INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460892 - I

H4.SMR/164 - 19

WORKSHOP ON CLOUD PHYSICS AND CLIMATE

23 November - 20 December 1985

SEMINAR/WORKSHOP 1982

INTERPRETATION OF NUMERICAL WEATHER PREDICTION PRODUCTS

S. TIBALDI
European Centre for Medium-range
Weather Forecasts
U.K.

SEMINAR/WORKSHOP 1982

INTERPRETATION OF NUMERICAL WEATHER PREDICTION PRODUCTS

Seminar : 13 - 17 September

Workshop : 20 - 24 September

September 1983

THE ECMWF ANALYSIS SYSTEM

Peter Lönnberg

European Centre for Medium Range Weather Forecasts

Reading, U.K.

1. INTRODUCTION

The ECMWF data assimilation system has been designed to provide initial states for the Centre's operational forecast model and to produce analyses from observations made during the First GARP Global Experiment (FGGE). The analysis system produces global fields of horizontal wind, geopotential height and humidity at six hourly intervals. Observed information from a six hour time window centered at the analysis time is combined with the forecast from the preceding analysis. The sea surface temperatures (SST) are regularly updated by SST analyses from the National Meteorological Center (NMC) in Washington.

Most of the observations available on the global telecommunication system (GTS) are used in the ECMWF analysis system. A brief description of the preprocessing of the observed information is given in Section 2.

The mass and wind analysis is based on the statistical interpolation method developed by Gandin (1963). Rutherford (1973) and Schlatter (1975) extended the method to a multivariate analysis of height and wind in two dimensions. The ECMWF scheme (Lorenc, 1981) is a complete 3-dimensional multivariate system in height, wind and thickness. This system has a horizontal resolution of 1.875° by 1.875° and 15 levels in the vertical (See Fig. 1). The statistical interpolation assigns weights to the first-guess, i.e. the six hour forecast, and any available observations in a statistically optimal way. The scheme gives consistent analyses for a wide variety of observation types and distributions. Temporal and spatial variations of the magnitude of the forecast errors are calculated in the system. The statistical

ANALYSIS	PREDICTION
ϕ,u,v,q(for p≥300)	T,u,v,q
p(mb)	σ
10	0.025(σ_1)
20	0.077
30	0.132
50	0.193
70	0.260
100	0.334
150	0.415
200	0.500
250	0.589
300	0.678
400	0.765
500	0.845
700	0.914
850	0.967
1000	0.996(σ_{15})

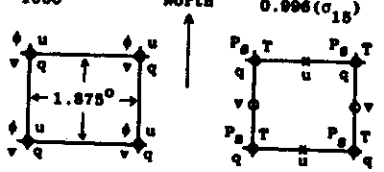


Fig. 1 The horizontal grid and the vertical levels of the ECMWF analysis and forecast systems.

interpolation method also offers an elegant technique of quality control of observations. The mass and wind analysis is discussed in detail in Section 3. A global balance between mass and wind fields is subsequently achieved by non-linear normal mode initialisation (Wergen,1982).

Section 4 describes the humidity analysis which is based on a two-dimensional distance weighted correction method (Lorenc and Tibaldi, 1979). The analysis of the water vapour content of the five lowest layers (Fig. 1) is formed as a combination of the predicted humidity and estimates of moisture from observations. The SST analysis (Section 5) is basically a transformation of the NMC SST analysis to the ECMWF grid.

Prior to the analysis, the model fields must be transformed to the analysis coordinates. The corrections made by the analysis are transformed back to model coordinates and added to the six hour prediction. These interpolation

methods are discussed in Section 6.

Some aspects of using the fields produced by the assimilation system for statistical interpretation of the weather are presented in the summary (Section 7).

2. OBSERVATIONS

All appropriate reports on the GTS are collected in the ECMWF Reports Data Base and ordered in groups covering a period within ± 3 hours of the analysis time. The observation types, which are presented to the analysis, are listed below with typical volumes per main synoptic hour:

Surface land and sea reports (SYNOPs/SHIPs ; 5000/800)

Radiosonde and pilot reports (TEMPs/PILOTs ; 600/200)

Satellite thickness reports (SATEMs ; 1000)

Satellite wind reports (SATOBs ; 800)

Aircraft reports (AIREPs and ASDARs ; 500)

Drifting buoy reports (DRIBUs ; 100)

Australian bogus reports (PAOBs ; 300)

The humidity analysis extracts information from TEMPs and SYNOPs only.

The observed information is checked at several stages. The ECMWF Reports Data Base checks code formats, the internal consistency between meteorological parameters within one observation, and compares the values against climatological extremes.

Each information item is presented to the analysis as a deviation from the first guess in nondimensional form, i.e. normalized by the expected forecast error for that variable and position. Four types of data are recognised by the statistical interpolation scheme, i.e. the east-west and the north-south wind components, the geopotential height at an analysis level, and the thickness between two neighbouring analysis levels. The analysis levels of the ECMWF scheme are shown in Fig. 1. Off-level observations are interpolated or extrapolated in the vertical to the nearest analysis level. Surface data are processed using the corresponding model surface fields. The sea surface pressure deviations are converted to height deviations using either the observed or the first-guess temperature and shifted to the closest analysis level. A similar procedure is applied to other pressure and height reports at non-analysis levels. The observed 10 m wind deviations are similarly assigned to the nearest analysis level. The frictional effects on the 10 m winds are assumed to be the same as in the first-guess. Winds from land stations above 500 m are not used unless the station is marked as important.

Orographic effects and the lack of detailed knowledge of land surface radiative properties degrades the quality of satellite wind and temperature measurements. Consequently, we exclude, over land, satellite thicknesses below 100 mb and all cloud winds.

An estimate of the observational error (as a function of pressure) is ascribed to each observing system and observed variable. The observation errors are assumed to be random, except for radiosonde heights and satellite temperatures. The radiosonde height errors are assumed to be vertically correlated. This enhances the accuracy of the reported gradient of height, i.e. temperature. The satellite thickness measurement errors are assumed to be both horizontally and vertically correlated.

Asynoptic surface pressure observations are corrected to the analysis time by the reported pressure tendency. For ships a correction is applied for its

movement, using the forecasted pressure gradient field. The error assigned to an asynoptic observation is the sum of the measurement error and the persistence error:

$$E'_{obs} = (E_{obs}^2 + E_{pers}^2)^{1/2} \qquad (1)$$

The growth rate of the persistence error is a function of latitude and season.

Each observation is then compared against the first-guess. A reliability flag is assigned to it as a function of the magnitude of the departure compared to the expected standard deviation (std) of the difference. An observation is rejected if the departure exceeds eight stds and no close report can support it. For example, the limit for rejection is about 30 mb for a ship pressure report in the Mid-Atlantic.

Frequently, compatible and close observations with small departures from the guess field are found. This redundant information is averaged to "super-observations" and presented in the compressed form to the analysis.

3. DESCRIPTION OF THE ECMWF MASS AND WIND ANALYSIS

3.1 Basic method

Statistical interpolation is a powerful technique of combining a first-guess and observations with different error characteristics. Linear constraints such as geostrophy can be built into a multivariate statistical interpolation scheme to give consistent analyses for several meteorological quantities. An efficient data checking method can also be devised in a statistical interpolation scheme.

The equations of statistical interpolation have been derived in many papers, e.g Lorenc (1981), but for completeness I will repeat the derivation in this section.

The following notations are used for the departures from the "truth"

$$a = A - T \tag{2a}$$

$$o = O - T \tag{2b}$$

$$p = P - T \tag{2c}$$

where T is the "true" value and A, O and P are the analysed, observed and predicted values, respectively. The "true" value is assumed to represent scales we are interested in or capable of analysing.

The associated errors are defined by

$$E^a = \langle a^2 \rangle^{1/2} \tag{3a}$$

$$E^o = \langle o^2 \rangle^{1/2} \tag{3b}$$

$$E^p = \langle p^2 \rangle^{1/2} \tag{3c}$$

The angle brackets indicate ensemble means over several similar realisations. As a consequence of this definition of the "truth", the errors also include the atmospheric variability of the unresolvable scale. The resolution of the ECMWF analysis system is of the order of several hundred kilometers.

The six hour forecast generally approximates the actual atmospheric state quite well and a substantial reduction in the analysis error is achieved from using a prediction as a guess instead of climatology. The observed departures from the six hour forecast are the quantities analysed. The analysed departure from the background field is calculated by combining the observed deviations in a statistically optimal way and adding it to the guess. All departures are nondimensionalised by the assumed error of the guess field (3c) to simplify the derivation and application of the analysis equations. The analysis is formed as a linear combination of all influencing observations.

$$\frac{A_k - P_k}{E_k^p} = \sum_{i=1}^{N} w_{ki} \frac{O_i - P_i}{E_i^p} \tag{4}$$

The subscript k denotes the analysis quantity, defined by level, horizontal position, and variable type. Similarly, subscript i defines the type and spatial position of observation i.

Define

$$\alpha^o = o/E^o$$

$$\alpha^p = p/E^p$$

$$\alpha^a = a/E^a \tag{5}$$

$$\epsilon^o = E^o/E^p$$

$$\epsilon^a = E^a/E^p$$

and insert into (4)

$$\alpha_k^a \epsilon_k^a = \alpha_k^p + \sum_{i=1}^{N} w_{ki} (\alpha_i^o \epsilon_i^o - \alpha_i^p) \tag{6}$$

Next, we want to determine the weights w_{ki} in such a way that the squared analysis error is minimized for an ensemble of similar situations, i.e. we take the ensemble mean of the square of (6). We also assume that the prediction and observation errors are uncorrelated:

$$\epsilon_k^{a\,2} = 1 - 2 \sum_{i=1}^{N} w_{ki} \langle \alpha_k^p \alpha_i^p \rangle$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ki} (\langle \alpha_i^p \alpha_j^p \rangle + \epsilon_i^o \langle \alpha_i^o \alpha_j^o \rangle \epsilon_j^o) w_{kj} \tag{7}$$

With vector (_) and matrix ( ) notations (7) can be written in the following form

$$(\varepsilon_k^a)^2 = 1 - 2\,\underline{W}_k^T\,\underline{P}_k + \underline{W}_k^T\,\underline{\underline{M}}\,\underline{W}_k \qquad (8)$$

where $\underline{W}_k$ is the column vector of weights w_{ki}, $\underline{\underline{P}}$ is the prediction error correlation matrix $[\langle\alpha_i^P\,\alpha_j^P\rangle]$, $\underline{\underline{O}}$ is the scaled observation error correlation matrix $[\varepsilon_i^o\langle\alpha_i^o\alpha_j^o\rangle\varepsilon_j^o]$ and $\underline{\underline{M}} = \underline{\underline{P}} + \underline{\underline{O}}$. $\underline{P}_k$ is the prediction error correlation vector between the gridpoint k and all observations; the components of the vector are $\langle\alpha_k^P\,\alpha_i^P\rangle$. We minimize (8) with respect to w_{ik}, i.e. we solve a set linear equations $\partial(\varepsilon_k^a)^2/\partial w_{ik} = 0$. The "optimal" weights are then

$$\underline{W}_k = \underline{\underline{M}}^{-1}\,\underline{P}_k \qquad (9)$$

The analysis and the error associated with it are

$$\frac{A_k - P_k}{E_k^P} = \underline{P}_k^T\,\underline{\underline{M}}^{-1}\,\underline{B} \qquad (10)$$

$$(\varepsilon_k^a)^2 = 1 - \underline{W}_k^T\,\underline{P}_k \qquad (11)$$

where $\underline{B}$ is the observation vector.

Formula (9) shows that the weights are independent of the observed values. The weights are determined by the geographical distribution of the selected observations and the assumed structures of the forecast and observation errors.

Formulae (9) and (10) show that the analysis depends on the horizontal grid and vertical coordinate system only through the prediction error correlations between the observation points and the gridpoint. Thus the analysis can be projected on any coordinates for which $\underline{P}_k$ is defined.

3.2 Forecast error statistics

The ECMWF system analyses simultaneously three fields (u, v and z) using four types of data (u, v, z and dz) in three dimensions. Consequently, prediction error correlations must be specified 3-dimensionally for all possible combinations of variables. The correlations are assumed to be separable in the horizontal and vertical, and are formed as products of a horizontal correlation function and a vertical correlation function. The horizontal correlation is modelled as a continuous function and the vertical correlation is given discretely, by a matrix, for all combinations of analysis levels.

Three constraints are modelled in the forecast errors. It follows from (10) that the analysis obeys the physical relationships built into the forecast error covariances in a given domain, only if a single observational data set is used for the analysis of all variables and points in that domain. If different variables (or gridpoints) are analysed with different sets of observations in a given domain, then the constraints will not be applied effectively. This means that all constraints are applied locally.

Currently, we ignore the error in the predicted divergent wind and determine the wind correlations (u-u, u-v, v-v) from an isotropic streamfunction correlation. This results in locally non-divergent wind changes.

The geopotential height and the wind are assumed to be in approximate geostrophic balance. By defining a correlation between the height and the streamfunction fields and assuming identical structures of these fields, we can determine the height-wind (z-u and z-v) correlations. A full coupling would make the height and wind increments geostrophic. In extratropical areas (poleward of 30° latitude) a height- streamfunction correlation of ± 0.95 is used. This gives in a situation of one height observation a wind analysis that is 95% of the geostrophic wind. For larger data amounts this

allows for large departures from geostrophy if the data defines it. The geostrophic balance is relaxed to zero at the Equator.

An illustration of a multivariate analysis is given in Fig. 2, where only one height observation, at the centre of the area, is available. A circular height analysis results from the isotropic z-z correlation. The analysed winds are non-divergent and have the direction of the geostrophic wind. The wind speed depends on the specified correlation between geopotential and streamfunction.

The second idealized example demonstrates the sensitivity of the analysis to the specified geostrophic coupling. The "observed" winds are as shown in Fig. 3a and at the same positions zero heights are specified. The resulting analysis is shown for three different values of the height-streamfunction correlation. If strict geostrophy is imposed, the analysis (Fig. 3b) is a bad compromise of the geopotential and wind "observations". A relaxation of the coupling to 0.95 (Fig. 3c) gives an analysis that returns a substantial part of the original ageostrophic wind. By decreasing the height-streamfunction correlation to 0.5 (Fig. 3d), an analysis faithful to both geopotential and wind data is obtained. However,a weak coupling is undesirable in extratropical latitudes wh e the flow is approximately geostrophic.

The third constraint on the analysis relates the thicknesses hydrostatically to the heights.

The horizontal height and streamfunction correlations are modelled by an exponential function of distance r

$$\mu(r) = e^{-1/2 \left(\frac{r}{b}\right)^2} \qquad (12)$$

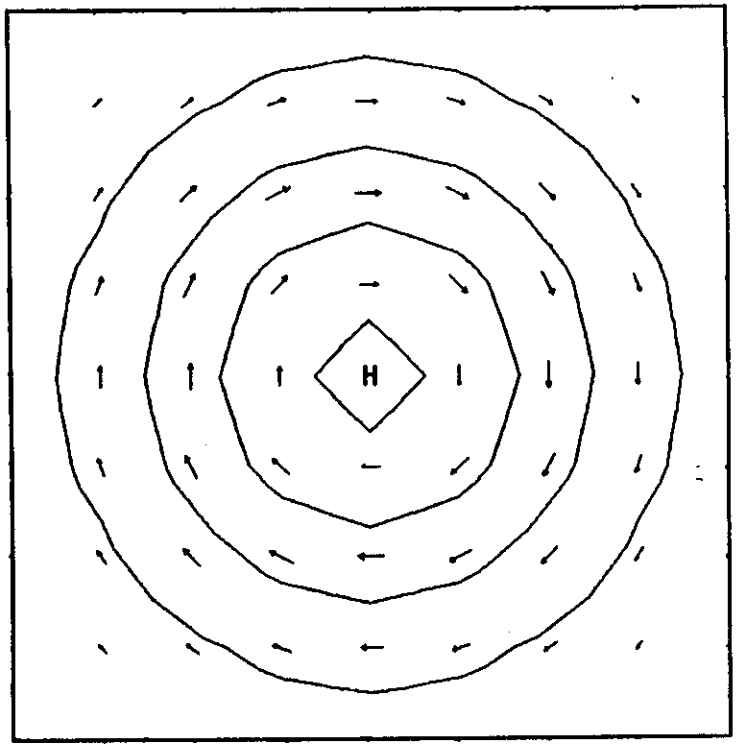


Fig. 2 A multivariate height and wind analysis using one height observation at the centre of the area.

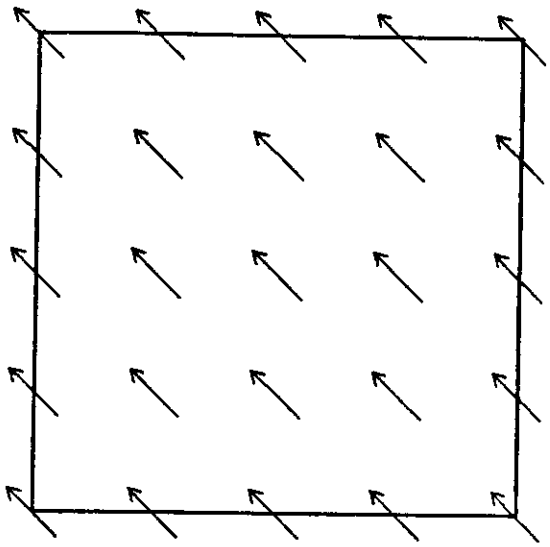


Fig. 3a The observed wind field in an idealized multivariate analysis with different height-streamfunction correlations. At each wind observation position a zero height datum is specified.

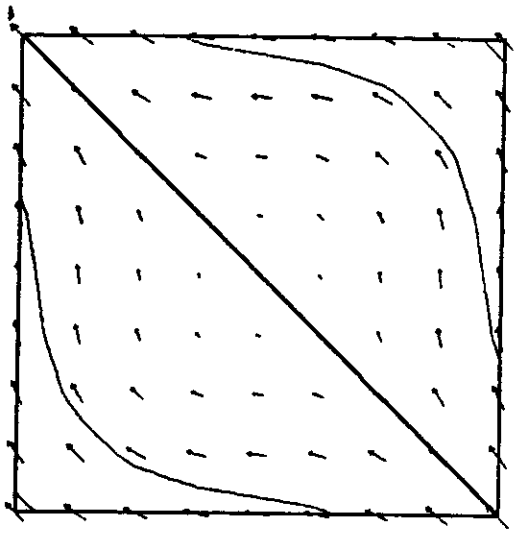


Fig. 3b The height and wind analysis from the observed fields of Fig. 3a with a height-streamfunction correlation $Y = 1.0$.

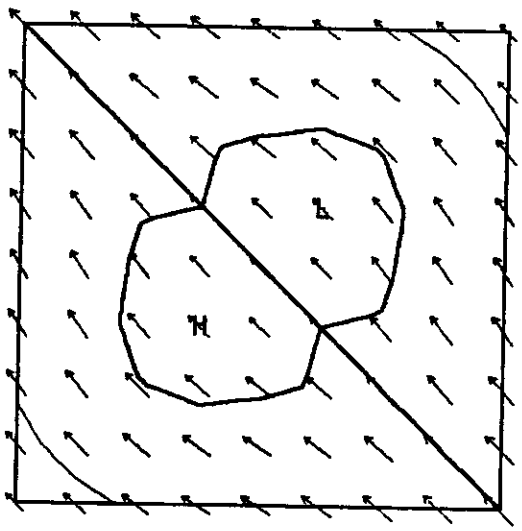


Fig. 3c Same as 3b but $Y = 0.95$.

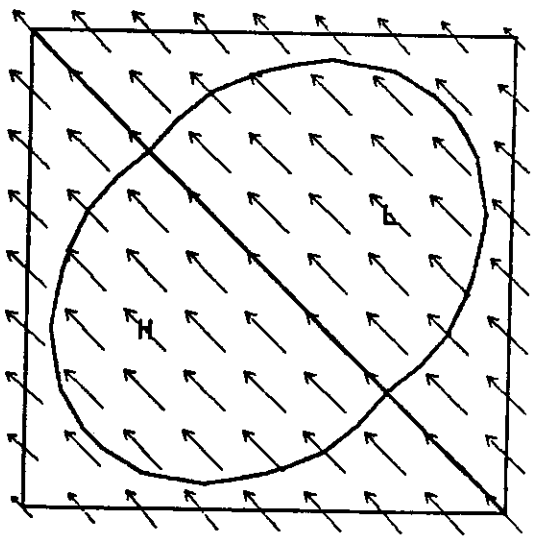


Fig. 3d Same as 3b but $Y = 0.5$.

where b defines the horizontal scale length of the forecast error. In the Northern Hemisphere b=600 km and in the Southern Hemisphere b=900 km.

The vertical correlations are based on several studies; the extratropical ones mainly on the thesis of Hollett (1975). Fig. 4 shows the impact in the vertical of one height observation at 1000 mb. The observation gives an almost constant (barotropic) increment up to 250 mb and above the tropopause its influence decreases rapidly. This is also reflected in the thickness as a weak change of the mean tropospheric temperature. A pressure fall (rise) at the surface causes a warming (cooling) at the tropopause and in the lower stratosphere.

The horizontal correlations are fixed in time, but the vertical correlations have an annual variation. The magnitude of the 6 hr. forecast error depends on the accuracy of the previous analysis. The analysis error is estimated according to (11) for all variables and levels on a coarse grid. This error is assumed to grow to the error of a random state in 36 hours, i.e. a sixth of the difference between the random and the analysis error is added to the analysis error. This apparently rapid error growth is caused not only by the pure forecast error but also from the fact that the analysis error is an underestimate of the "true" analysis error and that the initialisation moves the analysed state away from the "truth".

Fig. 5 shows the estimated forecast error of the 500 mb height on 1 November 1981 12 GMT. The North American and Asian continents as well as the Pacific have high forecast errors as a consequence of very few upper air observations in the previous analysis. The Equatorward decrease in the forecast error reflects the smaller climatological variability in the tropics.

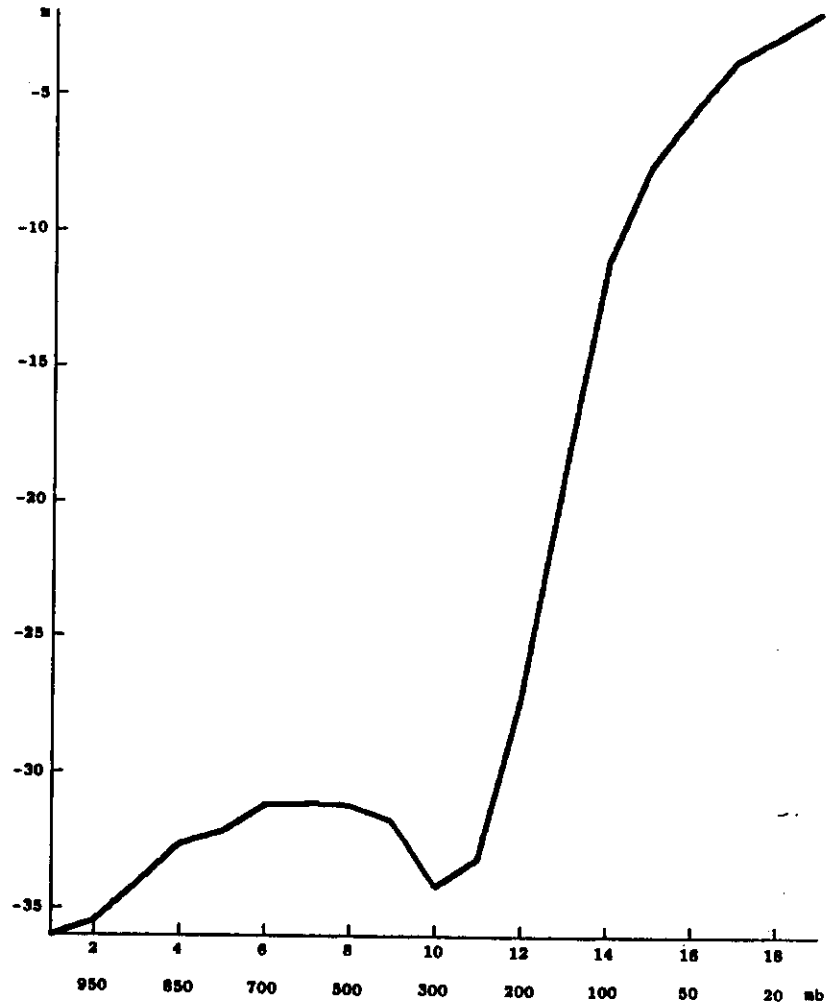


Fig. 4 The height analyses at different pressure levels from one height observation (-40 m) at 1000 mb. The assumed normalized observation error is 0.4.

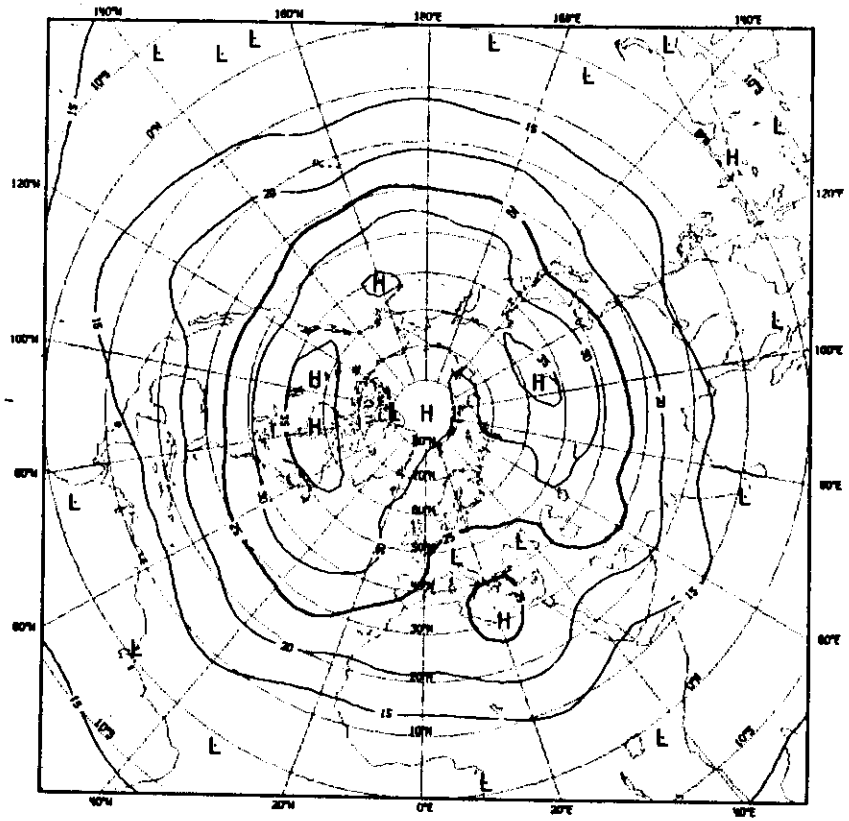


Fig. 5 An example of the estimated forecast error for 500 mb height at 12 GMT 1 November 1981. Unit: metre.

The estimated analysis error for the corresponding analysis, which has a normal data coverage, is shown in Fig. 6. The roughness of the field is caused by variations in the data coverage. Only a small reduction of the error of the first-guess is achieved over the oceans. Over the continents the error drops to a small fraction of the six hour forecast error.

3.3 Data selection and checking

After a coarse preliminary data check in the pre-analysis, each remaining item of information is checked by the statistical interpolation method. An analysis is calculated at each observation to be checked but without that observation. Averaged over a large ensemble, the mean of the squared difference between observations and independent analyses is the sum of the squares of the observation and analysis errors. An observation is most likely incorrect if its difference to the independent analysis exceeds the expected difference by a certain factor, say 4. Thus the observation is rejected if the following inequality is satisfied

$$(a_k^o - a_k^a) > c_1^2 \, (\epsilon_k^{a\,2} + \epsilon_k^{o\,2} + c_2) \qquad (13)$$

The constants c_1 and c_2 have been set to 4 and 0.1, respectively. The constant c_2 is added to the rejection limit to inflate the estimate of the analysis error which is based on the assumption that the statistics are perfect.

The discrete decision to reject or accept an observation can have enormous consequences for medium-range weather forecasting due to downstream propagation and amplification of the initial error (Cats, 1981).

The design of an efficient data selection algorithm is very difficult. Most analysis systems tend to chose no more than 10 observations for the analysis

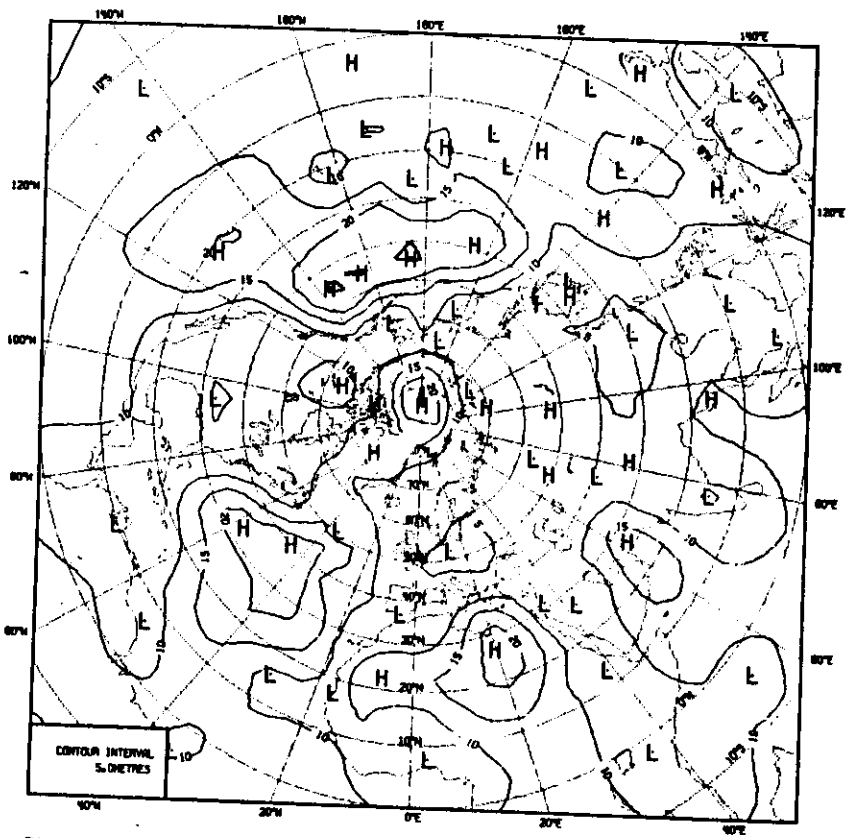


Fig. 6 The estimated analysis error for 500 mb height at 12 GMT 1 November 1981. Unit: metre.

of a particular point. On a fast vector computer it is practicable to invert matrices up to an order of 100-200. The ECMWF scheme selects all observations in an analysis volume and the selection proceeds outwards from the volume to a distance of 1 b-length (formula 12) until 5 data have been found for each analysis variable and level. An analysis volume is defined by the gridpoints and levels for which the analysis is calculated simultaneously using the same data set.

3.4 Organisation of the computation

The mass and wind analysis consists of three stages:-

- Preprocessing of observations as described in Section 2
- Data checking by the statistical interpolation method
- Calculation of the changes (increments) to the first-guess

As the matrix inversion is computationally expensive and the data selection for neighbouring gridpoints would in general be very similar, the analysed values for several gridpoints are calculated simultaneously.

The globe is divided into rectangular areas (boxes) of approximately equal size, 660 by 660 km (See Fig. 7). In data dense regions the atmosphere is divided into three layers in the vertical. For each such volume the appropriate data are selected and $\underline{\underline{H}}^{-1}\,\underline{\underline{B}}$ (formula 10) is solved. The analysis can then be projected on any grid by multiplication with $\underline{\underline{P}}_k$. In the data checking phase the analysis is only evaluated at the locations of the observations in the volume.

As the analysis might change dramatically between one box to its neighbour due to different data selection, analysis increments are also calculated to

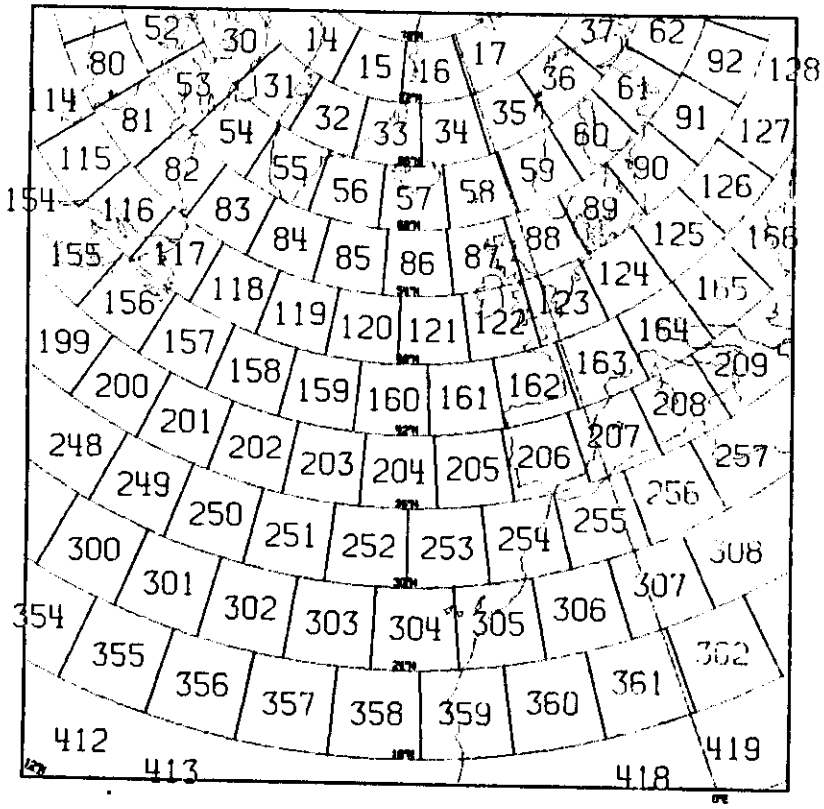


Fig. 7 The box structure of the ECMWF analysis system.

gridpoints in the neighbouring boxes. The final analysis increment is formed as a weighted mean of the different estimates of the analysis at that point.

3.5 Example of a mass and wind analysis

An example of height and wind increments fields produced by the statistical interpolation method is given in Fig. 8a. The corrections made by the analysis are generally less than 3-4 mb in the Northern Hemisphere and the winds and heights are in approximate geostrophic balance. The final height- and wind analysis is shown in Fig. 8b. The corresponding changes to the divergence are presented in Fig. 9a. A good agreement can be found between the divergence and the pressure tendency fields. Although the horizontal correlations are locally non-divergent the magnitude of the changes in the large-scale divergence field represent an appreciable part of the total analysed divergence (cf Figs. 9a and b).

4. HUMIDITY ANALYSIS

For the analysis of humidity one can justify a less sophisticated scheme than statistical interpolation. Consequently a two-dimensional correction method is considered adequate for the humidity analysis. As for the mass and wind analysis, the six hour forecast from the previous analysis provides the guess field which is modified by available observations. In anticipation of good quality satellite water vapour measurements, the vapour content between two analysis levels was chosen as the analysis variable. At present, the scheme extracts information only from TEMPs and SYNOPs. The pressure levels that define the analysis layers are 1000, 850, 700, 500, 400 and 300 mb.

The structure of the error of the forecasted humidity is poorly known. The forecast is generally of good quality in extratropical areas, but in the tropics where the coupling between mass and wind field is weak, updating of the predicted humidity field is important.

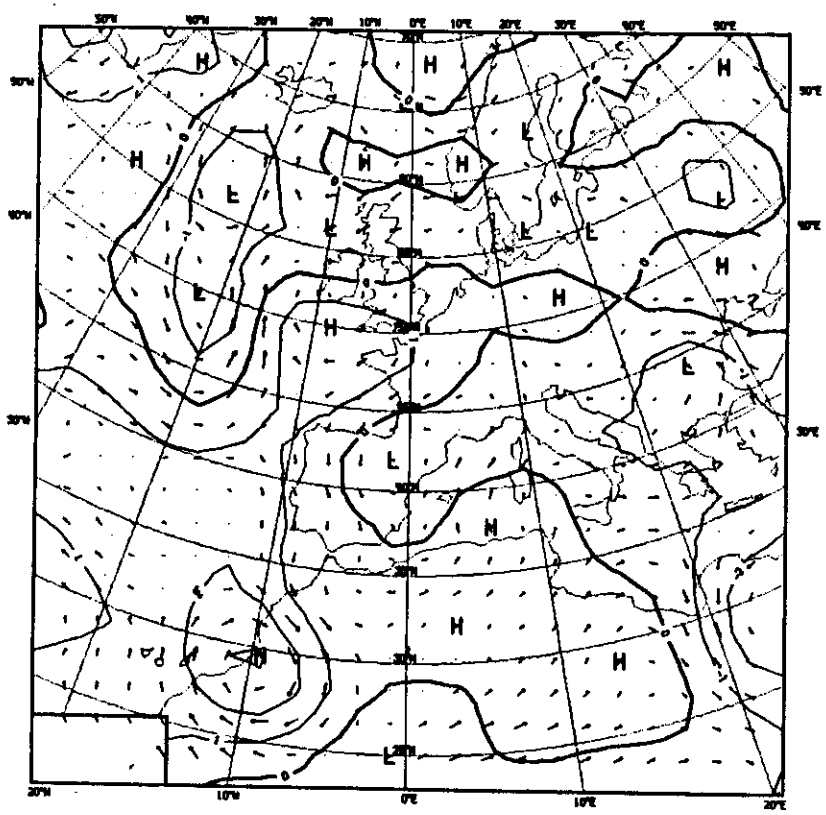


Fig. 8a The height and wind increments fields at 1000 mb at 12 GMT 1 November 1981. Unit for height is dm.

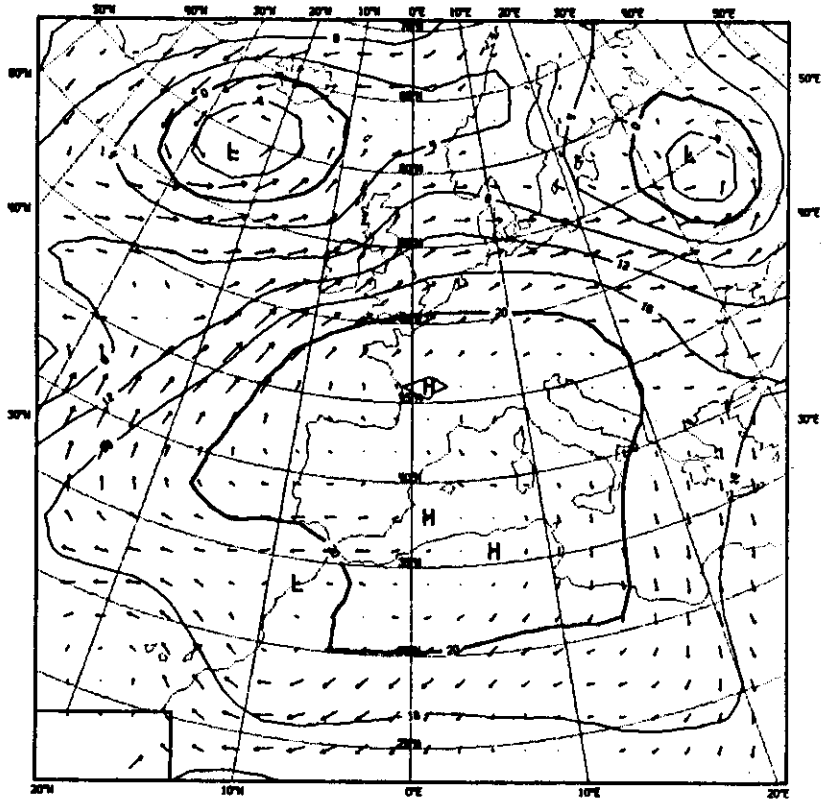


Fig. 8b Same as Fig. 8a but for height and wind analyses. The length of a wind vector is 0.4 of its magnitude in Fig. 8a.

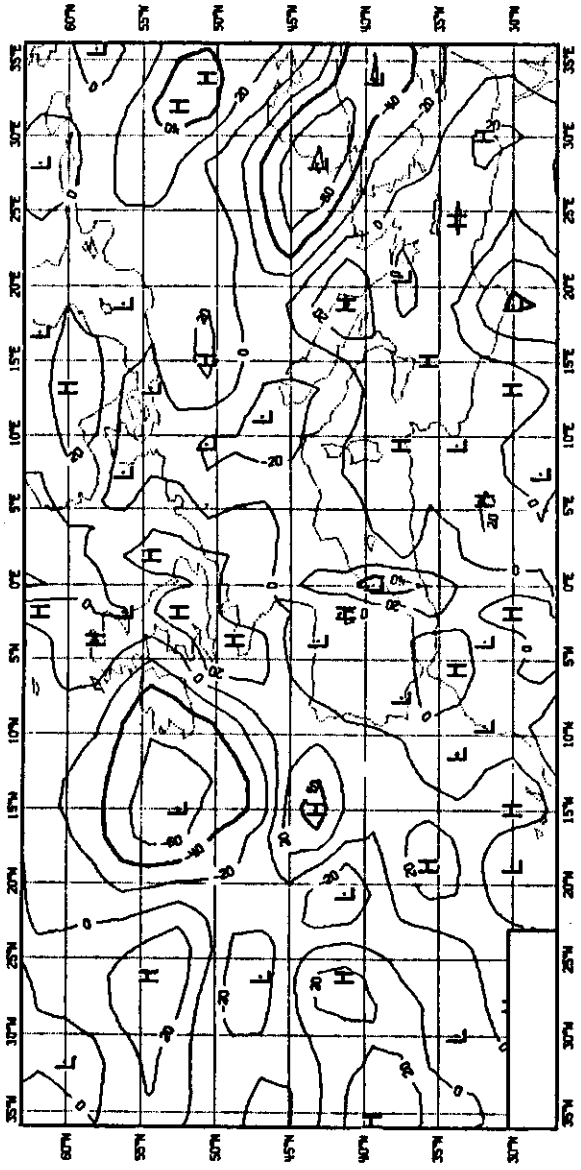


Fig. 9a The divergence of the wind increments in Fig. 8a. Units: 10^{-7} s^{-1}.

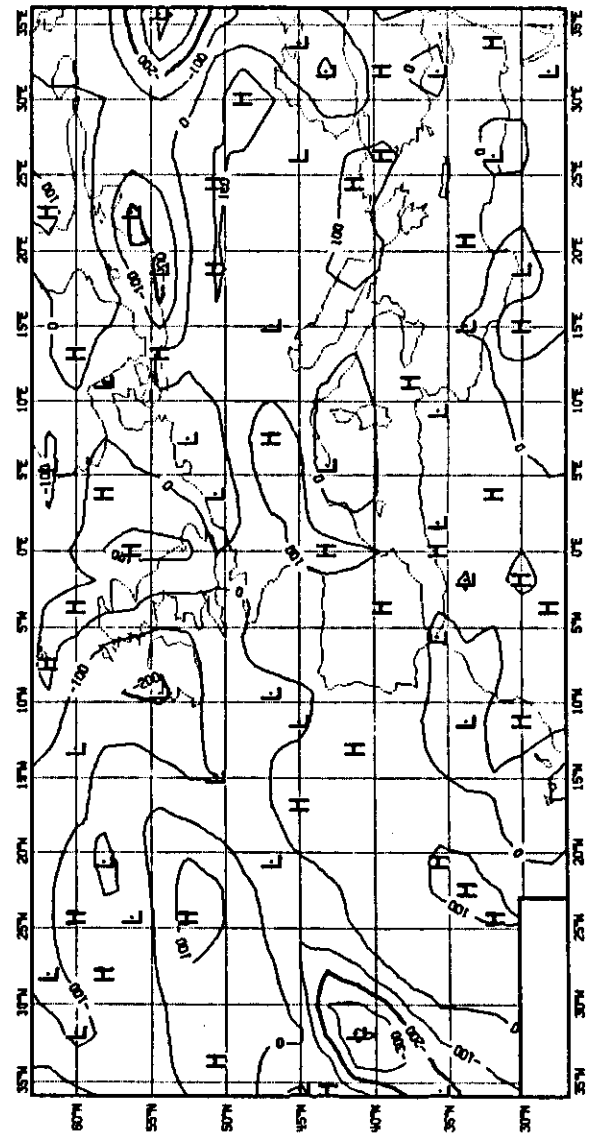


Fig. 9b The divergence of the final wind analysis in Fig. 9a. Units: 10^{-7} s^{-1}.

The two-dimensional scheme is a distance and error weighted mean of the observations. The weight given to an observation is

$$w_i = \mu(r_{ki})/[1 + \varepsilon_i^{o\,2} - \mu^2(r_{ki})] \cdot \qquad (14)$$

and the correction to the first-guess is given by

$$A_k - P_k = \sum_{i=1}^{N} w_i (O_i - P_i)/[1 + \sum_{i=1}^{N} w_i] \qquad (15)$$

where

$$\mu(r) = e^{-1/2 \left(\frac{r}{r_o}\right)^2} \qquad (16)$$

is the prediction error correlation for humidity as a function of distance r. r_0 = 250 km for the 1000-850 mb layer and increasing to 350 km for the 300-400 mb layer.

The temperatures and dew point temperatures are extracted from radiosonde reports and converted to mixing ratios, which are then integrated vertically to give the water vapour content in the analysis layers. In surface observations temperature and dewpoint, current weather and cloud amounts and types give estimates of the boundary layer and cloud level humidities. The weather and cloudiness information is converted to relative humidities according to a formulation based on Chu and Parrish (1977).

Above 300 mb the humidity is specified in the following way. A constant mixing ratio of $2.5*10^{-6}$ is assumed for the stratospheric humidity. Between 300 mb and the tropopause the relative humidity is assumed to decrease

linearly with respect to pressure. The relative humidities and temperatures then define the mixing ratios at any level. The evaluation of the mixing ratios above 300 mb takes place in the vertical interpolation (Section 6), where the temperatures are available.

The impact of the humidity analysis is greatest over the tropical continents and has a marked effect on the initial rainfall in the forecast. In extratropical regions the first-guess is usually close to the measured values.

5. SEA SURFACE TEMPERATURE ANALYSIS

ECMWF receives the sea surface temperature analyses of NMC on a 5 by 5 degree grid. The climatological monthly mean temperatures are subtracted from the NMC analyses and the anomalies are interpolated to the ECMWF grid using bicubic splines. The horizontally interpolated values are then added to the climatology on the ECMWF grid. Over ice areas the climatological field is used and a smooth transition is applied from ice to open water.

6. VERTICAL INTERPOLATION

The aim of the vertical interpolation is twofold. Firstly, it provides the analysis with a model forecast in the analysis coordinates. Secondly, the corrections made by the analysis to the first-guess are interpolated back to the model coordinates.

The numerical technique of the vertical interpolation depends on the meteorological parameter. The wind is interpolated linearly in ln p and shifted horizontally to coincide with the non-staggered analysis grid (cf Fig. 1). The model temperatures are integrated hydrostatically in sigma coordinates to give geopotential heights. These are then interpolated to pressure levels through cubic spline interpolation with respect to ln p. The upper stratospheric (30, 20 and 10 mb) background heights are determined by

adding a mix of persistence and climatological thicknesses to the 50 mb height. The stratospheric winds are determined through a similar procedure from the persistence and climatological wind shears and the 50 mb wind. The humidity first-guess is created through an interpolation of the model relative humidities to the pressure layers. The precipitable water content of the first-guess is then obtained from the first-guess relative humidity and layer mean temperature.

The changes or increments produced by the analysis are interpolated from pressure to sigma coordinates and added to the six hour forecast in model coordinates. This feature of the vertical interpolation preserves the model boundary layer structure. Surface pressure is not explicitly analysed, but calculated from the change of the geopotential at the first-guess surface pressure. The change in the geopotential at the surface is obtained from extrapolation or interpolation of the geopotential changes at the nearest analysis levels. Once the new surface pressure has been defined the pressure values of the sigma levels are determined.

The wind increments are staggered to the model's grid before the linear interpolation with respect to ln p. The height increments are first transformed into increments of mean virtual temperature. The temperature increments for the pressure layers are combined linearly to give the mean temperature change for each sigma layer. The weight given to each pressure layer temperature increment is proportional to its overlap with the sigma layer. A similar procedure is applied to the moisture interpolation. The vertically interpolated humidity variable is relative humidity increments calculated from the changes of layer water vapour content. This gives the sigma layer relative humidities which combined with the virtual temperatures is solved to give dry temperatures and mixing ratios.

7. SUMMARY

The ECMWF analysis system operates on 15 standard levels for mass and wind and 5 layers for humidity. Other parameters like the surface fields (land temperature,land soil moisture, surface pressure) either evolve during the data assimilation process or are a result of a purely numerical interpolation procedure applied to the analysed fields. Many important predictors used in statistical interpretation of weather depend strongly on the formulation in the forecast model of the influencing processes. Considerable errors occur in conjunction with high terrain, where the surface might be far from the closest analysis level.

References

Cats, G.J. 1981 Current problems in medium range weather forecasting at ECMWF: Data assimilation scheme. Proceedings of Seminar on Problems and prospects in long and medium range weather forecasting, 141-182.

Chu, R. and Parrish, D. 1977 Humidity analyses for operational prediction models at the National Meteorological Center, NWS/NOAA, Washington, D.C.

Gandin, L.S. 1963 Objective analysis of meteorological fields. Translated from Russian by Israeli Program for Scientific Translations, 1965, 242pp.

Hollett, S.R. 1975: Three-dimensional spatial correlations of PE forecast errors. MS thesis. Dept. of Meteorology, McGill University Montreal, Canada, 73pp.

Lorenc, A.C. 1981 A global three-dimensional multivariate statistical interpolation scheme. Mon.Wea.Rev. 109, 701-721.

Lorenc, A. and Tibaldi, S. 1979 The treatment of humidity in ECMWF's data assimilation scheme. ECMWF Technical Memorandum No. 7.

Rutherford, I. 1973 Experiments on the updating of PE forecasts with real wind and geopotential data. Preprints Third Symposium on Probability and Statistics, Boulder, Amer.Meteor.Soc., 98-201.

Schlatter, T. W. 1975 Some experiments with a multivariate statistical objective analysis scheme. Mon.Wea.Rev. 103, 246-257.

Wergen, W. 1982 Initialisation. ECMWF Seminar/Workshop on Interpretation of Numerical Weather Prediction Products.

Initialisation

W.Wergen

ECMWF

1. FOREWORD

This paper is not meant to be a strict mathematical review of nonlinear normal mode initialisation. This can be found elsewhere (Daley, 1980). It rather tries to explain in simple terms what initialisation is, and why it is needed. Frequent use of examples and analogies is made. This allows a straightforward physical explanation of normal mode initialisation and of its benefits compared to other methods. After these general ideas have been established, the performance of the scheme - in its recently introduced diabatic version - is discussed.

2. WHY INITIALISATION?

Once an analysis scheme has provided initial values for the prognostic variables of the model, one is in a position to start a forecast. However, as Fig. 1 shows, the result will be quite disappointing. The surface pressure at this particular point (0°E,52°N) shows unrealistic high frequency oscillations with considerable amplitude. By contrast, if the forecast is started from the initialised analysis (dashed), the behaviour is much better. When looking at the difference (initialised-uninitialised) between the two predicted surface pressure fields after 6 hours, a marked wave pattern is immediately evident (Fig. 2). In order to cause the pressure oscillations shown in Fig.1, these waves have to travel with a speed in the order of hundred meters per second, which is the characteristic speed of gravity waves.

The crucial role of initialisation becomes most evident within a data assimilation scheme. Before the actual analysis is performed, the observations are checked against the first guess (Lönnberg,1982). If the deviations exceed

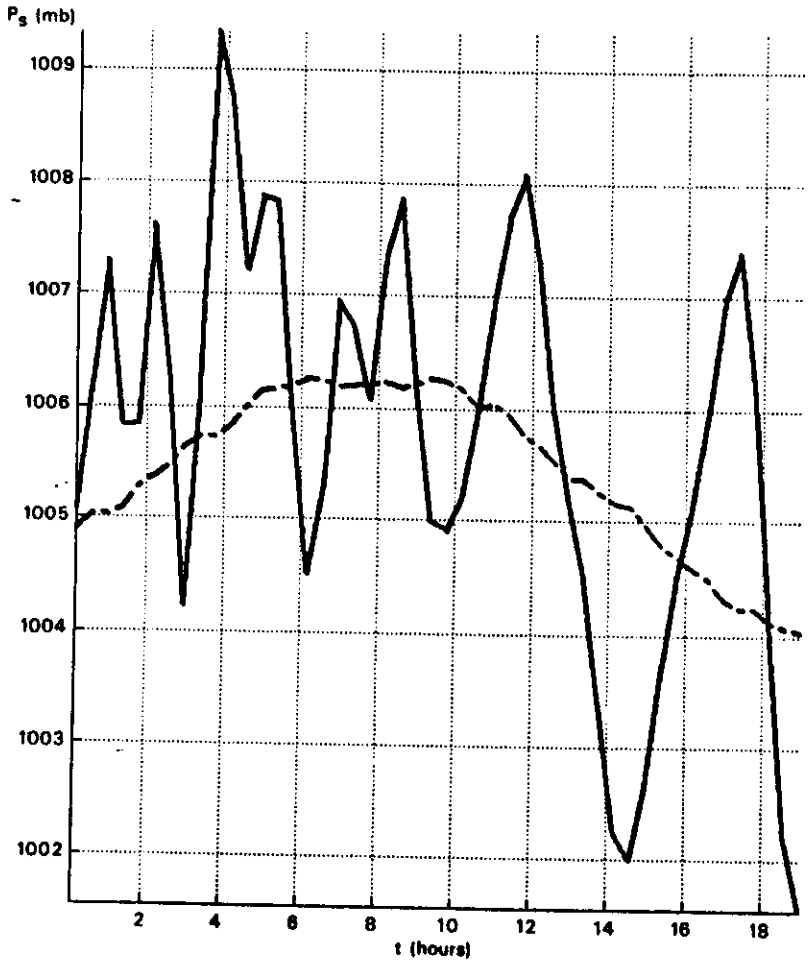


Fig. 1 Surface pressure evolution at gridpoint 0°E, 52°N for forecast started from uninitialised (full) and initialised (dashed) analysis for 12 GMT 6 September 1982.

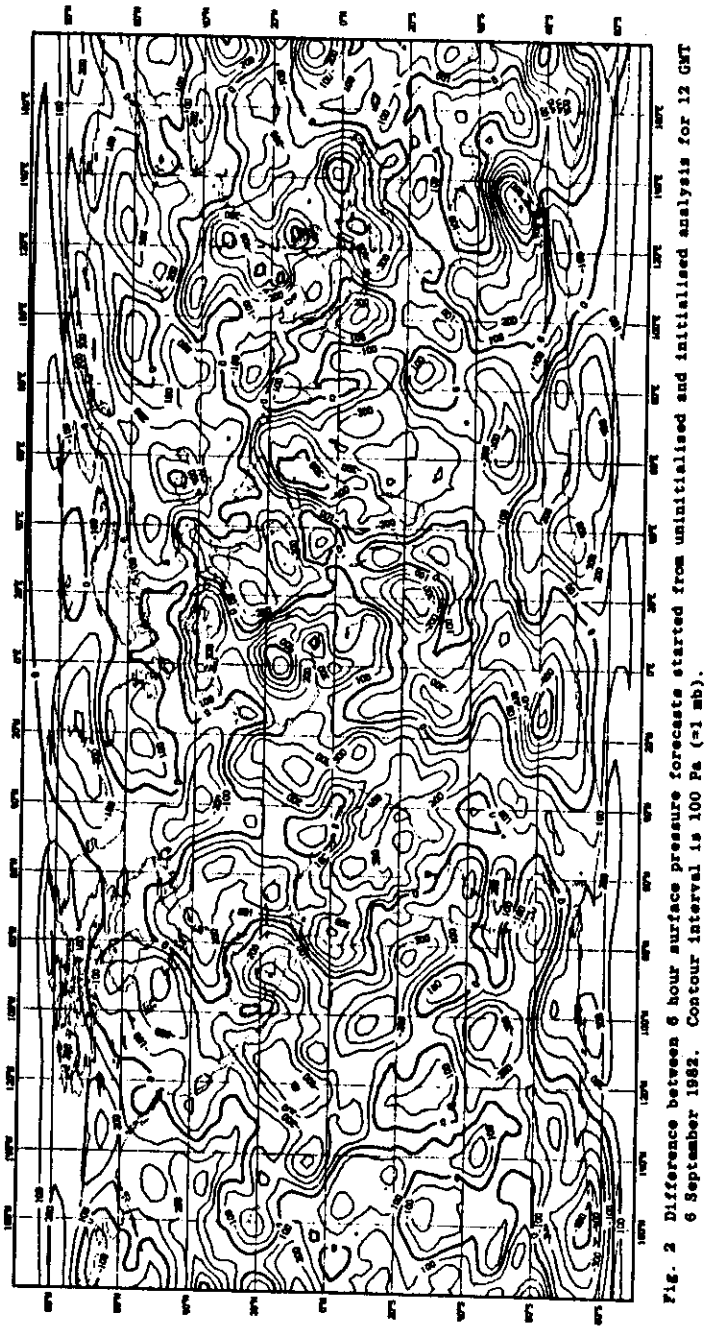


Fig. 2 Difference between 6 hour surface pressure forecasts started from uninitialised and initialised analysis for 12 GMT 6 September 1982. Contour interval is 100 Pa (=1 mb).

certain limits, the data are rejected. If the first guess is contaminated with spurious oscillations, this can easily lead to unjustified rejection or acceptance of data. Such a degraded analysis leads to a degraded forecast which will deteriorate the next analysis and so on.

To demonstrate this effect, the operational assimilation cycle for Sept.6, 1982 was repeated with the initialisation switched off. Although the absolute number of rejections did not change very much, a substantial number of data (around 30) were treated differently. Fig.3 shows the 850 mb observations superimposed on the first guess height and wind field. The two encircled wind observations near 25°E, 50°N and 60°E, 60°N were accepted in the re-run although their directions are clearly in error. The operational run with initialisation correctly rejected them. A number of other observations, which were perfectly valid, were rejected in the re-run because of spurious waves in the first guess field. The effect of assimilation without initialisation becomes more serious if one continues beyond one day.

3. WHAT ARE GRAVITY WAVES?

A simple example of gravity waves are the patterns originating from a stone thrown into a pond. The main processes leading to the propagation of these waves are sketched in Fig.4. Once the surface of an incompressible fluid has been elevated at an arbitrary point B, gravity tries to bring the particles back into their original lower level. This leads to mass convergence in regions A and C which, in turn, result in an elevation of the surface, as shown in the lower part of Fig.4. Two aspects are worth keeping in mind: the need for an initial disturbance and the importance of divergence.

For large scale gravity waves on a rotating planet, the Coriolis force cannot be neglected. Therefore, these waves are called inertia-gravity waves. Without them, the important geostrophic adjustment process cannot take place.

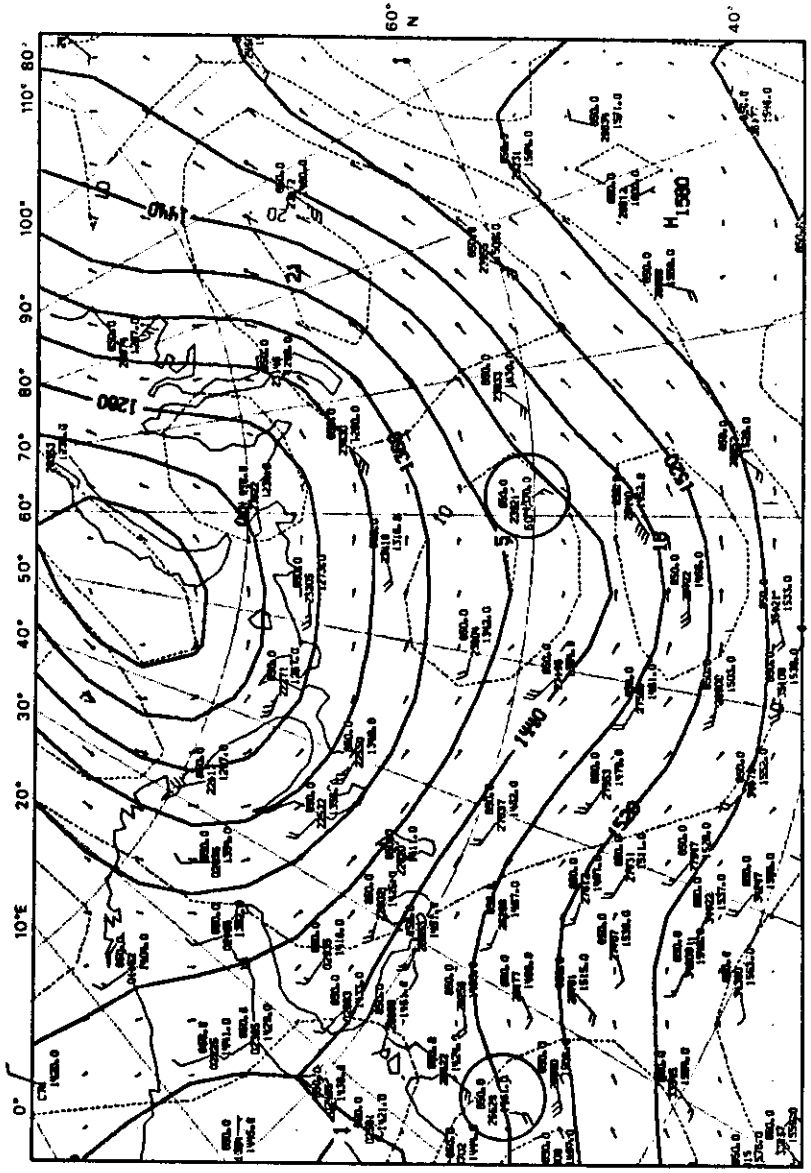


Fig. 3 850 mb observations and first guess height (full) and wind field. Contour inverval for isotachs (dashed) is 10 m/s. Circles indicate differently treated observations.

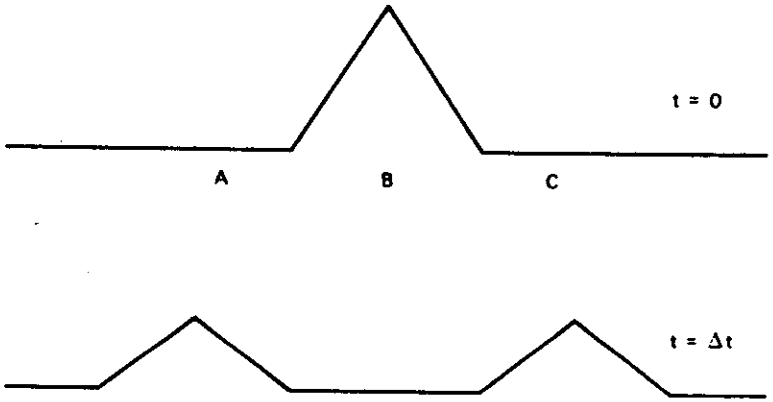


Fig. 4 Prototype gravity wave.

That is why their careful modelling has proven to be important for an atmospheric model. In a stratified fluid, gravity waves can travel not only along the surface, but also within the fluid. These internal waves are responsible for transports of momentum and energy, especially near mountains. They can sometimes be observed downstream of mountain ridges in the form of lenticular clouds which form in the wavetops, where the particles reach the condensation level.

The question now remains: why do atmospheric models show the excessive gravity wave activity as given in Figs.1 and 2. This question can best be answered by referring to Fig.4. Suppose we have a perfect model to predict the surface height of the fluid in a pond. To start the forecast, we need initial values of the surface height. Suppose, we have 3 observations in points A,B and C. If now the observation in point B has a positive error, the model will start with an initial state similar to the one sketched in Fig.4, thus immediately leading to excessive gravity wave activity. A further source for unrealistic initial amplitudes for gravity waves is the analyses system used to interpolate the irregularly spaced observations to regular grid points. Even with perfect observations an erroneous initial state might result only from errors in the analysis scheme. A third cause is the model itself, which - although presented with a perfect initial state - can generate gravity waves because it is only an approximate description of the fluid. Generally, all 3 causes contribute to unrealistic gravity wave amplitudes in the initial state.

4. WHAT IS INITIALISATION?

In quite general terms, initialisation tries to define a reasonable initial amplitude for inertia-gravity waves. (For the simple example in Fig.4 it would mean to set the amplitude at point B to zero). A number of methods have been used in the past. In the early days of numerical weather prediction,

gravity waves were excluded altogether from the models by modifying the governing equations. Therefore, no initialisation was required. However, it was soon realised that these filtered models are inferior to the primitive equation models. Early initialisation methods used for these models all tried to control gravity wave activity by a suitable diagnostic definition of the initial wind field using various kinds of approximate relations. The most simple approach is to use the geostrophic relation. A higher degree of approximation can be achieved by the "balance equation" which yields a non-divergent wind field from the analysed mass field. Thus, the observed wind information is not used. Furthermore, the diagnostic relations are extremely simplified. For instance, there are considerable divergent motions in synoptic systems, which are suppressed initially when the balance equation is used. A third shortcoming is the need sometimes to modify the analysed height field for mathematical reasons.

It is therefore no surprise, that alternative methods have been developed. One of them is dynamical initialisation, where the model itself is used to derive a suitable initial state. Usually, the model is integrated forward and backward in time and gravity waves are damped by heavy time and space filtering. However, the filter operators cannot damp gravity waves selectively, they also influence the meteorologically important flow. Because of the backward integration in time, irreversible physical processes can not, in general, be included. Furthermore, the method is expensive in terms of computer time. Recently, a third method, normal mode initialisation has successfully been applied. This method will be outlined in the following sections.

5. WHAT ARE NORMAL MODES?

Normal modes are the free motions of a system capable to vibrations. A typical example is a guitar string. Once it has been excited, it starts to vibrate in a way characteristic for that particular string. Therefore, these

vibrations are called "own" or "eigen" vibrations, or - in mathematical terms - normal modes. Another example are the surface waves on a pond. They are the eigenvibrations of the physical system "pond". Usually, there is an infinite number of eigenvibrations, each having a particular scale. If a stone is thrown into a pond, only a subset of these modes is excited, depending on the initial shape of the disturbance caused by the stone.

The eigenvibrations of atmospheric models fall into two classes: the Rossby modes and the gravity modes. Every mode shows a specific 3-dimensional structure of mass- and windfield. The vertical structures are identical for Rossby and for gravity waves. They differ, however, considerably in their horizontal structure and in the relation between mass- and windfield. Fig.5 gives some examples of vertical structures. The first vertical mode, - the external mode - represents fields which are nearly constant throughout the atmosphere. It therefore describes the barotropic component of the mass- and windfield. The second mode, also called the first internal, changes sign near the tropopause, thus accounting for the differences between stratosphere and troposphere. In general, vertical mode l has l-1 sign changes. With increasing order l, the region of maximum amplitudes moves towards lower levels.

The horizontal structure of an external Rossby mode is shown in Fig. 6. The prominent feature is the approximately geostrophic relation between geopotential and wind in the extra-tropics. In the tropics, where the geostrophic relation is not applicable, normal modes still define a relation between mass- and windfield. In fact, they are the only means to establish globally valid coupling. The Rossby mode shown in Fig. 6 can be characterised by 3 indices: - the vertical mode number - the zonal wavenumber - the meridional index. As Fig. 6 shows an external mode, it means that at every point the wind components and the geopotential all have the same vertical structure as the curve labelled 1 in Fig. 5. The meridional index in Fig. 6 is 1, which means the largest meridionally symmetric scale. Obviously, the zonal wavenumber is 1.

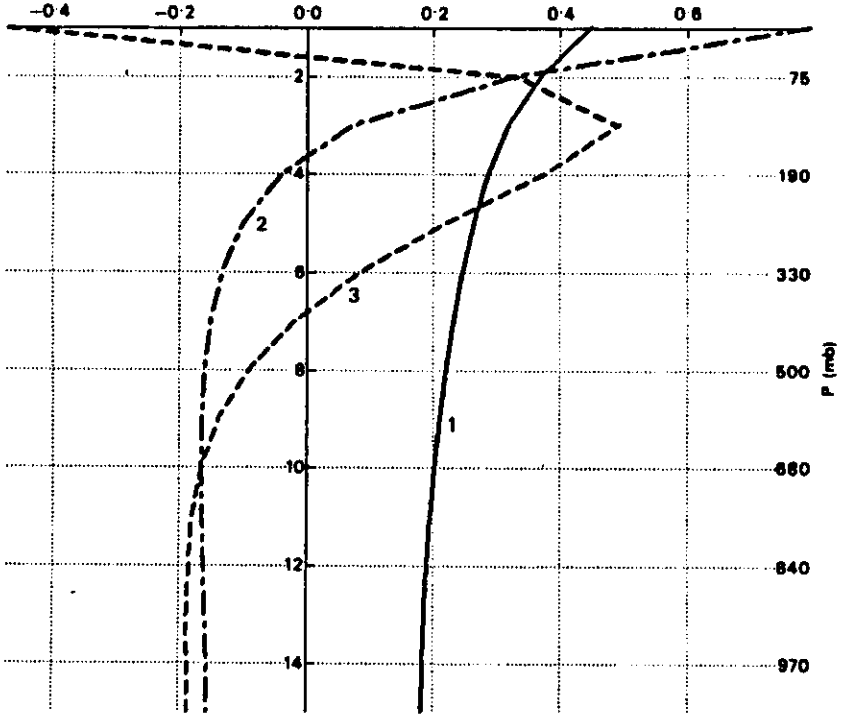


Fig. 5 First 3 vertical modes for isothermal (300 K) basic state. Bounded pressure values in mb are valid for p_s = 1000 mb.

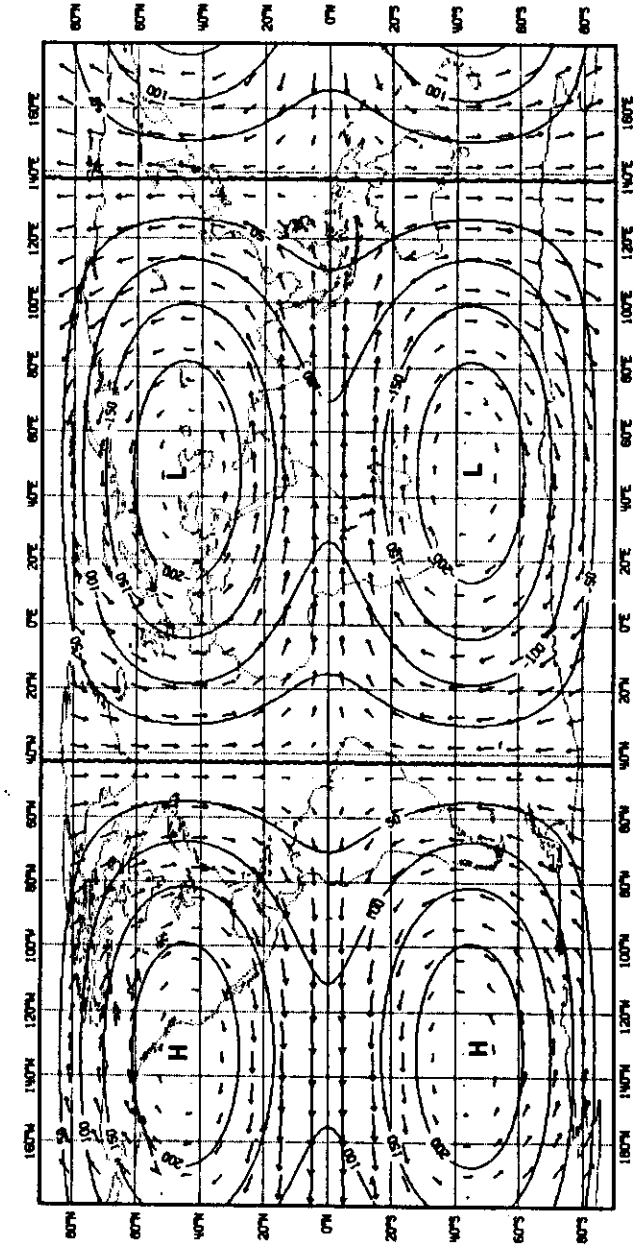


Fig. 6 Gravest asymmetric, zonal wavenumber 1, external Rossby mode.

When looking at a gravity mode (Fig. 7) for the same indices (vertical mode 1, zonal wavenumber 1, gravest symmetric), the ageostrophic structure of these modes is immediately evident. In the extra-tropics, strong winds are not supported by a height gradient. In the tropics, the flow is highly divergent. Gravity waves can further be separated into two groups: the eastward and westward travelling waves. Fig.7 shows a westward travelling gravity wave.

For the same combination of indices, Rossby and gravity waves have very different phase speeds (and periods). For example, the free period of the Rossby wave in Fig. 6 is 4.7 days, whereas the westward gravity wave in Fig.7 has a free period of only 12.7 hours. Apart from a few exceptions, gravity wave periods decrease with increasing zonal wavenumber and meridional index. In contrast, large scale Rossby waves have shorter periods than small scale ones. For both types, the periods decrease with increasing vertical mode number. Therefore, a large scale internal gravity mode may have the same period as an external small scale Rossby mode. That is why the period is not a useful tool to distinguish between Rossby and gravity modes; nor is the wavelength, as apparent from Figs. 6 and 7. The only way to make a proper distinction is to look at the 3-dimensional structure of both the mass- and the windfield.

6. WHAT IS NORMAL MODE INITIALISATION?

In normal mode initialisation, the unique features of normal modes are efficiently used to remove unwanted gravity wave amplitudes from the analysis.

First of all, the normal modes have to be computed. This is done by linearising the model equations around a simple basic state and solving the resulting set of equations. The details can be found in Temperton and Williamson (1981). The result of these computations are the Rossby and gravity modes discussed in the previous section. This computation needs to be done only once; the normal modes are then stored for subsequent use.

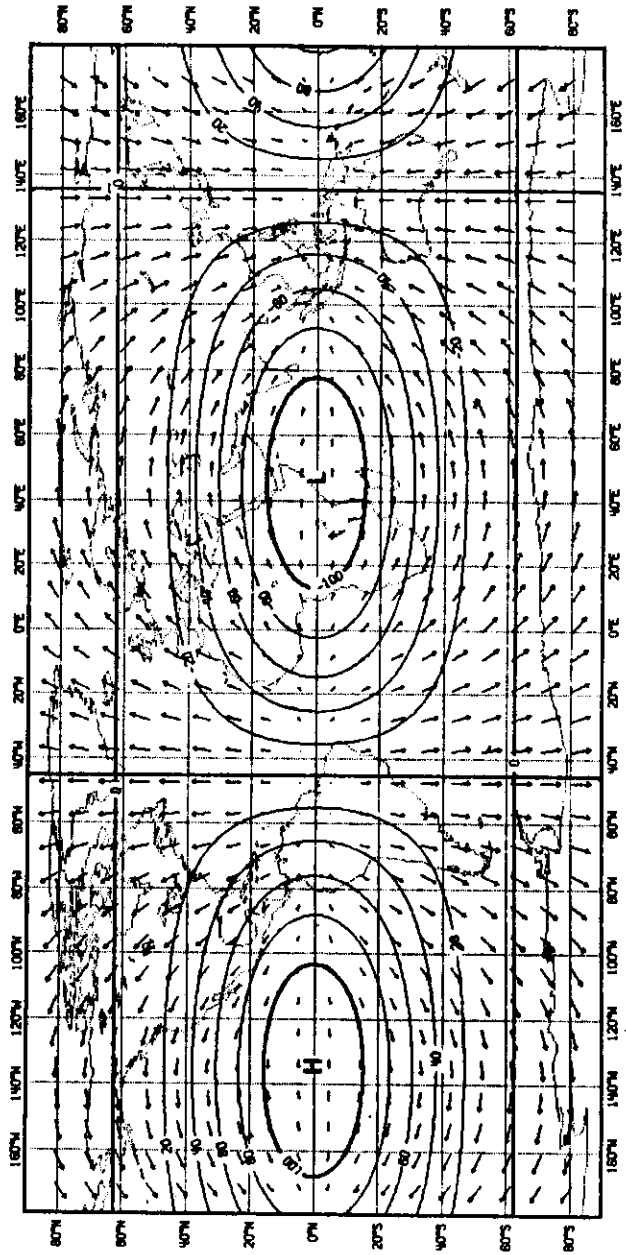


Fig. 7 Gravest symmetric, zonal wavenumber 1, westwards travelling, external gravity mode.

The first step in the initialisation procedure itself is the identification of those structures in the analysed fields which lead to the excitation of gravity waves. Speaking mathematically, the analysis is projected on the gravity modes. This guarantess, that the most important part described by the Rossby modes, is not touched at all.

Fig. 8 shows the gravity mode projection of the analysis increment (analysis - first guess) for 6.9.82, 12Z. As can be expected, the gravity mode projection of the wind field is mostly divergent. A closer inspection of the analysis off the Chilean coast (Fig. 9) reveals, that the analysis (first guess + increment) in this area is not in geostrophic balance. The isotachs (dashed) indicate a rapid decrease in wind speed towards the coast, which is not supported by a corresponding slackening of the height gradient (full lines). Similar results apply to all extra-tropical areas, where large gravity wave projections are invariably related to ageostrophic analysis increments.

In the second step, the identified gravity mode projections are modified. The simplest method would be to set them to zero (linear normal mode initialisation). However, this would not solve the problem, because the nonlinear advection processes and the parameterised physical processes would immediately start to excite new gravity waves. Machenhauer (1976) proposed to eliminate only those parts of the gravity wave components, which are not in balance with the nonlinear processes. As a result, the linear tendencies for the gravity waves will be exactly compensated by the nonlinear dynamical and physical tendencies, yielding vanishing total tendencies for the initialised gravity waves. The corresponding mathematical equations can easily be formulated in normal mode space. It is a system of nonlinear algebraic equations, which can be solved iteratively.

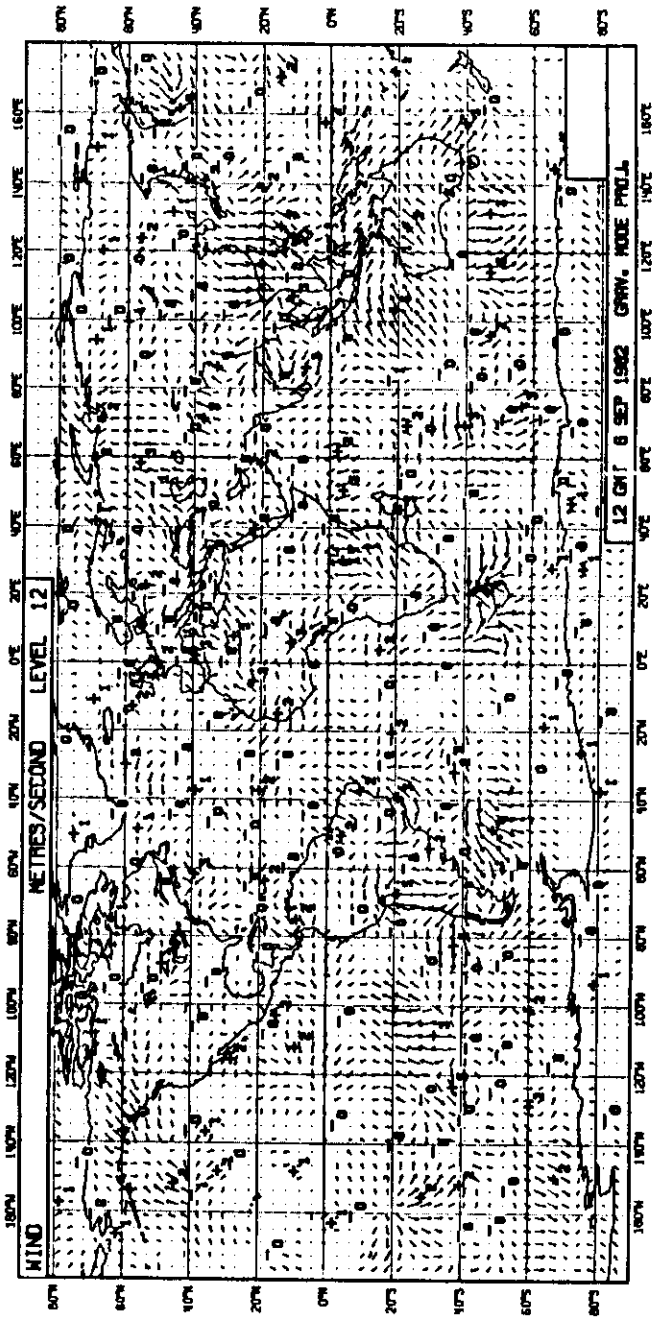


Fig. 8 Gravity mode projection of analysis wind-increments at model level 12 (around 850 mb) for 12 GMT 6 September 1982. Numbers indicate wind speeds in meters/second.

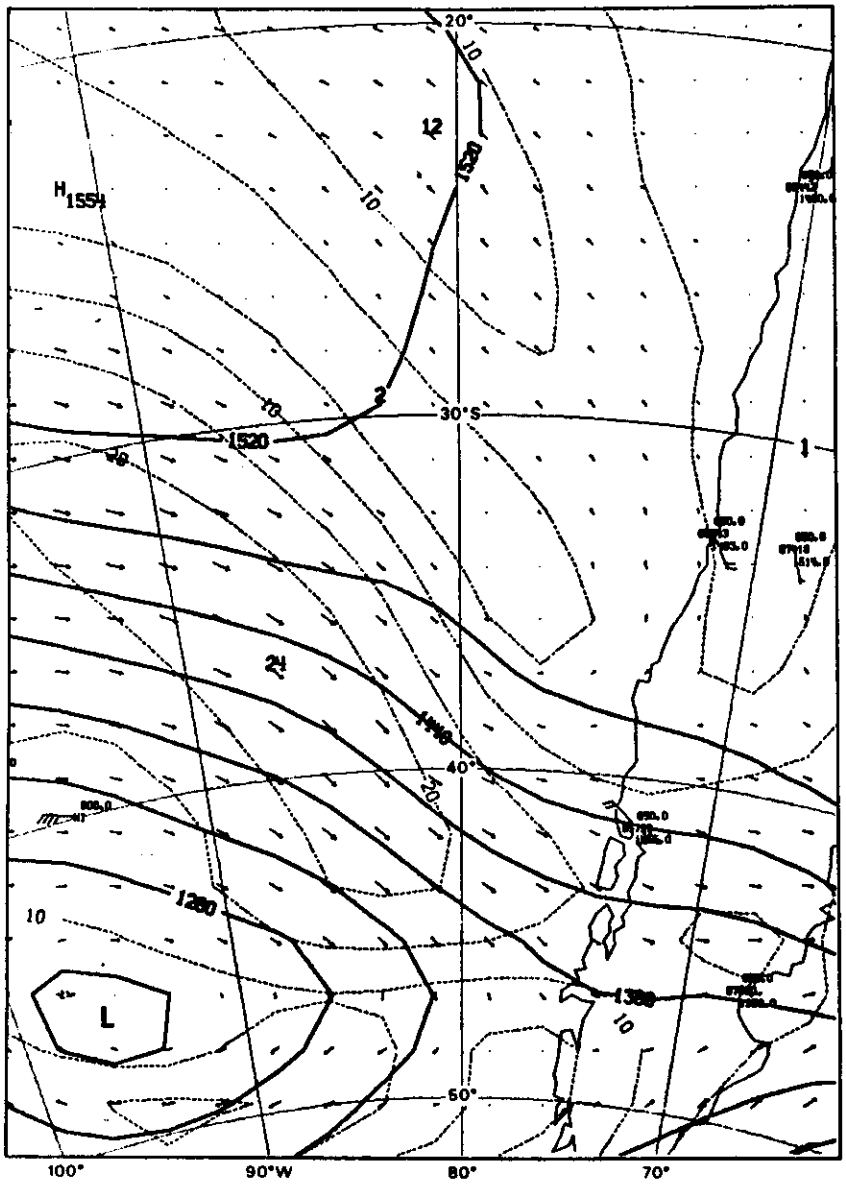


Fig. 9 850 mb height (full) and wind analysis plus observations. Isotachs (dashed) are in 10 m/s.

7. THE PERFORMANCE OF THE ECMWF SCHEME

a) Implementation

The normal modes used in the operational scheme are derived assuming an isothermal atmosphere (300K) at rest. Only the first five vertical modes are initialised. Two iterations are used to solve the nonlinear algebraic system. In an attempt to reduce surface pressure changes by initialisation, the analysed surface pressure is restored with a latitude-dependent weighting after the first iteration. An estimate of the diabatic forcing is included, but kept independent of the iteration because of convergence problems. This estimate is computed by time-averaging the diabatic tendencies during a 2-hour forecast started from the uninitialised analysis. Only those tendencies, which force inertia-gravity waves with periods longer than a certain cut-off period (11 hours) are included. In the calculation of the adiabatic tendencies, virtual temperature is used.

b) Typical initialisation changes

As neither the analysis scheme nor the data know the subtle nonlinear balance required to supress unwanted gravity wave activity, initialisation usually changes the analysed fields. Surface pressure and divergent wind are most affected, temperatue and rotational wind changes are neglegible. Fig. 10 shows the initialisation changes to surface pressure, averaged over all 122 cases for April 1981. Mean Changes are in the order of 1 mb. The initialisation changes to the wind field in 200 mb (Fig.11) are in the order of several meters/second. They are mostly divergent. With the adiabatic scheme in use in April 81, wind changes were largest in the tropics. As will be shown in the next section, diabatic initialisation results in smaller changes. In individual cases, initialisation changes can amount to 5 mb in surface pressure and up to 10 m/sec in upper tropospheric winds. Invariably, large initialisation changes point to grose imbalances in the analysis. In the extra-tropics, this is possible, because the mass wind relation for the

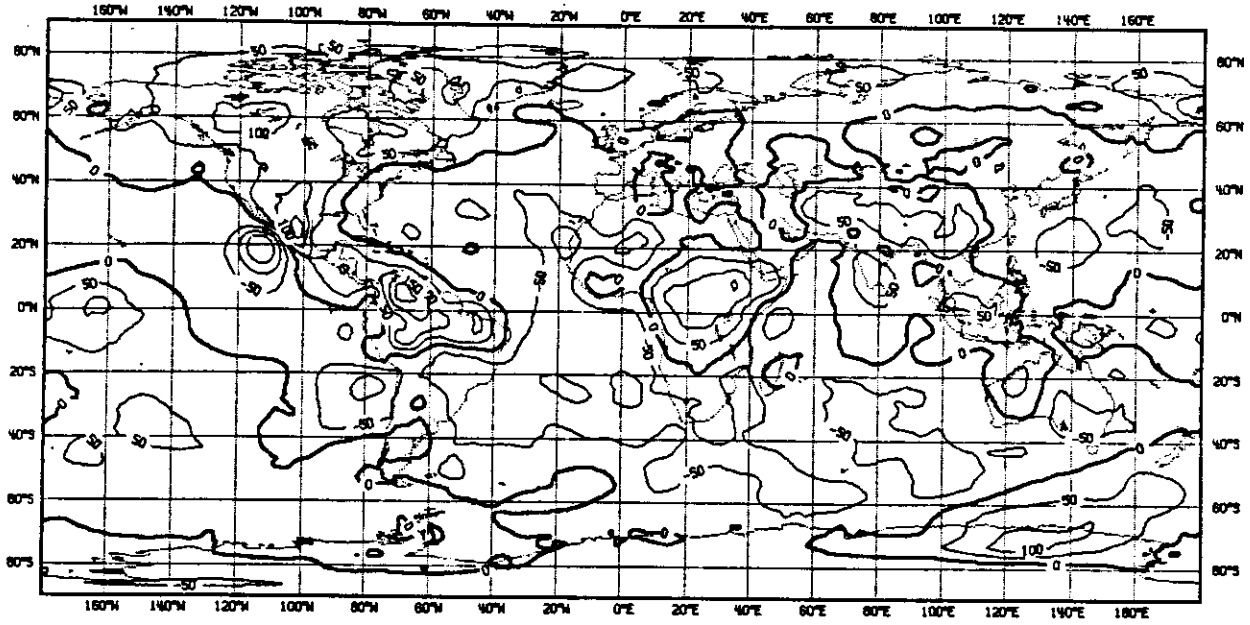


Fig. 10 Mean surface pressure initialisation change (initialised - unitialised) in 50 Pa (= 0.5 mb) for April 1981. (Averaged over 30 12 GMT cases).

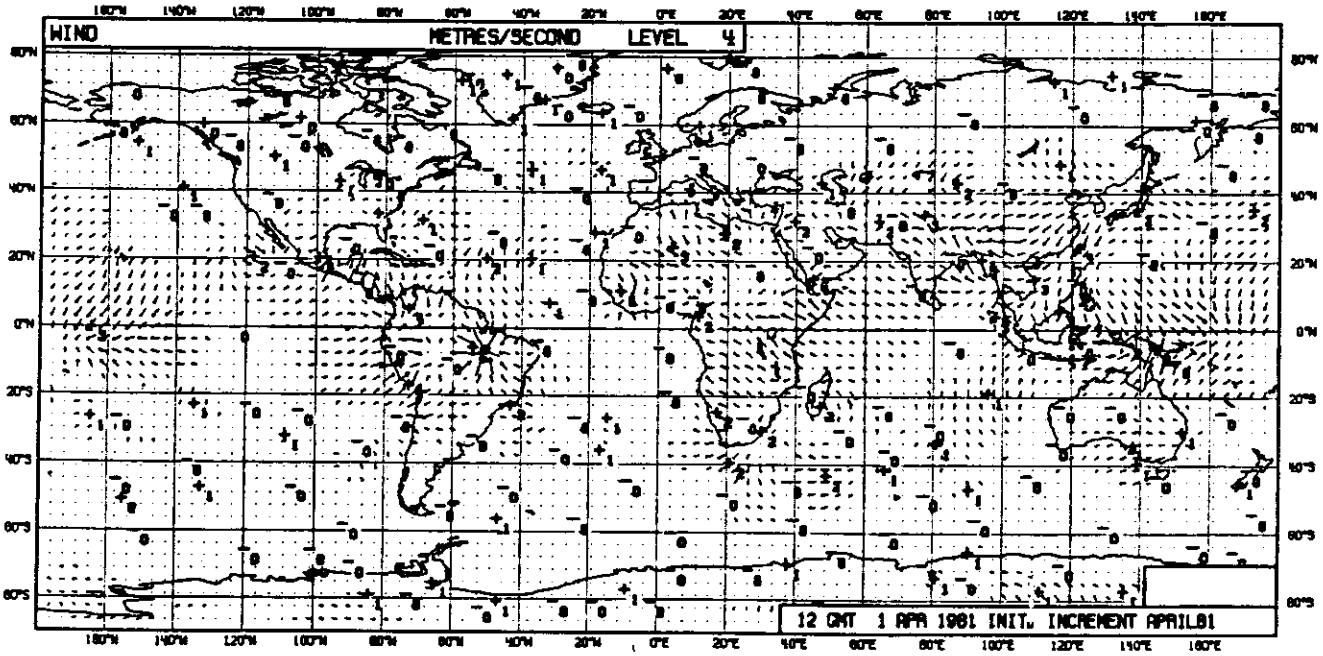


Fig. 11 Mean wind changes by initialisation at level 4 (around 200 mb) for April 1981. Numbers indicate wind speeds in m/s.

analysis increments is not exactly geostrophic (Lönnberg,1982) and because of different data selection for neighbouring analysis boxes. In the tropics, the present analysis scheme tends to alias a pure Rossby mode on a number of different Rossby- and gravity modes (Cats and Wergen, 1982).

c) The impact of diabatic processes

Adiabatic initialisation tends to supress diabatically driven circulations in the tropics (Bengtsson,1980). In order to test the impact of diabatic initialisation, 6 days of assimilation were repeated for the FGGE period 5. - 11.6.1979. Fig. 12 shows the zonally averaged, vertically integrated, initialised mass flux for 12.6.79, 12Z. This variable can be interpreted as a stremfunction for the flow in the ϕ-p plane. Adiabatic initialisation (top) supressed the upper tropospheric return flow in the Hadley cell as well as the mid-troposheric vertical motions. With diabatic tendencies included (bottom), the tropical circulation is more intense. A further indication of the positive impact is the time-averaged (6.6. - 11.6) initialised velocity potential at 200 mb (Fig. 13). With diabatic initialisation (bottom), divergent motions in the tropics are stronger than for the adiabatic scheme (top). The effect on the spin-up of the model is demonstrated in Figs. 14 and 15, which show the diabatic heating (averaged over 2 weeks) for the slab 700-300 mb in the northern hemisphere for the initialised analysis (left) and the 1 day forecast (right). For the method of calculation see White, 1983. Fig.14 gives the results for the first two weeks of September, when adiabatic initialisation was still in use. Initially, there is little diabatic heating in the tropics, whereas the 24 hour forecast tends to overshoot when compared with the level achieved in later stages of the forecast. After the introduction of diabatic initialisation into operations, the initial diabatic heating rates (Fig. 15, left) are closer to the 24 hour forecast values (Fig. 15, right). Even the diabatic heating fields for the storm tracks in the North-Atlantic are similar. The excessive convective activity over East-Africa is currently being investigated.

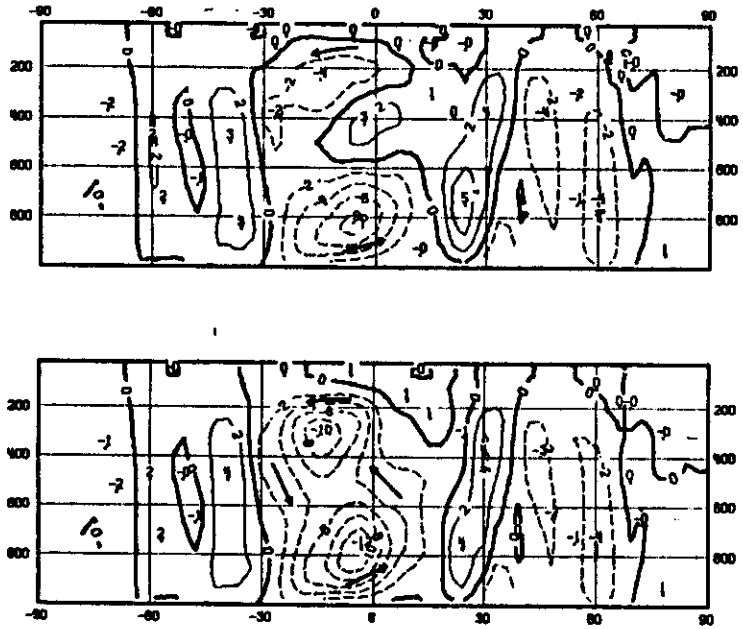


Fig. 12 Zonally averaged, vertically intergated mass flux in 10^{10} kg/s for 12 GMT 11 June 1979 for adiabatic (top) and diabatic (bottom) initisation.

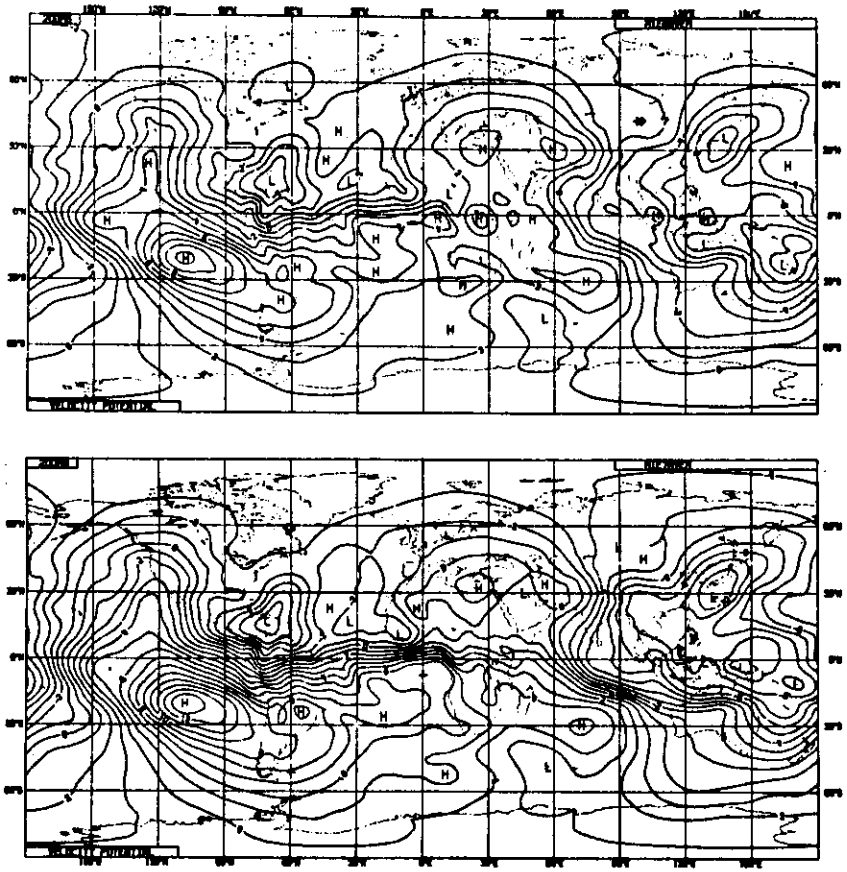


Fig. 13 Time-averaged (6-11 June 1979) initialised velocity potential at 200 mb in 10^6 m^2/s for adiabatic (top) and diabatic (bottom) initialisation.

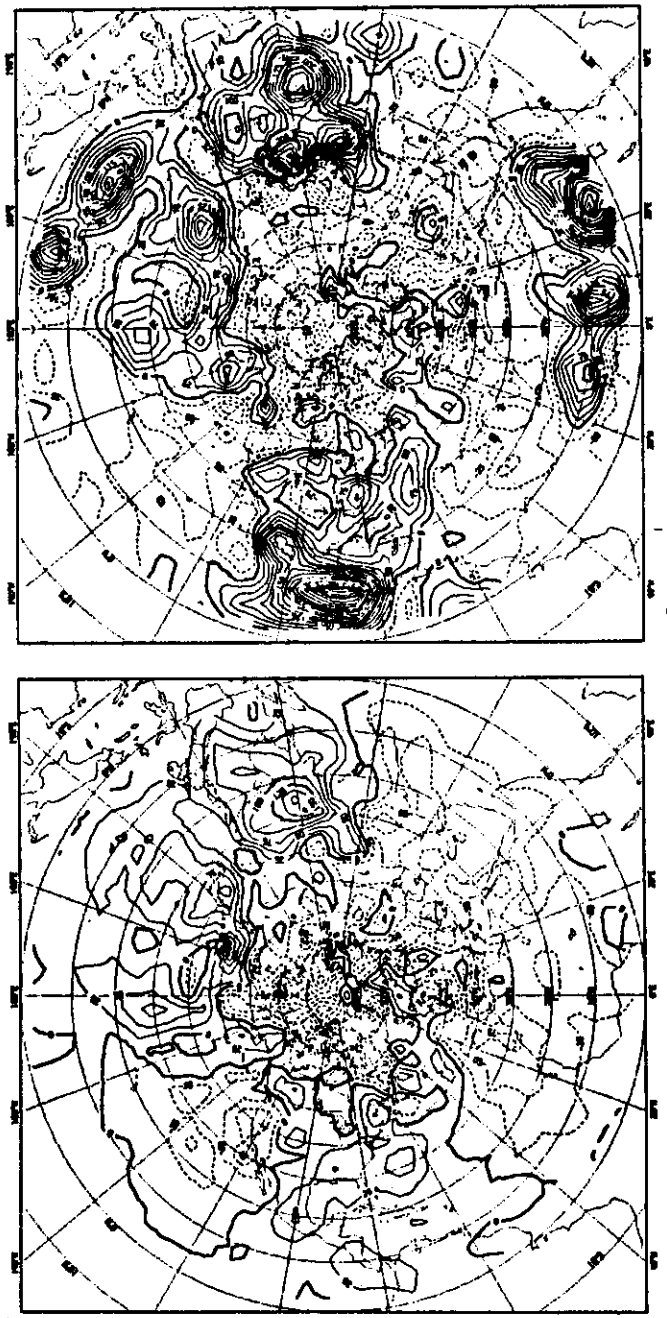


Fig. 14 Time averaged (1-15 September 1982) diabatic heating in W/m^2 for the slab 700 to 300 mb before the introduction of diabatic initialisation. Left for initialised analysis, right for 24 hour forecast.

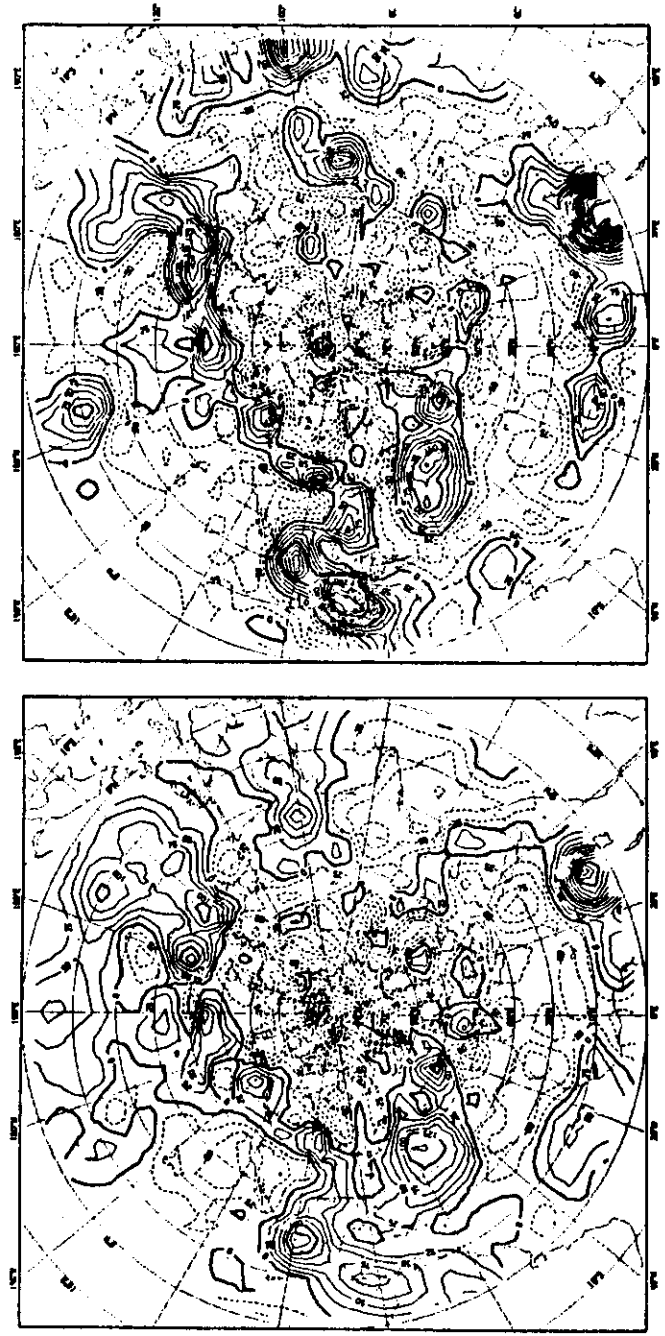


Fig. 15 Time averaged (1-15 October 1982) diabatic heating in W/m^2 for the slab 700 to 300 mb after the introduction of diabatic initialisation. Left for initialised analysis, right for 24 hour forecast.

Tropical scores for 200 mb winds (Fig. 16) show an improvement for the October 82 ensemble forecasts compared to the October 81 cases. The improvement is mainly due to improved skill in the large- and medium scale waves. As the forecasts are verified against the initialised analysis, part of the improvement is due to a more realistic verifying analysis.

8. SOME REMARKS FOR USERS

When using uninitialised or initialised analyses, some points should be kept in mind. The uninitialised analysis usually is closer to the data. It should therefore be used where a close fit to observation matters. However, it should not be used for calculations of divergence or derived quantities, such as vertical velocity or surface pressure tendency. The initialised analysis is much better suited for these purposes, although its balance - especially in the tropics - is model dependent. Also, a forecast should in general only be started from an initialised analysis. A close fit to observations does not necessarily mean that the observations are actually accepted by the model. Finally, when judging initialisation changes, the three-dimensional structure of both mass- and windfield should be taken into account. Invariably, big initialisation changes indicate problem areas in the analysis.

Acknowledgement

I should like to thank Glenn White, who allowed me to publish some of his figures. R.Daley and A.Hollingsworth carefully reviewed the manuscript.

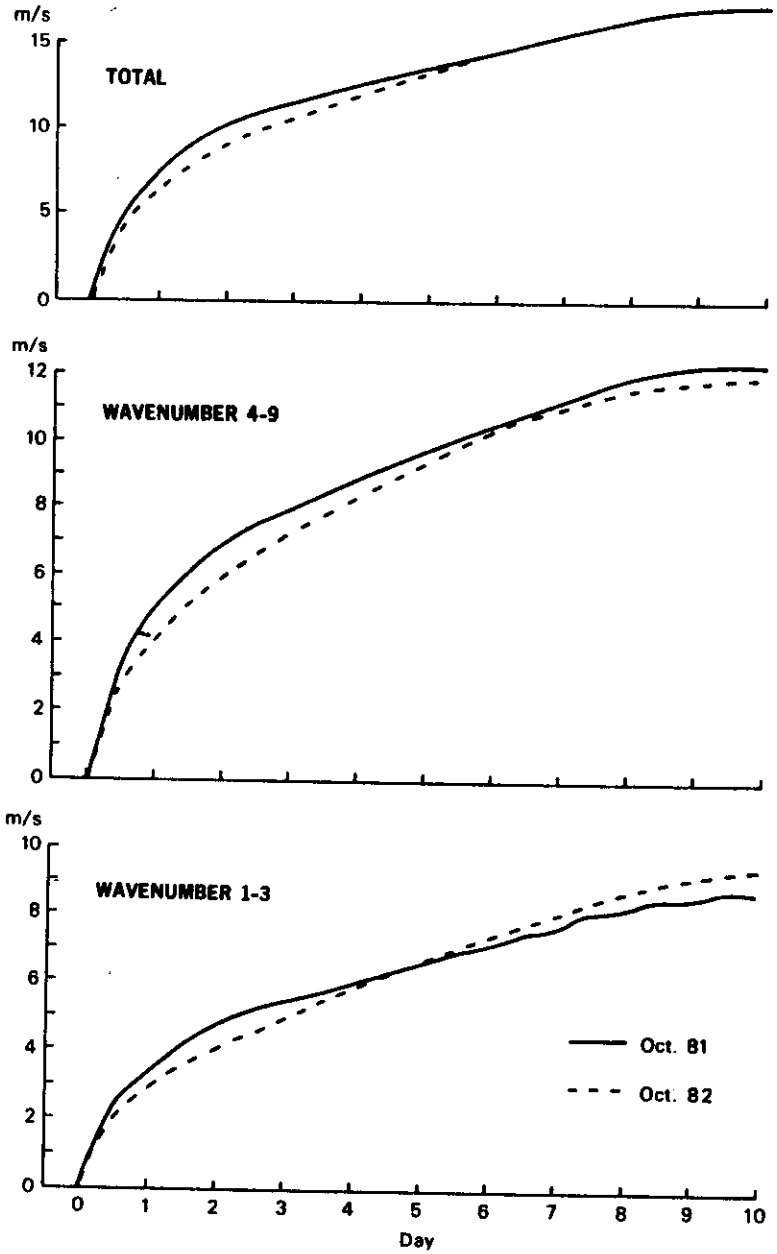


Fig. 16 RMS error of vector wind in the tropics (35°S to 32.5°N) at 200 mb averaged over all October 1981 forecasts (full) and October 1982 (dashed) forecasts.

REFERENCES

Bengtsson L., 1980: Current problems in four-dimensional data assimilation. ECMWF seminar proceedings on 'Data assimilation methods', 195-218. (Available from ECMWF).

Cats, G. and Wergen,W., 1982: Analysis of large scale normal modes by the ECMWF analysis scheme. To be published in ECMWF Workshop Proceedings on 'Current problems in data assimilation'.

Daley,R., 1980: Normal mode initialization. ECMWF Seminar Proceedings on 'Data assimilation methods', 107-158. (Available from ECMWF).

Lönnberg,P., 1982: The ECMWF analysis system. ECMWF Seminar Proceedings on 'Interpretation of numerical weather prediction products'.

Machenhauer,B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. Beitr.Phys.Atmos. 50, 259-271.

Temperton,C. and Williamson,D.L., 1981: Normal mode initialization for a multi-level grid-point model, Part 1: Linear aspects. Mon.Wea.Rev., 109,729-743.

White,G.H., 1983: The global circulation of the atmosphere December 1981 - November 1982 based upon ECMWF analyses. Technical Note, Dept.of Meteorology, Reading University, UK, in preparation.

ADIABATIC FORMULATIONS OF THE ECMWF FORECASTING SYSTEM

A. J. Simmons
ECMWF

1. INTRODUCTION

When charged with the task of presenting a summary of the adiabatic formulation of the operational ECMWF forecast model, it is appropriate at this particular point in time to discuss not one formulation, but two. Since August 1979, operational forecasting has been carried out at ECMWF using a second-order accurate finite-difference model with a regular latitude-longitude grid and resolution of 1.875°, a sigma-coordinate and 15-level resolution in the vertical, and a semi-implicit time scheme which allows a time step of 15 minutes in most cases. During this first operational period, one aspect of the Centre's research work has been directed towards the development and testing of alternative adiabatic formulations, and this has led to a new formulation which will shortly replace the current scheme. The new model is based on a spectral technique for the horizontal, a more general terrain-following vertical coordinate than the usual sigma-coordinate, and a semi-implicit time scheme that treats implicitly not only linearized gravity-wave terms, but also the linearized zonal advection of vorticity and moisture. In outlining the two different formulations, an account is thus given both of the model used to produce the results discussed in other contributions to these proceedings, and of the model which will be operational, or very close to becoming operational, by the time these proceedings are published. Mention will be made of differences which may influence the statistical interpretation of the model outputs.

The following section sets out the primitive equations for a moist atmosphere as adopted in both formulations, using a general vertical coordinate. Section 3 then discusses the horizontal discretization, with a summary both of the grid-point and spectral techniques and of the results of the comparisons between these techniques carried out at ECMWF. Aspects of the vertical and temporal discretizations are discussed in Sections 4 and 5. Finally, the incorporation of horizontal diffusion and the treatment of orography are described in Section 6.

2. THE PRIMITIVE EQUATIONS

We consider a general, terrain-following vertical coordinate, a monotonic function of pressure p and dependent on its surface values p_s:

$$\eta = \eta\ (p, p_s) \quad ,$$

where $\eta\ (0, p_s) = 0$ and $\eta\ (p_s, p_s) = 1$. The usual sigma coordinate

(Phillips, 1957) adopted for the adiabatic formulation of ECMWF's original operational model is a special case of this coordinate, with

$$\eta \equiv \sigma = p/p_s .$$

Kasahara (1974) has given the form of the primitive equations for a dry atmosphere using various coordinate systems. The η-coordinate form for a moist atmosphere is set down in this section. Prognostic variables are the horizontal wind components u and v, the temperature T, the specific humidity q and the surface pressure p_s. They are governed by the following equations.

Momentum equation

$$\frac{d\underline{v}}{dt} + f\,\underline{k} \times \underline{v} + \nabla\phi + R_d T_v \nabla \ell n p = \underline{P}_v + \underline{K}_v \qquad (1)$$

Thermodynamic equation

$$\frac{dT}{dt} - \frac{\kappa T_v \omega}{(1+(\delta-1)q)p} = P_T + K_T \qquad (2)$$

Moisture equation

$$\frac{dq}{dt} = P_q + K_q \qquad (3)$$

Continuity equation

$$\frac{\partial}{\partial\eta}\left(\frac{\partial p}{\partial t}\right) + \nabla\cdot\left(\underline{v}\frac{\partial p}{\partial\eta}\right) + \frac{\partial}{\partial\eta}\left(\dot{\eta}\frac{\partial p}{\partial\eta}\right) = 0 \qquad (4)$$

Hydrostatic equation

$$\frac{\partial\phi}{\partial\eta} = -\frac{R_d T_v}{p}\frac{\partial p}{\partial\eta} \qquad (5)$$

Here t is time, and $\frac{d}{dt}$ denotes the material derivative, which in η coordinates takes the form

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \underline{v}\cdot\nabla + \dot{\eta}\frac{\partial}{\partial\eta} .$$

$\underline{v}$ is the horizontal velocity vector, $\underline{v} = (u,v,0)$, and ∇ is the two-dimensional gradient operator on a surface of constant η. f is the Coriolis parameter, $\underline{k}$

the unit vertical vector, φ the geopotential, R_d the gas constant for dry air, and $\kappa = R_d/C_{pd}$, where C_{pd} is the specific heat of dry air at constant pressure. P_x and K_x denote the rates of change of variable x resulting respectively from parameterized processes (the subject of a separate contribution to these proceedings) and from horizontal diffusion.

An equation for the surface pressure, p_s, is obtained by integrating Eq. (4) from η = 0 to η = 1, using the boundary conditions $\dot{\eta} = 0$ at η = 0 and η = 1:

$$\frac{\partial p_s}{\partial t} = -\int_0^1 \nabla\cdot\left(\underline{v}\frac{\partial p}{\partial\eta}\right) d\eta \qquad (6)$$

while $\dot{\eta}$ and ω are given by

$$\dot{\eta}\frac{\partial p}{\partial\eta} = -\frac{\partial p}{\partial t} - \int_0^\eta \nabla\cdot\left(\underline{v}\frac{\partial p}{\partial\eta}\right) d\eta \qquad (7)$$

and

$$\omega = \frac{dp}{dt} = -\int_0^\eta \nabla\cdot\left(\underline{v}\frac{\partial p}{\partial\eta}\right) d\eta + \underline{v}\cdot\nabla p \qquad (8)$$

where $\frac{\partial p}{\partial t}$ is known in terms of $\frac{\partial p_s}{\partial t}$ from the definition of η.

Moisture effects appear in the momentum, thermodynamic and hydrostatic equations through the virtual temperature, T_v, which is given by

$$T_v = \left(1 + \left(\frac{R_v}{R_d} - 1\right)q\right) T$$

where R_v is the gas constant of water vapour. An additional term $(1+(\delta-1)q)$, where δ is the ratio of the specific heats at constant pressure of water vapour and dry air, is written in the thermodynamic equation. This term was neglected in the adiabatic formulation of the grid-point model, but as it is of the same order as the ratio of temperature and virtual temperature, it is included in the new model. Further detail will be given in the documentation manual of this model.

Equations (1) to (8) may readily be cast into their more familiar form for sigma coordinates by replacing $\partial p/\partial\eta$ by p_s, and $\frac{\partial p}{\partial t}$ by $\sigma\frac{\partial p_s}{\partial t}$. The pressure-gradient term $R_d T_v \nabla \ell n p$ becomes equal to $R_d T_v \nabla \ell n p_s$ at all levels. In general, the term has the latter value at the surface, and decreases to zero in the case in which coordinate surfaces become surfaces of constant pressure at upper levels. In this case η is independent of p_s for all pressures less than a certain value.

3. THE HORIZONTAL DISCRETIZATIONS

3.1 The grid-point model

A discussion of the adiabatic formulation of the finite-difference model has been given in a series of lectures by Burridge in the 1979 ECMWF Seminar and details will not be repeated here. The model uses a second-order accurate difference scheme based on the staggered grid of variables shown in Fig. 1, the grid known as the C-grid (Arakawa and Lamb, 1977). Choice of this grid was made mainly because of its low computational noise and the ease of implementation of a semi-implicit time scheme. Operationally, a grid interval of 1.875^{o} in latitude and longitude is used, and this resolution is referred to as N48, there being 48 grid intervals between equator and pole. Following the work of Arakawa (1966) and Sadourny (1975) the finite-difference scheme was designed to conserve, among other quantities, the potential enstrophy during vorticity advection by the horizontal flow. Further detail has been given by Burridge and Haseler (1977) and Burridge (1979).

3.2 The spectral model

A more detailed description of the spectral model will be given, although it largely follows the adiabatic formulation described by Baede et al. (1979) in an ECMWF Technical Report. The basic prognostic variables of the model are ξ, D, T, q and $\ln p_s$, where ξ and D are the vorticity and divergence computed on surfaces of constant η:

$$\xi = \frac{1}{a\cos\theta}\left\{\frac{\partial v}{\partial \lambda} - \frac{\partial}{\partial \theta}(u\cos\theta)\right\}$$

$$D = \frac{1}{a\cos\theta}\left\{\frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \theta}(v\cos\theta)\right\}$$

where a is the radius of the earth, λ is longitude and θ is latitude. Variables are represented in the horiozntal by truncated series of spherical harmonics:

$$X(\lambda, \mu, \eta, t) = \sum_{m=-M}^{M} \sum_{n=|m|}^{N(m)} X_n^m(\eta, t)\, P_n^m(\mu) e^{im\lambda} \qquad (9)$$

where X is any variable and μ is $\sin\theta$. The $P_n^m(\mu)$ are the Associated Legendre Functions, defined here by

$$P_n^m(\mu) = \sqrt{(2n+1)\frac{(n-m)!}{(n+m)!}}\ \frac{1}{2^{n+1}\,n!}\,(1-\mu^2)^{m/2}\,\frac{d^{n+m}}{d\mu^{n+m}}(\mu^2-1)^n\ ,\ m \geq 0, \qquad (10)$$

and

$$P_n^{-m}(\mu) = P_n^m(\mu)$$

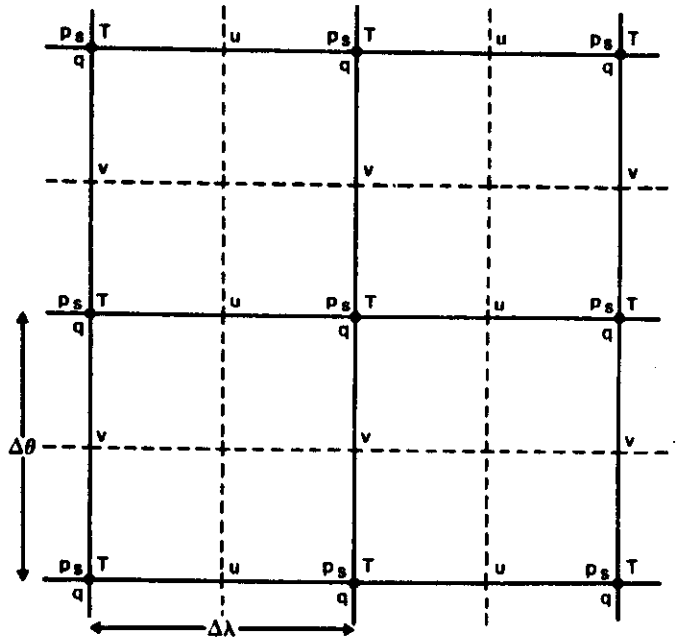


Fig. 1 Horizontal distribution of variables in the grid-point model. Operationally $\Delta\lambda = \Delta\theta = 1.875^{o}$.

and the normalization is such that

$$\frac{1}{2}\int_{-1}^{1} P_n^m(\mu)\, P_s^r(\mu)\, d\mu = \delta_{mr}\,\delta_{ns} \tag{11}$$

The X_n^m are the complex-valued spectral coefficients of the field X. Since X is real,

$$X_n^{-m} = (X_n^m)^*$$

where $(\ )^*$ denotes the complex conjugate. The model thus deals explicitly only with the X_n^m for $m \geq 0$.

The Fourier coefficients of X, $X_m(\mu,\eta,t)$, are defined by

$$X_m(\mu,\eta,t) = \sum_{n=|m|}^{N(m)} X_n^m(\eta,t)\, P_n^m(\mu) \tag{12}$$

with

$$X(\lambda,\mu,\eta,t) = \sum_{m=-M}^{M} X_m(\mu,\eta,t)\, e^{im\lambda} \tag{13}$$

Derivatives are given analytically by

$$\frac{\partial X}{\partial \lambda} = \sum_{m=-M}^{M} im\, X_m\, e^{im\lambda} \tag{14}$$

and

$$\left(\frac{\partial X}{\partial \mu}\right)_n^m = \sum_{n=|m|}^{N(m)} X_n^m \frac{dP_n^m}{d\mu} \tag{15}$$

where the derivative of the Legendre Function is given by the recurrence relation:

$$(1-\mu^2)\frac{dP_n^m}{d\mu} = -n\,\epsilon_{n+1}^m P_{n+1}^m + (n+1)\,\epsilon_n^m P_{n-1}^m \tag{16}$$

with

$$\epsilon_n^m = \left(\frac{n^2-m^2}{4n^2-1}\right)^{\frac{1}{2}}$$

As in the first ECMWF spectral model (Baede et al., 1979) the model is programmed to allow for a flexible pentagonal truncation, depicted in Fig. 2. This truncation is completely defined by the three parameters J, K and M illustrated in the Figure. The common truncations are special cases of the pentagonal one:

Triangular	M = J = K
Rhomboidal	K = J + M
Trapezoidal	K = J, K > M

The spectral calculation utilizes the transform technique pioneered by Eliasen et al. (1970) and Orszag (1970). It follows that of the early multi-level spectral models described by Bourke (1974) and Hoskins and Simmons (1974), and the ECMWF spectral model reported by Baede et al. (1979), although it differs in its use of an advective rather than a flux form for the temperature and moisture equations. The objective of the calculation is to compute spectral tendencies $\left(\frac{\partial X}{\partial t}\right)_n^m$ for each prognostic variable, from which new values may be computed using the time differencing discussed in Section 5. The orthogonality of the spherical harmonics is such that these spectral tendencies are related to grid-point tendencies by

$$\left(\frac{\partial X}{\partial t}\right)_n^m = \frac{1}{4\pi}\int_{-1}^{1}\int_{0}^{2\pi} \left(\frac{\partial X}{\partial t}\right) P_n^m(\mu)\, e^{-im\lambda}\, d\lambda\, d\mu \tag{17}$$

An outline of the model's computation of spectral tendencies can now be given. First, a grid of points covering the sphere is defined. Using the basic definition of the spectral expansions (9) and the linear equations relating wind components with vorticity and divergence, values of ζ, D, u,v,T,q and $\ln p_s$ are calculated at the grid points, as also are the required derivatives $\frac{\partial T}{\partial \lambda}$, $\frac{\partial T}{\partial \mu}$, $\frac{\partial q}{\partial \lambda}$, $\frac{\partial q}{\partial \mu}$, $\frac{\partial \ln p_s}{\partial \lambda}$ and $\frac{\partial \ln p_s}{\partial \mu}$ using (12) - (15). The resulting grid-point values are sufficient to calculate the required grid-point contributions to adiabatic tendencies, and also the parameterized tendencies since prognostic surface fields associated with the parameterization are defined and updated on the same grid. The integrands of the prognostic equations of form (17) are thus known at each grid-point, and approximate spectral tendencies are calculated by numerical quadrature. Integration by parts is used to avoid computation of some derivatives:

$$\int_{-1}^{1} \frac{\partial A}{\partial \mu} P_n^m\, d\mu = -\int_{-1}^{1} A \frac{dP_n^m}{d\mu}\, d\mu$$

where $\frac{dP_n^m}{d\mu}$ is known from (16).

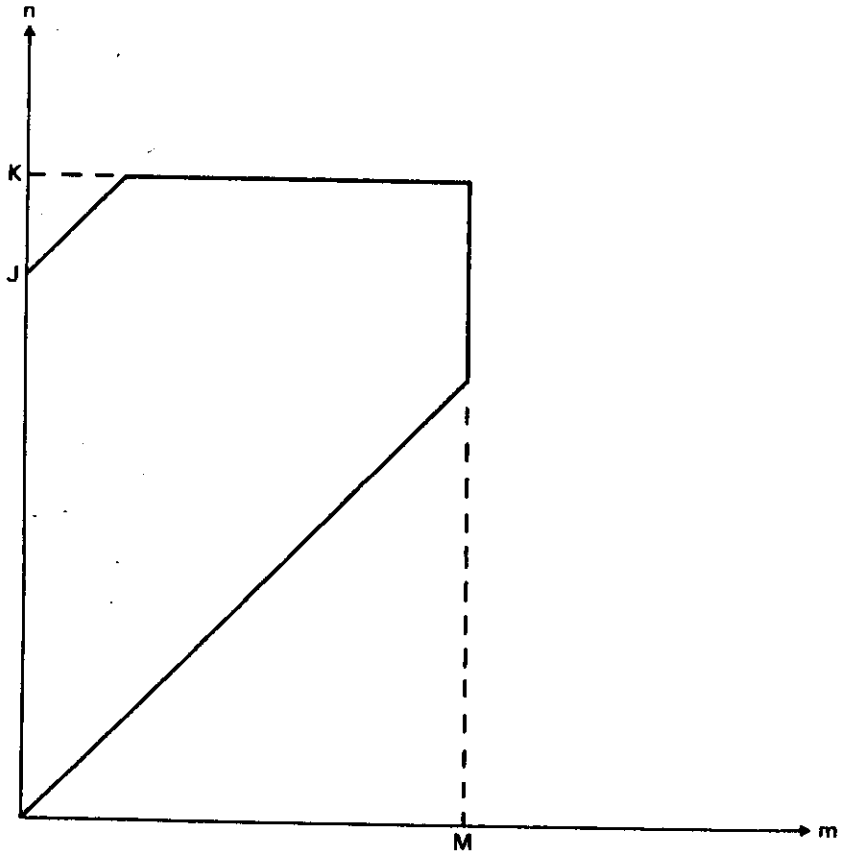


Fig. 2 Pentagonal truncation.

The grid on which the calculations are performed is in fact determined to give an exact (given the spectral truncation of the fields, and within round-off error) contribution to spectral tendencies from quadratic non-linear terms. The integrals with respect to λ involve the product of three trigonometric functions, and as shown by Machenhauer and Rasmussen (1972) they may be evaluated exactly using a regularly-spaced grid of at least 3M+1 points. For the latitudinal integrals, Eliasen et al. (1970) showed that quadratic non-linear terms lead to integrands which are polynomials in μ of a certain order. They may thus be computed exactly using Gaussian quadrature with points located at the (approximately equally-spaced) latitudes which satisfy $P_L^0(\mu) = 0$, for a sufficiently large integer N_G. These latitudes form what are referred to as the "Gaussian latitudes". For triangular truncation, the minimum value of N_G is (3M+1)/2.

It is likely that triangular truncation with M = 63 will be adopted for the first operational version of the spectral model. The associated grid of 192 longitude points and 96 latitude points is a very close equivalent of the regular N48 grid used by the operational grid-point model. Detail in addition to that given here will be found in the documentation manual for the new forecast model.

3.3 The quasi-operational comparison of grid-point and spectral techniques

The primary factor influencing the decision to change operationally to the spectral technique was the better performance of the technique in an extended experiment comparing forecasts performed once per week for a complete year (Girard and Jarraud, 1982). In this experiment, the operational grid-point model forecasts were compared with spectral forecasts using triangular truncation at total wavenumber 63 (T63). The two models used identical parameterization schemes, and required a similar amount of computing resources. Although the models often gave a very similar forecast, some clear differences in overall performance were found. An indication of this is given by Fig. 3, while Fig. 4 presents one example (out of by no means few) of a markedly better local forecast by the spectral model.

A question central to the theme of this particular seminar is to what extent will the change to the spectral model influence statistical forecasts of local weather made using the MOS technique with statistics derived from earlier grid-point forecasts. Insofar as the two models give generally similar large-scale forecasts, with substantial differences found mainly in the medium range in places and cases where the grid-point model is subject to a significant error in its prediction of the synoptic scale, the use of spectral forecasts in conjunction with statistics produced using the grid-point model output should not cause

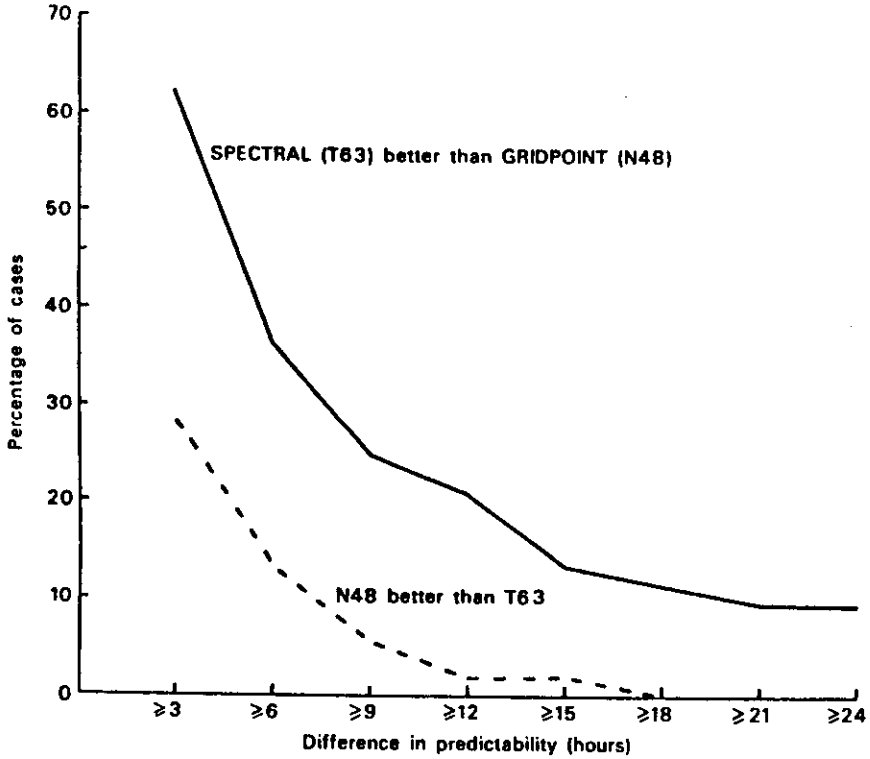


Fig. 3 The difference in predictability (measured by the length of the forecast period for which the anomaly correlation of the 1000 mb height over the extratropical Northern Hemisphere remains above 60%) between spectral (T63) and grid-point (N48) models. Results are expressed in terms of the percentage of cases for which one or other model gave better results.

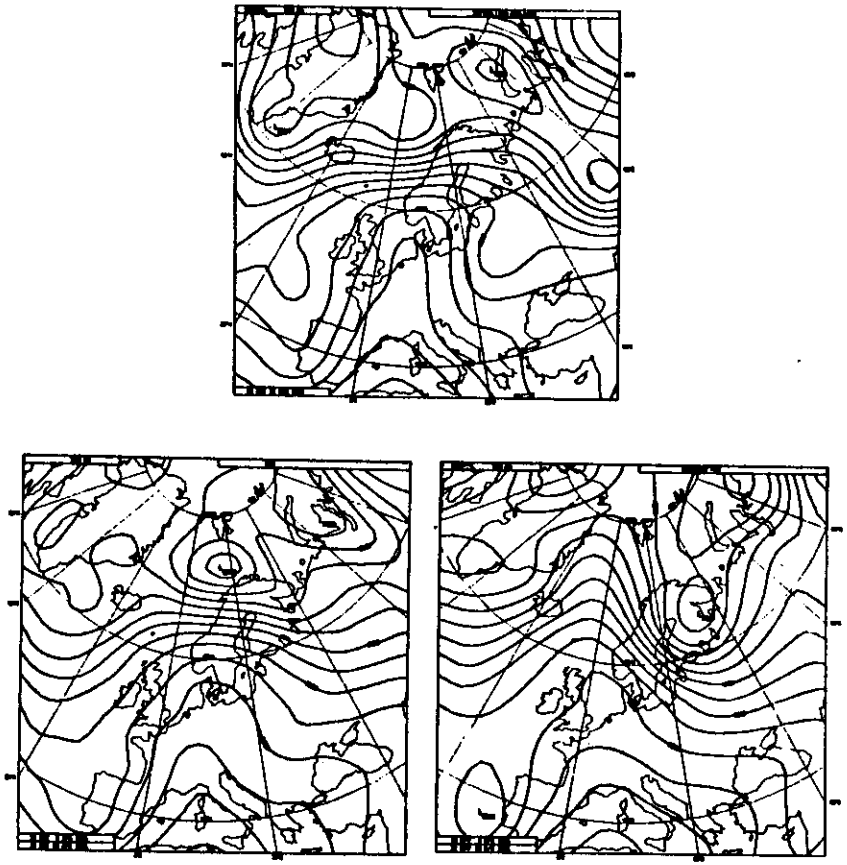


Fig. 4 The analyzed 500 mb height for 10 April, 1981 (upper) and 5-day forecasts for this date by the T63 spectral model (lower left) and the N48 grid-point model (lower right).

particular problems, with the statistical forecasts benefiting from the improved model prediction of the large-scale flow. Two points should, however, be borne in mind.

The first concerns some systematic differences in phase speed found between the two models. Statistics presented in Table 1 from the quasi-operational comparison show phase speeds to be generally better represented by the spectral model, at least in the short range (for which an unambiguous identification of analyzed and forecast lows was possible). In view of such differences, use of model predictors at times shifted from the forecast time of interest in order to compensate for systematic phase errors in synoptic-scale systems should evidently be treated with caution.

Table 1 Errors in the displacement (in degrees longitude) of surface lows between day 1 and day 2 of the forecasts for spectral (T63) and grid-point (N48) model forecasts.

Displacement (D)	Cases	Error (Degrees) T63	N48
$D < 5^{\circ}$	64	+ .6	+1.0
$5^{\circ} < D < 10^{\circ}$	39	+ .3	+ .2
$10^{\circ} < D$	89	-1.8	-2.6
$15^{\circ} < D$	44	-1.8	-3.3
$20^{\circ} < D$	16	-2.9	-4.5

The second point concerns the use of model surface and near-surface parameters. If the parameter in question is particularly sensitive to the nature of the model surface (whether it be land or sea, its height, etc.), then caution is again called for, since the different location of grid-points in the new model may give rise to differences in surface and near surface parameters interpolated from neighbouring grid-points in the vicinity of coastlines and steep orography.

4. THE VERTICAL DISCRETIZATION

The vertical variation of the dependent variables is represented by dividing the atmosphere into a number (NLEV) of layers as illustrated in Fig. 5. In general these layers are defined by the pressure of the interfaces between them (the "half-levels"). Prognostic variables are defined at intermediate levels (the "full-levels"). The precise location of these full levels is not required by the adiabatic formulation (apart from the topmost level) since it generally uses only half-level pressures. Full-level pressures need to be specified, however, for the initial analysis of data and for use in the parameterization schemes.

In the operational grid-point model a sigma coordinate is used, with half levels

$$p_{k+\frac{1}{2}} = \sigma_{k+\frac{1}{2}} p_s \quad , k = 0,1,2,\ldots NLEV, \tag{18}$$

and full levels

$$p_k = \sigma_k p_s \tag{19}$$

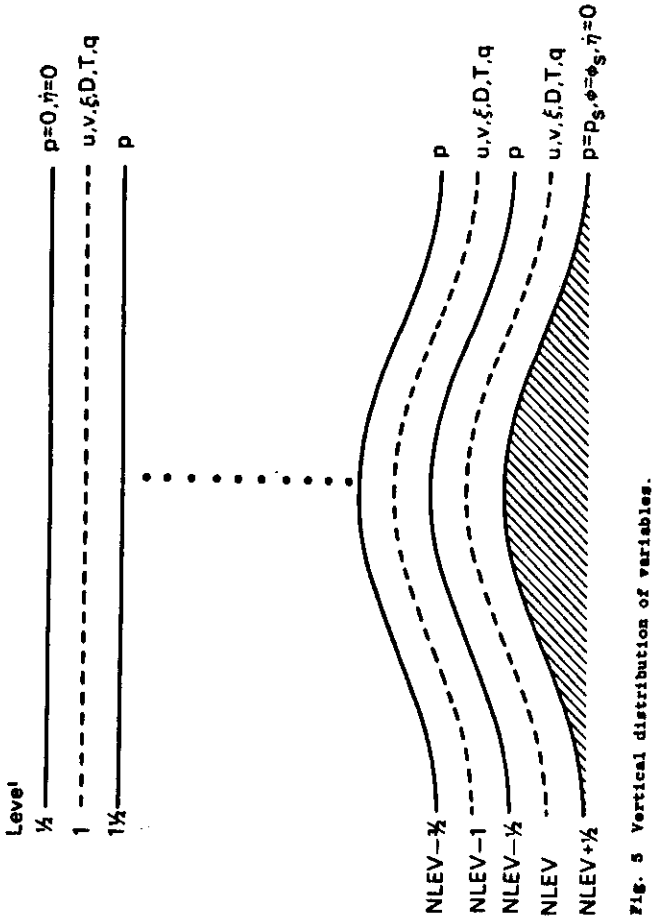


Fig. 5 Vertical distribution of variables.

15 levels are used, and the sigma values are given for both half and full levels by

$$\sigma_k = .75\ S_k + 1.75\ S_k^3 - 1.5\ S_k^4 \qquad (20)$$

where $S_k = (k-\frac{1}{2})/15$. Full-level pressures are given in the left-hand column of Table 2.

In the new model, a more general specification of half-level pressures is adopted:

$$P_{k+\frac{1}{2}} = A_{k+\frac{1}{2}} + B_{k+\frac{1}{2}}\ P_s \qquad (21)$$

with full-level values given by

$$P_k = \tfrac{1}{2}\ (P_{k-\frac{1}{2}} + P_{k+\frac{1}{2}}) \qquad (22)$$

Necessary values are

$$A_{\frac{1}{2}} = B_{\frac{1}{2}} = A_{NLEV+\frac{1}{2}} = 0 \quad , \quad B_{NLEV+\frac{1}{2}} = 1$$

The sigma-coordinate form is reproduced by setting all the $A_{k+\frac{1}{2}}$ to zero, while zero upper-level values of $B_{k+\frac{1}{2}}$ imply that the vertical coordinate is locally a pressure coordinate. Advantages of a hybrid coordinate which transforms smoothly from a sigma coordinate at low levels to a pressure coordinate at upper levels have been discussed by Simmons and Burridge (1981), and Simmons and Strüfing (1981).

Final details of the operational implementation of the new coordinate remain to be finalized, but a radical change in the number or location of levels is unlikely. In particular, the stratospheric resolution over the sea will be essentially unchanged in the first instance, although the topmost one or two levels will be constant-pressure levels if final testing proves satisfactory. Once the system is established operationally, testing of alternative resolutions coincident with the resolution of the stratospheric pressure-level analysis, a possibility which is a potential advantage of the new formulation, will take place.

A minor change in resolution may occur, however, and a possible distribution of 16 full-level pressures is given in Table 2. Above the planetary boundary layer, levels differ little from those of the operational grid-point model, the extra level being used to give a less rapid variation in vertical resolution close to the ground.

Table 2 Full-level pressures (mb) for the original operational model and possible alternatives for the new operational model. Values are for a surface pressure of 1000 mb.

LEVEL	PRESSURE (mb)	
	Original Operational Model	New Operational Model
1	25	25
2	77	75
3	132	128
4	193	185
5	260	250
6	334	324
7	415	406
8	500	496
9	589	589
10	678	681
11	765	769
12	845	846
13	914	909
14	967	955
15	996	982
16		996

The motivation behind the proposed change is to give an unambiguous treatment of the lowest model level. The form of the function (20) defining both full and half-level sigma values is such that the lowest full-level is defined to have a value of sigma approximately equal to

$$1 - \Delta\sigma_{15}/4$$

where $\Delta\sigma_{15}$ is the difference in σ between the ground ($\sigma_{15\frac{1}{2}} = 1$) and the next level ($\sigma_{14\frac{1}{2}}$). In particular, this value is used to define the height of the lowest full level in the boundary layer parameterization. Conversely, the vertical finite-difference scheme for the adiabatic model effectively assumes the lowest level to be at

$$\sigma = 1 - \Delta\sigma_{15}/2$$

There is thus an ambiguity in the treatment of the lowest model level, and this can only be removed by adding an additional level if the resolution is constrained to be essentially unaltered in the free atmosphere and to give an unchanged height of the lowest model level in the boundary-layer scheme. The impact of this possible change on the large-scale forecast is unlikely to be large, although it remains to be seen whether near-surface model fields exhibit an improved behaviour.

The vertical finite-difference scheme for the first operational model has also been described in detail in an earlier seminar (Burridge, 1979), and details will not be repeated here. The representation of the $\frac{RT_v\omega}{p}$ term in (2) is such that the change in potential energy associated with it balances the change in kinetic energy due to the term $\underline{v}.(\nabla\phi + R_d T_v \nabla \ln p)$ that arises in the kinetic energy equation derived from (1). In addition, energy conservation is preserved by the formulation of the vertical advection terms: the representation of $\dot{\eta}\frac{\partial X}{\partial \eta}$ for any variable X is such that

$$\sum_{k=1}^{NLEV} X_k \Delta p_k \quad \text{and} \quad \sum_{k=1}^{NLEV} X_k^2 \Delta p_k$$

are not changed due to finite-difference errors in the treatment of this term. Here X_k denotes the value of X at level k, and $\Delta p_k = p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}}$.

The vertical finite-difference scheme used in the new formulation (21) is a straightforward extension of the sigma-coordinate scheme, and has been discussed by Simmons and Burridge (1981) and Simmons and Strüfing (1981). The only point of difference lies in the choice of a representation of the $\nabla\phi + R_d T_v \nabla \ln p$ term in (1) which ensures no spurious generation or dissipation of angular momentum due to vertical truncation error. The impact of this change on sigma

coordinate forecasts has been found to be extremely small, but idealized calculations of pressure-gradient error over sloping ground for a temperature field dependent only on pressure have indicated that this error is reduced for a hybrid coordinate by choosing the angular-momentum conserving scheme rather than simpler alternative finite-difference approximations (Simmons and Strüfing, 1981).

Simmons and Strüfing (loc. cit.) have also reported on forecast tests using the new vertical scheme. Overall, new and old schemes gave a very similar performance, although the differences that were found generally favoured the new system. The new scheme was not tested in data assimilations, where a further small benefit might be anticipated.

5. TIME SCHEMES

Burridge (1979) has also discussed the semi-implicit time-stepping scheme adtoped for the operational grid-point model, and only the barest outline of the scheme, which derives from the work of Robert et al. (1972), will be given here. If X is a model variable satisfying the equation

$$\frac{\partial X}{\partial t} = \dot{X}$$

the time-scheme for adiabatic terms is formally written

$$X(t+\Delta t) = \tilde{X}(t-\Delta t) + 2\Delta t \{\dot{X}(t) + \tfrac{1}{2}[\dot{X}_g(t+\Delta t) + \dot{X}_g(t-\Delta t) - 2\dot{X}_g(t)]\} \qquad (23)$$

with

$$\tilde{X}(t) = X(t) + \alpha\{X(t+\Delta t) + \tilde{X}(t-\Delta t) - 2X(t)\} \qquad (24)$$

In Eq.(23), X_g represents that component of X associated with linear gravity wave motion about a resting basic state with temperature $T_r(\sigma)$, and the implicit treatment of X_g terms ensures that the time-step criterion is not determined by the rapid ($\approx 300\ ms^{-1}$) phase speed of the model's fastest gravity wave. Eq.(24) describes the time filter analyzed by Asselin (1972), which acts to inhibit the growth of the spurious computational mode associated with the leap-frog scheme.

Operationally, an isothermal reference temperature, with $T_r = 300$ K, is used, a choice governed by the computational stability properties of the semi-implicit technique (Simmons et al., 1978). The value of the time-filtering parameter α is 0.05. A timestep Δt of 15 minutes is generally used with the model, although very strong winds in the polar-night jet of the Southern Hemisphere stratosphere have necessitated a reduction to 12 minutes in September both in 1981 and in 1982.

The extension of the semi-implicit method to the hybrid vertical coordinate discussed in the preceding section has been described by Simmons and Burridge (1981) and Simmons and Strüfing (1981), who also discuss how additional care must be taken in the choice of reference state for this coordinate. Also in the context of the new operational model, an extension of the semi-implicit technique will be introduced. Following results obtained by Robert (1981), who showed that in a semi-implicit shallow-water equation model the time-step limit was determined by the explicit treatment of the vorticity equation, an implicit treatment of the linearized zonal advection of vorticity and moisture will be included. The time-step to be used remains a matter for experimentation.

6. HORIZONTAL DIFFUSION AND THE PRESCRIPTION OF OROGRAPHY

Ideally, the horizontal diffusion that is represented by the K_x terms on the right-hand sides of Eqs. (1)-(3) would be regarded as representing the influence of unresolved scales of motion on the explicitly forecast scales and treated with a physically based parameterization scheme. In practice, since the smallest scales in a model are inevitably subject to numerical misrepresentation, it is common to chose empirically a computationally convenient form for horizontal diffusion and adjust it to ensure that fields of interest do not become excessively noisy. Such an approach has been adopted at ECMWF, and some results may be found in Technical Memoranda by Jarraud and Cubasch (1979) and Strüfing (1982).

The diffusion scheme used operationally since March 1980 may be written in the form

$$K_x = k\, D^4 X + C_x \qquad (25)$$

The operator D^4 is given for the grid-point model by

$$D^4 = \frac{1}{a^4}\left(\frac{1}{\cos^4\theta}\,\delta_\lambda^4 + \delta_\theta^4\right) \qquad (26)$$

where

$$\delta_x X = \frac{1}{\Delta x} \{X(x + \Delta x/2) - X(x - \Delta x/2)\}$$

and the diffusion coefficient k has the value 4.5×10^{15} $m^4 s^{-1}$ in the operational forecasts, a value twice as large being used in data assimilation cycles. The operator δ_θ^4 is computed on values at time-step t-Δt, while δ_λ^4 is applied on the value for t+Δt using Fourier analysis and synthesis. This implicit treatment enables the model to be integrated without any additional spatial filtering to counteract the influence of the convergence of meridians on the time-step criterion.

The term C_x in (25) represents a correction connected with the forecasting of precipitation in mountainous areas. After the introduction of a new steeper orography in April 1981, the uncorrected scheme was found to alad to highly unrealistic precipitation patterns and amounts near mountains. Since the diffusion scheme mixed temperatures on model sigma surfaces, it tended to warm spuriously the mountain tops, and this leads to spurious convection and precipitation.

The ideal way of preventing this happening would be to apply the diffusion on the quasi-horizontal surfaces of constant pressure, but this would not be straightforward to implement, and would be computationally expensive. As a compromise, the following correction operators on temperature and humidity were introduced operationally:

$$C_T = -\left(\frac{\partial T}{\partial \ell n \sigma}\right) D^4 \ell n p_s$$

$$C_q = \frac{q}{q_s} \frac{\partial}{\partial \ell n \sigma} q_s \; D^4 \ell n p_s$$

where q_s denotes the saturation specific humidity. This appears to have largely solved the problem, although it is evident that forecast precipitation must be treated with particular caution in mountainous areas.

Horizontal diffusion in the spectral model also is in the form (25), but with D^4 now representing the ∇^4 operator. Thus in the absence of a correction,

$$K_x = - k \nabla^4 x$$

and

$$(K_x)_n^m = - \frac{k}{a^4} n^2 (n+1)^2 x_n^m$$

It is also applied at time-step t+Δt. The diffusion coefficient used for most past experimentation is smaller than that used in the grid-point model: $k = 7 \times 10^{14}$ $m^4 s^{-1}$. This lessens the liklihood of spurious precipitation near mountains, and this problem may also be lessened by use of the hybrid vertical coordinate (Simmons and Strüfing, 1981). If necessary, a correction of the type used in the grid-point model may be applied operationally, but details remain to be finalized.

A further remark about the prescription of the orography is also appropriate. A study by Wallace et al. (1983) has indicated that a significant part of the systematic error in the operational forecasts of the extratropical height field may be due to inadequate orographic forcing of the large-scale flow, and has shown use of a higher "envelope" orography to result in significant improvements in the medium-range forecasts for a mid-winter period. It thus appears likely that another change in the prescription of orography will take place when further research has been completed. This should be noted when using model output statistics for mountainous areas.

REFERENCES

Arakawa, A. 1966 Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part 1. J. Comp. Phys., 1, 119-143.

Arakawa, A. and V.R. Lamb 1977 Computational design of the basic dynamical processes of the UCLA general circulation model. Methods in Computational Physics, 17, J. Chang, Ed., Academic Press, 337 pp.

Asselin, R. 1972 Frequency filter for time integrations. Mon. Wea. Rev., 100, 487-490.

Baede, A.P.M., M. Jarraud and U. Cubasch 1979 Adiabatic formulation and organization of ECMWF's spectral model. ECMWF Technical Report No. 15, 40 pp.

Bourke, W. 1974 A multi-level spectral model: I Formulation and hemispheric integrations. Mon. Wea. Rev., 102, 688-701.

Burridge, D.M. 1979 Some aspects of large scale numerical modelling of the atmosphere. Proceedings of ECMWF Seminar on Dynamical Meteorology and Numerical Weather Prediction, Vol. 2, 1-78.

Burridge, D.M. and J. Haseler 1977 A model for medium range weather forecasting - Adiabatic formulation. ECMWF Technical Report No. 4, 46 pp.

Eliasen, E., B. Machenhauer and E. Rasmussen 1970 On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Inst. of Theor. Met., Univ. of Copenhagen, Report No. 2.

Girard, C. and M. Jarraud 1982 Short and medium range forecast differences between a spectral and a grid-point model. An extensive quasi-operational comparison. ECMWF Technical Report No. 32. 178 pp.

Hoskins, B.J. and Simmons, A.J. 1975 A multi-layer spectral model and the semi-implicit method. Quart. J.R. Met. Soc., 101, 637-655.

Jarraud, M. and U. Cubasch 1979 Horizontal diffusion experiments with the ECMWF spectral model. ECMWF Technical Memorandum No. 8, 17 pp.

Kasahara, A. 1974 Various vertical coordinate systems used for numerical weather prediction. Mon. Wea. Rev., 102, 509-522.

Machenhauer, B. and E. Rasmussen 1972 On the integration of the spectral hydrodynamical equations by a transform method Inst. of Theor, Met., Univ. of Copenhagen, Report No. 4.

Orszag, S.A. 1970 Transform method for calculation of vector coupled sums: application to the spectral form of the vorticity equation. J. Atmos. Sci., 27, 890-895.

Phillips, N.A. 1957 A coordinate system having some special advantages for numerical forecasting. J. Met., 14, 184-185.

Robert, A.J. 1981 A stable numerical integration scheme for the primitive meteorological equations. Atmos. Ocean, 19, 35-46.

Robert, A.J. 1982 A semi-Lagrangian and semi-implicit numerical integration scheme for the primitive meteorological equations. J. Met. Soc. Japan, 60, 319-325.

Robert, A.J., J. Henderson and C. Turnbull 1972 An implicit time integration scheme for baroclinic models of the atmosphere. Mon. Wea. Rev., 100, 329-335.

Sadourny, R. 1975 The dynamics of finite difference models of the shallow-water equations. J. Atmos. Sci., 32, 680-689.

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

Simmons, A.J. and R. Ströfing 1981 An energy and angular-momentum conserving finite-difference scheme, hybrid coordinates and medium-range weather prediction. ECMWF Technical Report No. 28, 68 pp.

Simmons, A.J., B.J. Hoskins and D.M. Burridge 1978 Stability of the semi-implicit time scheme. Mon. Wea. Rev., 106, 405-412.

Ströfing, R. 1982 Some comparisons between linear and non-linear horizontal diffusion schemes for the ECMWF grid-point model. ECMWF Technical Memorandum No. 60, 42 pp.

Wallace, J.M., S. Tibaldi and A.J. Simmons 1983 Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. Submitted to Quart. J. Roy. Met. Soc.

PARAMETERISATION OF SUB-GRID SCALE PROCESSES.

J.F.Louis

ECMWF

1. INTRODUCTION

The role of the parameterization in the forecast model is to take into account the physical forcings which make the equations not purely adiabatic. This includes the exchange of momentum, heat and moisture at the earth's surface, the horizontal and vertical eddy fluxes of the same quantities, the effect of precipitation and radiative exchanges.

I shall not expand on the horizontal diffusion. Although it should represent the effect of the sub-grid scale horizontal fluxes that arise from the discretization of the model, it is in fact used mainly as a mathematical device to control numerical noise in the integration. In the operational model it is done with a linear, fourth order scheme, which is implicit in the east-west direction.

My purpose, in this paper, is not to give a detailed mathematical description of all parts of the parameterization schemes of the ECMWF model. Such a description can be found in the Forecast Model Documentation Manual and in other published papers (Louis,1979, Geleyn and Hollingsworth,1979). I shall try, instead, to give the reader an idea of the general principles on which our methods are based, with only enough details to understand the forecast products which are derived from the sub-grid scale parameterization, (namely the near surface temperature and winds, the precipitation and the forecast cloudiness).

2. BOUNDARY LAYER FLUXES

Under this heading I include the surface exchanges as well as the eddy fluxes in the atmosphere. Even though these fluxes are most important in the boundary layer, they are computed throughout the model atmosphere since they can provide a fair amount of dissipation and mixing in region of large vertical shear such as near the jet stream. Results from the surface fluxes computations are also used to determine the two-metre temperature and the ten-metre wind which are among the experimental products of the Centre.

The distribution of levels in the forecast model is such that we can resolve some structure in the boundary layer. As can be seen on figure 1, we have four levels in the lowest 1500 m of the atmosphere, and the lowest level is at about 30 m. Note that the heights shown on the figure are approximate heights above the ground, computed for a standard atmosphere.

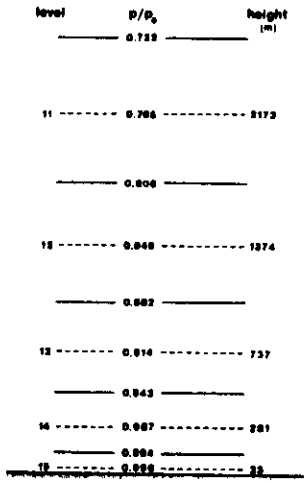


Fig. 1 Vertical distribution of the model levels near the ground

We use the winds and temperature of the lowest level, as well as the ground surface temperature, to compute the fluxes at the surface, according to a drag law. It is well known, however, that surface fluxes are not only dependant on the wind shear at the surface, as in the simplest drag law, but are a strong function of stability. We have used the Monin-Obukhov similarity theory to determine the behaviour of the surface fluxes in terms of the stability.

This theory assumes that, near the ground, the local vertical gradients of potential temperature and wind depend only on the height, the surface fluxes of heat and momentum, and the expansion coefficient of the air. Then, by integrating these flux-gradient relationships between the roughness length z_o and the lowest model level h, it can be shown that the surface fluxes are related to the wind at the lowest model level V_h and the potential temperature difference $\Delta\theta_h$ in the lowest layer through universal functions of h/z_o and the Richardson number R_i given by

$$Ri = g\, h\, \Delta\theta_h \,/\theta \left| V_h \right|^2 \qquad (1)$$

Hence the expressions for the surface fluxes can be written

$$F_M = \left[k \,/\, \ln(h/z_o) \right]^2 \cdot F(h/z_o, Ri) \left| V_h \right| V_h$$
$$F_H = \left[k \,/\, \ln(h/z_o) \right]^2 \cdot G(h/z_o, Ri) \left| V_h \right| \Delta\theta_h \qquad (2)$$

in a way which shows the logarithmic profiles in neutral conditions (i.e. when F=G=1). The constant k is von Karman's constant.

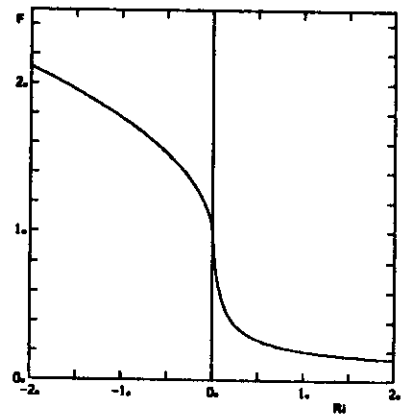


Fig. 2 Variation of the normalized drag coefficient in terms of the Richardson number, for $h/z_o = 50$.

The analytical form of the functions F and G is not uniquely defined. In the course of two years of operational forecasts, we have in fact changed them twice, on the basis of observations of the performance of the model. The variation of F with the stability parameter Ri in the current operational model is shown on figure 2 for $h/z_o=50$. This curve represents the ratio between the drag coefficient under varying stability to the one for a neutral atmosphere.

Above the surface layer we have extended the similarity arguments by assuming that, if the model layers are thin enough, the wind and temperature gradients should depend only on the fluxes through the layer, the expansion coefficient, and a mixing length which is a function of height only. The diffusion coefficients are then related to the Richardson number of the layer through formulae very similar to (2).

It should be noted that the effect of moisture on stability must be taken into account since water vapour is lighter than dry air. This is done by using the virtual potential temperature in the definition of the Richardson number.

The two-metre temperature and the ten-metre wind which are disseminated as experimental products of the Centre's forecast should be consistent with the computation of the fluxes described above. In principle (1) and (2) form a system of equations which can be solved for $\mathbf{V}$ and θ at any height, given the surface fluxes F_M and F_H. However the analytical form of these equations make the solution of this system rather difficult. Hence we do an approximate computation where we first calculate the equivalent roughness length z_o' which would produce the same surface fluxes assuming the logarithmic profile, i.e. z_o' is such that:

$$[\, k / \ln(h/z_o') \,]^2 = [\, k / \ln(h/z_o) \,]^2 \cdot F(h/z_o, Ri) \qquad (3)$$

Then we use the logarithmic profiles (i.e. Equ.2 with F=G=1) with this equivalent roughness length to compute $\mathbf{V}$ at 2 and 10 m and θ at 2 m. A similar interpolation of the humidity is used to compute the dew-point temperature at 2 m.

It should be noted that, so far, the diurnal variation of the solar radiation has not been included in the operational model: we use, as solar input, the average over 24 hours. This means that one should compare the near-surface variables produced by the model with the averaged observed values and not instantaneous ones. In addition, the absence of diurnal cycle is likely to produce too small a depression of the dew-point temperature.

3. MOIST PROCESSES

We distinguish two kinds of clouds in the forecast model: stratiform and convective.

The treatment of stratiform clouds is quite simple. It is assumed that, whenever the total humidity of a grid point becomes greater than its saturation value, the excess moisture condenses and the corresponding latent heat is released. This excess moisture, however, is not necessarily removed as rain immediately. It only precipitates if either the top level of the cloud is cold enough (below -12 C) or the total liquid water content of the column is large enough (greater than 2 mm). The first criterion takes into account the greater efficiency of ice nuclei at low temperature. The second condition crudely simulates the fact that in a deep cloud the droplet distribution has a wider spectrum than in a thin cloud and the coalescence of droplets into rain drops is more rapid. Finally, as the rain falls, it can re-evaporate in the drier layers below the cloud.

The main weakness of this scheme is the assumption that condensation can occur only when the grid point is entirely saturated, even though it is fairly obvious that stratiform clouds can exist which do not entirely fill a volume nearly 200 km on the side and up to 100 mb thick. In order to overcome this difficulty one would need some additional information such as the mixing ratio of liquid water in the clouds and the variance of humidity in the volume. A drawback of the present scheme is that we only have one prognostic variable for moisture in the model: the total amount of water. Hence we cannot advect separately the water vapour and the liquid water and, at each time step, the liquid water is diagnosed as the amount which is supersaturated. We have preferred to choose 100% relative humidity as the critical value rather than some arbitrary lower value. In a future model we shall have a prognostic variable for the amount of liquid water and we may also be able to relax this 100% relative humidity criterion.

Our convective precipitation scheme follows closely the method proposed by Kuo (1965,1974). A convective cloud exists when there is a net convergence of moisture into a conditionally unstable layer. The amount of cloud air is computed as the ratio of the latent heat contained in the converging water vapor to the excess moist static energy of the cloud. Part of the moisture contained in the cloud is then mixed with the environment, while the rest rains out, releasing its latent heat. This fraction is determined by the relative humidity of the environment. Again, some of the rain can re-evaporate below the cloud.

One should be aware that the division between stratiform and convective precipitation is somewhat arbitrary and model-dependent. As an illustration of this statement we show in figure 3 two forecasts done with the limited area version of the Centre's model. One was done with the operational model. In the other forecast the Kuo convection scheme was replaced by a method proposed by Arakawa and Schubert (1974). It can be seen that, although the

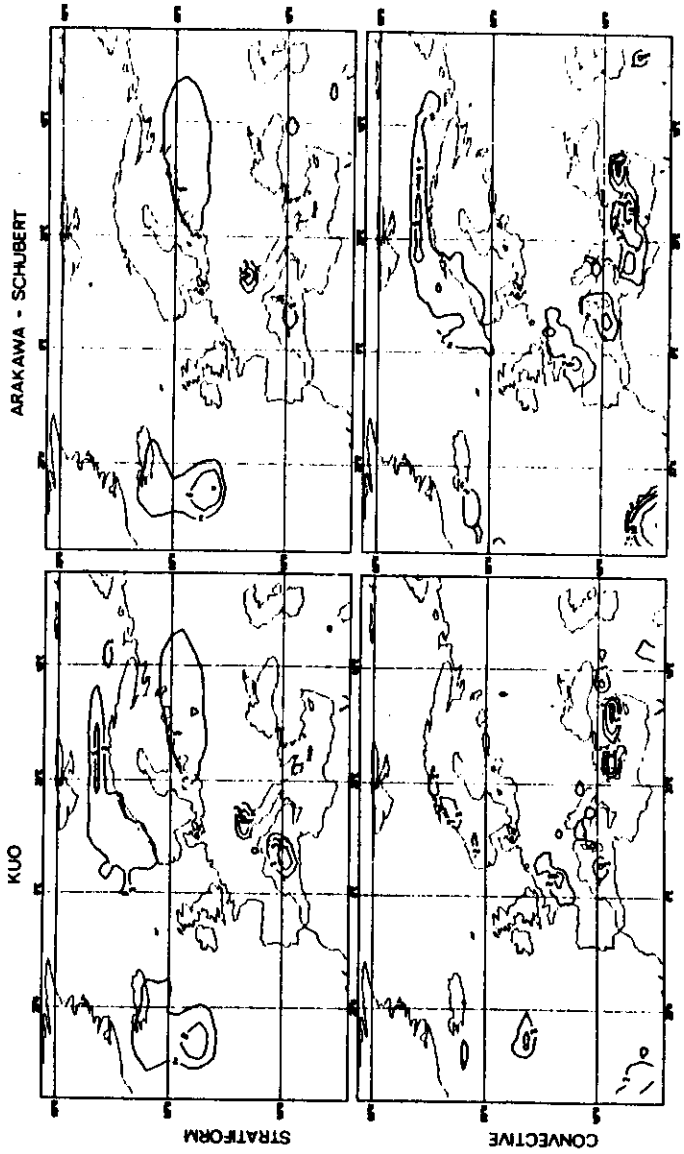


Fig. 3 Comparison of two 24-hour precipitation forecasts

total precipitation is nearly the same in the two forecasts, the partition between stratiform and convective precipitation is very different in the polar region. Another thing which might be worth mentioning is that, at the moment, no distinction is made between snow and rain within the cloud. It is only when the precipitation reaches the ground that it is assumed to be snow if the surface temperature is below 0 C, or rain otherwise. This means that there is a slight energetic inconsistency between the treatment of the clouds and that of the surface processes where we do take into account the heat necessary to melt the snow. This inconsistency will be removed in the future.

4. RADIATION

One can look at the role of radiation in the atmosphere from two different points of view. Globally, radiation provides the primary source of energy for the atmosphere and maintains the thermal gradient between the equator and the poles. Locally, radiation has a large effect on the weather by driving the boundary layer processes and affecting the development of clouds. The importance of the cloud-radiation interactions is obvious for the local weather, but it is no less important for the global heat balance.

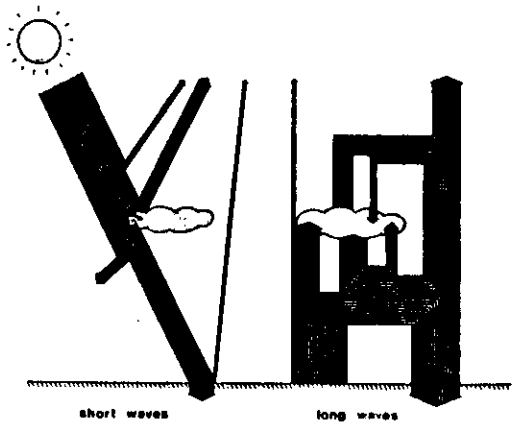


Fig. 4 Qualitative description of the radiative fluxes in the earth/atmosphere system

The diagram of figure 4 illustrates this statement. It describes, in a very schematic way, what happens to the radiative fluxes in the atmosphere. The left side of the figure represents the short wave fluxes, while the right side shows the long wave fluxes. I have not put any numbers on the figure because some of the components are still rather uncertain, but the thickness of the arrows is approximately proportional to the magnitude of the fluxes. The short wave diagram is fairly straightforward, showing that, although little of the sun light is absorbed by the clouds, about one third of the incident solar radiation is reflected by them. The long wave fluxes are a bit more difficult to represent in such a diagram because a lot of absorption and re-emission takes place within the atmosphere, but since the clouds are nearly black bodies in the infra-red, they are clearly important in the transfer of long wave radiation.

In view of the importance of cloud-radiation interactions in both long and short term processes, we have placed a high emphasis on the treatment of the clouds in the radiation scheme. Hence the grey processes are computed first: absorption and scattering by clouds and aerosols, and Rayleigh scattering by the air molecules. In this calculation clouds are allowed to occur in any layer of the model and multiple scattering is taken into account. Then the absorption and emission by the gases (CO_2, H_2O, and O_3) modifies the fluxes resulting from the first part of the computation.

The main problem for the parameterization is to determine the cloudiness of each layer of the model. We cannot, unfortunately, use the information from the parameterization of the moist processes since, with the 100% relative humidity criterion for condensation, we would never have partial cloudiness. Also in its present implementation the Kuo convection scheme does not give us any information on the lateral extent of the cumulus clouds.

We have then used statistics of Pham and Rousseau (1976) on the relationship between relative humidity and the frequency of moist adiabatic lapse rate to evaluate a regression curve between cloudiness and relative humidity. This curve is shown on figure 5 for the 600 mb level. The critical relative humidity below which no clouds are assumed to exist is a inverse function of height. In each layer of the model a fraction of the area corresponding to the curve in figure 5 is assumed to be entirely filled with clouds, and the radiative fluxes are computed separately in the cloudy and clear parts. When clouds exist in two adjacent layers, maximum overlap of the cloudy parts is assumed.

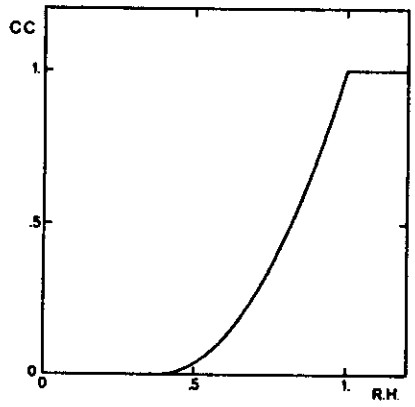


Fig. 5 Assumed relationship between cloud cover and relative humidity at $p/p_s = 0.6$.

It is evident that such a parameterization is very crude, and likely to be poor in the regions which have either tall cumuli penetrating into dry layers, or thin clouds which would not entirely fill a whole layer of the model. The latter problem is particularly bad in the case of thin stratus clouds developing at the top of the boundary layer. Radiation is important in the development of these clouds, with strong cooling at the top and warming within the cloud which destabilises the cloud layer. In the model, however, with a relatively coarse vertical resolution and our assumption that the cloud occupies the whole model layer, the effect of radiation tends to destabilise the whole boundary layer and produce a wrong feed-back involving vertical diffusion and condensation. In order to avoid an exaggerated precipitation in this situation we have been forced to suppress the clouds at the top of an unstable boundary layer, as far as the radiation scheme is concerned.

Once we introduce a prognostic variable for the liquid water content we hope to be able to connect in a more consistent fashion the cloudiness in the radiation scheme and the liquid water determined by the moist processes. At any rate this cloudiness parameter is a quantity which can be directly compared to observations. This is why we routinely produce pseudo-satellite pictures which are simply this radiation cloudiness, integrated over all the levels, and plotted as various shades of grey. An example is given in figure 6, compared with the actual satellite photograph.

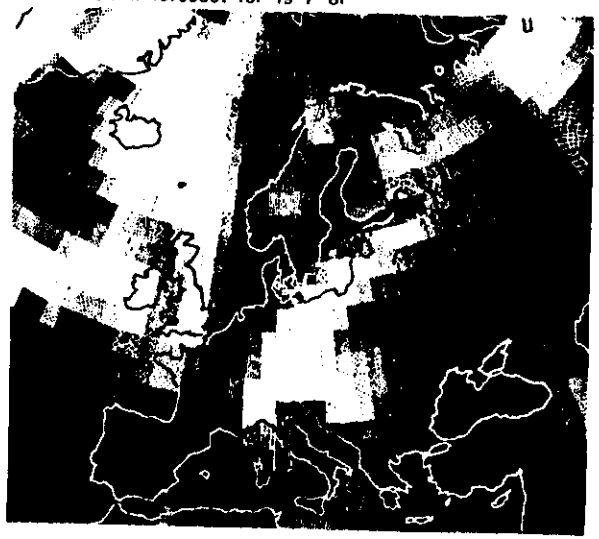
5. GROUND PROCESSES

The model has prognostic equations for the surface temperature and moisture. At the time of writing, the method of computing these variables is about to be modified, and I shall describe the new scheme. It is designed so that we can switch on the diurnal cycle which, as already mentioned, is now suppressed. The scheme is a straightforward simulation of the diffusion equation in the ground, using a finite difference scheme with three layers. The top layer is thin enough to react to the daily cycle. The second layer responds to changes with time scales of about a month. Finally, a climate value is imposed in the bottom layer. The boundary condition at the top is the net flux at the surface: radiative, sensible and latent heat fluxes for the temperature; precipitation and evaporation for the moisture.

Nearly a year ago, a bulk parameterization scheme was introduced, following the ideas of Deardorff (1978). Various problems with our implementation of this scheme have delayed a subsequent introduction of the diurnal cycle. We hope that these difficulties will be avoided with this new scheme so that we can turn on the diurnal cycle soon.

Another element of the parameterization which is important for the radiation is the treatment of the snow, since there is a strong feed-back between the

ECMWF 24 hr forecast for 19·7·81



NOAA-6 satellite photograph [visible]

Fig. 6 An example of a 24-hour cloudiness forecast, with the corresponding satellite photograph

albedo and the ground temperature in the presence of snow. Snow is assumed to have an albedo of 0.8. However, in order to avoid shocks which might produce noise in the forecast, we assumed that the albedo changes smoothly between the value of the bare ground and that of the snow, depending of the snow cover which is taken as a monotonic function of snow depth. The albedo value of 0.8 is reached when the snow depth is about 1 m.

There are a number of weaknesses in the treatment of the snow in the operational model. I have already mentioned that precipitation is assumed to be snow only when the ground temperature is below freezing. Another problem, which is due to the fact the we do not carry information about vegetation in the model, is that we do not take into consideration the fact that forest can have a relatively low albedo, even with deep snow, if it is windy. The change of albedo with the age of the snow is not considered either. Finally, and probably most important, is the fact that we do not change the heat capacity and conductivity of the ground when snow is present. This results in frequent overestimation of the snow covered ground temperature and too rapid melting of the snow. We hope to correct this latter problem some time in the near future.

6. CONCLUSION

I hope that this quick overview of the parameterization schemes of our forecast model has given the reader an idea of the methods we use. It is evident that all the schemes are fairly simple, even though they nonetheless represent the current state of the art in large scale model parameterization. One reason for this simplicity is the need for fast computation. In the present operational model the parameterization takes up about 50% of the whole computing time. We would not want this figure to increase too much with more sophisticated methods. Already we are forced to perform the radiation calculation only twice per forecast day because it would be prohibitively expensive to do it more often in its present form.

In addition to the need for fast computation, we also prefer to keep our schemes simple, (with few arbitrary parameters) in order to understand their behaviour more easily; but at the same time we try to include in the schemes all the significant interactions between the various processes, in order to simulate as well as possible the effect of all the feed-back loops. In future development, we shall try first to remove some of the inconsistencies which still exist, such as the different definition of cloudiness in the condensation and radiation, problems with the snow, and the absence of diurnal cycle.

REFERENCES

Arakawa,A., Schubert,W.H., 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, part I. J.Atmos.Sci.,31,674-701.

Deardorff,J.W., 1978: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. J.Geophys.Res.,83, 1889-1903.

Geleyn,J-F., Hollingsworth,A., 1979: An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. Beitr.Phys.Atmos.,52,1-16.

Kuo,H.L., 1965: On formation and intensification of tropical cyclones-through latent heat release by cumulus convection. J.Atmos.Sci.,22,40-63.

Kuo,H.L., 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J.Atmos.Sci.,31,1232-1240.

Louis,J-F., 1979: A parametric model of the vertical eddy fluxes in the atmosphere. Boundary-Layer Meteor.,17,187-202.