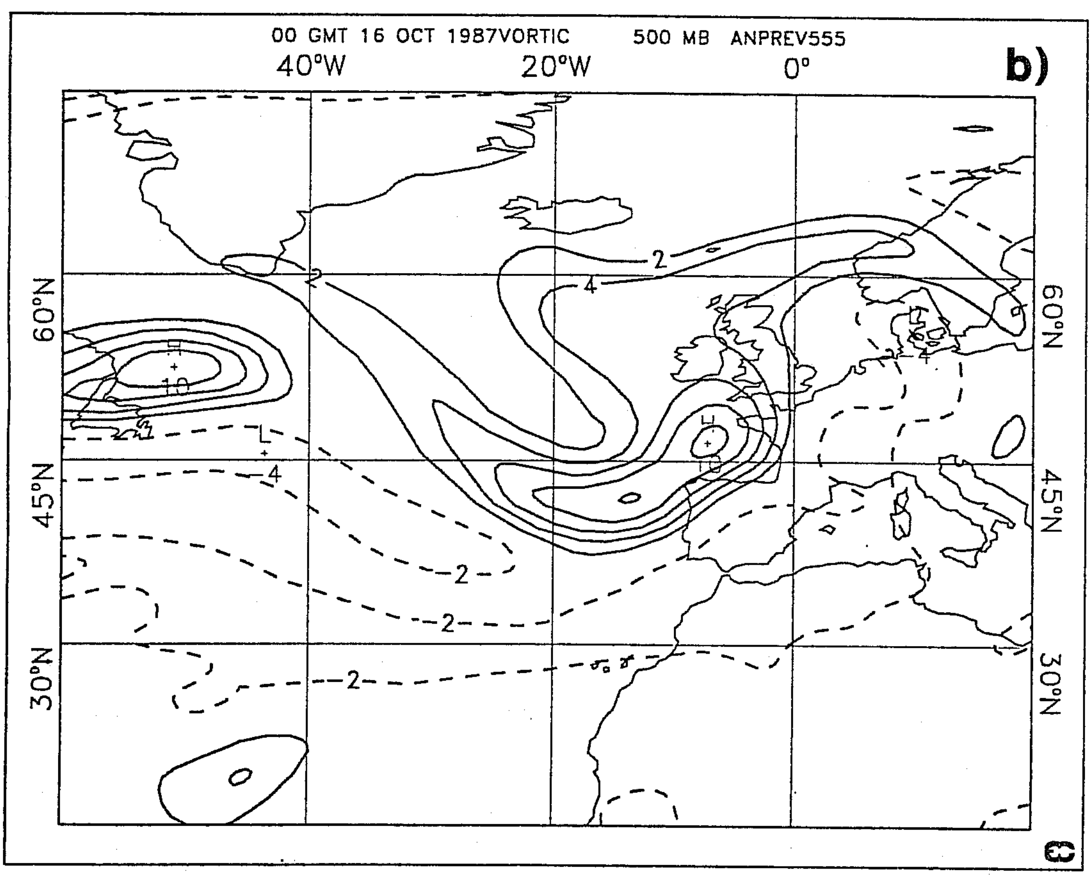


Fig. 17 *36-hour forecasts for the height and temperature at 850 hPa, valid for 16/10/87,0 UTC, and corresponding T63 analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis and panel c) the corresponding analysis.*


00 GMT 16 OCT 1987VORTIC 500 MB ANPREV444
40°W
20°W
0°
a)
60°N
45°N
30°N
L
-1
2
4
6
10
H
L
-6
-4
-2
-4
-2
-2
2


00 GMT 16 OCT 1987VORTIC 500 MB ANPREV555
40°W
20°W
0°
b)
60°N
45°N
30°N
2
4
10
H
L
-4
-2
-2
-2

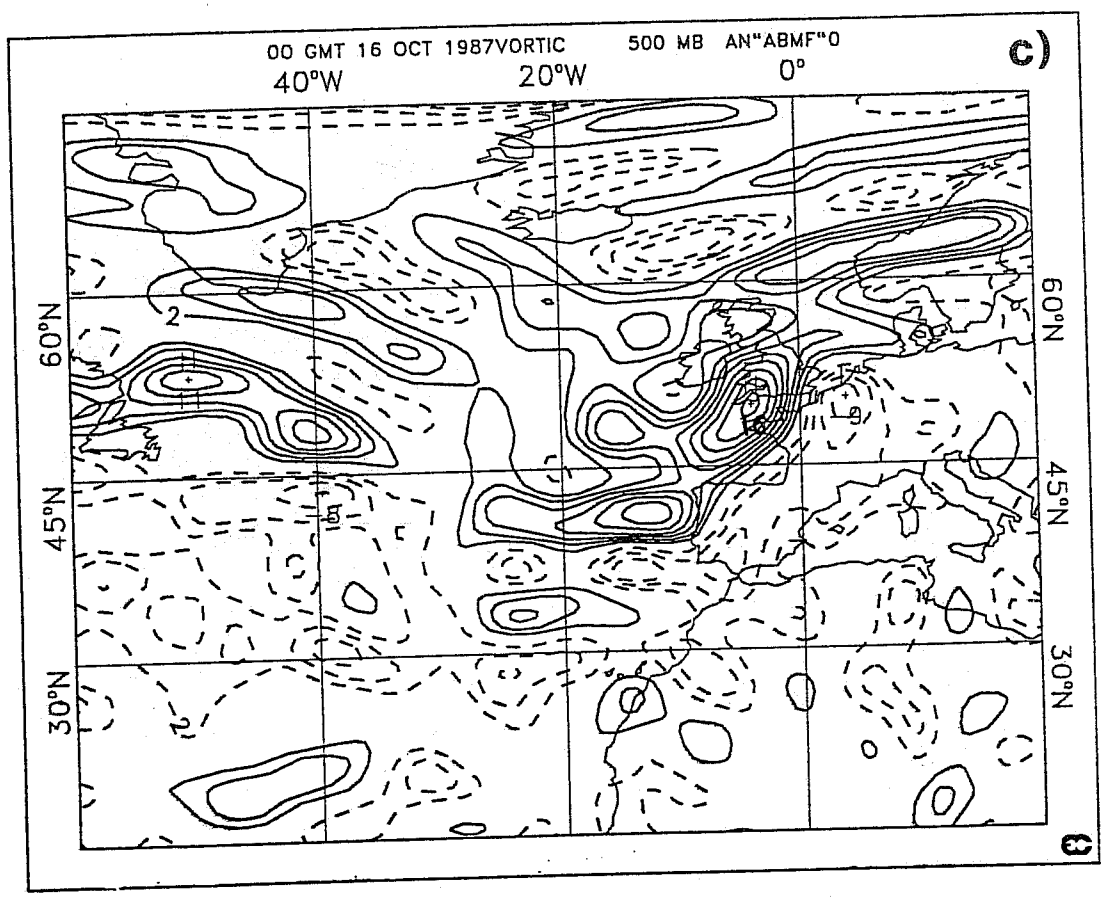
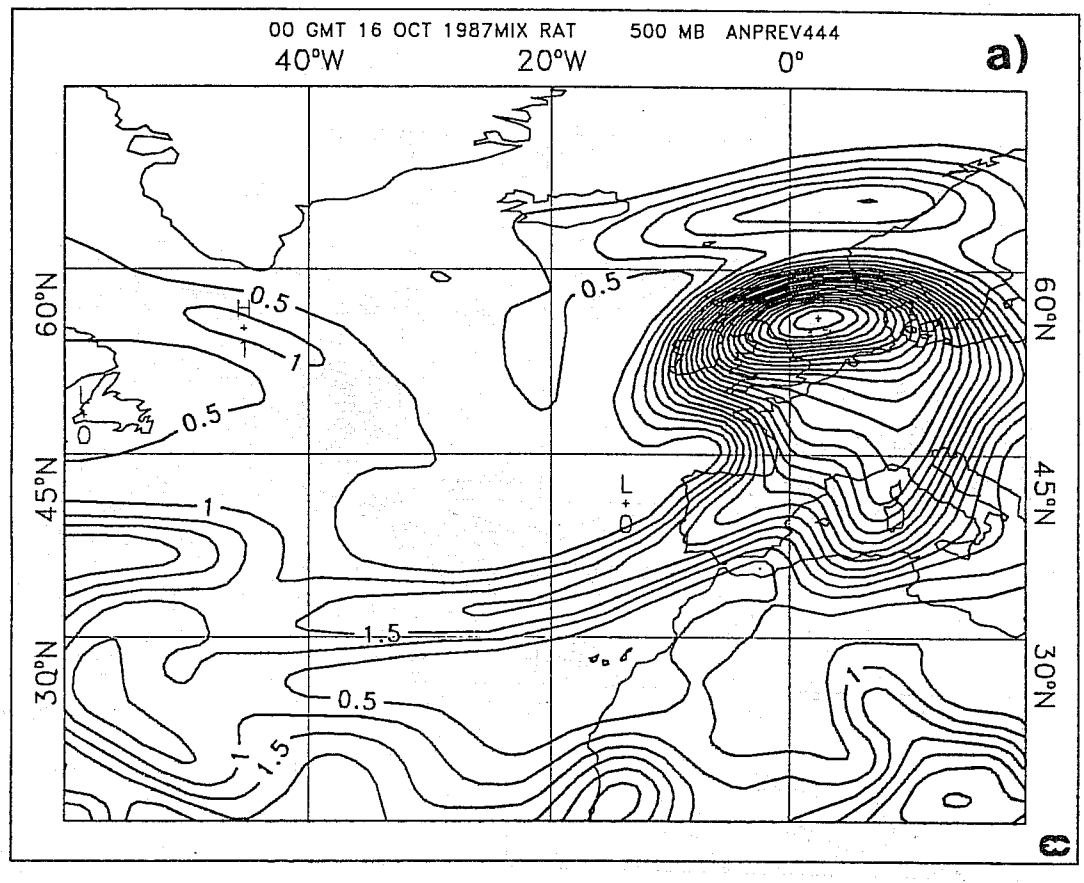
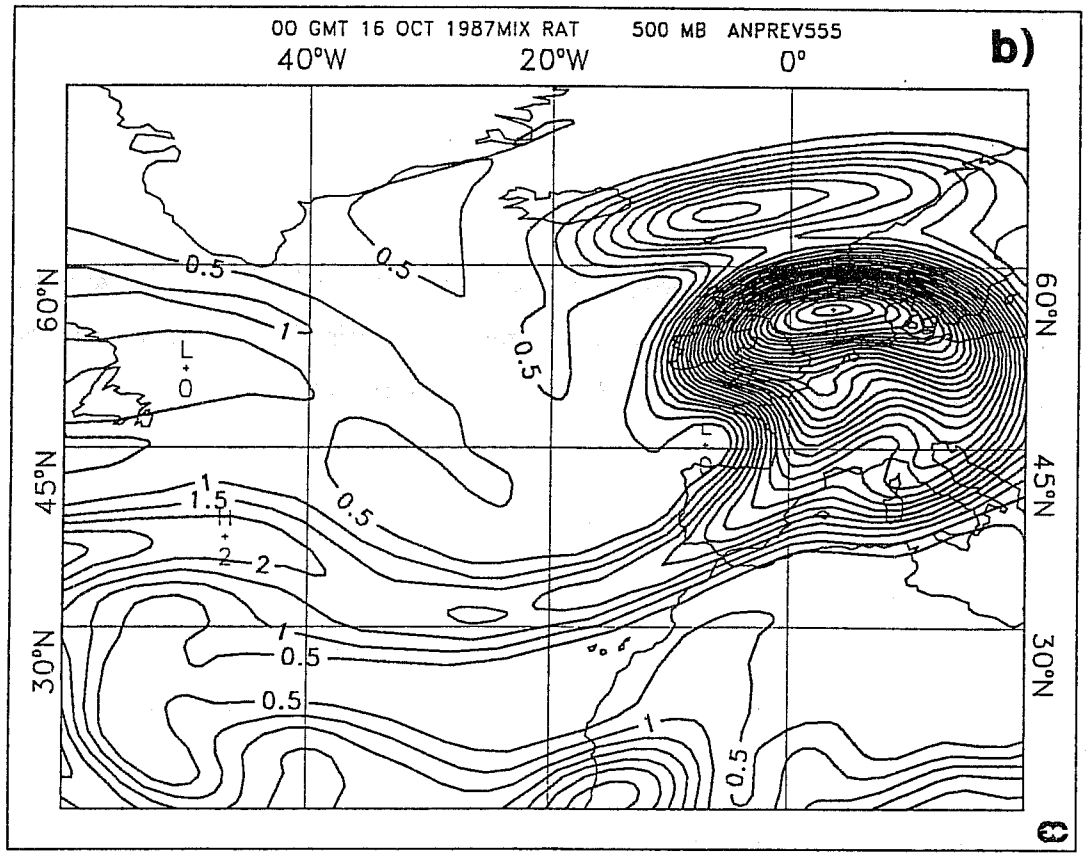


Fig. 18 *36-hour forecasts for the vorticity at 500 hPa, valid for 16/10/87,0 UTC, and corresponding T63 analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis and panel c) the corresponding analysis. The units are* $10^{-5}s^{-1}$.


00 GMT 16 OCT 1987MIX RAT 500 MB ANPREV444
40°W
20°W
0°
a)
60°N
45°N
30°N
0.5
1
1.5
L
0


00 GMT 16 OCT 1987MIX RAT 500 MB ANPREV555
40°W
20°W
0°
b)
60°N
45°N
30°N
0.5
1
1.5
2
L
0

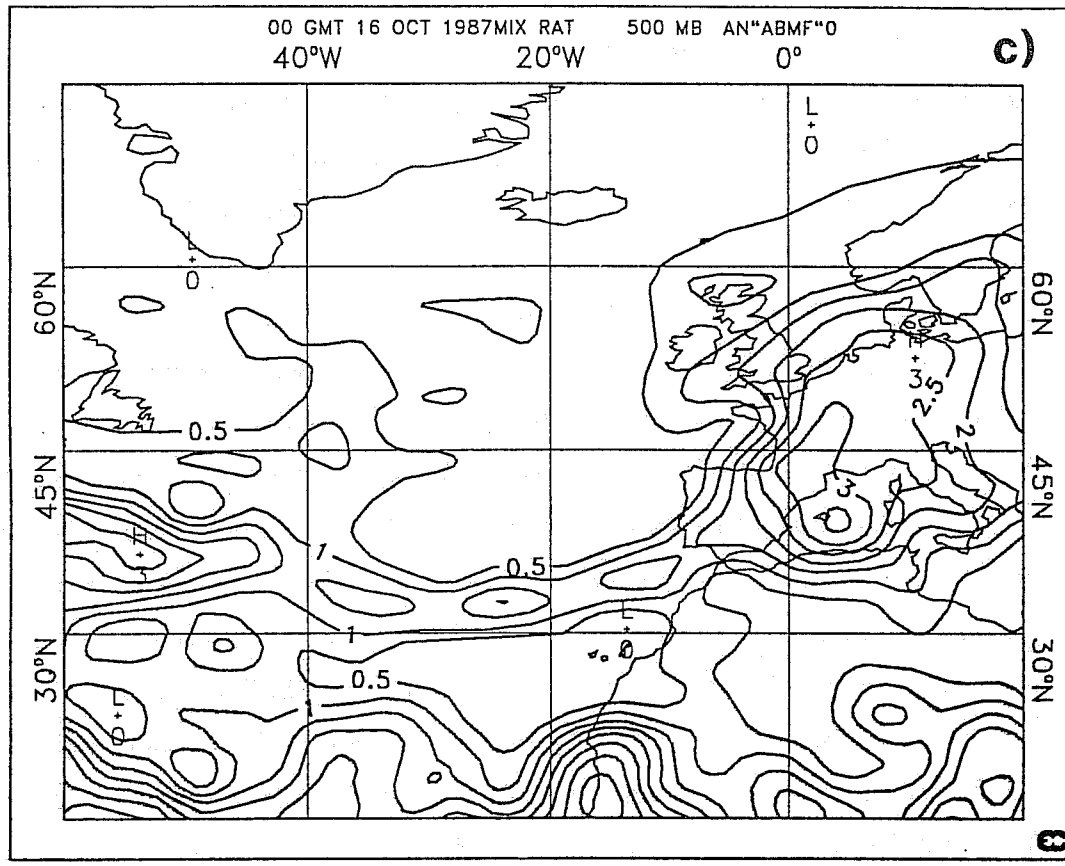


Fig. 19 *36-hour forecasts for the specific humidity at 500 hPa, valid for 16/10/87,0 UTC, and corresponding T63 analysis. Panel a) represents the forecast from the OI analysis, panel b) the forecast from 4D-VAR analysis and panel c) the corresponding analysis. The units are in g/kg.*

structure of the vorticity field at 500 hPa in the case of the forecast from the 4D-VAR analysis, were also present in the previous purely adiabatic case.

A hint that at least the large-scale condensation process is lacking in these experiments is given by Fig. 19, showing the specific hunidity fields at 500 hPa. It is obvious, from the presence of huge unrealistic maxima in the North sea, that both forecasts are oversaturating the atmosphere. It is well-known that latent-heat release is an important part of marine cyclogenesis, and Hoskins and Berrisford (1988) stressed the role of this physical process in this particular situation.

As a preliminary conclusion of this study, it can be said that in this explosive cyclogenesis, the model used both for the assimilation stage and the forecast stage at truncation T63 without physical processes or with only vertical diffusion, is not adapted to analyse and forecast such an intense meteorological situation. It is clear that a disadvantage was given to 4D-VAR for two reasons. The first one is that the weight given to the first-guess was probably overestimated. The second one is that, for the satellite observations, we did not try to incorporate directly in the variational scheme the radiances (which were not directly available for this period) but used the retrieved SATEMs profiles. This study could be extended in several ways. One could imagine using more physical processes together with their adjoints in the assimilation stage, and of course in the forecasts. Another promising approach is to use the full model (high resolution, sophisticated physics) for the trajectory and use a simplified model for solving a minimization problem in the vicinity of this trajectory.

7 DISCUSSION

Apart from its conceptual simplicity and its adaptability, four-dimensional variational assimilation possesses many advantages. In particular, it can be shown to be equivalent to the optimal Kalman Filter in the linear context and under the hypothesis of a perfect model. In that case, its main advantage upon the Kalman Filter is its lower cost. Moreover, in nonlinear cases, it has been shown to be able to take advantage of dynamical nonlinearities (Rabier and Courtier, 1992; Lorenc, 1988a). However it is still rather expensive and extensive testing has to be carried out before its possible operational implementation (Courtier and Thépaut, same volume).

It is of primary importance to both analyse and quantify the potential benefit of using such a method relative to more classical ones such as a simplified sequential algorithm.

The main part of this study was intended at providing a preliminary answer to this question and we deliberately investigated the problem in a simplified context in order to facilitate the interpretation of the results. Firstly, the meteorological situation chosen as the framework for comparison is a baroclinic instability case typical of mid-latitude dynamics where no physical processes are included, and the

knowledge of the such-defined truth is available at any time. Secondly, the comparison was performed within the conditions of validity of the tangent-linear hypothesis, for which results for the analysis given at the end of the assimilation period can be expected to be very similar between 4D-VAR and an EKF. The basic principle underlying the set-up of the experiments is to give exactly the same information to both algorithms on the 24-hour period under consideration. Not only the observations are the same, but also 4D-VAR is given the same background information as the one needed to initialize the sequential algorithm, i.e. a background at the beginning of the assimilation period and its associated error covariance matrix. The major weakness of current assimilation systems being precisely the approximate definition of forecast error covariance matrices, the same initial handicap affects both algorithms.

The major global conclusion which can be drawn from the results is the significantly better performance of the 4D variational scheme, as it gives consistently better analyses than its sequential counterpart and as the error grows in the subsequent forecasts at a lesser rate.

In the context of perfect data, the propagation of information due to the coherent treatment of the dynamics is clearly visible in an experiment where additional reports at 250 hPa over one latitude row are found to affect all variables and a large geographical area in a manner perfectly consistent with the dynamics.

When the differential updating of forecast error variances at each analysis step due to different data coverage is inserted in a crude manner in the sequential algorithm, it leads to a spatial homogeneity of errors comparable to that of the 4D variational scheme.

The prevailing strength of 4D-VAR is that these two types of information (reduction of forecast error by previous insertion of data or complicated structure functions depending on the dynamics) do not have to be explicitely prescribed, with necessary assumptions: they are automatically provided by the algorithm itself.

In a first step to test the robustness of both methods to various specifications of their implementation, 4D-VAR proves to be much less sensitive to the way the slow manifold constraint is applied. But both algorithms are rather sensitive to the specification of the forecast error covariance matrices, which means that an accurate definition of those should not be neglected, even in the four-dimensional scheme. In any case, whatever configuration was chosen for implementation of the experiments, the quality of 4D-VAR analyses and forecasts was always higher than that of the results of the sequential assimilation scheme, especially in the Southern hemisphere.

The second part of this paper was dealing with a comparison between 4D-VAR and a simplified operational assimilation scheme, for the October 87' Storm over France and England. The model

used was a T63L19 model with very simplified physics (only vertical diffusion). We performed two 36-hour forecasts from 4D-VAR analysis and Optimal Interpolation Analysis, but the impact of using a more elaborate assimilation scheme was hardly noticeable. For this explosive cyclogenesis, a more sophisticated model would be needed in order to be able to obtain significant results. In any case, this situation should be used as a test-case to evaluate the improvements brought by refinements in the four-dimensional assimilation system (introduction of physics, increase in resolution). It is essential to investigate further this four-dimensional variational scheme on such extreme events.

8 ACKNOWLEDGEMENTS

The authors would like to acknowledge all the ARPEGE/IFS team, for their constant support. The study was partially supported by the Groupement de Recherche Méthodes Variationnelles en Météorologie et Océanographie, funded by Centre National de la Recherche Scientifique and Météo-France.

9 REFERENCES

Cohn, S.E. and D.F. Parrish, 1991. The behavior of forecast error covariances for a Kalman Filter in two dimensions. *Mon. Wea. Rev.*, **119**, 1757-1785.

Courtier, P., and O. Talagrand, 1990. Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42A**, 531-549.

Courtier, P., C. Freydier, J-F Geleyn, F. Rabier and M. Rochas 1991. the ARPEGE Project at Météo-France. *ECMWF Seminar proceedings, September 1991.*

Courtier, P. and J.-N. Thépaut, 1992. A strategy for operational implementation of 4D-VAR. *Proceedings of ECMWF Workshop on variational assimilation, with special emphasis on three-dimensional aspects, 9-12 November 1992.*

Daley, R., 1991. Atmospheric data analysis. *Cambridge Atmospheric and Space Science series*, (eds Cambridge University Press).

Daley, R., 1992. The lagged innovation covariance: a performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.*, **120**, 178-196.

Dee, D. P., 1991. Simplification of the Kalman filter for meteorological data assimilation. *Q. J. R. Meteorol. Soc.*, **117**, 365-384.

Derber,J. C., 1989. A variational continuous assimilation technique. *Mon. Wea. Rev*, **117**, 2437-2446.

ECMWF Research Manual 1, 1992: ECMWF Data Assimilation scientific documentation

Gauthier, P., P. Courtier and P. Moll, 1992. Data assimilation with an Extended Kalman-Bucy Filter. *submitted to Mon. Wea. Rev.*.

Gelb, A., 1974. Applied optimal estimation. *Cambridge: MIT press.*

Ghil, M., S. Cohn, J. Tavantzis, K. Bube and E. Isaacson, 1981. Applications of Estimation Theory to Numerical Weather Prediction. In : *Dynamic meteorology. Data assimilation methods* (eds. L. Bengtsson, M. Ghil and E. Källén). Springer-Verlag, New-York, 139-224.

Ghil, M. and P. Malanotte-Rizzoli, 1991. Data assimilation in Meteorology and Oceanography. *Adv. in Geophys.*,**33**, 141-266.

Gilbert, J-C. and C. Lemaréchal, 1989. Some numerical experiments with variable storage quasi-Newton algorithms. *Mathematical Programming*, **B 25**, 407-435.

Heckley, W., P. Courtier, J. Pailleux and E. Andersson, 1992. On the use of first-guess information on the variational analysis at ECMWF. *Proceedings of ECMWF Workshop on variational assimilation, with special emphasis on three-dimansional aspects, 9-12 November 1992.*

Hollingsworth, A., D. Shaw, P. Lönnberg, L. Illari, K. Arpe and A. Simmons, 1986. Monitoring of observation and analysis quality in a data assimilation system. *Mon. Wea. Rev.*, **114**, 871-879.

Hollingworth, A. and P. Lönnberg, 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I : The wind field. *Tellus*, **38A**, 111-136.

Hoskins, B., and P. Berrisford, 1988. A potential vorticity perspective of the storm of 15-16 October 1987. *Weather*, **43,3**, 122-129.

Jarraud, M., Goas, J. and C. Deyts, 1989. Prediction of an exceptional storm over France and Southern England. *Wea. For.*, **4**, 517-536.

Jazwinski, A. H., 1970. Stochastic processes and filtering theory. Academic Press, New-York.

Lacarra, J.F., and O. Talagrand, 1988. Short range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81-95.

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

Lions, J., 1971. Optimal control of systems governed by partial differential equations. *Berlin: Springer-Verlag.*

Lönnberg, P. and A. Hollingsworth, 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part II : the covariance of height and wind error. *Tellus*, **38A**, 137-161.

Lorenc, A.C., 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177-1194.

Lorenc, A.C., 1988a. Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.*, **114**, 205-240.

Lorenc, A.C., 1988b. A practical approximation to optimal four-dimensional objective analysis. *Mon. Wea. Rev.*, **116**, 730-745.

Lorenc, A. C. and O. Hammon, 1988. Objective quality control of observations using Bayesian methods. *Q. J. R. Meteorol. Soc.*, **114**, 515-543.

Lorenc, A.C., R.S. Bell, T.Davies and G.J.Shutts, 1988. Numerical forecast studies of the October 1987 storm over southern England. *Meteorol. Magazine*, **117**, 118-130.

Machenhauer, B., 1977. On the dynamics of gravity oscillations in a shallow-water model with applications to normal mode initialization. *Beitr. Phys. Atmos.*, **50**, 253-271.

Mitchell, H. L., C. Charette, C. Chouinard and B. Brasnett, 1990. Revised interpolation statistics for the Canadian data assimilation procedure : Their derivation and application. *Mon. Wea. Rev.*, **118**, 1591-1614.

Morris, R. M. and A. J. Gadd, 1988. Forecasting the storm of 15-16 October 1987. *Weather*, **43-3**, 70-90.

Navon, I. M. and X. Zou, 1991. Variational data assimilation with the direct and adjoint primitive equations and normal mode initialization. *Preprints of the Ninth Conference on Numerical Weather Prediction, October 14-18,1991. Denver, Colorado*, **13B.4**, 455-458.

Pailleux, J., W. Heckley, D. Vasiljevic, J-N. Thépaut, F. Rabier, C. Cardinali and E. Andersson, 1991. Development of a variational assimilation system. *ECMWF Research Department Technical Memorandum 179.*

Parrish D. P. and J. C. Derber, 1992. The National Meteorological Center's Spectral Statistical-Interpolation Analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Rabier, F. and P. Courtier, 1992. Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.*, **118**, 649-672.

Rabier, F., P. Courtier and O. Talagrand, 1992. An application of adjoint models to sensitivity analysis. *Beitr. Phys. Atmos.*, in press.

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

Simmons, A., and D.Burridge, 1981. An energy and angular momentum conserving vertical finite difference scheme and hybrid vertical coordinate.*Mon. Wea.Rev.*, **109** 758-766.

Thépaut J.-N. and P.Courtier, 1991. Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.*, **117**, pp 1225-1254.

Thépaut J.-N., D. Vasiljevic, P. Courtier and J. Pailleux, 1992. Variational assimilation of meteorological observations with a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.*, **119**, pp 153-186.

Thorncroft, C.D. and B.J. Hoskins, 1990.Frontal cyclogenesis. *J. Atmos. Sci.*, **47** 2317-2336.

Todling, R., 1992. The Kalman Filter for two-dimensional stable and unstable atmospheres. *Ph.D. dissertation, U.C.L.A.*.

Vukicevic, T., 1991. Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.*, **119** 1602-1611.

Zou, X. and I. M. Navon, 1991. On the effect of gravity

oscillations, boundary control and incomplete observations in a shallow water equations model with applications to variational data assimilation. *Preprints of the Ninth Conference on Numerical Weather Prediction, October 14-18,1991. Denver, Colorado*, **13B.4**, 459-462.

Zupanski, M., 1992. Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Submitted to Mon. Wea. Rev.*.