TECHNICAL REPORT No. 6

THE ECMWF ANALYSIS AND DATA - ASSIMILATION SCHEME: ANALYSIS OF MASS AND WIND FIELDS

by

A. Lorenc, I. Rutherford and G. Larsen

December 1977
First reprint June 1980

CONTENTS PAGE NUMBER

A B S T R A C T

This report sets out the overall plan for ECMWF's analysis and data assimilation scheme, and describes in detail the analysis of the mass and wind fields.

Also a first test analysis of mass and wind using a climatological first guess is presented.

1. Introduction

The main objectives of the analysis and data-assimilation scheme which is being developed at ECMWF are to provide initial states for the Centre's operational forecast model, and to produce analyses from the observations made during the First GARP Global Experiment (FGGE), for use by the international scientific community.

The scheme must thus produce global analyses in numerical form using all types of available observations, including those from the satellite observing systems in use and being developed. Furthermore it has to run efficiently, with minimal human intervention, on a large fast vector-processing computer. Finally it should be flexible, so that new types of observations can be used, and new techniques of analysis tried, while retaining the possibility of reverting to the older, tested methods if the new ones prove inferior.

Such a scheme is currently being developed at ECMWF, and the purpose of this report is to describe the part of the system which is most advanced: the analysis of the dynamic variables (mass and wind fields). However, in order to place this and subsequent reports in context, we commence in the following section by briefly describing the entire scheme as currently envisaged.

2. Envisaged overall structure of the assimilation scheme

The scheme has two major stages:- The analysis of the statistical corrections to a predicted field as implied by observations and various empirical or dynamical relationships such as geostrophy, - and the subsequent assimilation of these corrections into a prediction model in order to produce the next predicted field. A six hour time period will probably be used for the analysis forecast cycle; longer would make less useful the observations not made at the main synoptic hours, 0 and 12 GMT, while a shorter period might make the scheme undesirably dependent on the behaviour of the particular numerical prediction model used for the assimilation.

The prediction model serves the purpose of carrying information forward in time and space to be used in regions where the coverage of observations at a particular time is not sufficient to define the atmospheric state. In this way it assimilates the observations distributed in space and time into a dynamically consistent picture of the atmosphere, and through its simulation of the atmospheric processes it may induce reasonably accurate values for parameters which are not or only poorly observed, such as vertical velocity and relative humidity.

It is presently planned that a simplified version of the Centre's medium-range forecast model, a global, sigma co-ordinate primitive equation model, will be used as assimilation model (Burridge and Haseler, 1977). In the analysis (interpolation) stage meteorologically consistent corrections to the predicted fields are calculated. If gross relationships such as geostrophy are satisfied by the analysed increments then assimilation of the data by the prediction model and induction of unobserved parameters is improved.

It is convenient for the analysis to be done using pressure co-ordinates, since most of the observations, and most of the analyses are represented in these co-ordinates. However this means that grid transformations to and from the prediction model grid are necessary. We intend to lessen the smoothing and unbalancing effects of these by interpolating, where possible, only the corrections to a field, rather than the complete field (Rutherford, 1977a). A high vertical and horizontal resolution will furthermore reduce such effects.

To maintain flexibility, for example the ability to change to a different prediction model, the scheme has been designed with a modular structure. Largely independent computer programs (or groups of programs) perform separate functions, with data files acting as the interface between them. Similarly each program is split into subroutines, with common blocks acting as interfaces. To help standardize programming techniques the Olympus system is used (Roberts, 1974).

The relationships between the major functions and files are illustrated in Figure 1. Shown dashed is a possible link between the analysed fields and the subsequent predicted fields in pressure co-ordinates, necessary if the σ to p grid transformation is of the forecast changes only, as proposed by Rutherford (1977a).

Another link which should be implemented in time is between the estimated analysis errors and the prediction error statistics. Once the error growth characteristics of the prediction model are known, forecasted values can be used instead of seasonal averages.

3. General analysis methodology

3.1 Statistical interpolation of deviations between observed and predicted values

The analyses method is an extension of optimum interpolation (Gandin, 1963) to a multivariate three-dimensional form.

This technique enables consistent use to be made of observations with different error characteristics, and takes into account their spatial distribution. The equations used are set out in Section 4.

Because of the various assumptions made in using linear regression and error covariance modelling the interpolation is not truly optimal and the name "Statistical interpolation" is preferred.

3.2 Grid independent global analysis structure

The interpolation equations and the error covariance modelling are expressed in latitude-longitude co-ordinates and also do not depend intrinsically on any assumptions of limited global applicability, such as the geostrophic relationship; the scheme can therefore be used to produce global analyses. (However a geostrophic type relationship is assumed in mid and high latitudes; see Section 5).

The program organisation, data selection and the bulk of the computation are furthermore independent of the analysis grid. At present any global or limited-area regular latitude-longitude grid may be used; to modify the program for other grids is very simple. We intend to run the program using a horizontal grid similar to that of the prediction model used.

At present the program is restricted by the vertical covariance statistics (Section 5) to analysis on predefined pressure levels. Since this is where the vertical covariance statistics in general will be available, and since the observational data either is or can be assigned to such levels - provided they possess a reasonably high vertical resolution - it is not envisaged to relax this restriction in the future.

3.3 Multiple use of matrix inverses

Statistical interpolation requires more computation than many other interpolation methods, since the solution of a set of simultaneous equations is necessary for each value interpolated, the order of the equations being equal to the number of observational data used. Previous schemes have kept this number low (about 15) by careful selection. This scheme takes advantage of the fact that vector-processing computers can solve larger systems without greatly increasing the computation time. Thus less time in the program need be spent selecting observations since most observations near the area being analysed are used. Such a simple selection scheme has the added advantage that the choice is much less dependent on the precise position being analysed, and the same selection of observations may be

used for all analysis points within an area. Since only the right-hand sides of the simultaneous equations depend on the analysis point, the left-hand side covariance matrix can be inverted once and for all, leading to a large saving in computation. This furthermore makes the computation time much less dependent on the analysis grid density.

3.4 Data checking by analysis

Statistical interpolation can be used to obtain both the interpolated value as well as the estimated interpolation error at each observation point, and hence to detect those observations which do not agree within a reasonable tolerance with the analysed values obtained without them. If slightly modified interpolation equations are used (Section 4.3) it is possible to do this using the same matrix inverses as are needed for the grid-point analysis.

The analysis program is used in a preliminary run to perform this data checking, and to write a temporary file containing the matrix inverses. The grid-point analysis is then done re-using these matrix inverses. Slightly modified equations (Section 4.4) are used in order to omit data which failed the check.

3.5 Dynamical relationships imposed by multivariate interpolation

Statistical interpolation provides a convenient framework for imposing simple dynamical relationships on the interpolated increments to the prediction. By using a particular relationship to derive the error covariances between different variables the scheme forces the analysed corrections to obey the relationship. Schlatter (1975) and Rutherford (1976) have used the geostrophic relationship in this way, Rutherford (1976) the hydrostatic relationship, and Schlatter et al. (1976) the streamfunction-wind relationship (in the tropics). It should be noted that it is only the corrections which are constrained. If the predicted fields to which they are added do not obey the relationships then nor will the analysis. Also they apply exactly only while the selection of observations used remains the same (i.e. within the volume of applicability of one of the matrices discussed in Section 3.3).

The relationships used in the first version of this scheme are described fully in Section 5. They imply non-divergent, approximately geostrophic, constraints on wind and height increments with the geostrophic relationship relaxed near the equator. The hydrostatic relationship is used in the pre-analysis to convert temperature observations to thicknesses so it is not needed in the actual analysis.

3.6 Three-dimensional use of data

Radiosondes produce relatively accurate vertically consistent observations of height and wind at each level of the atmosphere. When producing an analysis at one level, only the data for that level from each radiosonde need be used; data from other levels does not significantly improve the analysis. Other types of observation need to be used three-dimensionally however. Single level observations must be used at all near-by analysis levels to maintain vertical consistency in the analysed fields. Satellite temperature soundings provide thickness data whose usefulness is enhanced by reference-level data at any level of the atmosphere. For example only by three-dimensional use of the data can optimum use be made of a surface pressure observation, a set of satellite temperature soundings, and a high level wind. The thickness and thickness gradient (thermal wind) information in the soundings increases the region of effectiveness of the pressure and wind data.

Because of this the analysis scheme generally uses the data from all observations except radiosondes at all levels. In data sparse regions where there are only few radiosonde soundings all levels from these are also used, and the same matrix inverse is used to produce analyses at all levels as explained in Section 3.3. In regions with many radiosondes the analysis proceeds level by level, using the radiosonde data for that level only.

3.7 Data selection by boxes

Because the number of data chosen for one analysis matrix can be quite large, and since the same choice of data is to be used for both height and wind analyses over quite a large analysis volume, the selection algorithm used does not evaluate and select observations individually. Instead the globe is divided conceptually into boxes of such a size that when analysing for a box in a typical area of relatively high observation density only data from the nearest neighbouring boxes are needed. This box size has tentatively been set to ($\sim 6^{\circ} \times 6^{\circ}$) latitude (Figure 9).

A preanalysis program reads the observations box by box, and detects situations of high observational redundancy. Such groups of close data values are then compressed into "super observations" of improved accuracy.

Then the preanalysis program selects sufficient "primary" observations to represent the box, when analysing each of its neighbouring boxes. Primary observations are defined as observations needed for the analysis of points in neighbouring boxes.

The selection is thus dependent on the amount and types of data available, on proximity to neighbouring box, and on the geographical position of the analysis area. In this way it is possible to select and flag observations in a neighbouring box, effectively setting up a minimum influence distance within which all data is used. This ensures the necessary overlapping of data to avoid discontinuities at box boundaries. Any remaining "secondary" observations are used only to define the small scales when analysing the box itself.

The analysis program treats the analysis points within one box at a time, using its primary and secondary observations plus the relevant primary observations from its neighbouring boxes. If this is not sufficient then observations from the neighbours' neighbours are used, and so on. In this way data selection is extended in data sparse directions, when needed.

4. Analysis equations

Three different applications of the well known statistical (optimum) interpolation technique are used in the scheme: to form "super-observations", to check observations, and to produce grid point values while excluding rejected observations (without repeating the large matrix inversion). It is convenient to include the derivation of the three sets of equations in this section together with a derivation of the basic statistical interpolation equations.

The interpolated variables are at the present geopotential height, geopotential thickness between two specified isobaric levels, and eastward and northward components of the wind.

4.1 Statistical Interpolation

The basic technique used is a three-dimensional, multivariate, statistical interpolation of normalized deviations from a predicted field, where the normalization factors are the estimated root mean square errors of the predicted values.

Using A to represent any scalar variable, and E to represent its estimated rms error, with superscripts i, p, o and t respectively denoting interpolated, predicted, observed and true values, the basic interpolation equation is thus

$$\frac{A_k^i - A_k^p}{E_k^p} = \sum_{i=1}^{N} W_{ki} \frac{A_i^o - A_i^p}{E_i^p} \qquad (4.1.1)$$

where subscript k denotes the point and variable being analysed and subscripts i = 1,N range over all points and variables of observation.

W_{ki} are the weights to be determined.

Writing $\alpha_i^o = (A_i^o - A_i^t) / E_i^o$

$\alpha_i^p = (A_i^p - A_i^t) / E_i^p$

$\alpha_k^i = (A_k^i - A_k^t) / E_k^i$

$\varepsilon_i^o = E_i^o / E_i^p$

$\varepsilon_k^i = E_k^i / E_k^p$

equation (4.1.1) becomes

$$\alpha_k^i \, \varepsilon_k^i = \alpha_k^p + \sum_{i=1}^{N} W_{ki} \, (\alpha_i^o \, \varepsilon_i^o - \alpha_i^p) \qquad (4.1.2)$$

Squaring this, and taking the ensemble average (denoted by < >) gives

$$\varepsilon_k^{i\,2} = 1 + 2 \sum_{i=1}^{N} W_{ki}(<\alpha_k^p \alpha_i^o>\varepsilon_i^o - <\alpha_k^p \alpha_i^p>)$$

$$+ \sum_{i=1}^{N} \sum_{o=1}^{N} W_{ki}(<\alpha_i^p \alpha_j^p> + \varepsilon_i^o <\alpha_i^o \alpha_j^o> \varepsilon_j^o - \varepsilon_i^o <\alpha_i^o \alpha_j^p>$$

$$- <\alpha_i^p \alpha_j^o> \varepsilon_j^o) \, W_{kj} \qquad (4.1.3)$$

To simplify subsequent algebra we assume at this point that the correlations of prediction and observation errors are zero, i.e.

$$<\alpha_i^o \alpha_j^p> = <\alpha_i^p \alpha_j^o> = 0$$

This assumption is reasonable for the types of observation currently available; if necessary it could be relaxed. We also introduce a vector and matrix notation:- $\underset{\sim}{W}_k$ is the column vector of weights W_{ki}, $\underset{\sim}{P}$ is the prediction error correlation matrix $[<\alpha_i^p \alpha_j^p>]$, $\underset{\sim}{O}$ is the scaled observation error correlation matrix $[\varepsilon_i^o <\alpha_i^o \alpha_j^o> \varepsilon_j^o]$, and $\underset{\sim}{M} = \underset{\sim}{P} + \underset{\sim}{O}$.

Equation (4.1.3) becomes

$$\varepsilon_k^{i\,2} = 1 - 2\,\underset{\sim}{W}_k^T\,\underset{\sim}{P}_k + \underset{\sim}{W}_k^T\,\underset{\approx}{M}\,\underset{\sim}{W}_k \qquad (4.1.4)$$

The "optimal" weight vector is that which minimizes the estimated normalized interpolation error variance $\varepsilon_k^{i\,2}$. This is then given by

$$\underset{\sim}{W}_k = \underset{\approx}{M}^{-1}\underset{\sim}{P}_k \qquad (4.1.5)$$

The minimum value of $\varepsilon_k^{i\,2}$ corresponding to these weights is given by

$$\varepsilon_k^{i\,2} = 1 - \underset{\sim}{W}_k^T\,\underset{\sim}{P}_k \qquad (4.1.6)$$

4.2 Super-Observation formation

It is convenient to combine nearby compatible observations at the pre-analysis stage to form a "super-observation". This super-observation is then used in the analysis as an ordinary observation of increased accuracy.

If the standard optimum interpolation technique were used to create the super-observation then the interpolated value would contain information from, and hence be correlated with, the predicted value A_k^p. This is inconvenient since for ordinary observations it is assumed that $\langle \alpha_i^p\ \alpha_j^o \rangle = 0$. So the interpolation equations used for super-observation formation are modified by imposing a constraint that no information is taken from the local predicted value. Using the notation $\hat{}$ to indicate those values which differ in the modified method, the constraint can be expressed :

$$\langle \hat{\alpha}_k^i\ \alpha_k^p \rangle = 0 \qquad (4.2.1)$$

Using (4.1.2) to expand this, and using the matrix notation gives

$$\hat{\underset{\sim}{W}}_k^T\,\underset{\sim}{P}_k = 1 \qquad (4.2.2)$$

The interpolation error is still given by (4.1.4) :

$$\hat{\varepsilon}_k^{i\,2} = 1 - 2\hat{\underset{\sim}{W}}_k^T\,\underset{\sim}{P}_k + \hat{\underset{\sim}{W}}_k^T\,\underset{\approx}{M}\,\hat{\underset{\sim}{W}}_k \qquad (4.2.3)$$

Minimizing this with the constraint (4.2.2) gives the following equation for the constrained weights :

$$\hat{\underset{\sim}{W}}_k = (1 + \lambda)\underset{\approx}{M}^{-1}\underset{\sim}{P}_k \qquad (4.2.4)$$

Comparing this with (4.1.10) we see that the constraint forces a renormalization of the interpolation weights:

$$\hat{\underset{\sim}{W}}_k = (1 + \lambda)\underset{\sim}{W}_k \qquad (4.2.5)$$

with λ defined by substitution back in (4.2.2).

The error in the interpolated super-observation is given by substituting (4.2.2) and (4.2.4) in (4.2.3) :

$$\hat{\varepsilon}_k^{i\,2} = \lambda \qquad (4.2.6)$$

4.3 Observation checking

When checking a datum k we compare its normalized deviation from the predicted value $(A_k^o - A_k^p) / E_k^p$ with an interpolated deviation using nearby data $(A_k^i - A_k^p) / E_k^p$. Hence it is appropriate when deriving the equations for this interpolation to minimize the expected variance of the difference between the two deviations, rather than the deviation from the true value. Thus instead of (4.1.4) we minimize :

$$\varepsilon_k^{io\,2} = \varepsilon_k^{o\,2} + 1 - 2\underset{\sim}{W}_k^T \underset{\sim}{M}_k + \underset{\sim}{W}_k^T \underset{\approx}{M} \underset{\sim}{W}_k \qquad (4.3.1)$$

If the datum being checked is also used for the interpolation then $\underset{\sim}{M}_k$ is a column of $\underset{\approx}{M}$ and minimizing (4.3.1) leads to the trivial result

$$\underset{\sim}{W}_k = \underset{\sim}{D}_k \qquad (4.3.2)$$

where D_k is defined as a vector whose k'th element is one and other elements are zero, i.e. since we are trying to interpolate a value for the observation including its observational error, the best estimate is naturally the observation itself. What we must do is minimize (4.3.1) subject to strong constraints that certain data (datum k and perhaps others) are given zero weight. Using $\hat{}$ to indicate values obtained with the constraints, the constraints can be expressed:

$$\underset{1(i)}{D}^T \hat{\underset{\sim}{W}}_k = 0 \quad (i = 1, \text{ n constraints}) \qquad (4.3.3)$$

Minimizing (4.3.1) with these constraints gives

$$\hat{\underset{\sim}{W}}_k = \underset{\sim}{D}_k + \sum_{i=1}^{n\ \text{constr.}} \lambda_i \underset{\approx}{M}^{-1} \underset{\sim}{D}_{1(i)} \qquad (4.3.4)$$

The multipliers λ_i can be found by multiplying by $\underset{\sim}{D}^T_{1(j)}$ and using (4.3.3):

$$\sum_{i=1}^{n \text{ constr.}} \underset{\sim}{D}^T_{1(j)} \underset{\approx}{M}^{-1} \underset{\sim}{D}_{1(i)} \lambda_i = -\underset{\sim}{D}^T_{1(j)} \underset{\sim}{D}_k \qquad \text{(for j=1, n constraints)} \qquad (4.3.5)$$

Substituting (4.3.3) and (4.3.4) back in (4.3.1) gives:

$$\varepsilon_k^{io^2} = \varepsilon_k^{o^2} + 1 - \underset{\sim}{\hat{W}}^T_k \underset{\sim}{M}_k \qquad (4.3.6)$$

4.4 Grid point analysis

The matrix inverse $\underset{\approx}{M}^{-1}$ used in the observation checking equations is expensive to compute, so it is advantageous to re-use it when performing the grid-point analysis rather than computing a slightly different one. However, data which have been rejected in the checking are included in the inverse, and it is therefore necessary to apply constraints that these rejected data are given zero weight. Thus we need to minimize (4.1.4) with constraints like (4.3.3). Again using ^ to indicate values obtained with the constraints, we get:-

$$\hat{\varepsilon}_k^{i^2} = 1 - 2 \underset{\sim}{\hat{W}}^T_k \underset{\sim}{P}_k + \underset{\sim}{\hat{W}}^T_k \underset{\approx}{M} \underset{\sim}{\hat{W}}_k \qquad (4.4.1)$$

$$\underset{\sim}{D}^T_{1(i)} \underset{\sim}{\hat{W}}_k = 0 \quad (i=1, \text{ n constraints}) \qquad (4.4.2)$$

This gives:-

$$\underset{\sim}{\hat{W}}_k = \underset{\sim}{W}_k + \sum_{i=1}^{n \text{ constr.}} \lambda_i \underset{\approx}{M}^{-1} \underset{\sim}{D}_{1(i)} \qquad (4.4.3)$$

The multipliers λ_i can be found as in (4.3.5):

$$\sum_{i=1}^{n \text{ constr.}} \underset{\sim}{D}^T_{1(j)} \underset{\approx}{M}^{-1} \underset{\sim}{D}_{1(i)} \lambda_i = -\underset{\sim}{D}^T_{1(j)} \underset{\sim}{W}_k \qquad (4.4.4)$$

(for j=1, n constraints)

The interpolation error is given by:

$$\hat{\varepsilon}_k^{i^2} = 1 - \underset{\sim}{\hat{W}}^T_k \underset{\sim}{P}_k \qquad (4.4.5)$$

5. Prediction error modelling

5.1 General considerations

To solve the above equations we need values for the expected rms prediction errors of all variables and positions which are observed or analysed, and for the prediction error correlations $\langle \alpha_i^p \alpha_j^p \rangle$ between them. Even if empirical data for these were available it would be impossible to tabulate values for every possible combination, so they must be modelled. This has the effect that relationships assumed when modelling the correlations are imposed on the analysed increments. For example, correlations and errors of the thickness Δh_{12} are calculated from those of the heights h_1 and h_2 using the equation $\Delta h_{12} = h_2 - h_1$. This means that as long as the same data are used for each, the analysed increments to the predicted field will obey the same equation, even if there are data for the three variables which do not obey it. Thus our choice of a correlation model was governed by the relationships which we wished the analysed increments to obey, as well as by the available empirical data and convenience of use.

In our present program the data used are geopotential heights and winds at a set of pressure levels, and geopotential thicknesses between the levels. In the near future it is planned to use also normal modes of the satellite thickness soundings. The analysed variables are height and wind at a set of pressure levels. Thus estimated rms prediction errors are needed for, and prediction error correlations are needed between all of these. Since only predefined pressure levels are used the vertical correlations are stored in matrix form, while a continuous functional representation (at present Gaussian) is used for the horizontal correlations. We assume furthermore that correlations between points at different levels and horizontal positions can be expressed as a product of the corresponding vertical and horizontal correlations.

5.2 Empirical prediction errors and correlations

When our analysis and prediction are operational we plan to accumulate statistics for the errors of our particular model. Until then it is necessary to use estimates either based on the few published data concerning other models or, in the case of a climatological "prediction" as was used for the test analysis presented in section 8, for deviations from climatology.

The rms height deviations from climatology and their vertical correlations were based on data for a north European radiosonde station, with a latitudinal profile

based on that of Oort and Rasmusson (1971), and the Gaussian horizontal correlations were based on those of Schlatter (1975). When the prediction is from a forecast model we use values based on those for a Canadian forecast model (Rutherford 1976) with a similar latitudinal profile and a shorter range of horizontal correlations.

5.3 Height and thickness correlations

The height prediction error correlations are assumed to be separable, isotropic and locally homogeneous. That is for two heights at horizontal positions i and j and levels L and M the correlation is assumed to be given by $\mu_v(L,M)\mu_h(i,j)$, where μ_v is specified by an empirically based table, and μ_h is a function of the distance r_{ij} with empirically derived parameters. At present we take μ_h to be $\exp(-br_{ij}^2)$ with $b = 0.61 \times 10^{-12}m^{-2}$ for climatological "predictions" and $b = 2.0 \times 10^{-12}m^{-2}$ for model forecasts. Provisions have been made for letting μ_v and b vary geographically (although they should be kept constant locally).

Errors of, and correlations involving thicknesses are calculated from those for the corresponding heights as mentioned in section 5.1. More complicated linear combinations of height such as normal mode expansions of thickness soundings can be treated similarly.

5.4 Wind correlations

In order to have a formulation for wind correlations which is valid globally and which is not dependent on one particular coordinate system, the normalized wind increments are expressed in terms of a streamfunction ψ and velocity potential χ. Separable homogeneous isotropic correlation models are then assumed for these scalar parameters in the same way as for heights. Correlations for longitudinal (i.e. along the great circle joining points i and j) and transverse wind components can then be found by differentiating μ_h with respect to r_{ij}, and correlations for wind components along any other axes by a simple rotation of coordinates.

At present, for simplicity and because of the lack of empirical correlations, the streamfunction correlations are assumed to be identical with the height correlations, while the velocity potential prediction errors are assumed to be zero, making velocity potential correlations redundant. This last assumption has the effect that as long as the same data are used (i.e. within one analysis box) the normalized corrections to the predicted wind field also have zero velocity potential and are thus non-divergent. Since our normalizing rms prediction errors

vary only slowly with latitude the absolute increments are also approximately non-divergent within one analysis box. However if the data imply divergences on scales greater than this it is still possible to analyse them.

The horizontal correlations which follow from these assumptions are:-

$$\mu_h(\psi_i, u_j^{transv}) = -\mu_h(u_i^{transv}, \psi_j) = \sqrt{2br}\exp(-br^2)$$

$$\mu_h(\psi_i, u_j^{long}) = \mu_h(u_i^{long}, \psi_j) = 0$$

$$\mu_h(u_i^{transv}, u_j^{transv}) = (1-2br^2)\exp(-br^2)$$

$$\mu_h(u_i^{long}, u_j^{long}) = \exp(-br^2)$$

$$\mu_h(u_i^{transv}, u_j^{long}) = (u_i^{long}, u_j^{transv}) = 0$$

5.5 Height-wind correlations

It remains to specify the cross correlations between height and streamfunction $\mu_{h\psi}$, and if velocity potential errors are also to be modelled, between height and velocity potential $\mu_{h\chi}$ and between streamfunction and velocity potential $\mu_{\psi\chi}$. These last two are related to amount of cross isobar flow. $\mu_{h\psi}$ is directly related to geostrophy; if $\mu_{h\psi} = 1$ and the ratio of the assumed rms height and wind errors is $f/g\sqrt{2b}$ then, neglecting horizontal derivatives of f and the assumed rms errors, the corrections to the heights and winds are in geostrophic balance within one analysis box.

At present we assume that $\mu_{h\psi} = .95$ north of 30^oN, $-.95$ south of 30^oS, and .95 sin $(\frac{\pi}{2}$ latitude$/30^o)$ in between. A correlation between a height and a wind is thus modelled as the product of $\mu_{h\psi}$ at the latitude of the height and the correlation between ψ, at the position of the height, and the wind.

5.6 Mathematical constraints

It should be mentioned that the various parameters of this prediction error correlation model are not completely independent, and care must be taken that all the correlation matrices M are positive definite and not ill-conditioned, otherwise unrealistic analysis increments can result. Thus the matrix μ_v and the function μ_h must be positive definite, and if these are allowed to differ for different variables they must be compatible with each other and with $\mu_{h\psi}$, $\mu_{h\chi}$ and $\mu_{\psi\chi}$. Our present simplified correlation model is well behaved, and with realistic values for the observation error correlations, which form $\underset{\approx}{O}$, the matrix $\underset{\approx}{M}$ is not ill-conditioned.

6. Data flow, program structure and files

Figure 2 shows the programs and files which make up the analysis of dynamic variables step of Figure 1. A brief description of each is given below.

6.1 Sorted observations

ECMWF will maintain a data-bank of observations for its operational purposes, and a similar format is planned to be used for the FGGE data. Before each analysis a program will search this for observations in the appropriate time period, and produce a file of these sorted into the data boxes described in Section 3.7, for input into the pre-analysis program.

6.2 Predicted fields

These are specified on the horizontal analysis grid which is similar to that of the prediction model, but on standard pressure levels. The methods used to interpolate these from the prediction model's sigma levels are discussed by Rutherford (1977a).

6.3 Prediction error statistics

Root mean square prediction errors and vertical and horizontal multivariate prediction error correlations are obtained and modelled as described in Section 5.

6.4 Observation error statistics

Assumptions are made about the rms observation error (including the pseudo-error due to scales smaller than

those being analysed) for each type of observation. Details are given in Section 7, along with details about the spatial correlations assumed (in most cases zero).

As comparably little change is expected within this data it is included in the program code as data statements, so in fact there is at present no separate observation error statistics file.

6.5 Pre-analysis

This program performs all possible preliminaries to the analysis in order to simplify the analysis program. The observations for a box are read from the observation file (Section 6.1) a predicted value for each is interpolated from the predicted fields (Section 6.2) and observation and prediction errors for each estimated (using Section 6.3 and Section 6.4). Observations in the same box which are too close to give useful independent information about the scales being analysed are combined to form "super-observations" (see Section 4.2).

If there are enough observations in the box then sufficient primary observations are selected to represent the box when analysing neighbouring boxes (see Section 3.7). The information is then written to a work file in a form convenient for the analysis program.

6.6 Work file

This is a random access disc file, with one record per box for the primary observations plus one record for each box which has secondary observations (nearly 2000 records with the present box size). Each record contains the observation positions, types, errors, flags and values, as well as counts of the data and the appropriate prediction error statistics. The values are stored as observed minus prediction differences, normalized by the estimated prediction error.

6.7 Data checking analysis

This program is the same as the grid-point analysis program (Section 6.10), except that it calculates interpolated values at the data points for checking purposes, using the equations of Section 4.2. Flags indicating the results of the check are written back to the work file.

6.8 List of rejected data

A list of probably bad observations is created in the data checking analysis (Section 6.7). Facilities for the human monitoring of this list and amendment of the work file flags are being planned.

6.9 Matrix inverses

There is an option for saving the matrix inverses and other partial results during the data checking analysis, in order to avoid repeating identical computations in the grid-point analysis. The size of this file depends on the matrix orders used, the numbers of observations and grid-points, and any packing performed before output. It might be as much as 10^8 words.

6.10 Grid-point analysis

The analysis program uses normalized differences from the prediction to calculate normalised increments to the field. The actual predicted values are only used in the pre-analysis and post-analysis. The program reads from the random access work-file the records containing primary and secondary data for a box, and the appropriate primary data from its neighbours. It then sub-divides the box into analysis sub-volumes if necessary, forms a list of the data to be used for the analysis within each sub-volume, and calculates and inverts the error covariance matrix for these data (this need not be done if the matrix inverse file is used). The matrix inverse is now used to calculate interpolation weights for each data value for each variable and analysis point within the analysis sub-volume, and hence the normalised interpolated increments and the estimated analysis errors (see Section 4.4).

6.11 Increments and analysis errors

This temporary file is written by the analysis program as the values are calculated, and therefore needs to be reordered for convenient use.

6.12 Post-analysis

This program sorts the normalized analysis increments and errors into a convenient order, and, using the predicted fields and their assumed errors (the normalization factors), produces the analysed fields.

7. Observational data and errors

For ECMWF's operational analysis data exchanged internationally on the global trunk circuit (GTS) will be obtained via a telecommunications link to the British Meteorological Office at Bracknell.

The FGGE analyses (the FGGE Level III-b data set) will be produced on the basis of the complete Level II-b delayed observation collection data sets produced at the Surface-based Data Centre in USSR, and the Satellite and Special Observing System Data Centre in Sweden (Larsen, 1977). In addition to the conventionally transmitted data on the GTS these data sets include observations collected in a delayed mode that would not normally be available (many ship reports, Southern Hemisphere observations etc.) as well as observations from special observing systems (research satellites, buoys, etc.).

For the test analysis described in Section 8 a data-set made available by the U.S. National Meteorological Centre (NMC) was utilised. This was collected during a data systems test (DST) designed to test some of the special observing systems planned for the FGGE.

It consists of the data available operationally at NMC via the GTS plus data from experimental US satellites similar to those planned for the FGGE.

If they prove to be successful during the FGGE it is likely that these satellites will become part of the operational observing system. Hence it is possible that the DST data-set, as well as being a test for the FGGE, is also indicative of what will be available operationally to ECMWF in the future.

The different observation types in the data-set (relevant to the mass and wind field analysis) and the observation errors and error correlations assumed for each are presented in Tables 1, 2 and 3. The observational errors are estimates of the accuracy with which each observation measures the bulk properties of the atmospheric state which is to be analysed, rather than of the accuracy of the measuring instrument alone, since effects of phenomena below the scale which can be analysed (e.g. thunder storms) are not desirable.

8. A trial analysis

In this section we present the results of an analysis made during the testing and development of the scheme. The observations used were from the DST data-set for the 6 hour period centred on 00Z of the 10th February 1976. The trial analysis was performed on a $3\frac{3}{4}^{o}$ regular lat./long. grid (for the operational system a somewhat finer resolution will be used). As the "predicted" field we used February climatology, with the somewhat arbitrarily chosen prediction error statistics described in section 5. A very simple data selection

algorithm was used; super-observations were formed between compatible observations in the same analysis box which were closer than 100 km, then, when analysing each box, data were taken from the best ten observations (Larsen et al., 1977) in that box, followed by data from the nearest level of all radiosonde soundings in the influencing area, followed by other data, up to a limit of 151 data values or 50 of any one variable (height, thickness, u wind component or v wind component). All data used were checked using the equations of section 4.3. If any datum disagreed with the interpolated value by more than four times the estimated error in the interpolated value then it was not used for the analysis, nor for checking subsequent data. Out of 45300 data values used, 918 were rejected. Data were used from, and height and wind analyses produced at, ten pressure levels from 1000 mb to 100 mb. Surface data were converted to the nearest pressure level.

Example fields from this test analysis are shown in Figures 3, 5 and 7, while Figures 4, 6 and 8 show the same fields from an analysis made using the same DST data-set by NMC with their operational analysis scheme (available on a $2\frac{1}{2}^{o}$ grid). Figures 9,10 and 11 show the normalized error variance estimated by our scheme, together with the observations available for this analysis. (Figure 9 also shows the analysis box grid used). The normalized error variance is a measure of the information taken from the prediction by our analysis scheme. Thus for values greater than 0.5 (the thick contour) less than 50% of the analysed value came from observations. Since we used climatology as the prediction (the 500 mb height field of which is shown in Figure 12), while NMC used a 6 hr. forecast, we must expect the analyses to differ in such areas.

At least for the scales visible on this size of chart, the 1000 mb height analyses (Figures 3 and 4) are very similar in the northern hemisphere. In the southern hemisphere however there are differences; although they agree on the positions of most major systems, they disagree somewhat about their strengths and shapes except over the few regions with adequate observations, such as Australia and New Zealand.

At 500 mb in the northern hemisphere the height fields are similar in the two analyses (Figures 5 and 6). In the southern hemisphere the positions of the major features agree while their shapes and intensities differ somewhat. The wind analyses (Figures 7 and 8) differ slightly more than the heights; the major jet stream (shown in Figures 5 and 6 by the 30 m/s isotach) differ noticably even in the northern hemisphere. Another interesting difference is the kink at 180^{o}W 40^{o}N in both our height and wind field at 500 mb, which is not present in NMC's analysis. This feature is actually supported by the few satellite

temperature soundings, cloud winds and surface observations available, although the feature seems somewhat exaggerated by the present crude selection scheme.

In order to give some insight into the more detailed properties of our analysis, we now present 1000 mb and 500 mb charts for Europe, with geopotential heights and wind vectors superimposed. Comparing our analysis (Figure 13) with NMC's (Figure 14) several differences show up, particularly in the wind fields. At 1000 mb our analysis has significant cross isobar flow not present in NMC's analysis, and at 500 mb our analysed winds in the trough over Scotland are slightly weaker.

These differences can be attributed to the fact that our scheme places less emphasis on the geostrophic relationship. Since we assume that the maximum height-streamfunction prediction error correlation $\mu_{h\psi}$ is less than 1.0 (see section 5), in regions with sufficient wind observations the wind analysis will depend on these rather than the height gradient. This effect is exaggerated in Figure 15, which shows the analysis which results if $\mu_{h\psi}$ is assumed to be zero everywhere. In this case the height and wind analyses are completely independent. Figure 16 shows the analysis which results if the standard value of $\mu_{h\psi}$ (0.95) is taken, but no wind data is used. This leads to a height analysis identical to that of Figure 15, while the analysed wind increments to the climatological winds are 95% of the geostrophic ones implied by the height increments. Exaggerated wind speeds can thus clearly be seen in relation to troughs, and at the surface over land. Obviously there is scope for tuning our analysis scheme using the assumed prediction error correlations, to give a realistic amount of coupling between the height and wind analyses.

Figures 17 and 18 illustrate another aspect which needs some tuning; the data checking. Figure 17 shows the analysis given by our scheme with no data checking, while Figure 18 shows the 500 mb geopotential height analysis and a hand drawn mean-sea level pressure analysis from Deutscher Wetterdienst. The human analyst, using his experience and two SYNOP observations from the south east coast of Greenland, has drawn a very small cut off low on the mean sea level chart which is not even hinted at by our analysis (Figure 13) or NMC's analysis (Figure 14). The two SYNOP pressures were in fact rejected by our data checking scheme; using them does improve the analysis in this region somewhat (Figure 17). (The low is a real development; 12 hours later it is approaching Iceland). Part of the problem is in the logic used for the checking, one observation was rejected only after the other had been excluded. But the main problem lies with the prediction

error correlations used in this experiment. The value of b used for the horizontal correlations, $0.61 \times 10^{-12} m^{-2}$, is just not suitable for features of this size, since the correlation still has a value of 1/e at a distance of 1280 km.

9. Further development

Our analysis and data assimilation scheme is expected to be in a state of continuous development and improvement for several years. In this section we describe briefly the most important projects currently in progress or planned.

9.1 Extension to other variables

Other parameters such as humidity and sea surface temperature need also be analysed operationally as well as for FGGE. A system for humidity analysis is presently under development and different ways of analysing sea surface temperatures are being investigated.

9.2 Link with prediction model

This is an area of immediate interest and a simple link involving grid transformations of the complete fields is being tested. We then intend to test the more attractive scheme outlined by Rutherford (1977a) where only the analysis increments are transformed to the prediction model's levels.

9.3 Initialization

If analysed fields are used directly as initial conditions for a forecast model, the integration may be contaminated by spurious high-frequency gravity-wave oscillations caused by imbalances between the variables (particularly between the mass and wind fields). These imbalances may be removed by a variety of initialization procedures. There is disagreement, even between groups involved in operational forecasting, on the impact of initialization on the accuracy of the subsequent forecast, and on the best means of achieving the required balance. In the context of the assimilation/forecast cycle, the link between the analysis and the forecast, and the constraints imposed during the analysis, might ideally render initialization unnecessary. On the other hand, with a cycle time as short as six hours, any imbalances remaining after an analysis are likely to be present in the six-hour forecast used as the background field for the next analysis. Lack of an initialization procedure may thus degrade the quality of the analyses.

A research project is being mounted at ECMWF to investigate these questions. A simple dynamic initialization scheme (Temperton 1976) is already available for the ECMWF grid-point model. A more sophisticated normal mode

initialization procedure (Machenhauer 1977) is also under development.

9.4 Prediction error statistics

Obviously empirical prediction error statistics for the forecast model used in our operational scheme will not be available until the scheme has been running for some time. We intend to collect such statistics as a matter of routine. In the interim a more heuristic approach is necessary. Possible developments of the present scheme are:-

1) the "forecasting" of prediction errors from the preceding analysis errors.

2) the incorporation of velocity potential errors into the model, so that the wind increments are not forced to be strictly non-divergent within each analysis box.

3) the specification of correlations which vary in space according to the meteorological situations, so as to incorporate objectively the sort of experience which made the human analyst draw the small low discussed at the end of section 8.

9.5 Data checking

The criterion for data rejection needs tuning. Inappropriate prediction error statistics can lead to undue confidence in the interpolated value. Hence the present limit of four times the estimated interpolation error is too severe. The logic for coordinating the checking of different data from the same observation, and of neighbouring observations when some have already been rejected, needs more development.

9.6 Data-selection

Plans are being made for improving present simple data selection algorithm.

10 Acknowledgements

We would like to thank the US National Aeronautics and Space Administration and their GARP DST operations group for making available the DST data-set and NMC analyses, and for their kind assistance in connection with the use of the data.

Dr L. Bengtsson has given most inspiring guidance and criticism throughout the development of the system.

Mr C. Little and Mr C. Clarke wrote the pre-analysis and plotting programs without which this work would not have been, and gave valuable assistance in producing the analyses; Mrs E. Edberg wrote the observation decoding program and likewise assisted with the processing of the DST data tapes.

11. References

Burridge, D.M. and Haseler, J. (1977) 'A model for medium range weather forecasts - Adiabatic formulation'. ECMWF Technical Report No. 4.

Gandin, L.S. (1963) 'Objective analysis of meteorological fields'. Translated from Russian by the Israeli Program for Scientific Translations 1965.

Larsen, G., Little, C., Lorenc, A. and Rutherford, I. (1977) 'Analysis error calculations for the FGGE'. ECMWF Internal Report 11.

Larsen, G. (1977) 'Plan for the FGGE Level III-b data set production'. ECMWF, Oct. 1977.

Machenhauer, B. (1977) 'On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialisation'. Contributions to Atmospheric Physics 50, 253-271.

Oort and Rasmusson (1971) 'Atmospheric Circulation Statistics'. NOAA professional paper 5, US Dept. of Commerce.

Roberts, K.V. (1974) 'An introduction to the OLYMPUS system'. Computer Physics Communications 7, 237-244.

Rutherford, I.D. (1976) 'An operational 3-dimensional multi-variate statistical objective analysis scheme'. The GARP programme on numerical experimentation. Report No. 11 WMO/ICSU.

11. References (continued)

Rutherford, I.D. (1977a) 'Interfaces between the prediction model and the analysis program in the data assimilation system; interpolation questions'. ECMWF Working Paper 1/12/E/RD2/046/197.

Rutherford, I.D. (1977b) 'Separation of prediction and observation error covariances'. ECMWF Working Paper 1/6/E/RD2/048/1977.

Schlatter, T.W. (1975) 'Some experiments with a multi-variate statistical objective analysis scheme'. Mon. Wea. Rev. 103, 246-257.

Schlatter, T.W., et al. (1976) 'Testing a global multi-variate statistical objective analysis scheme with observed data'. Mon. Wea. Rev. 104, 765-783.

Temperton, C. (1976) 'Dynamic initialisation for barotropic and multi-level models'. Q.J. Roy. Met. Soc. 102, 297-312.

TABLE 1 - Observation error characteristics -

COMPONENT	VARIABLES	ERROR CHARACTERISTICS
Sondes	V	random $3ms^{-1}$
	Z	vert. corrn. (Table 2)
Surface	V	random $6ms^{-1}$
	Z	random 12 m
Aircraft	V	random $4ms^{-1}$
Satellite Temperature Soundings	Z	horiz.corrn. $\exp(-16\times10^{-12}r^2)$ vert. corrn. (Table 3)
Satellite Cloud-Track Winds	V	850 mb: $3ms^{-1}$ random, 300 mb: $6ms^{-1}$

TABLE 2 - Radiosonde height errors and vertical correlations - $\mu_v(Z,Z)$ -

	1000	850	700	500	400	300	250	200	150	100
σ_Z(m)	6.0	7.8	9.3	13.1	16.1	21.7	27.3	33.9	35.6	32.2
1000	1.00									
850	.70	1.00								
700	.50	.72	1.00							
500	.39	.56	.74	1.00						
400	.34	.48	.69	.91	1.00					
300	.32	.45	.63	.84	.92	1.00				
250	.29	.42	.60	.80	.85	.94	1.00			
200	.27	.38	.56	.76	.79	.88	.96	1.00		
150	.27	.38	.56	.76	.80	.81	.92	.96	1.00	
100	.27	.41	.56	.71	.76	.79	.75	.74	.83	1.00

TABLE 3 - Satellite thickness errors and vertical correlations - $\mu_v(\Delta Z,\Delta Z)$ -

	1000 -850	850 -700	700 -500	500 -400	400 -300	300 -250	250 -200	200 -150	150 -100
$\sigma_{\Delta Z}$(m)	13.3	14.2	21.2	13.4	18.1	12.2	15.9	21.9	33.2
σ_T(°C)	2.8	2.5	2.15	2.05	2.15	2.28	2.43	2.60	2.80
1000 - 850	1.00								
850 - 500	.50	1.00							
700 - 500	.35	.55	1.00						
500 - 400	.20	.35	.60	1.00					
400 - 300	.10	.20	.35	.60	1.00				
300 - 250	0	.10	.20	.30	.55	1.00			
250 - 200	0	0	.05	.05	.15	.52	1.00		
200 - 150	0	0	0	0	0	.05	.50	1.00	
150 - 100	0	0	0	0	0	0	0	.50	1.00

LIST OF FIGURES

16. Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from an ECMWF analysis using height data only.

17. Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from an ECMWF analysis made without rejecting any data.

18. Deutscher Wetterdienst hand drawn analysis of 500 mb geopotential height and mean-sea-level pressure.

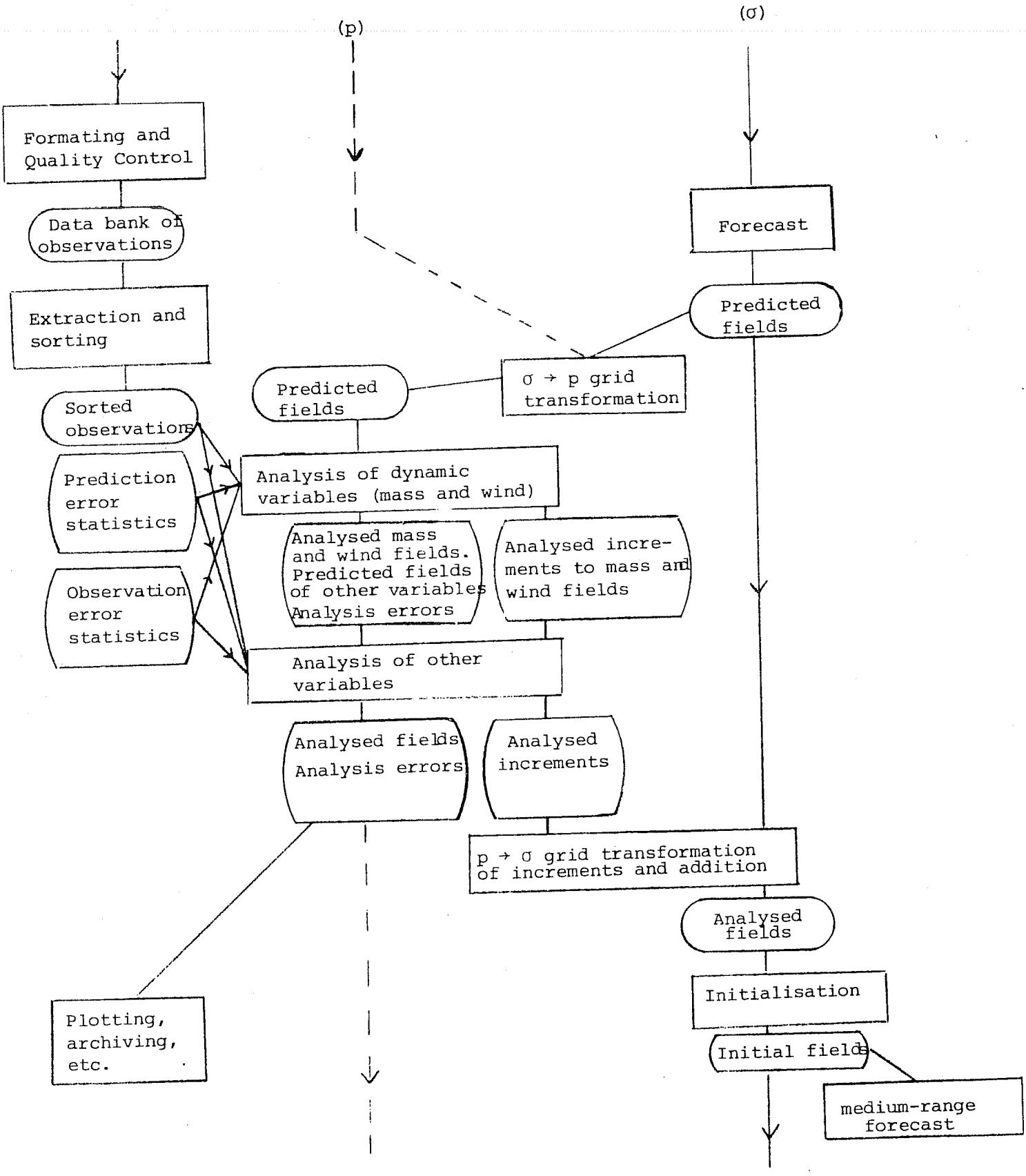


Fig. 1 Major functions (square boxes) and files (round boxes) in one cycle of the analysis and data assimilation scheme

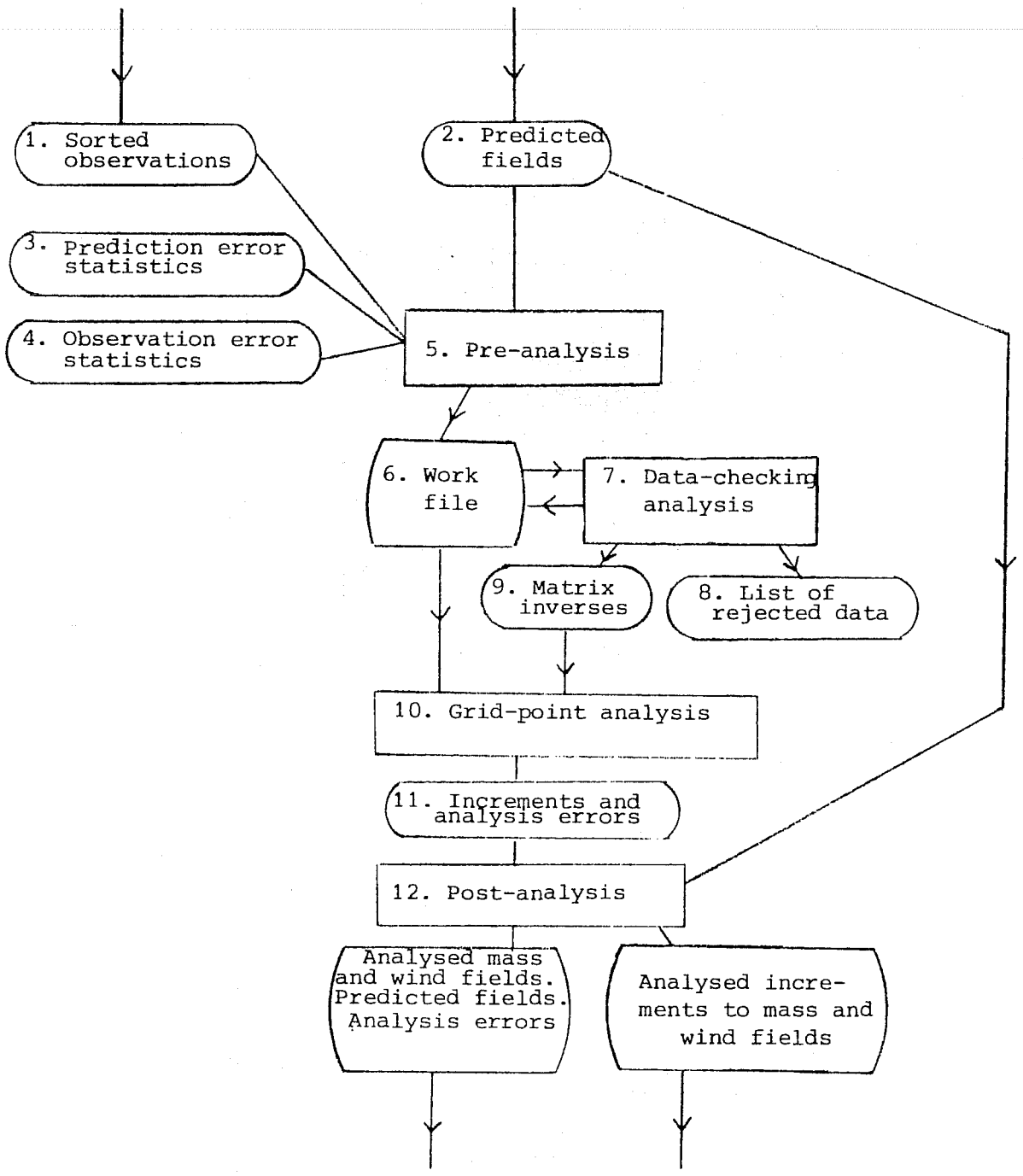
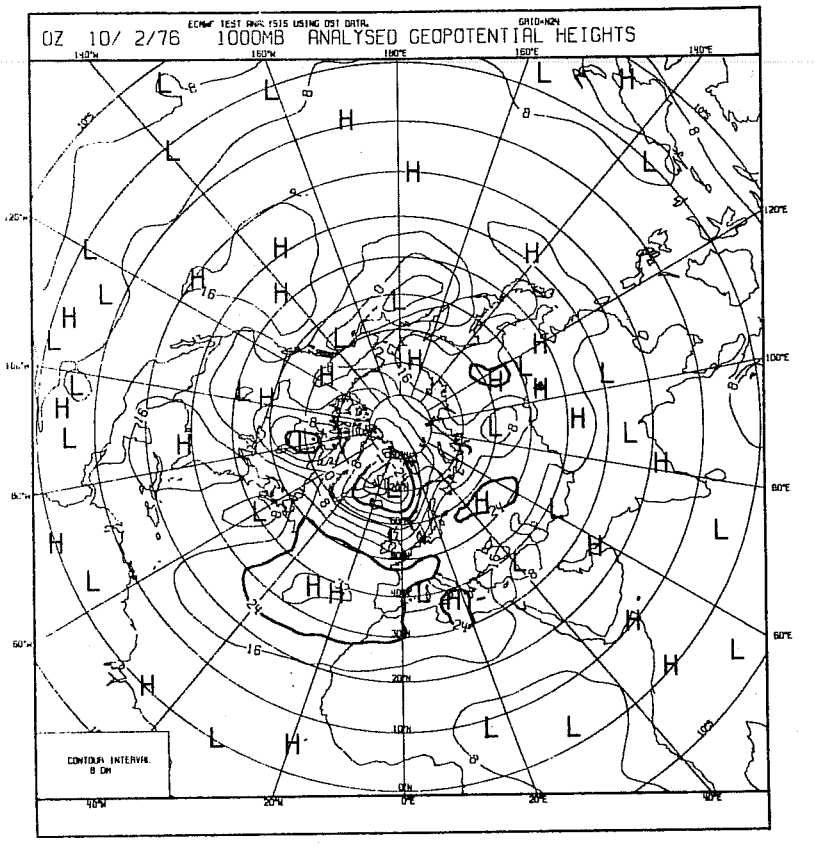


Fig. 2 Programs and files in the analysis of the dynamic variables (numbers refer to sub-sections of Section 6).



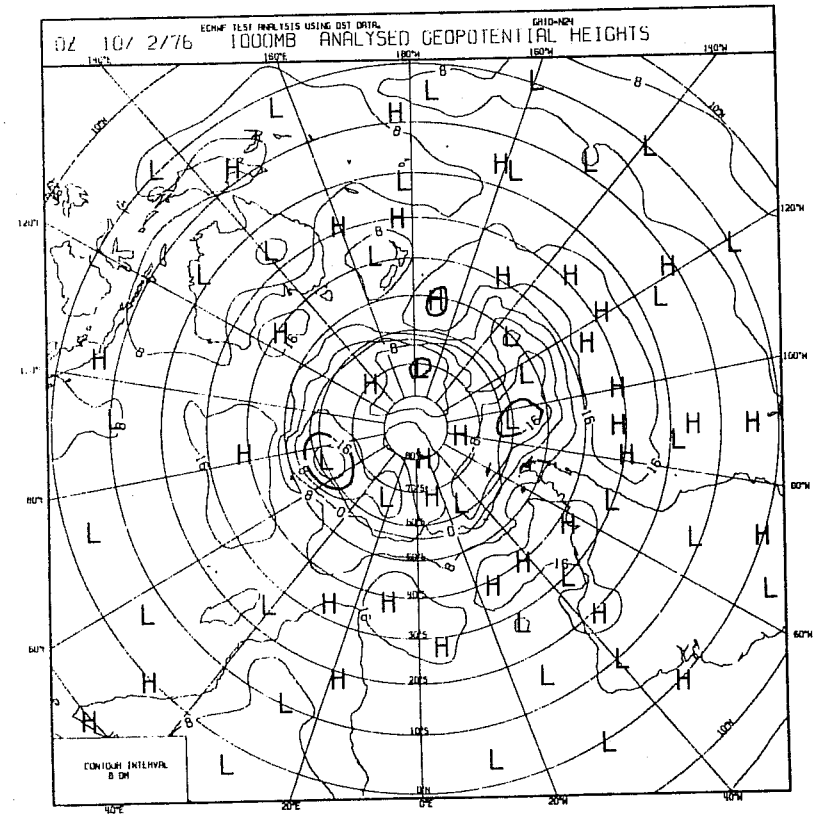
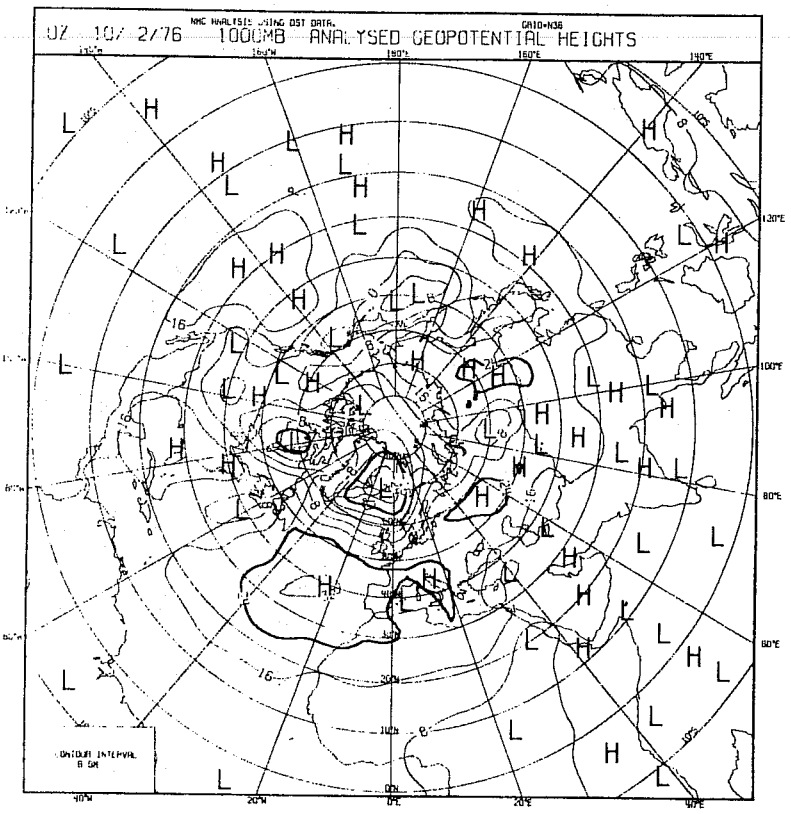


Fig. 3 ECMWF test analysis: 1000 mb geopotential heights.



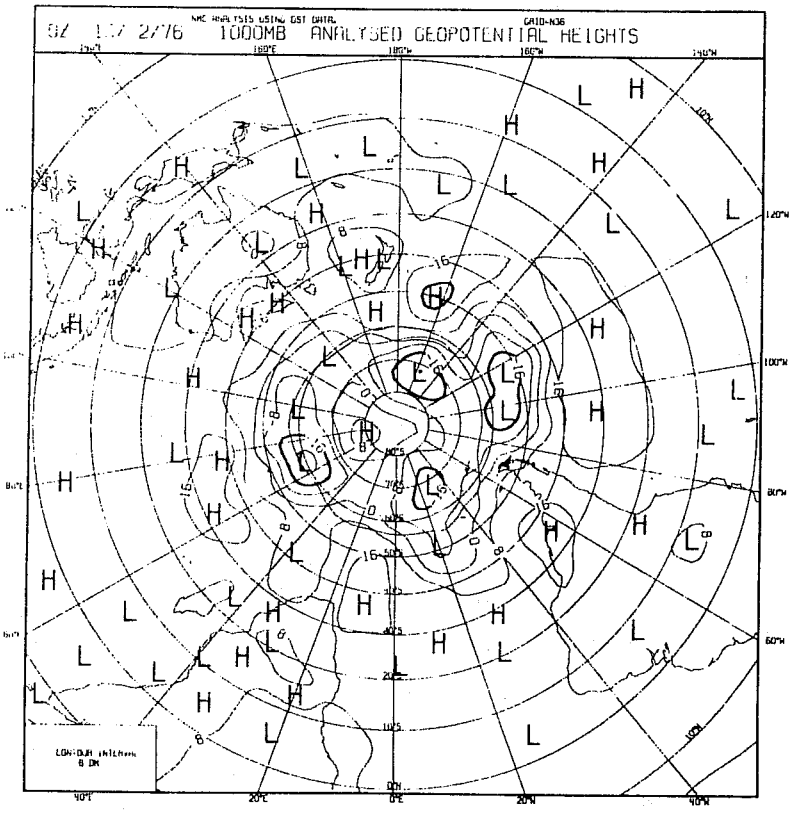


Fig. 4 NMC analysis: 1000 mb geopotential heights.

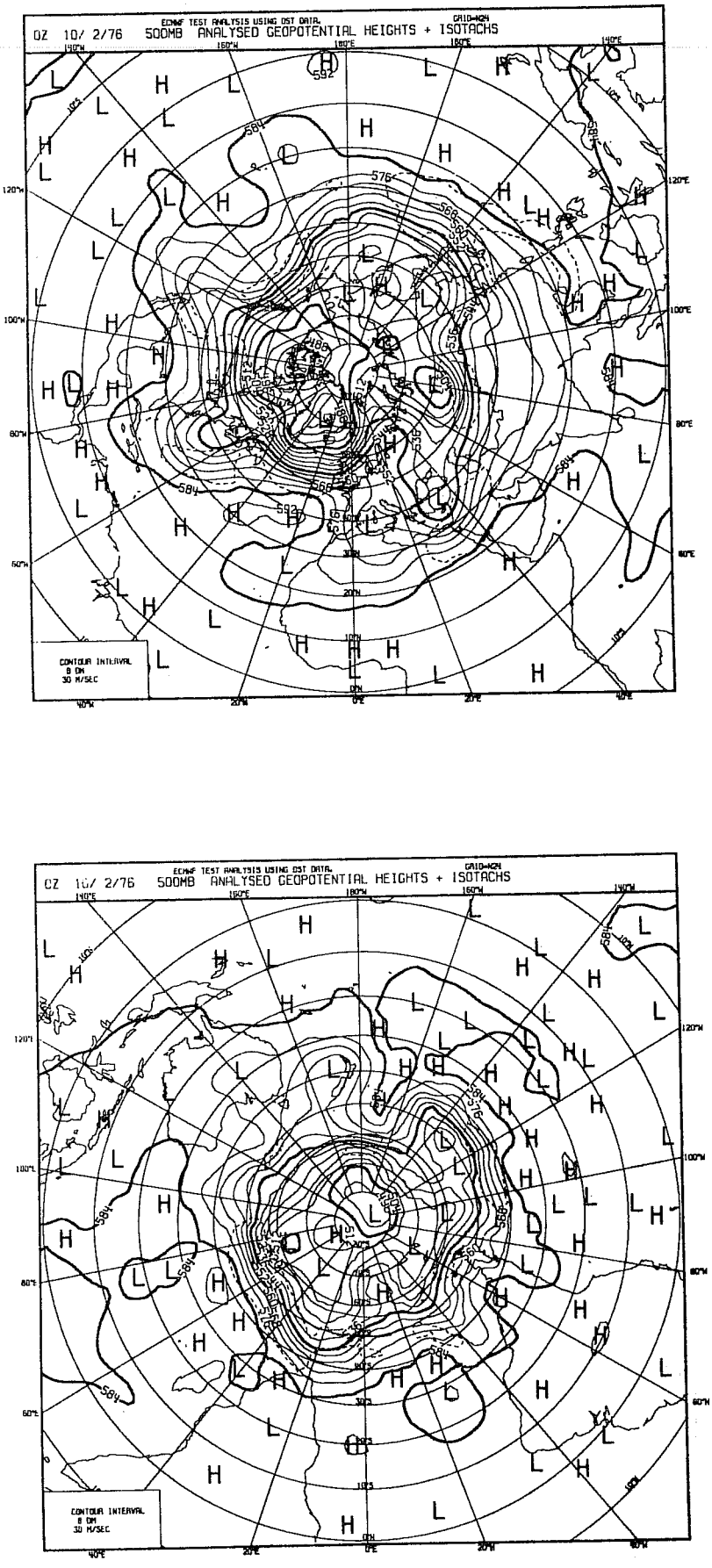
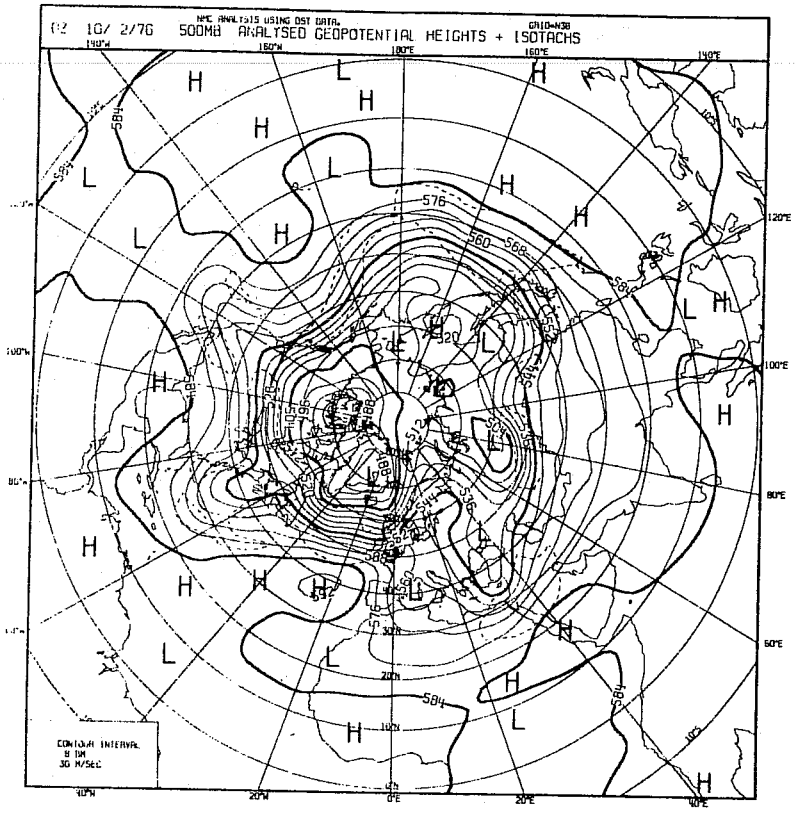


Fig. 5 ECMWF test analysis: 500 mb geopotential heights and the 30 m/s isotach.



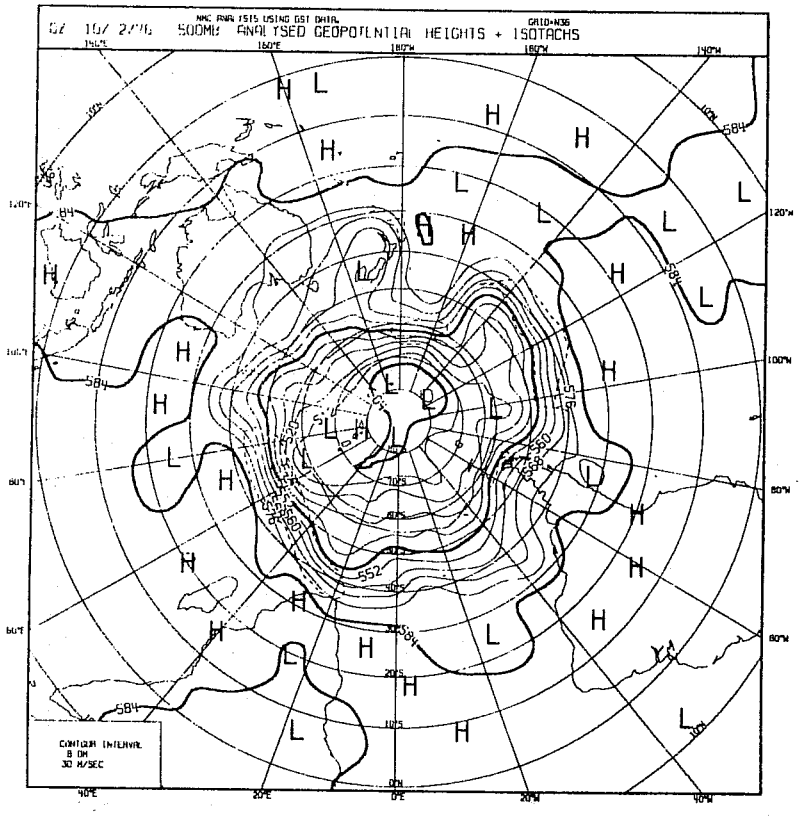
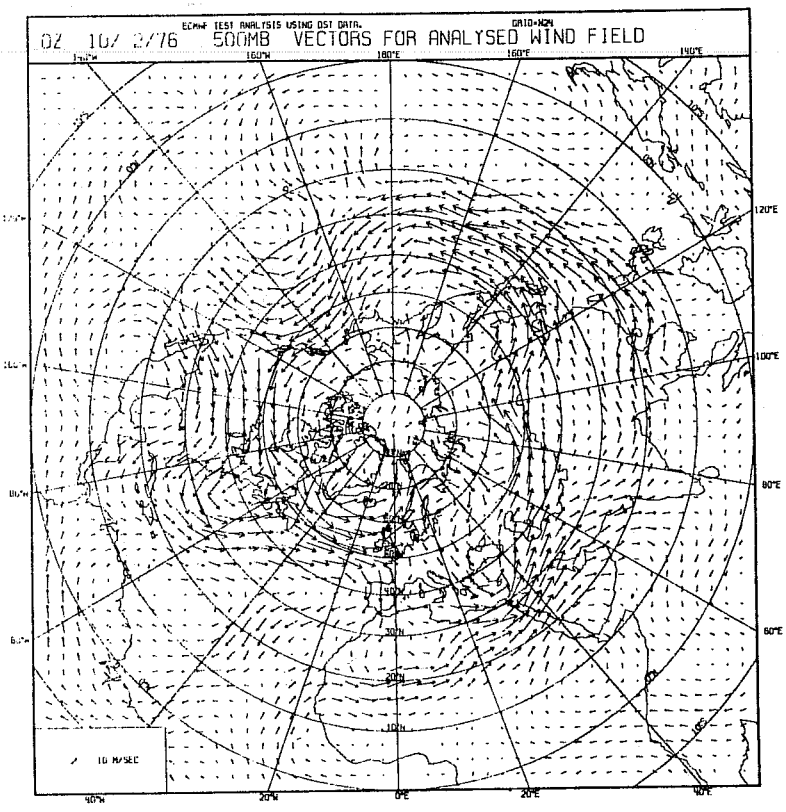


Fig. 6 NMC analysis: 500 mb geopotential heights and the 30 m/s isotach.



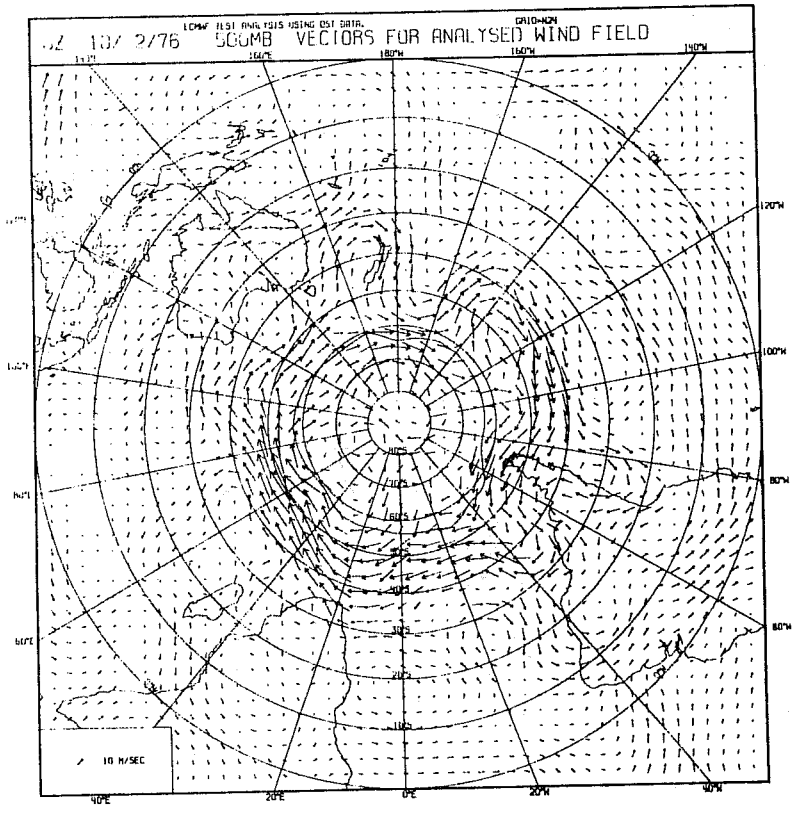


Fig. 7 ECMWF test analysis: 500 mb wind vectors.

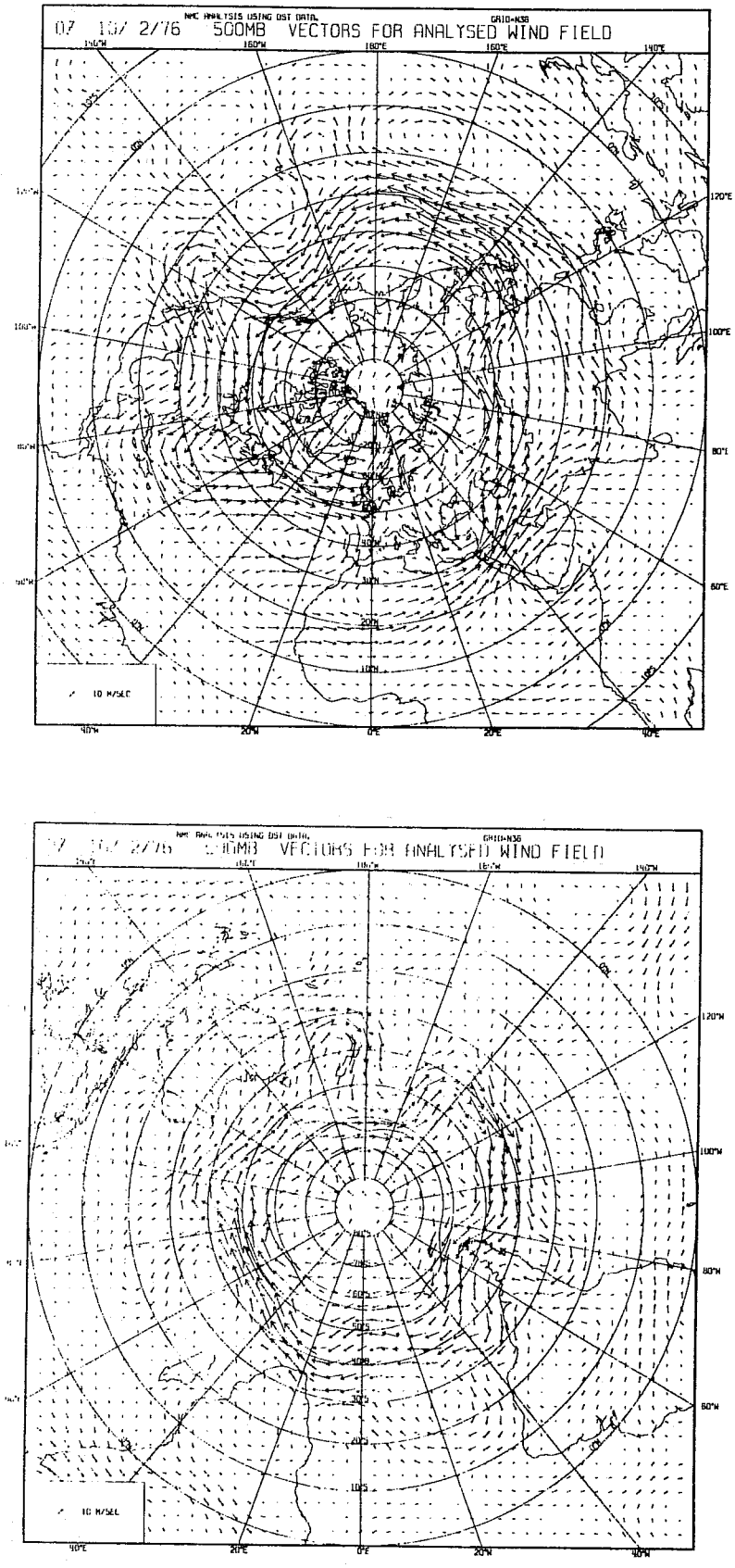
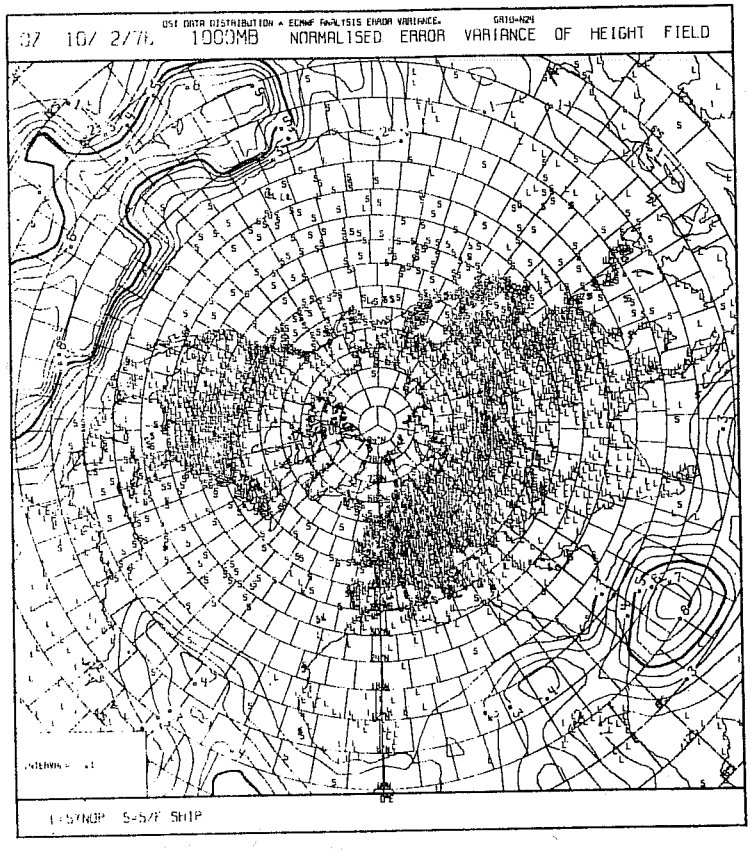


Fig. 8 NMC analysis: 500 mb wind vectors.



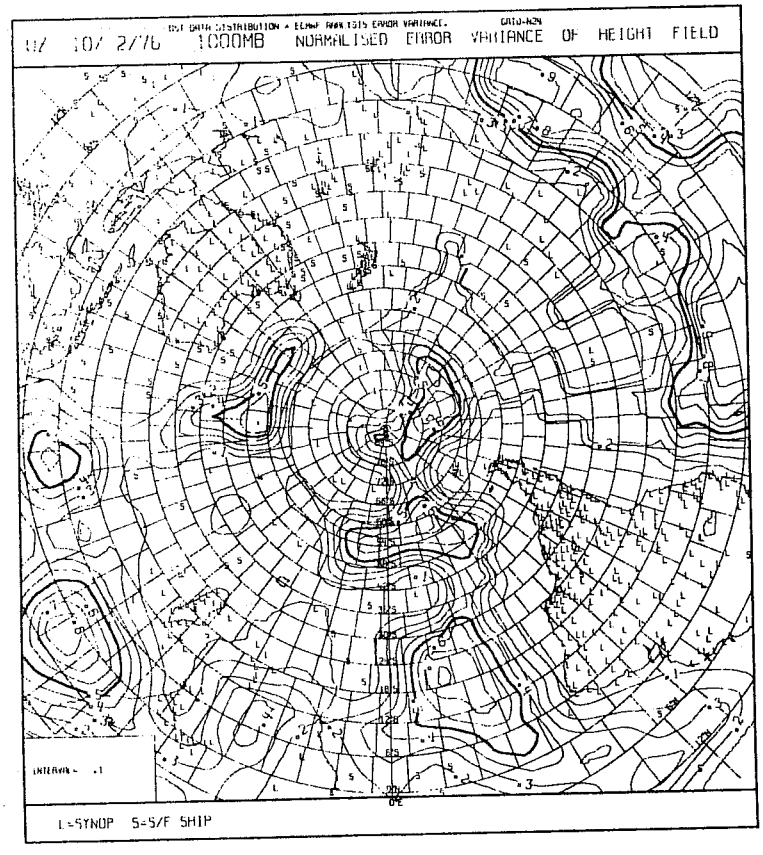
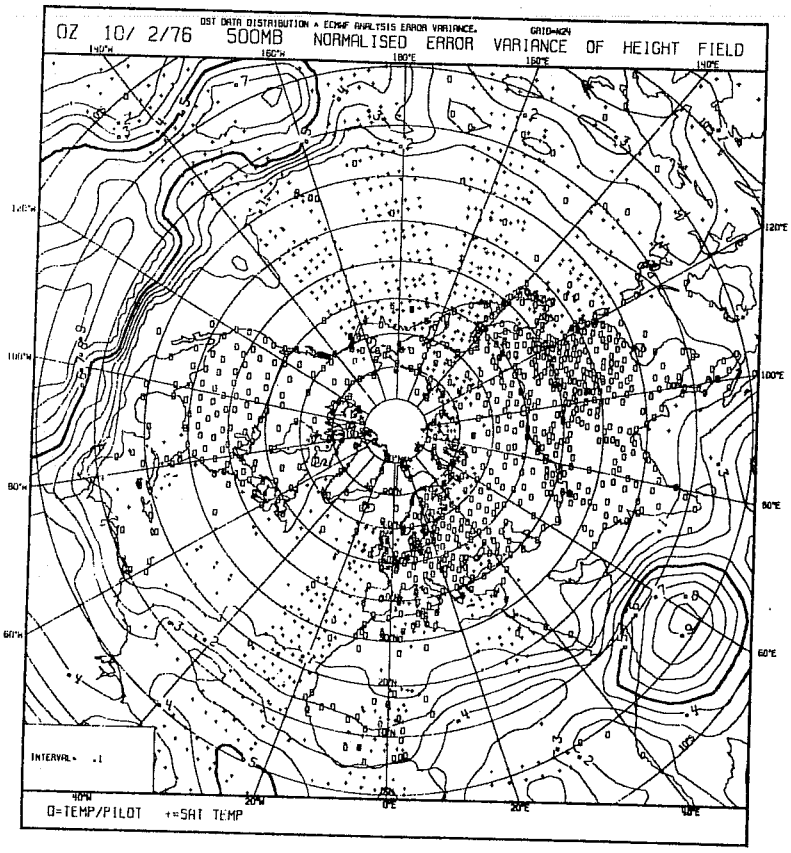


Fig. 9 DST surface observations, the estimated normalized error variance of the ECMWF 1000 mb geopotential height analysis, and the analysis box grid



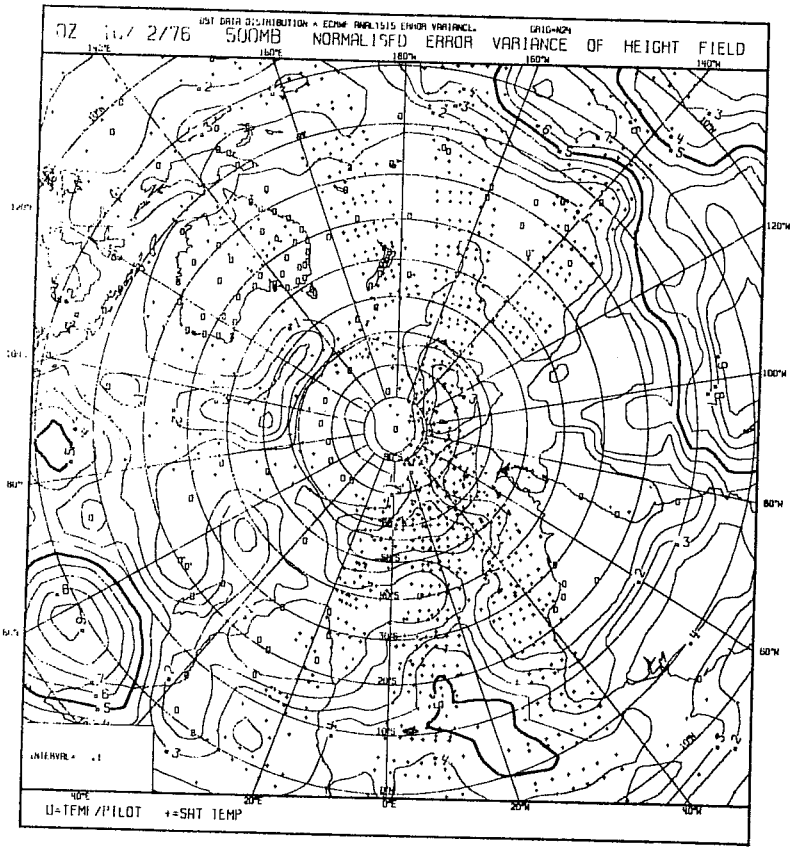
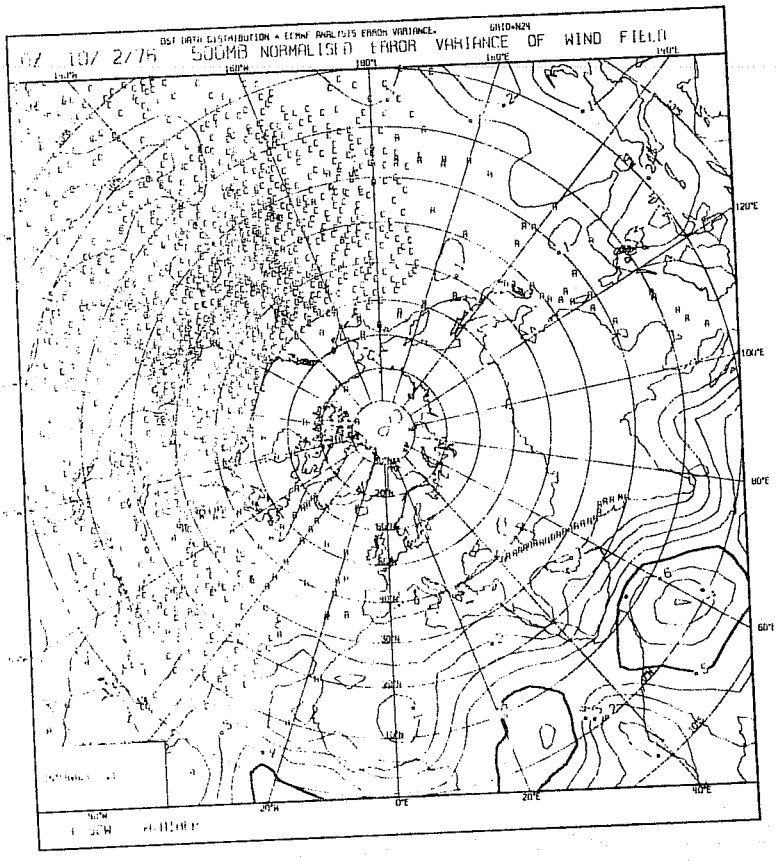


Fig. 10 DST soundings, and the estimated normalized error variance of the ECMWF 500 mb geopotential height analysis.



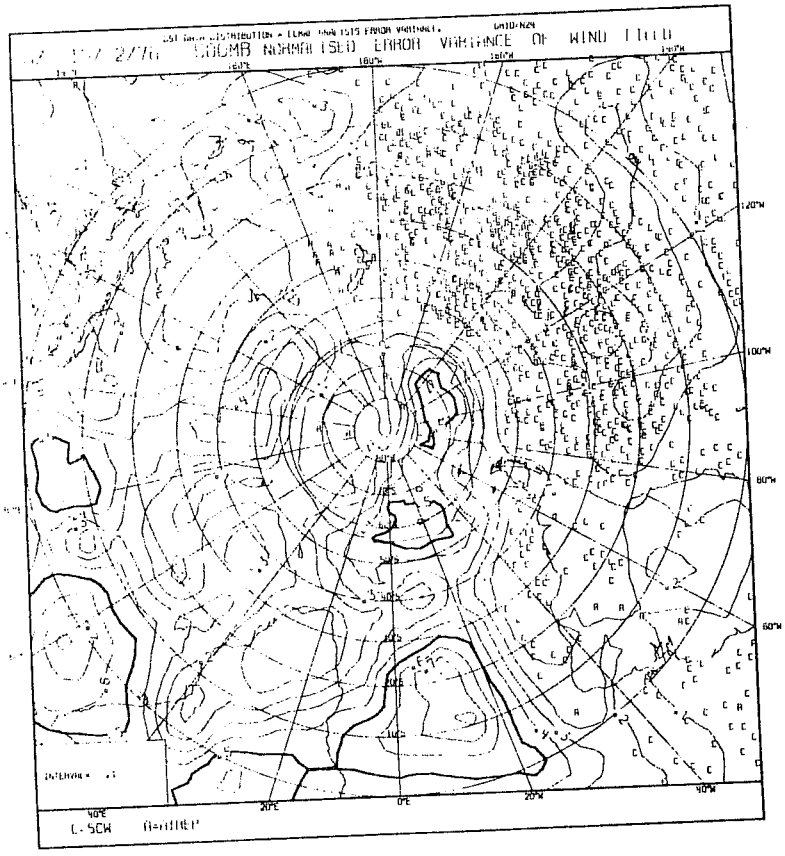
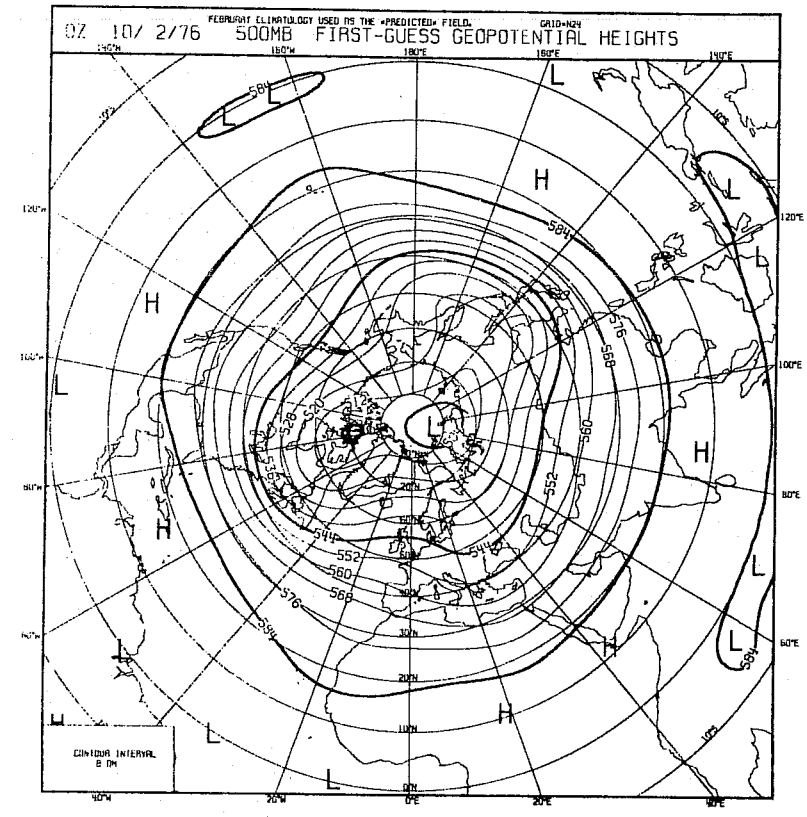


Fig. 11 DST single level upper air observations, and the estimated normalized error variance



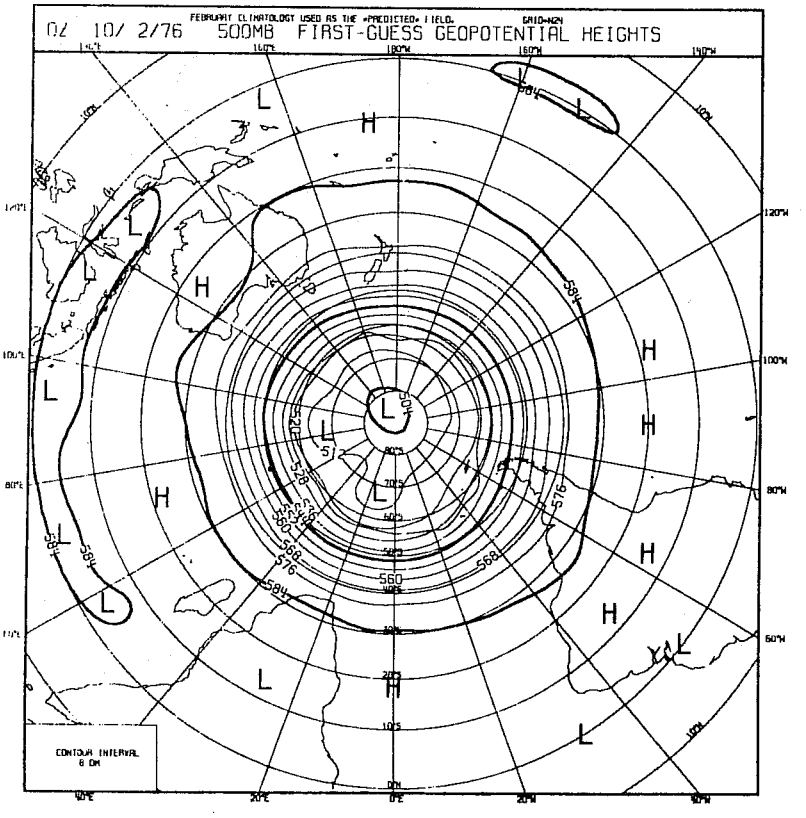
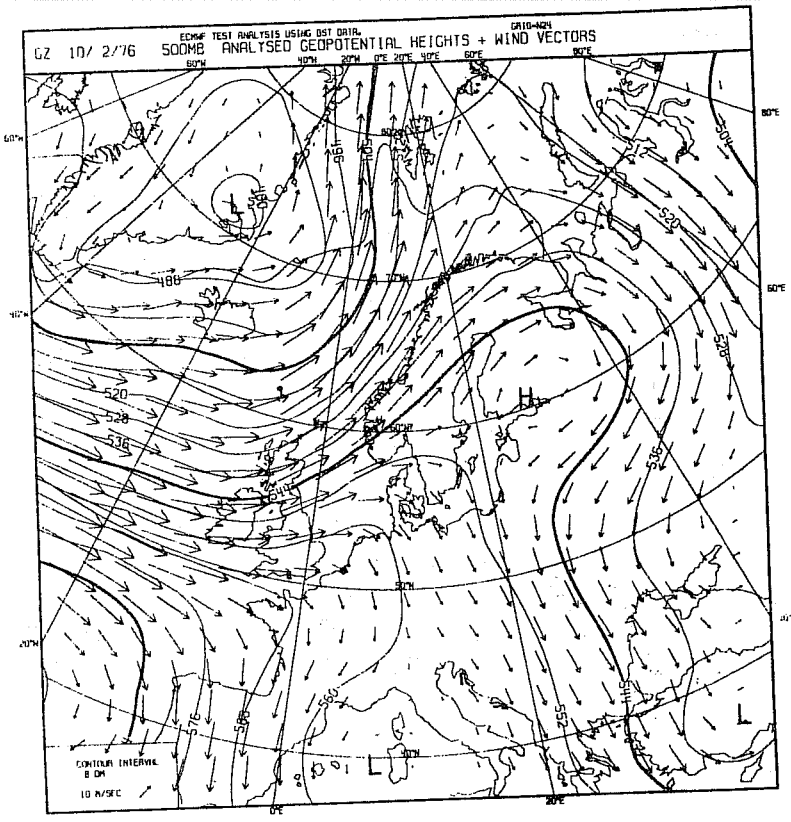


Fig. 12 500 mb geopotential heights of February climatology.



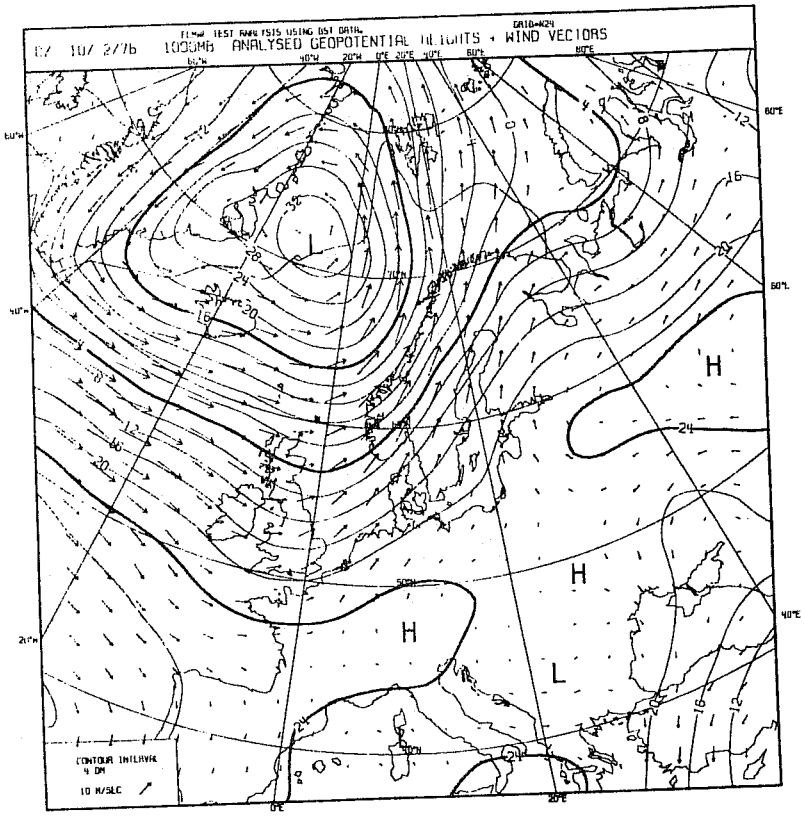
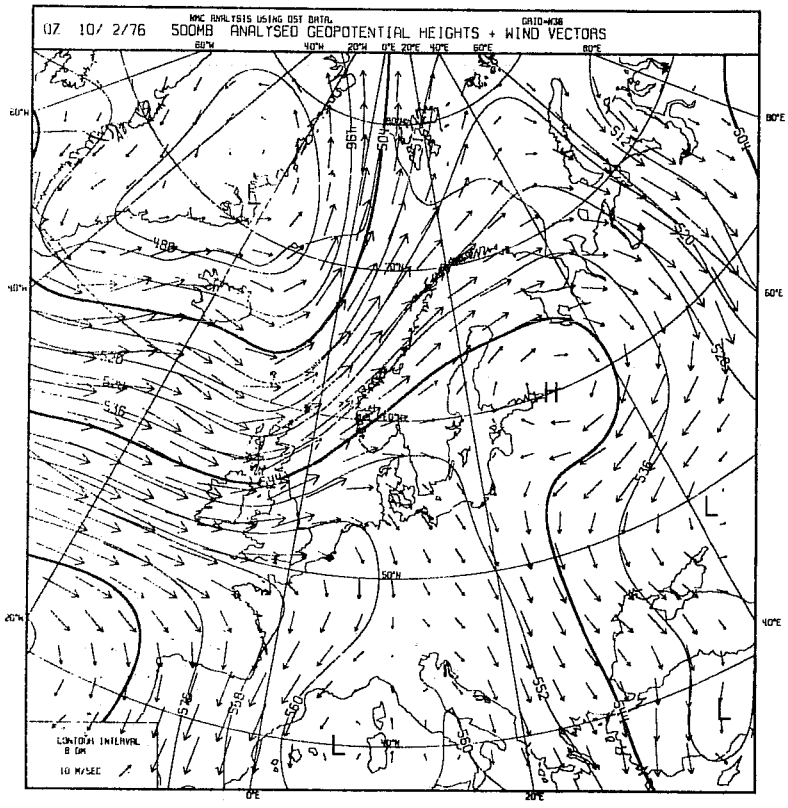


Fig. 13 Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from the ECMWF test analysis.



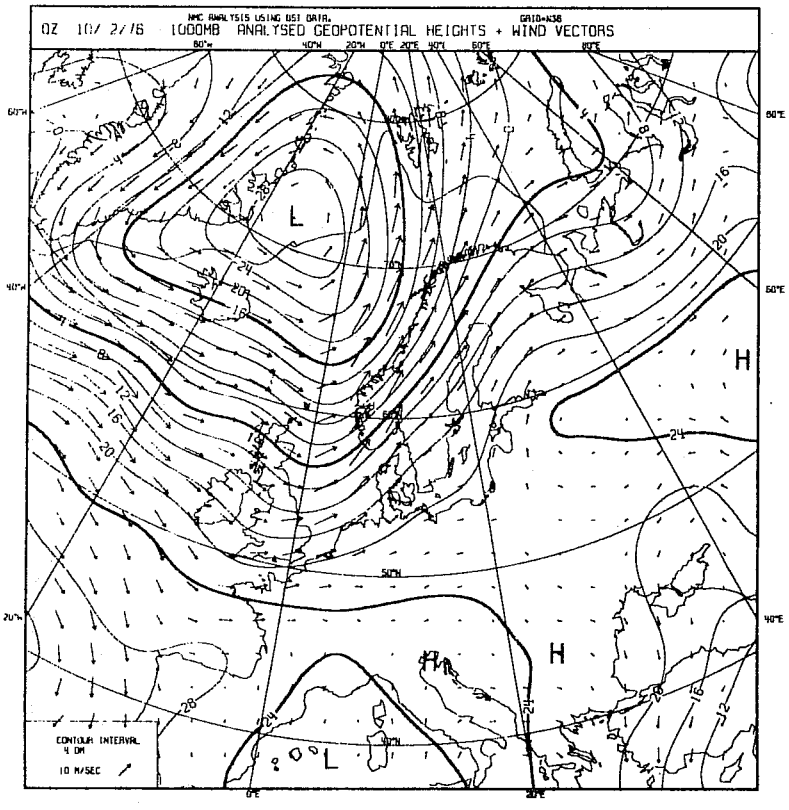
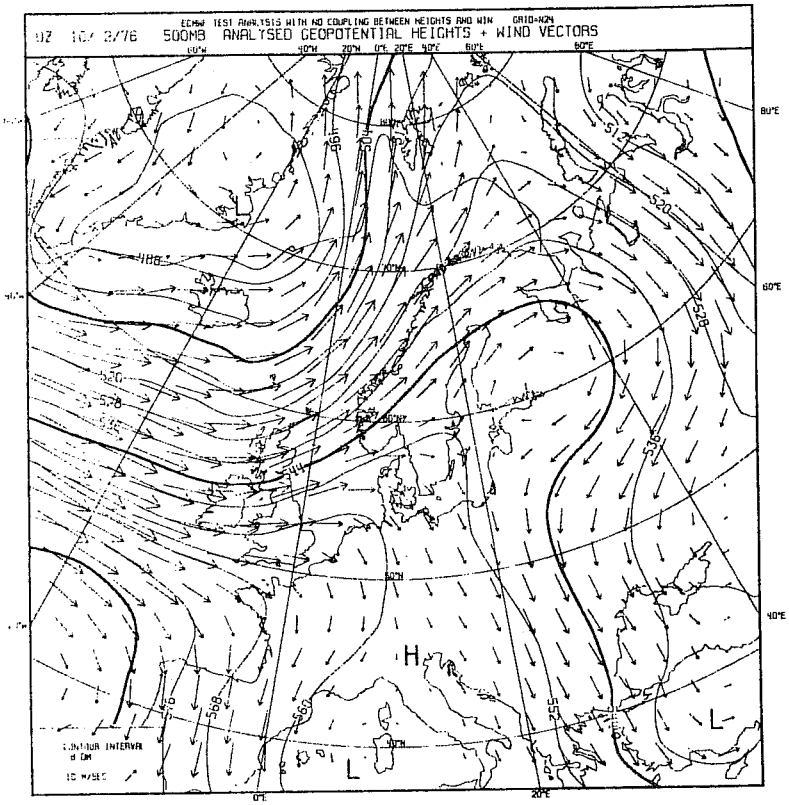


Fig. 14 Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from the NMC analysis.



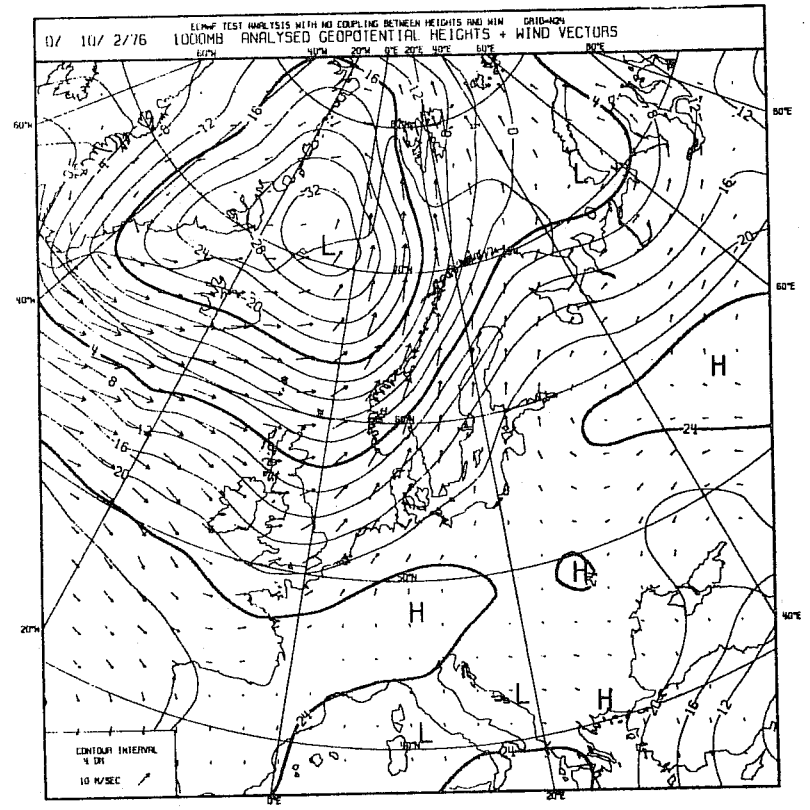


Fig. 15 Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from an ECMWF analysis with heights and winds uncoupled.

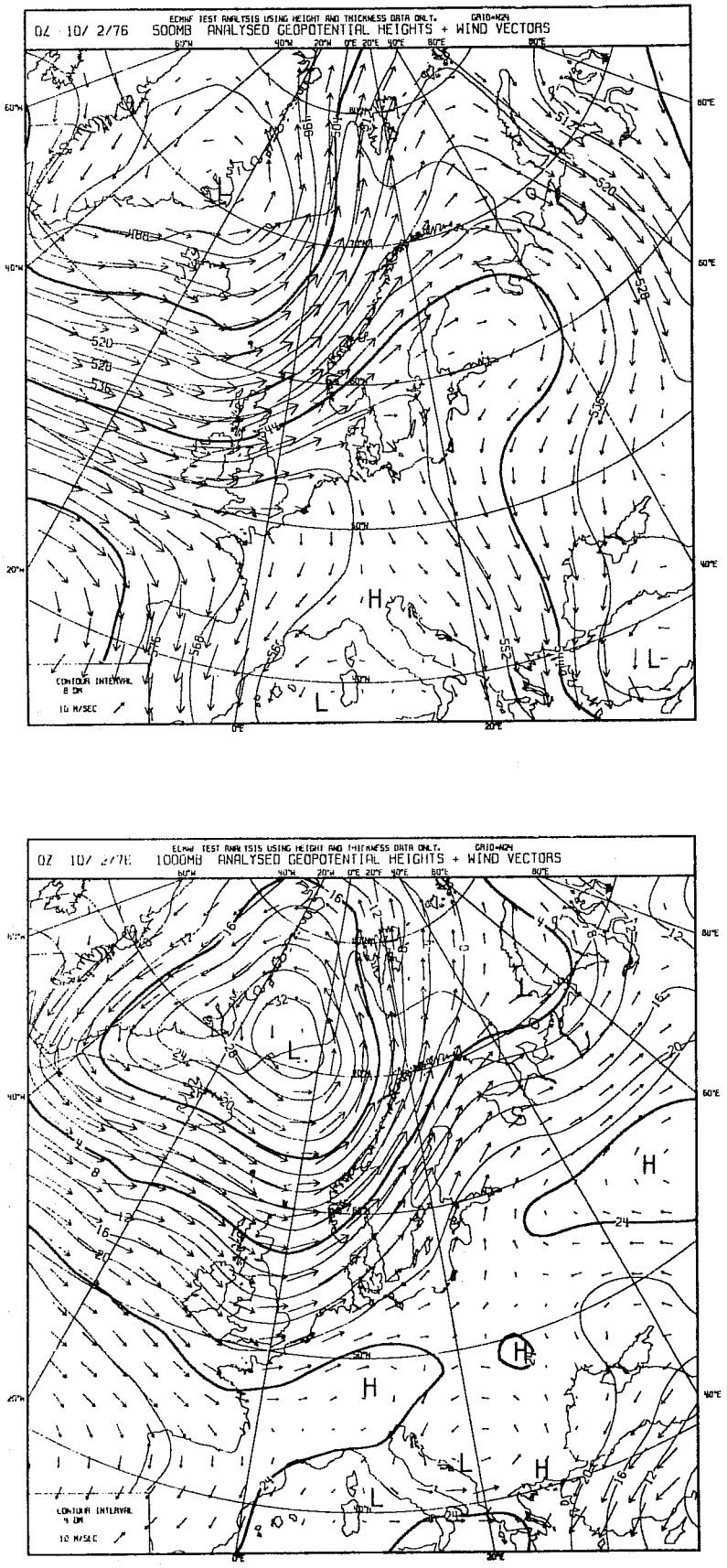
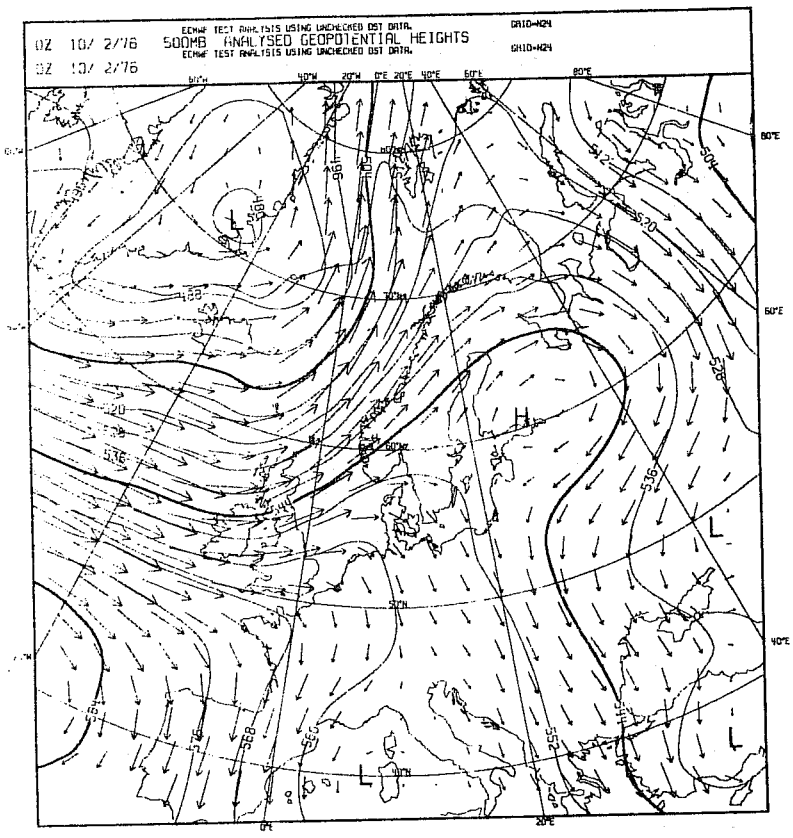


Fig. 16 Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from an ECMWF analysis using height data only.



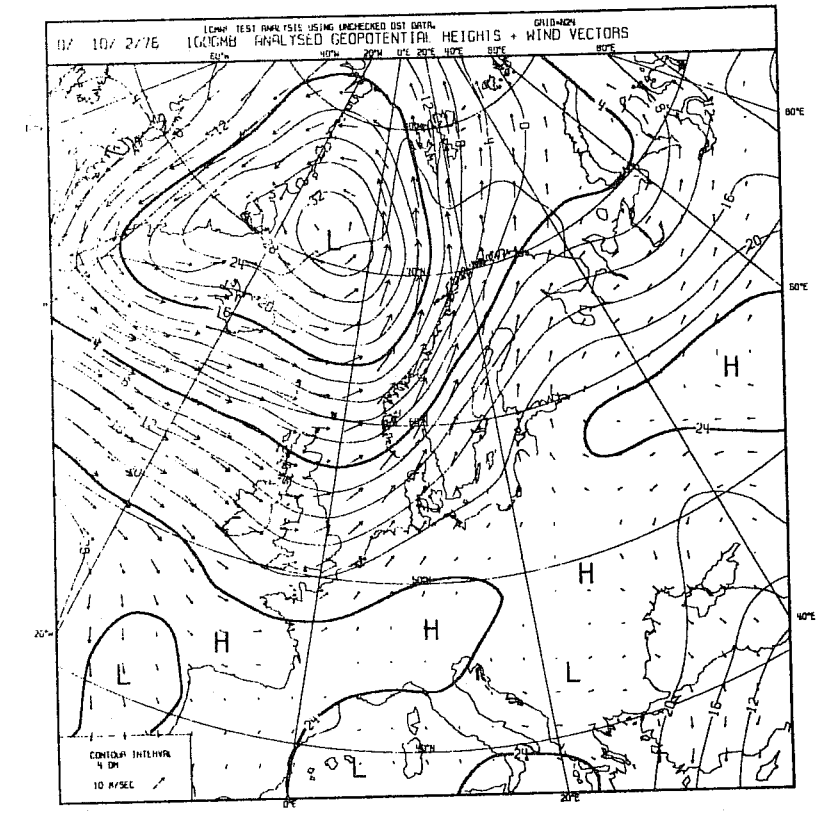


Fig. 17 Analysis for Europe of 1000 mb and 500 mb geopotential heights and winds, from an ECMWF analysis made without rejecting any

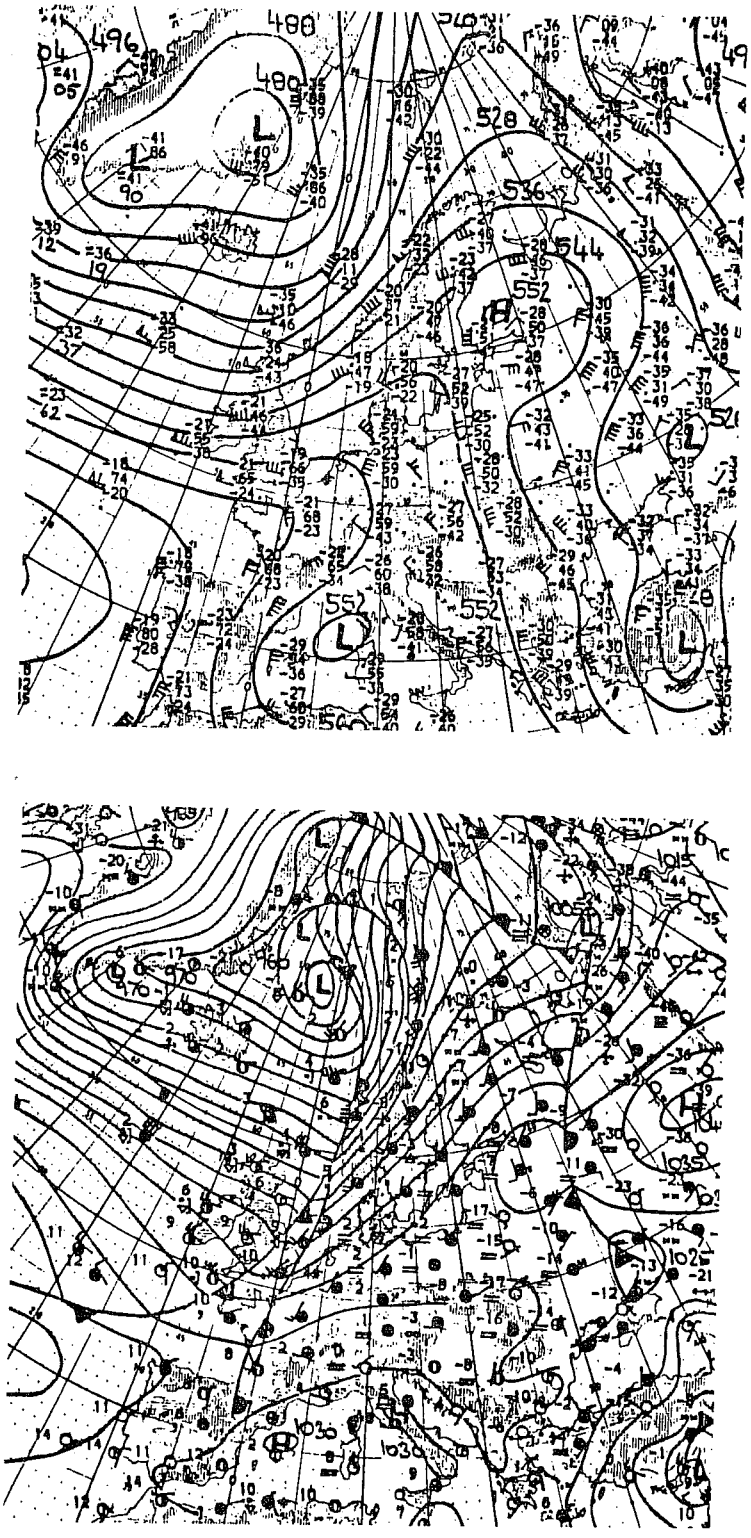


Fig. 18 Deutscher Wetterdienst hand drawn analysis of 500 mb geopotential height and mean-sea-level pressure.

EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

Research Department (RD)

Technical Report No. 6

No. 1 A Case Study of a Ten Day Prediction

No. 2 The Effect of Arithmetic Precision on some Meteorological Integrations

No. 3 Mixed-Radix Fast Fourier Transforms without Reordering

No. 4 A Model for Medium Range Weather Forecasting - Adiabatic Formulation -

No. 5 A Study of some Parameterisations of Sub-Grid Processes in a Baroclinic Wave in a Two-Dimensional Model

No. 6 The ECMWF Analysis and Data Assimilation Scheme - Analysis of Mass and Wind Fields

No. 7 A Ten Day High Resolution Non-Adiabatic Spectral Integration: A Comparative Study

No. 8 On the Asymptotic Behaviour of simple Stochastic-Dynamic Systems