TECHNICAL MEMORANDUM

7

The treatment of humidity in ECMWF's data assimilation scheme

A. Lorenc and S. Tibaldi

Research Department

1979

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

ABSTRACT

The treatment of humidity data in ECMWF's data assimilation scheme is described. The scheme uses a six-hourly intermittent analysis-initialization-forecast cycle, the forecast being made using ECMWF's global 15 level N48 primitive equation model with parameterization of physical processes. The humidity analysis consists of a simple correction to the forecast precipitable water content field, using radiosonde observations of temperature and dew point, and humidities deduced from surface observations of temperature, dew point, current weather and cloud. Code to use satellite soundings of precipitable water content has been written but has yet to be tested.

Preliminary tests indicate that distinguishable structures in the humidity analysis are derived largely from the forecast model.

1. INTRODUCTION

Humidity is certainly an important participant in the interactions governing the motion of the atmosphere and the weather, and any numerical model intended for predictions other than short term predictions of pressure and wind needs to consider its effects. However it is less clear whether it is necessary to specify the humidity field as an independent input to the model. Smagorinsky et al (1970) suggested that, north of 20 $^{\circ}$N, the input humidity field is in fact redundant, the model generating its own in less than a day. However some detailed studies of selected cases do show a positive impact for humidity data, particularly on rainfall amounts during the first few hours (e.g. Atkins 1974). Although in most cases this seems to have little effect on the subsequent forecast, it is probable that instances exist where this latent heat release is important to the subsequent development of a system.

In the tropics the coupling of the mass and wind fields is less strong, making the relaxation time for generating reasonable vertical motion fields, and hence humidity fields and rainfall nearer 3 days. Humidity data is thus more important.

At ECMWF we are concerned with making forecasts for Europe up to 10 days ahead. On such a time scale the influence of the tropics and of developing depressions is considerable, so a humidity analysis has been incorporated in our operational scheme. The same scheme is also being used to make level IIIb analyses, including humidity, using the data gathered during the First GARP Global Experiment.

2. DATA ASSIMILATION

An analysis, if it is to be as accurate as possible, must supplement information from the currently available observations by two other data sources:-

1. Information from earlier observations.
2. Knowledge of the likely structure and scales of atmospheric motion, and of the balance which is usually observed between the various fields (mass, wind, humidity) of the atmosphere.

In a data-assimilation scheme both of these are provided by a numerical model of the atmosphere, which can update information from past observations to the current analysis time, and assimilate all the data into a consistent multivariate three dimensional analysis which represents the atmospheric state in a realistic way. When, as at ECMWF, the main use of the analysis is to provide initial conditions for a numerical forecast, the advantage of using a numerical model for this outweighs the main disadvantage, which is that biases and inaccuracies in the model's formulation and limitations to its resolution mean that the final analysis does not always accurately represent all the detail available in the observations.

Ideally, observations should be inserted into the assimilating model at the valid model time. However, this is difficult to organise, particularly if sophisticated analysis methods are used to help ensure that the information is inserted into realistic scales of motion, with approximate balance between the various fields. At ECMWF a compromise 6 hourly intermittent data-assimilation is used, illustrated in Fig. 1.

The analysis is performed in two stages : mass-wind, and humidity. The mass and wind analysis is a 3-dimensional multivariate statistical interpolation of deviations from the model prediction (Lorenc et al 1977). The humidity analysis is described in Section 3 below.

The non-linear normal mode initialization (Temperton and Williamson 1979) effectively sets to zero the tendencies of selected gravity wave modes in the forecast model. This results in meteorologically reasonable vertical velocities and surface pressure tendencies throughout the subsequent forecast.

The 6-hour forecast in the data assimilation cycle is performed using the same model as is used for ECMWF's 10-day forecasts: a global 15 level N48 primitive equation model, with quite sophisticated parameterizations of physical processes. The adiabatic formulation is described by Burridge and Haseler (1977). The parameterizations include large-scale and convective rain processes, turbulent and surface fluxes of momentum, moisture and sensible heat, and a radiative heating/cooling scheme which take account of the actual model cloud and humidity distribution (Tiedtke et al 1979).

The forecast model and initialization are performed using $\sigma = p/p_{surface}$ as vertical coordinate, while the analysis uses the standard pressure levels at which most observations are conventionally reported. This necessitates vertical interpolations before and after each analysis step, which are at present done using cubic splines in log pressure.

The variable interpolated for humidity is the vertically integrated precipitable water content, so as to preserve the total water content of each column.

3. HUMIDITY ANALYSIS

The analysis scheme in general, and the humidity analysis in particular was designed on the principle that the background field provided by the model forecast is generally quite accurate. Its estimated error and the estimated observational errors of the various observation types are taken into account when determining the weights given to each observation. The interpolated value in the analysis is the deviation from the background field, so that if no datum disagrees with it the background field is unchanged.

The analysis variable is the integrated specific humidity of layers (called analysis layers) enclosed between successive analysis levels. At present the levels used are the standard levels up to 300 mb:- 1000, 850, 700, 500, 400, 300 mb.

$$ {}_{1-2}\Delta Q = \int_{p_1}^{p_2} q \, dp $$

Above 300 mb q is extrapolated towards a climatological value (Harries 1976).

Three types of observation can be used:-

i) TEMPS and TEMPSHIPS (radiosondes)

ii) SYNOPS and SHIPS (surface observations)

iii) SATEMS (satellite soundings).

We shall now describe how information about the analysis variable is derived from each type.

i) Radiosondes report pressure, temperature T, and dew point T_D at a number of levels. (Standard and Special). At each level the specific humidity q is calculated, and its estimated rms observation error δq, using the Clausius-Clapeyron relationship and an estimate for δT_D, which is assumed to vary linearly with ln p from 0.5°K at 1000 mb to 1.0°K at 300 mb. Next ΔQ and its estimated observational error $\delta \Delta Q$ are calculated for the layers between successive observation levels using :-

$$ {}_a\Delta_b Q = \frac{q_a + q_b}{2} (p_a - p_b) $$

$$ \delta_a\Delta_b Q = \frac{\delta q_a + \delta q_b}{2} (p_a - p_b). $$

Data up to 100 mb is used in this procedure and if two consecutive levels are separated by more than 200 mb the sounding is intercepted. The background ΔQ values are interpolated to the same observation layers and subtracted. The observation increments so obtained, and their estimated errors $\delta \Delta Q$ are finally interpolated back to the analysis layers.

ii) Surface observations, as well as providing direct humidity information by way of T and T_D, also report the current weather and cloud amounts and types from which humidity information can be deduced. There have been several efforts to do this in the past (Chisholm et al 1968, Atkins 1974, Kaestner 1974, Jonas 1976, Tuller 1976); we have based our first attempt on a method used by NMC (Chu and Parish 1977).

Each observation is used to provide an estimate of the average relative humidity in four layers roughly equivalent to the planetary boundary layer and the classification of low medium and high cloud. The four layers are defined by the five levels:-

$$p_1 = p_* \quad \text{(the analysed model surface pressure)}$$
$$p_2 = p_1 - 50 \text{ mb}$$
$$p_3 = p_2 - \frac{1}{3}(p_2 - p_5)$$
$$p_4 = p_3 - \frac{1}{3}(p_2 - p_5)$$
$$p_5 = 300 \text{ mb}$$

Three estimates are usually available for the lowest layer - from the observed T and T_D, from the reported current weather using a table given by Chu and Parish (1977), and from the low cloud amount if the cloud base is below 600 m. Their average value is used. For the other layers the relative humidity is a function of cloud amount

$$RH_x = M_x - A_x \cos\left(\frac{\pi}{8} \text{ OKTAS}_x\right) .$$

Again we use values of M and A for each layer derived from those originally proposed by Chu and Parish (1977).

The background values of ΔQ are converted to relative humidities, for the analysis layers, and integrated to give predicted mean relative humidities in each observation layer, assuming the relative humidity to be constant in each analysis layer. This is subtracted to give an observed increment of relative humidity. These are then interpolated back to the analysis layers and then converted to increments for ΔQ. $\delta\Delta Q$ is calculated assuming an observational error of .15 in the relative humidities, except in the analysis layer using the observed surface values, where a lower value is assumed.

iii) Satellite soundings provide precipitable water content data between a reference level and standard pressure levels. This is converted to ΔQ between observation levels and the predicted ΔQ is subtracted. The resulting deviation is partitioned amongst the overlapping analysis layers. This part of the code has not yet been tested for lack at ECMWF of suitable decoded observations. The observation errors for these data have yet to be decided.

The observed increments are finally interpolated horizontally to the analysis grid points using a simple "weighted average" method, and added to the background field to give the analysis. The interpolation equations used are designed to give statistically optimum weights, minimizing the expected interpolation error, for isolated data, while remaining well behaved for dense data and being reasonably cheap to compute. Defining -

$$D_i = \frac{\Delta Q_i^{\text{observed}} - \Delta Q_i^{\text{predicted}}}{\delta \Delta Q_i^{\text{predicted}}}$$

$$E_i = \frac{\delta \Delta Q_i^{\text{observed}}}{\delta \Delta Q_i^{\text{predicted}}}$$

The prediction error correlation is assumed to have the form

$$\mu_{ki} = e^{-r_{ki}/r_o}$$

r_o is currently assumed to be 300 km for the lowest analysis layer increasing to 400 km at 300 mb.

Then the interpolation weights are given by

$$W_{ki} = \frac{\mu_{ki}/(1+E_i-\mu_{ki})}{1+\sum_j \mu_{kj}/(1+E_j-\mu_{kj})}$$

and the interpolated value by

$$\frac{\Delta Q_k^{\text{analysed}} - \Delta Q_k^{\text{predicted}}}{\delta \Delta Q_k^{\text{predicted}}} = \frac{\sum_j W_{kj} D_j}{1 + \sum_j W_{kj}}$$

$\delta \Delta Q^{\text{predicted}}$ is estimated as a fraction of the climatological ΔQ.

4. TEST RESULTS

A test experiment to evaluate the impact of the humidity analysis has been run using data from the FGGE end to end test. Starting from a January climatological field the full data assimilation scheme was run from 06Z 13th January 1979 to 00Z 16th January 1979 (experiment A). The German hand drawn sea level pressure analysis for 00Z on the 16th is shown in Fig. 2. It can be compared with our 1000 mb height and wind analysis (Fig. 3) since 5 mb is approximately equivalent to 4 dm. To show the quality of the 6 hour forecast used as first guess this is shown in Fig. 4. It is necessary to study the details of the analysis to evaluate its quality; for instance the analysis has drawn the wave developing over Scotland better than the first guess. Similarly it is necessary to study the details of the humidity analysis to detect changes from the first guess. The mean relative humidity analysis for the 850-700 mb layer is shown in Fig. 5, and the model forecast first guess in Fig. 6. There are differences, for example the humidity associated with the front across the centre of the figures is reduced in the analysis. Unfortunately it is less easy to verify that the analysis is better. One can readily confirm that the overall pattern, common to the first guess and the analysis, is reasonable. The band of high humidity just mentioned is well correlated with the analysed fronts, and its well defined western edge shows up clearly on Meteosat cloud pictures. The band of drier air over western Europe agrees quite well with the observations of clear skies. However none of these comparisons are precise enough to differentiate between the forecast and the analysis.

The drying of the atmosphere by the analysis, mentioned above for the frontal band, is apparently a global effect, the model's average integrated specific humidity dropping over the 3 days of experiment A from 283 to 259 pascals as shown in Fig. 7. For all the analysis layers and for both sounding and surface observations the global mean observation first guess differences were usually negative. Clearly a 3-day period is not long enough for the assimilation scheme to reach an equilibrium for the atmospheric water content.

Global root mean square difference statistics for 00Z on the 16th are given in Table 1. Note that the analysis fits the sounding data better than that deduced from the surface observations, owing to the much smaller assumed observational errors of sounding data. However, the analysis does not manage to fit the soundings to within their assumed error. This indicates that the assumed error, based on an optimistic estimate of the accuracy of dew point temperature data, is unrealistically low. The assumed observational error should also contain a component due to scales which cannot be resolved by the analysis grid. The rms observed-analysis statistics given are calculated in the way most favourable to the analysis, since the same data, treated in the same way, are used for the analysis and the verification. Results from a verification package independent of the analysis program, which compared spot values of relative humidity from the forecast model with observed values from soundings, shows a much smaller difference between the first guess and analysis statistics.

In order to assess the impact of the humidity analysis on the assimilation scheme a second experiment (experiment B) was started from 00Z on the 15th, in which the humidity field in the forecast model was unchanged by the analysis. The total humidity curve for this experiment is also shown on Fig. 7. Evidently it is the unsettling effect of the humidity analysis, or more likely the vertical interpolation before and after the analysis, which causes the forecast model to increase the humidity. As soon as this is stopped the humidity levels off and even starts dropping slightly. The rms verification statistics for the 6 hour forecast of experiment B valid at 00Z on the 16th are also shown in Table 1. As we would hope these are worse than those for experiment A. The 1000 mb height and wind forecast of experiment B is almost exactly identical with that shown in Fig. 3. The 850-700 mb relative humidity forecast is shown in Fig. 8. This appears to be just as organised and correlated with the frontal and cloud features as Fig. 5, the main difference being that the moist regions are more uniformly moist while the dry regions are more uniformly dry.

5. CONCLUSIONS

No definite conclusions can be drawn from this limited study. We can say that our data assimilation scheme produces horizontal humidity structures which appear meteorologically realistic, but that most of the features in the analysis may be due to the forecast model.
For this single short period of tests the humidity analysis had little impact on the quality of the subsequent 6 hour forecasts of the mass and wind field. The humidity data assimilation scheme needs tuning in several aspects, the most important being perhaps the methods of vertical interpolation between the forecast and analysis layers.

6. PLANS FOR FURTHER WORK

Research is continuing on several aspects of the scheme:-

i) The methods of vertical interpolation are being studied, with three lines of approach; (a) to interpolate the analysis increments to sigma coordinates before adding them to the forecast, so that in data voids the forecast field will be precisely unaltered by the analysis process, (b) to interpolate vertically relative humidity, a more vertically uniform variable, (c) to study vertical soundings and cross sections of the analysis and observations, to see whether any vertical coupling would be desirable in the analysis.

ii) When they become available at ECMWF we plan to test the impact of satellite humidity soundings.

iii) The method of infering humidity data from surface observations can be refined, for instance to give relatively greater weight to cases where a clear inference can be made, such as during continuous heavy rain.

iv) The model parameterization scheme can perhaps be tuned, for instance the criterion for large scale rainfall, so as to reduce the bias between forecast and observations.

v) Further impact studies on fields such as the rainfall and on longer forecasts are planned.

REFERENCES

Atkins, M.J.	1974	'The objective analysis of relative humidity.' Tellus, 26, 663-671.
Ball, J.T. Veigas, K.W.	1968	'The analysis of upper-level humidity.' J. Appl. Met., 7, 620-625.
Burridge, D.M. Haseler, J.	1977	'A model for medium range weather forecasting - Adiabatic formulation.' ECMWF Technical Report No. 4, 46p.
Chisholm, D.A. Ball, J.T. Veigas, K.W. Luty, P.V.	1968	'The diagnosis of upper-level humidity.' J. Appl. Met., 7, 613-619.
Chu, R. Parrish, D.	1977	'Humidity analyses for operational prediction models at the National Meteorological Center.' NOAA NMC Office Note 140.
Harries, J.E.	1976	'The distribution of water vapor in the stratosphere.' Rev. Geophys. and Space Phys., 14, 565-575.
Jonas, P.R.	1976	'The use of surface synoptic data to estimate upper-level relative humidity over the sea.' Met. Mag., 105, 44-56.
Kaestner, A.	1974	'A procedure for objective analysis of relative humidity.' Arch. Met. Geoph. Biokl., Ser.A, 23, 137-148.

Lorenc, A.
Rutherford, I
Larsen, G. 1977 'The ECMWF analysis and data assimilation scheme - Analysis of mass and wind fields.' ECMWF Technical Report No. 6, 47p.

Smagorinsky, J.
Mikakoda, K.
Strickler, R.F. 1970 'The relative importance of variables in initial conditions for dynamical weather prediction.' Tellus, <u>22</u>, 141-157.

Temperton, C.
Williamson, D.L. 1979 'Normal mode initialization for a multi-level grid point model.' ECMWF Technical Report No. 11, 91p.

Tiedtke, M.
Geleyn, J-F.
Hollingsworth, A.
Louis, J-F. 1979 'ECMWF Model - Parameterization of sub-grid scale processes.' ECMWF Technical Report No.10, 46p.

Tuller, S.E. 1977 'The relationship between precipitabl water, vapor and surface humidity in New Zealand.' Arch. Met. Geoph. Biokl., Ser.A, <u>26</u>, 197-212.

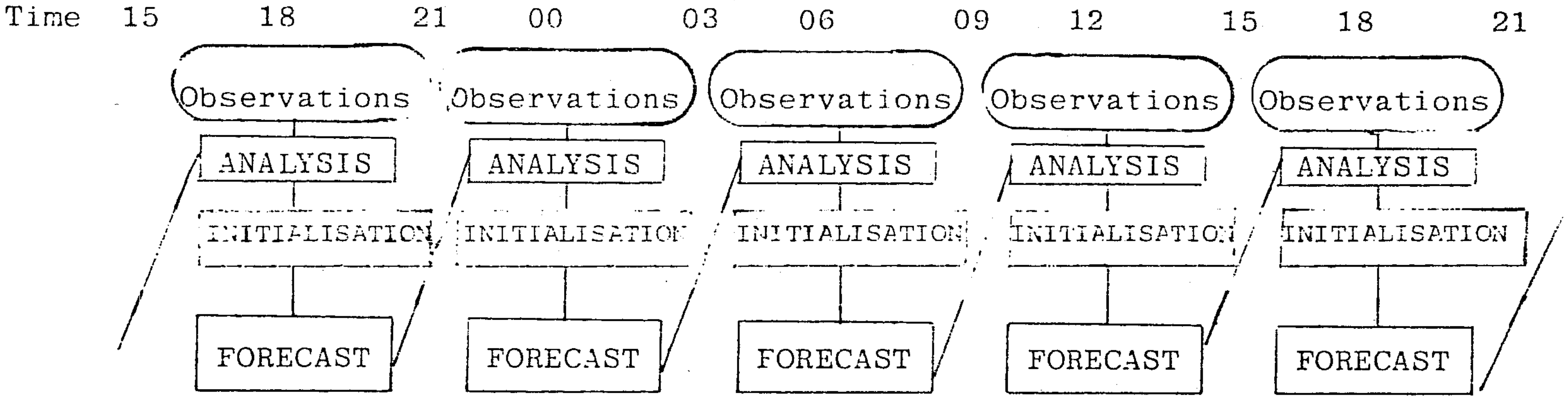

Fig. 1 INTERMITTENT DATA ASSIMILATION

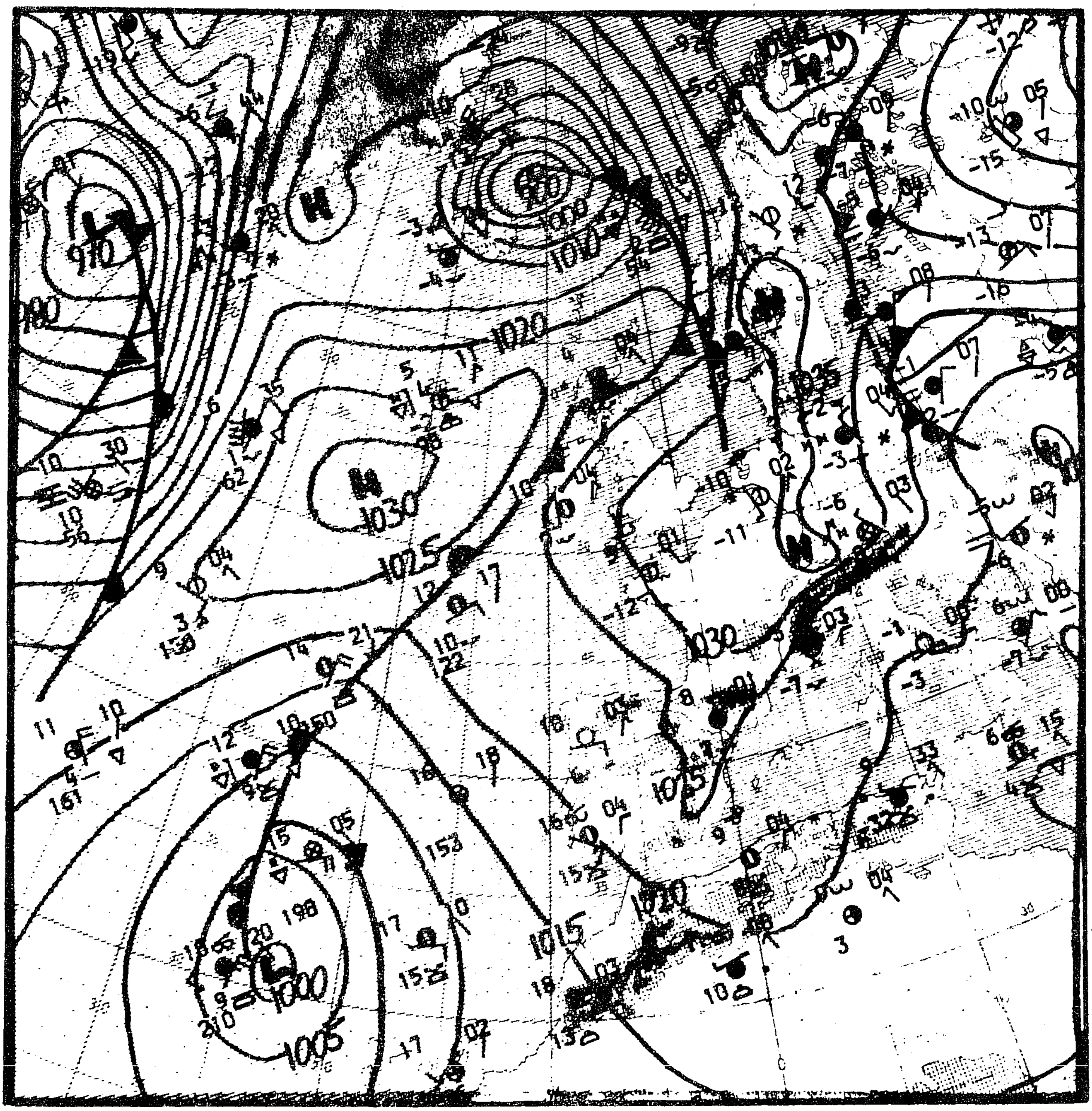

Fig. 2 Europäisher Wetterbericht hand drawn sea level pressure analysis for 00Z 16 January 1979.

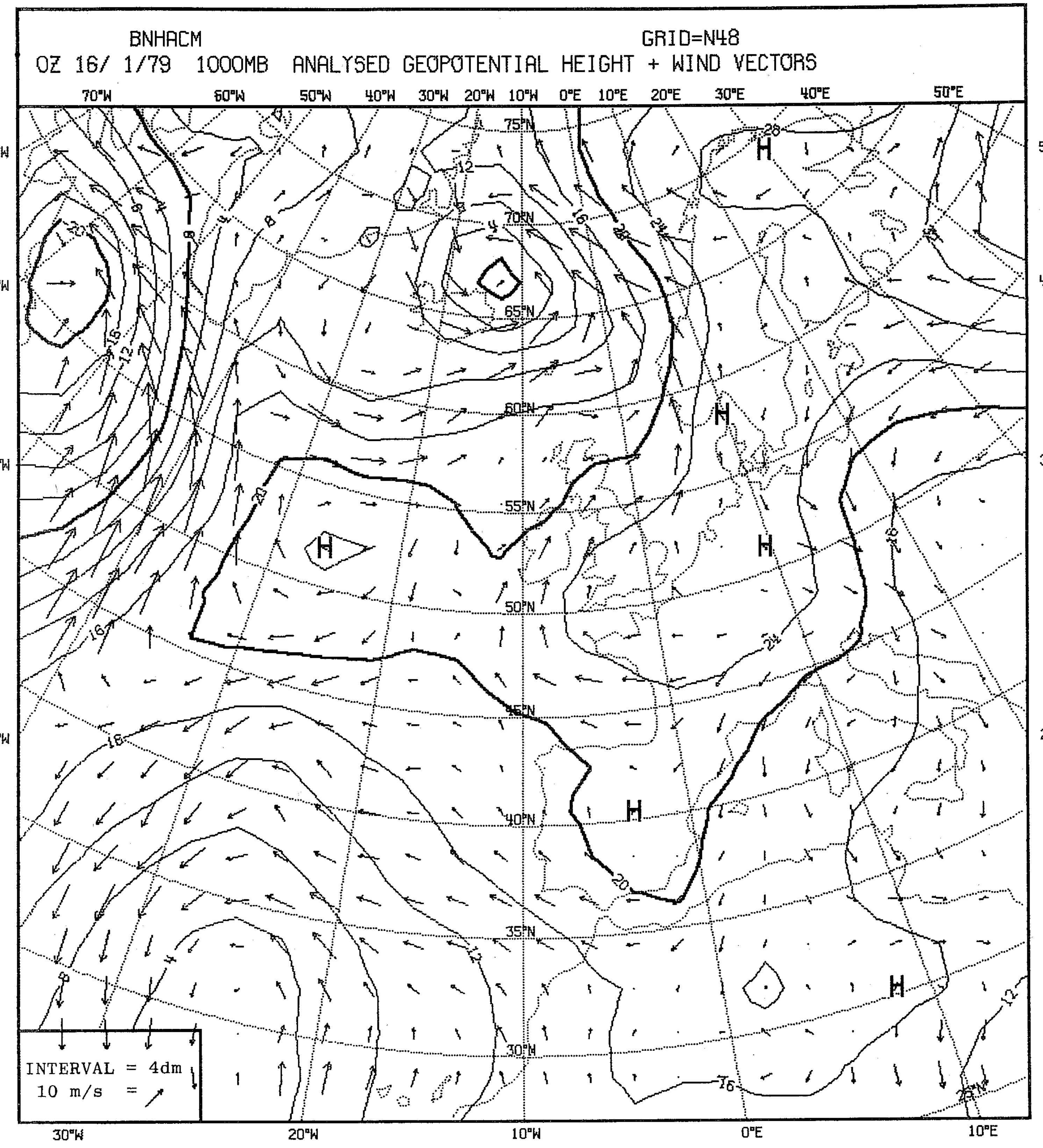

Fig. 3 ECMWF 1000mb height and wind analysis for 00Z 16 January 1979 (experiment A)

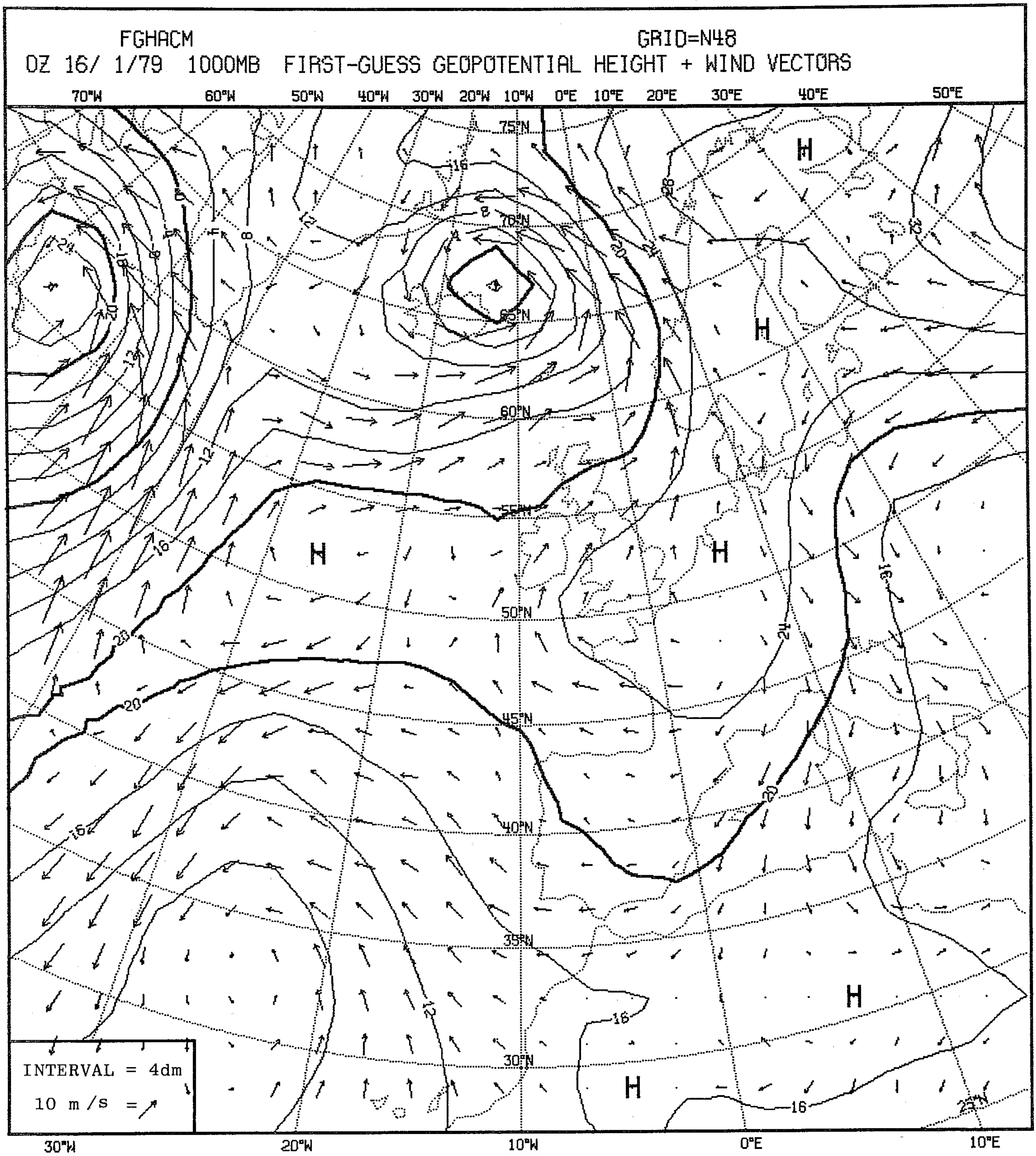

Fig. 4 ECMWF 1000mb height and wind 6 hour forecast valid

00Z 16 January 1979

(experiment A)

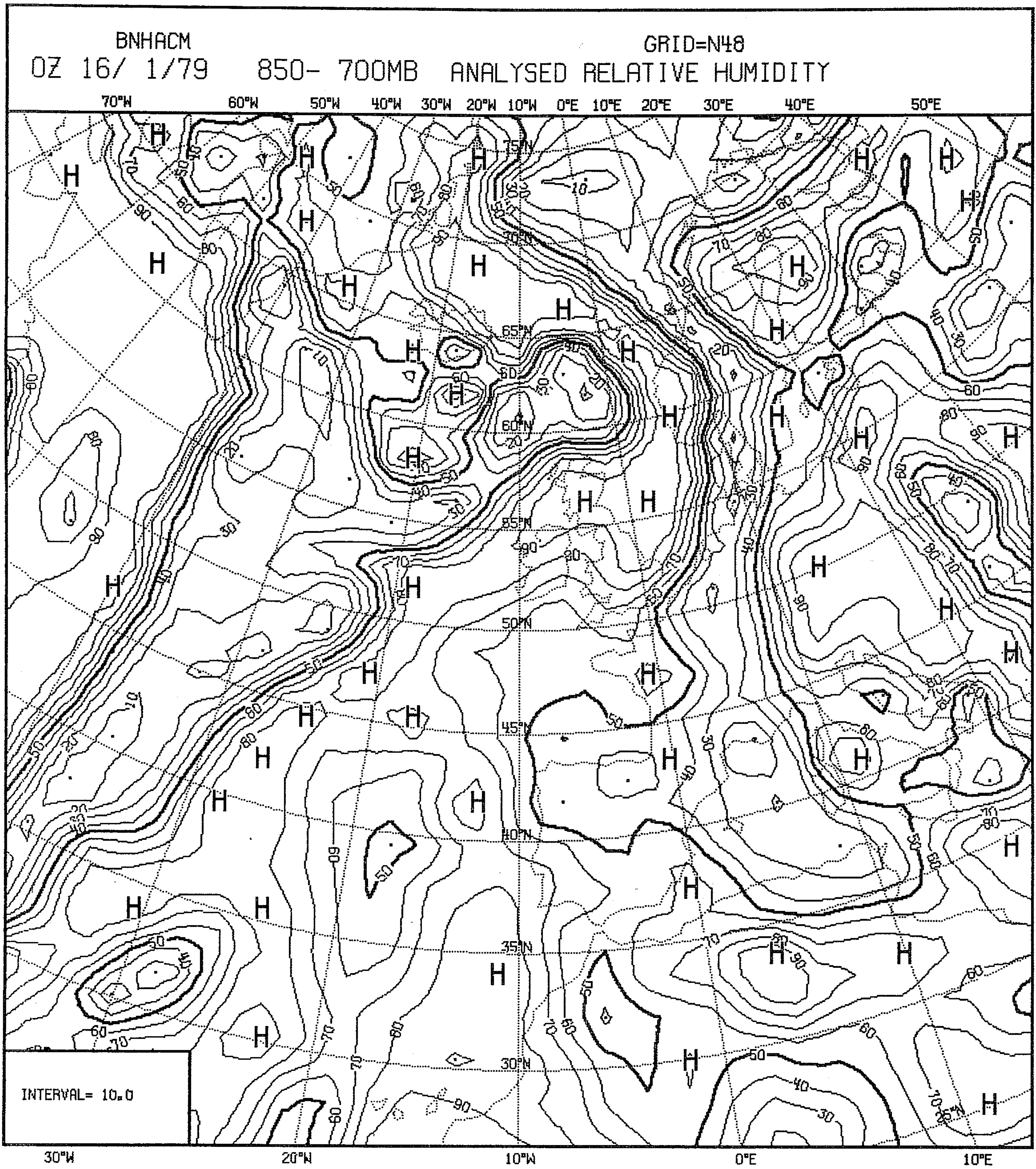

Fig. 5 ECMWF 850-700mb mean relative humidity analysis for

00Z 16 January 1979

(experiment A)

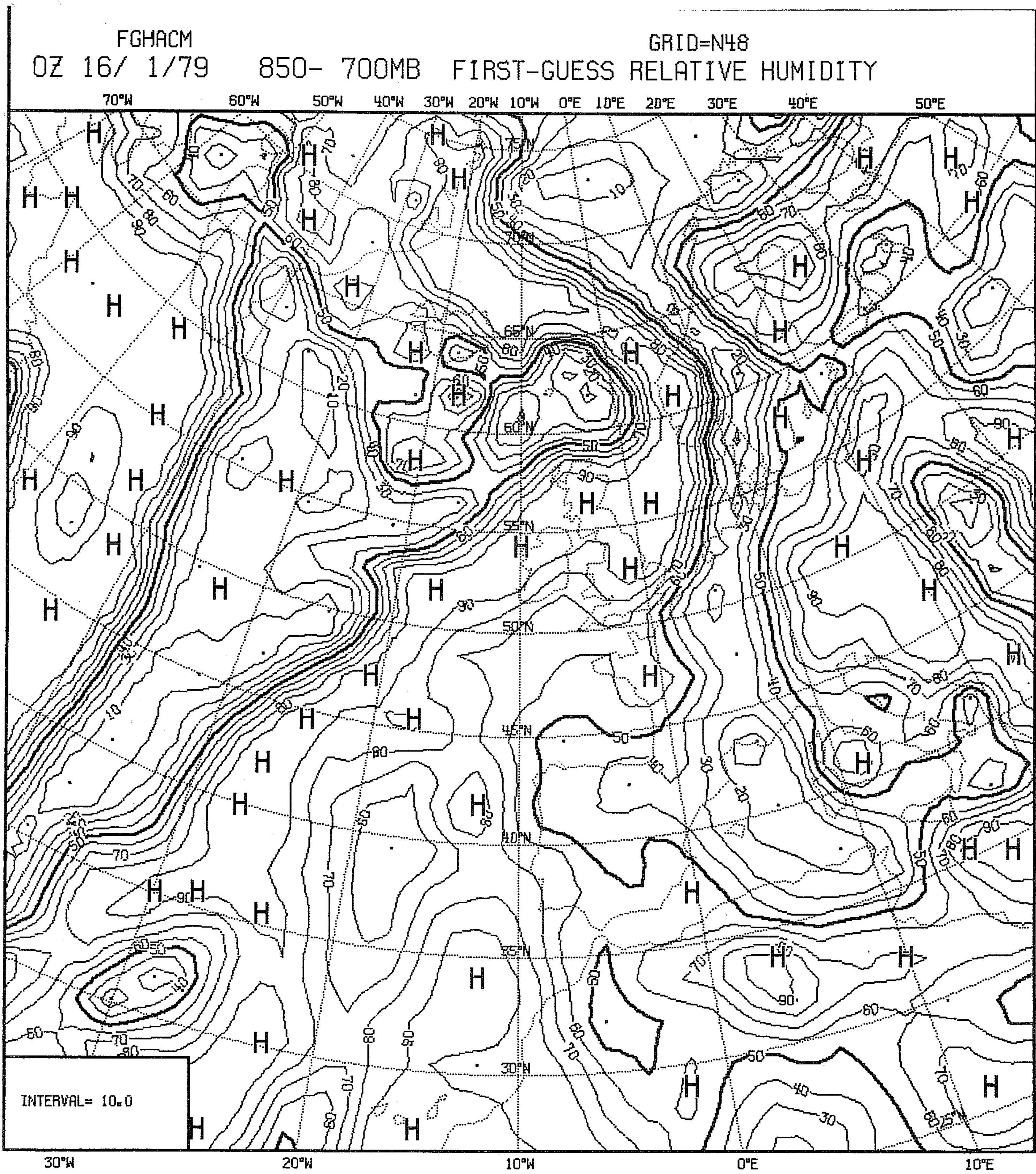

Fig. 6 ECMWF 850-700mb mean relative humidity forecast valid

00Z 16 January 1979

(experiment A)

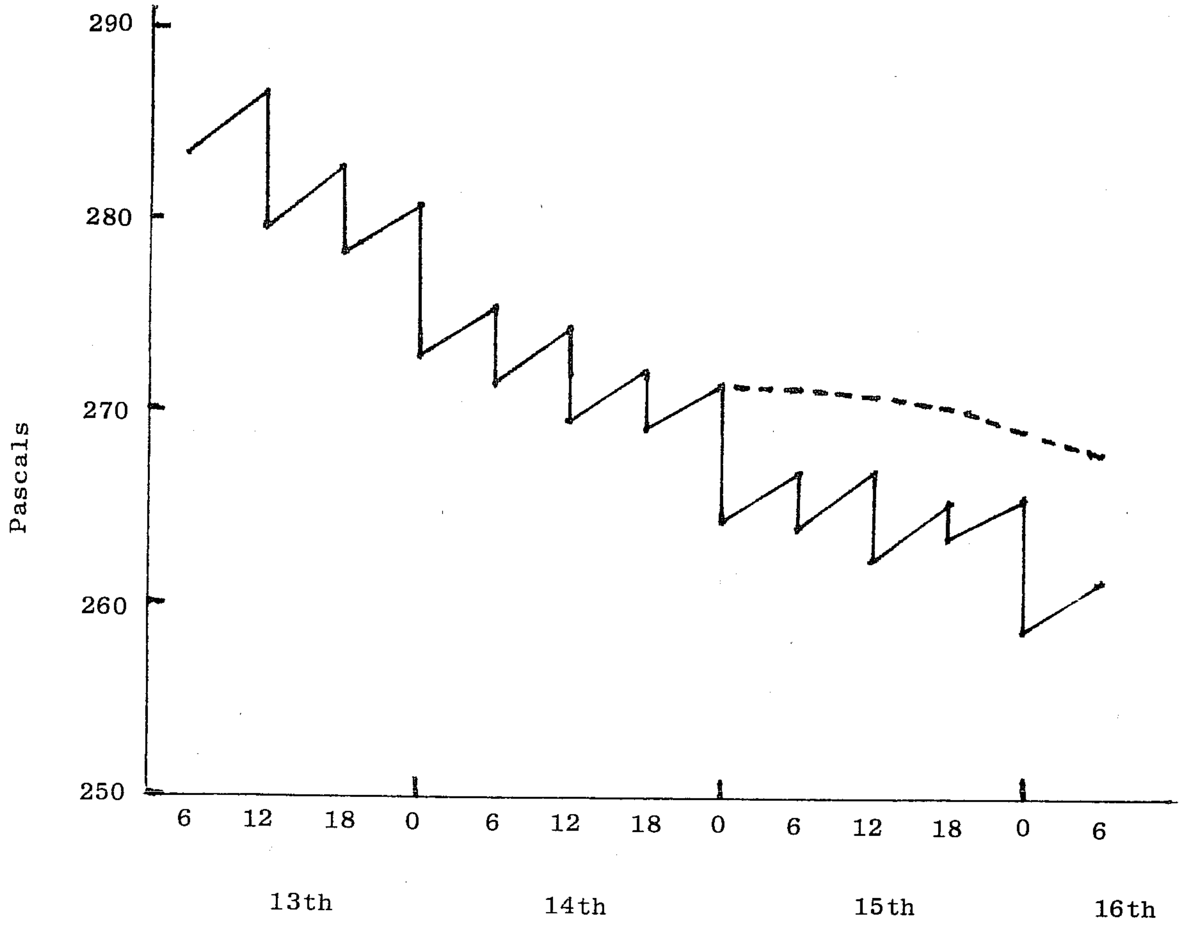

Fig. 7 Global mean vertically integrated specific humidity for experiments A (solid line) and B (dashed line).

	1000	850	700	500	400 300mb
Surface observations No.	2761	2865	2894	2894	2894
(O-P) expt. B	30.3	22.9	21.4	5.1	2.2
(O-P) expt. A	28.0	19.2	18.0	4.4	1.9
(O-A) expt. A	23.9	17.7	16.4	3.9	1.7
(P-T) assumed	17.3	11.9	5.9	3.4	1.1
(O-T) assumed	20.8	20.4	15.4	3.7	1.5
Soundings No.	663	648	562	591	528
(O-P) expt. B	30.4	24.6	17.4	5.9	2.5
(O-P) expt. A	24.7	20.0	10.7	5.0	1.8
(O-A) expt. A	12.3	14.3	3.7	4.7	1.5
(P-T) assumed	14.3	9.9	4.8	2.8	0.9
(O-T) assumed	3.1	2.4	2.0	0.6	0.3

Table 1. Global root mean square deviations of integrated specific humidity ΔQ between observed (O), predicted (P) and analysed (A) values, and the corresponding assumed prediction errors (P-T) and observations errors (O-T) for 00Z 16 January 1979.

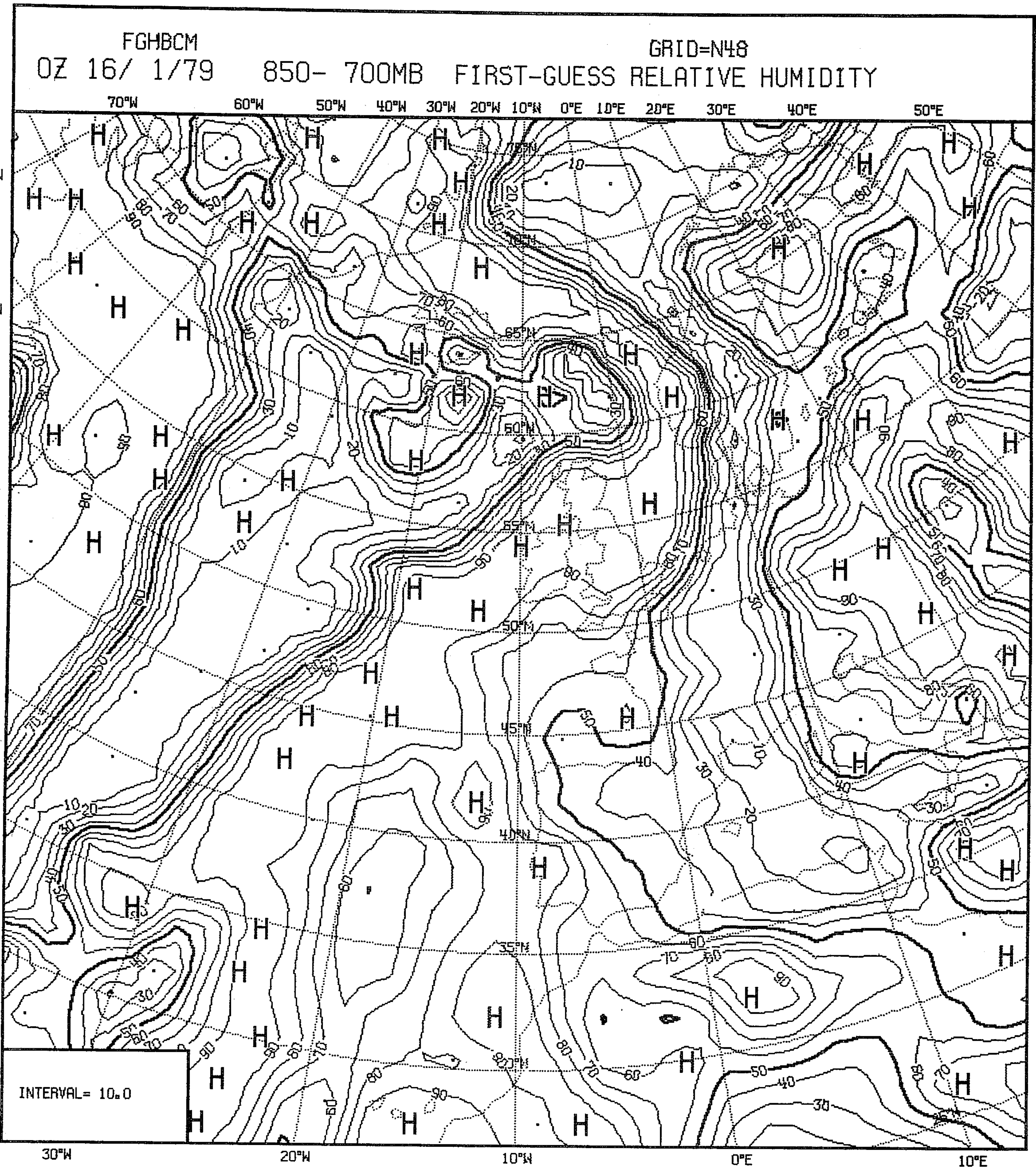

Fig. 8 ECMWF 850-700mb mean relative humidity forecast valid

00Z 16 January 1979

(experiment B)