KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

VERSLAGEN
V-334

Jan van Maanen

# Description of some properties of the humidity analysis scheme in use at the European Centre for Medium Range Weather Forecasts

De Bilt, 1979

Publikatienummer: K.N.M.I. V-334 (MO)

Koninklijk Nederlands Meteorologisch Instituut,
Meteorologisch Onderzoek,
Postbus 201,
3730 AE De Bilt,
Nederland.

U.D.C.: 551.509.38 :
551.571.7

# DESCRIPTION OF SOME PROPERTIES OF THE HUMIDITY ANALYSIS SCHEME IN USE AT THE EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

by Jan van Maanen

## 1. Introduction

In March 1979 several experiments were carried out to investigate the general properties and to tune the humidity analysis program of the European Centre for Medium Range Weather Forecasts. The problem was approached in two ways: the function that determines the weight of an observation was changed in several ways, and some forecasts were made with exactly the same forecasting model but with different initial states. One should bear in mind that the experiments were done with an analysis program and a model that have not yet reached their final state; however, we do not believe that this will decrease the usefulness of our results.

Several papers have dealt with the problem of humidity analysis. Atkins (1974) has described a successive correction method with two scans. The weight function she has used depends not only on the distance between the gridpoint being analysed and the observation, but also on the magnitude and direction of the gradient of the background humidity field. This serves the purpose of keeping details in the analysis, especially near fronts with a strong humidity gradient.

Kästner (1974) also has described an analysis made with the correction method. Not only direct humidity observations by means of radiosondes are used, but also the surface observations and calculations of various flow parameters like vertical velocity are taken into account, in order to provide an estimate of the upper-air humidity.

In section 2 a description is given of the humidity analysis program used in our experiments, in section 3 the result of changing the weight function is presented. The three forecasts made from a different initial state are presented in section 4, and section 5 consists of a summary and some conclusions.

## 2. Description of the humidity analysis scheme developed at ECMWF

Although the mass and wind analyses are made with optimum interpolation, the water vapour analysis is carried out with a correction method. The underlying philosophy is that the correction method takes less computer time as compared with an optimum interpolation analysis, and it is not proven whether the extra effort in doing water vapour analyses with the optimum interpolation method influences the forecast enough to make it worth to spend this extra amount of computer time.

The water vapour is analysed as follows. In the vertical five layers are analysed: 1000-850, 850-700, 700-500, 500-400 and 400-300 mbar. The total amount of water within a layer is expressed in pressure units. From the data of the multi-level observations (radiosondes) the water content is computed in the five analysis layers. The surface observations can also be used for providing an estimate of the upper-air humidity. The procedure for making these estimates is equal to the one used at the National Meteorological Centre (U.S.A.), except for the addition of a final step to interpolate these estimates to the five analysis layers in use at ECMWF.

The next step consists of interpolating and extrapolating the first guess of the model from $\sigma$-levels to the analysis levels. Then, the first guess is subtracted from the observations, and the observations are made dimensionless by dividing them by the (estimated) error of the first-guess humidity field. The correction method is applied to these dimensionless deviations. An analysed gridpoint value is determined by summing weighted deviations which are near the gridpoint being analysed. The weights are a function of distance and the (estimated) observation error, and are discussed in more detail in the next section. Also in the next section some examples are given of analyses with several weight functions.

The next step of the analysis is multiplying all analysis values by the first-guess error, to get dimensional quantities, and the final steps consist of adding the analysed deviations to the first

guess field and interpolating the values at p-levels to values at $\sigma$-levels.

The examples of analyses shown in the next section are the results on the p-levels. For the first-guess errors no exact information is available. Therefore these errors were specified from data of Oort and Rasmusson (1971). The first-guess errors are assumed to be a function of season, latitude and height in such a way that the first-guess error is roughly 35% of the climatological value of the humidity field.

## 3. The effect of the weight function

In this section some examples of analyses are shown, all valid for 1976 Feb. 6th, 00 GMT. The guess field was constructed by starting from the NMC-analysis 48 hours earlier, and by running the ECMWF gridpoint model and analysis program in cycles of 6 hours up to Feb. 6th. So it may be assumed that the first guess for the humidity, mass, and wind fields are of the same quality as can be expected in the operational mode.

### 3.a The weight functions used

In the humidity analysis several formulae were used for determining the correction to the first-guess value. They are denoted by $W_1$, $W_2$, $W_3$ and $W_4$ and are given by

$$w_{1i} = \begin{cases} 1 & \text{if } d_i < 300 \text{ km}, \\ 0 & \text{if } d_i > 300 \text{ km} \end{cases} \qquad W_1 = \frac{1}{\Sigma\, w_{1i}} \sum w_{1i}\, d_i$$

$$w_{2i} = \frac{a^2}{r_1^2 + a^2} \frac{b^2}{\varepsilon_i^2 + b^2} \quad , \qquad W_2 = \frac{\Sigma\, w_{2i}\, d_i}{(\Sigma\, w_{2i}) + w_{fg}}$$

$$w_{3i} = \frac{\rho_i}{\varepsilon_i^2 + n} \quad , \qquad W_3 = \sum w_{3i}\, d_i$$

$$w_{4i} = \frac{\dfrac{\rho_i}{1 + \varepsilon_i^2 - \rho_i^2}}{1 + \sum \dfrac{\rho_i}{1 + \varepsilon_i^2 - \rho^2}} \quad , \qquad W_4 = \sum w_{4i}\, d_i$$

where

$\rho_i$ is the correlation between the deviations from the guess field at the locations of the gridpoint and the observation;

$d_i$ is the normalized difference between the observation and the guess field;

$\varepsilon_i$ is the normalized observation error;

n is the number of observations that can influence the gridpoint;

a, b, $w_{fg}$ are the numerical constants, which may be level-dependent. For the 1000-850 mbar layer they have values of $(100\ \text{km})^2$, 1, and 1, respectively;

$w_{fg}$ is introduced to give the first guess a weight that can easily be controlled.

For $\rho_i$ as a function of $r_i$ was chosen

$$\rho_i = \exp(-r_i/r_o) \quad , \qquad r_o = (800\ \text{km})^{-1} \quad .$$

Use of the function $W_1$ simply means averaging all observations within a distance of 300 km. The function $w_{2i}$ consists of two factors: one takes into account the decreasing influence of a station with increasing distance, and is 1 if the distance is zero; the second factor decreases the influence of the station when the observation error is high. The function $W_2$ has the property that if there is only one observation, the weight given to the observation is always less than $\frac{1}{2}$. Function $W_3$ avoids this difficulty. If there is only one observation, the weight $w_{3i}$ is equal to the weight that would have been determined by the optimum interpolation method.

The last function, $W_4$, is derived in the following way. Suppose there are n observations near the gridpoint under consideration. The first step is to make from each of these observations n

independent estimates at the gridpoint, such that the errors of these estimates are independent of the guess-field value at the gridpoint.

Defining for every station i, i = 1, ..., n:

$a_i^o$ is the observed value minus the true value.

$a_{ki}^e$ is the estimated value at gridpoint k from observation i, minus true value at gridpoint k.

$a_k^p$ is the guess-field value at gridpoint k, minus the true value.

$\varepsilon_i^2$ is the mean square observation error, normalized with the prediction error: $\varepsilon_i^2 = \langle a_i^{o\,2} \rangle / \langle a_i^{p\,2} \rangle$ .

$a_i^p$ is the guess-field value at station i, minus the true value.

$\rho_i$ is the correlation of the prediction errors:
$\rho_i = \langle a_i^p a_k^p \rangle / [\langle a_k^{p\,2} \rangle \langle a_i^{p\,2} \rangle]^{\frac{1}{2}}$ .

All quantities are now replaced by their normalized equivalents, which implies $\langle a_k^{p\,2} \rangle = 1$, and $\rho_i = \langle a_i^p a_k^p \rangle$ .

$a_{ki}^e$ is determined according k

$$a_{ki}^e = a_k^p + V_i(a_i^o - a_i^p),$$

and $V_i$ such that

$$\langle a_{ki}^e \, a_k^p \rangle = 0 .$$

This leads to (assuming $\langle a_i^o \, a_k^p \rangle = 0$)

$$\langle a_{ki}^e \, a_k^p \rangle = \langle a_2^{p\,2} \rangle + V_i[\langle a_i^o \, a_k^p \rangle - \langle a_i^p \, a_k^p \rangle] = 0 =$$

$$= 1 + V_i(-\rho_i) ,$$

so

$$V_i \rho_i = 1.$$

The mean square error of this estimate is:

$$\langle a_k^{e\,2} \rangle = \langle a_k^{p\,2} \rangle + 2\, V_i [\langle a_k^p a_i^o \rangle - \langle a_k^p a_i^p \rangle]$$

$$+ V_i^2 [\langle a_i^{o\,2} \rangle + \langle a_i^{p\,2} \rangle - 2\langle a_i^o a_i^p \rangle] =$$

$$= \langle a_k^{p\,2} \rangle - 2\langle a_k^{p\,2} \rangle + \langle a_k^{p\,2} \rangle \frac{1}{\rho_i^2} [\varepsilon_i^2 + 1] =$$

$$= \langle a_k^{p\,2} \rangle \left[ \frac{1 + \varepsilon_i^2 - \rho_i^2}{\rho_i^2} \right] = \frac{1 + \varepsilon_i^2 - \rho_i^2}{\rho_i^2}$$

We now have $n + 1$ independent estimates for the value at the gridpoint, i.e. $a_{ki}^e$ , $i = 1, \ldots, n$, and the guess-field value $a_i^p$ . Combining these estimates to one, taking into account their respective errors, one gets for the weight to be given to the i-th observation:

$$w_i = \frac{\dfrac{\rho_i}{1 + \varepsilon_i^2 - \rho_i^2}}{1 + \sum_{j=1}^{n} \dfrac{\rho_j^2}{1 + \varepsilon_j^2 - \rho_j^2}} \quad .$$

To increase the stability of the analysis it is better to restrict the weights to $0 \leqslant w_i \leqslant 1$, therefore a change is made to obtain the definitive expression for $w_i$:

$$w_i = \frac{\dfrac{\rho_i}{1 + \varepsilon_i^2 - \rho_i^2}}{1 + \sum_{j=1}^{n} \dfrac{\rho_i}{1 + \varepsilon_j^2 - \rho_j^2}}$$

If there are two stations at such a distance from the grid-point that $\rho_1 = \rho_2 = \exp(-1) = 0.37$, the effect of the change is a decrease in each $w_i$ of 25-30%.

3.b Examples of analyses

The objective analysis of the 1000 mbar field of February 6th, 1976 is shown in fig. 3.1. Interesting features are the frontal trough stretched along 25°W, the blocking high over Scandinavia, and a depression near Spain.

In fig. 3.2 the guess field used in all experiments is shown. This guess field was constructed by starting from the NMC analysis of February 4th, 00 GMT, and running the ECMWF analysis/initialization/forecast cycle for two days. The standard ECMWF analysis program was used, and in the humidity analysis weight function $W_2$ was used. Only results for the lowest layer (1000-850 mbar) are displayed. In figs. 3.4 to 3.7 the analyses of relative humidity are shown, expressed as a percentage, using respectively weight functions $W_1$, $W_2$, $W_3$ and $W_4$. In figs. 3.8 to 3.11 are shown the absolute humidities using, in the same order, the same weight functions. The absolute humidity is the total amount of water vapour in the layer, expressed in pascal. When many data are available for the analysis, the effect of changing the weight function will be less pronounced, as the analysis system has no option but to follow the data closely. Therefore, only radiosonde data were used. The observations of the radiosondes at 850 mbar are shown in fig. 3.12.

The humidity analyses show some interesting features. The front stretching along 20°W is marked by a band of high humidity. Probably this band is analysed entirely from information already present in the guess field. The quality of the analysis over Europe is more difficult to judge. There clearly are differences between the analyses, but these are not very large, and it seems that the quality of a humidity analysis cannot be determined by merely looking at the maps. The analysis quality can probably be better determined by looking at independent information. There is the possibility of running a forecast model from several analyses, and decide to selct the analysis scheme that consistently gives the best forecasts, giving special attention to the humidity and precipitation forecasts. Another possibility is to omit a number of observations over Europe, to make an analysis, and to

compare this analysis with the omitted observations. The analysis scheme which has the best fit with the omitted observations is assumed to be the best one. These lines of thought are not pursued here any further, however.

## 4. The effect of the initial state on the forecast

The importance of the quality of the humidity analysis was studied by making three forecasts from a different initial state. The date of February 6th 1976 was selected because of its meteorological interest. The meteorological situation on that date is described in the previous section; during the following five days the high over Scandinavia moved to the south, while the depression over the Mediterranean disappeared. Air from the ocean invaded Western Europe. The forecast model used was the ECMWF spectral model. The initial states were (a) the NMC wind, mass and humidity field, (b) the ECMWF analysis for the same variables, (c) the ECMWF analysis in which the humidity was changed to 60% at all levels. The forecast model was run for five days.

In fig. 4.1 the initial state for experiment (b) is shown. The three forecasts tend to have more zonal flow in comparison with the initial situation, and display a breakdown of the blocking high over Scandinavia. Note that the map projection used for the forecast maps is not conformal, so the zonality of the flow and isolines is less strong as it appears to be.
In figs. 4.2, 4.3 and 4.4 three 1000 mbar forecast charts are presented. In experiment (a) a trough has developed over Scandinavia, which is much less pronounced in experiment (b) and can hardly be distinguished in experiment (c). In experiment (c) - initial humidity 60% - a trough over Britain has developed, which is almost absent in the other forecasts. The 500 mbar developments (not shown) are in good agreement with the surface developments.

The humidity charts of figs. 4.5, 4.6 and 4.7 show more mutual differences than the pressure charts. Comparing the humidity forecasts, we note that the area of low humidity over the North Sea, present in experiment (a), is absent in experiment (b), while the

distribution of humidity in this region is again very different in experiment (c). In all forecasts the humidity is low over Italy and Greece and rather high (80 or 90%) north of 60° North. The subjective analysis of surface pressure of February 11th is shown in fig. 4.8.

The three forecasts seem to be roughly of the same good quality. The effect of changing the humidity to 60% in the initial state is of the same size as the effect of the difference between the NMC and the ECMWF analysis. If this conclusion holds in general and not only for the case of February 6th, this implies that the specification of the initial humidity field is important for the quality of a 5-day forecast. Of course the humidity field will adjust to the other variables of the model, so after some time the humidity field will not contain much independent information. Nevertheless, it seems that in the first stages of the forecast the mutual adjustment can influence the mass and wind fields in such a way that differences are observed in the 5-day forecast.

## 5. Summary and conlusions

The properties of the objective analysis of the humidity field and its relevance for a numerical forecast were studied in two independent ways. The first was to change the weight function used in the correction method and to investigate the resulting analysis; the second consisted of studying the result of different initial fields on the forecast.

Four weight functions were derived and used in the correction method, and with them humidity analyses were constructed. The absolute quality of the various analyses turned out to be difficult to estimate because of the fundamental uncertainty in the true atmospheric state. Probably the best test will be a comparison with independent data obtained by dividing the observations into two sets, using one for the analysis and the other as a reference for the analysis. From a theoretical point of view, function $W_4$ of section 3.b is preferable because of its superior mathematical properties. (It is continuous, it is equal to one in the limiting

case of an errorless observation coinciding with the gridpoint, and it parameterizes the correlations between nearby stations).

In the second part of this study we made 5-day forecasts with the ECMWF spectral model, using different initial states, especially with regard to the humidity field. The resulting forecasts were found to differ substantially in the humidity fields; also differences in the pressure field could easily be observed on large-scale maps. From this we conclude that a good specification of the initial humidity is essential for numerical forecasting, especially in connection with predictions of rainfall amounts and cloudiness. Further investigation in this direction seems therefore advisable.

Acknowledgments. I am greatly indebted to Mr. A. Lorenc and Dr. S. Tibaldi for introducing me into the ECMWF analysis system and for stimulating discussions. One of the weighting functions of section 3 was based on an idea of A. Lorenc. Also I thank Dr. N. Jarraud and Dr. M. Capaldo for running the forecast model.

I thank the European Centre for Medium Range Weather Forecasts for their hospitality.

This research was sponsored by the European Centre for Medium Range Weather Forecasts and the Royal Netherlands Meteorological Institute.

-o-o-o-

References

Oort, A.H. and E.M. Rasmusson, Atmospheric Circulation Statistics, NOAA Professional Papers, 5, 1971.

Atkins, M.J., The objective analysis of relative humidity, Tellus, 26, 1974, pp. 663-671.

Kästner, A., Ein Verfahren zur Numerischen Analyse der relativen Feuchte, Arch. Met. Geoph. Biokl., Ser. A, 23, 1974, pp. 137-148.

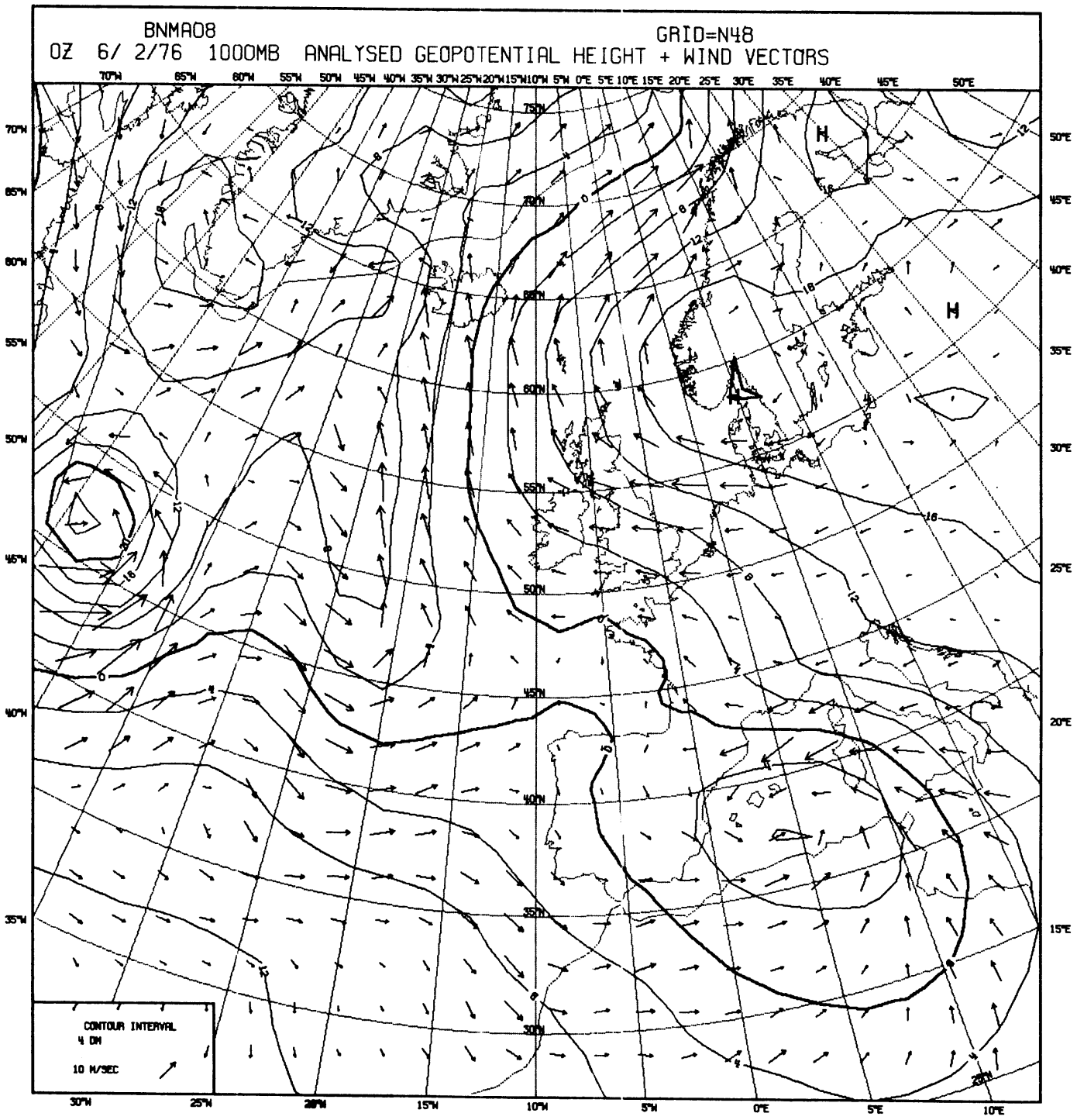


Fig. 3.1 Objective analysis of 1000 mbar height and wind of 1976 Feb. 6th, 00 gmt.

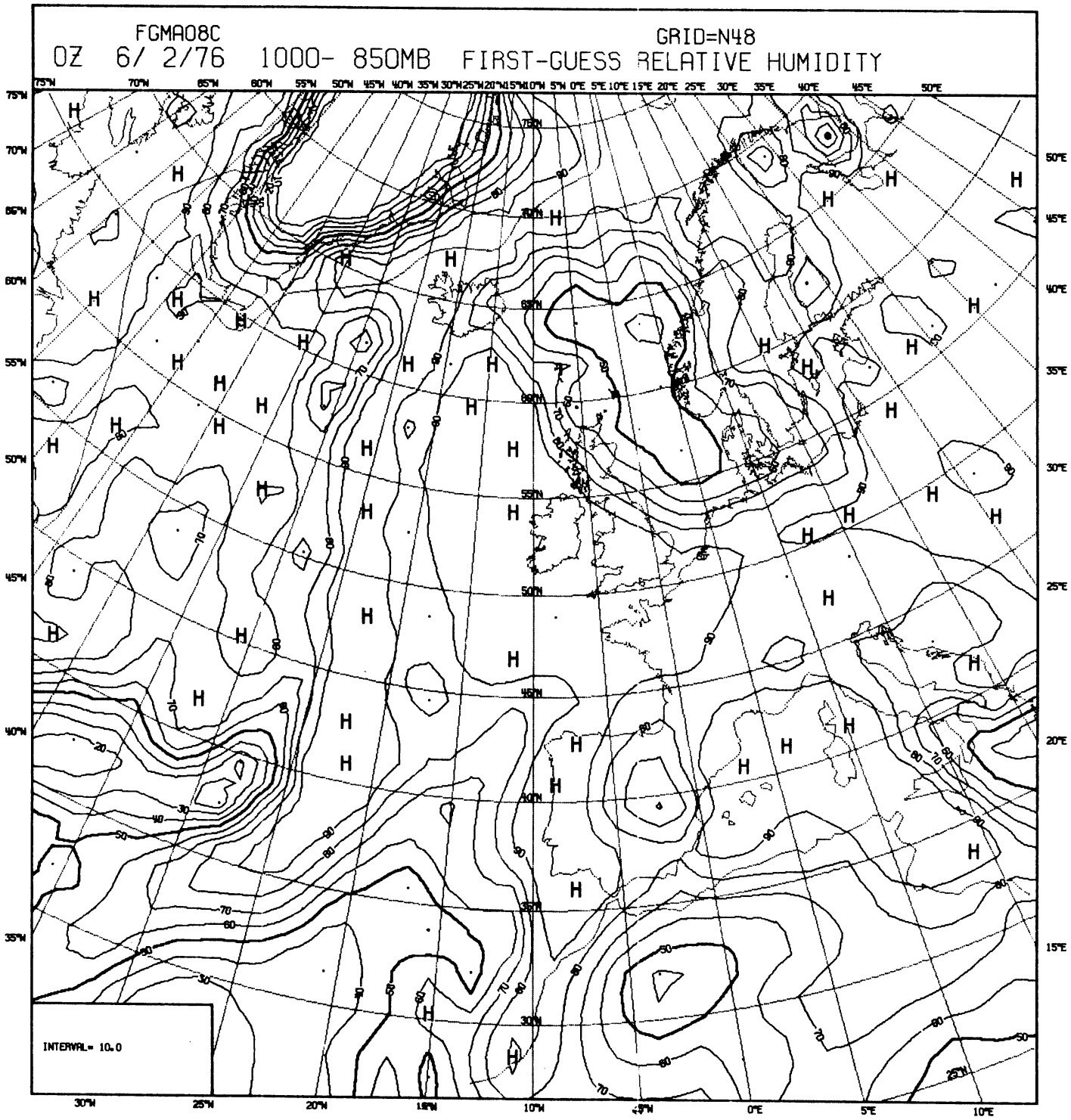


Fig. 3.2 Relative humidity guess field for analysis of 1976 Feb. 6th, 00 gmt. The relative humidity is expressed as a percentage.

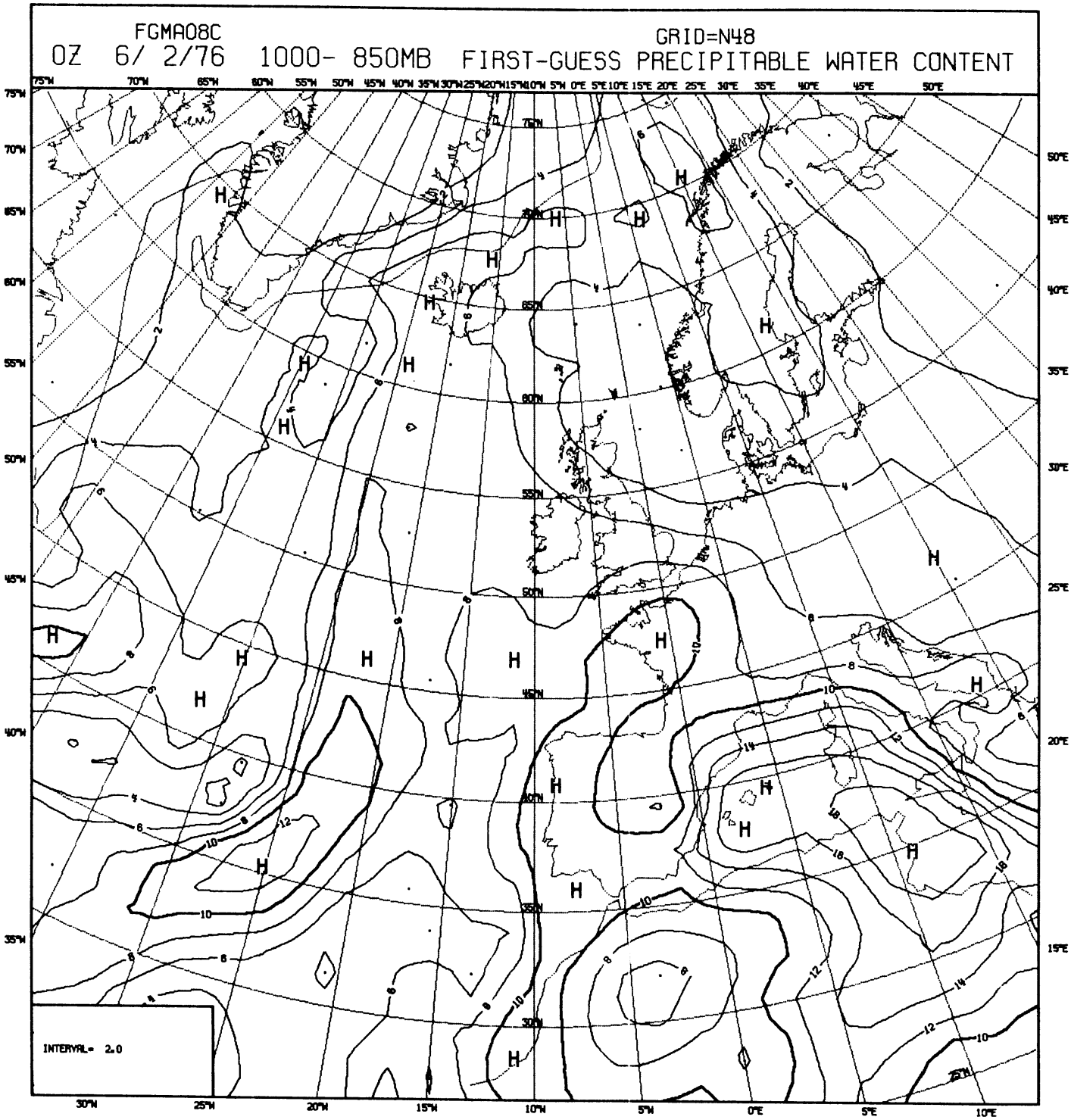


Fig. 3.3 Absolute humidity guess field for analysis of 1976 Feb. 6th, 00 gmt.
The absolute humidity of the 1000-850 mbar layer is expressed in pascal.

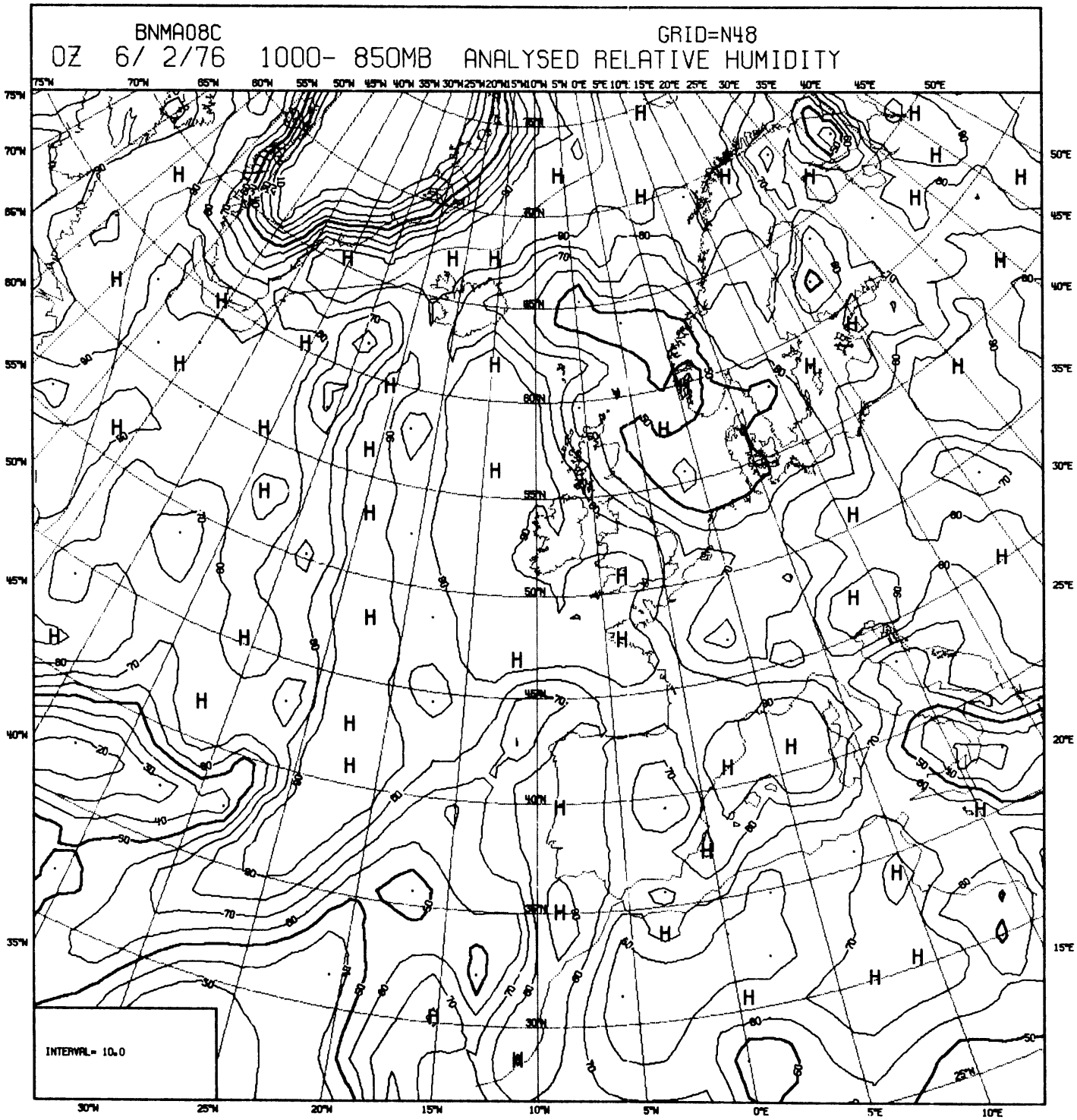


Fig. 3.4 Analysis of relative humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_1$.

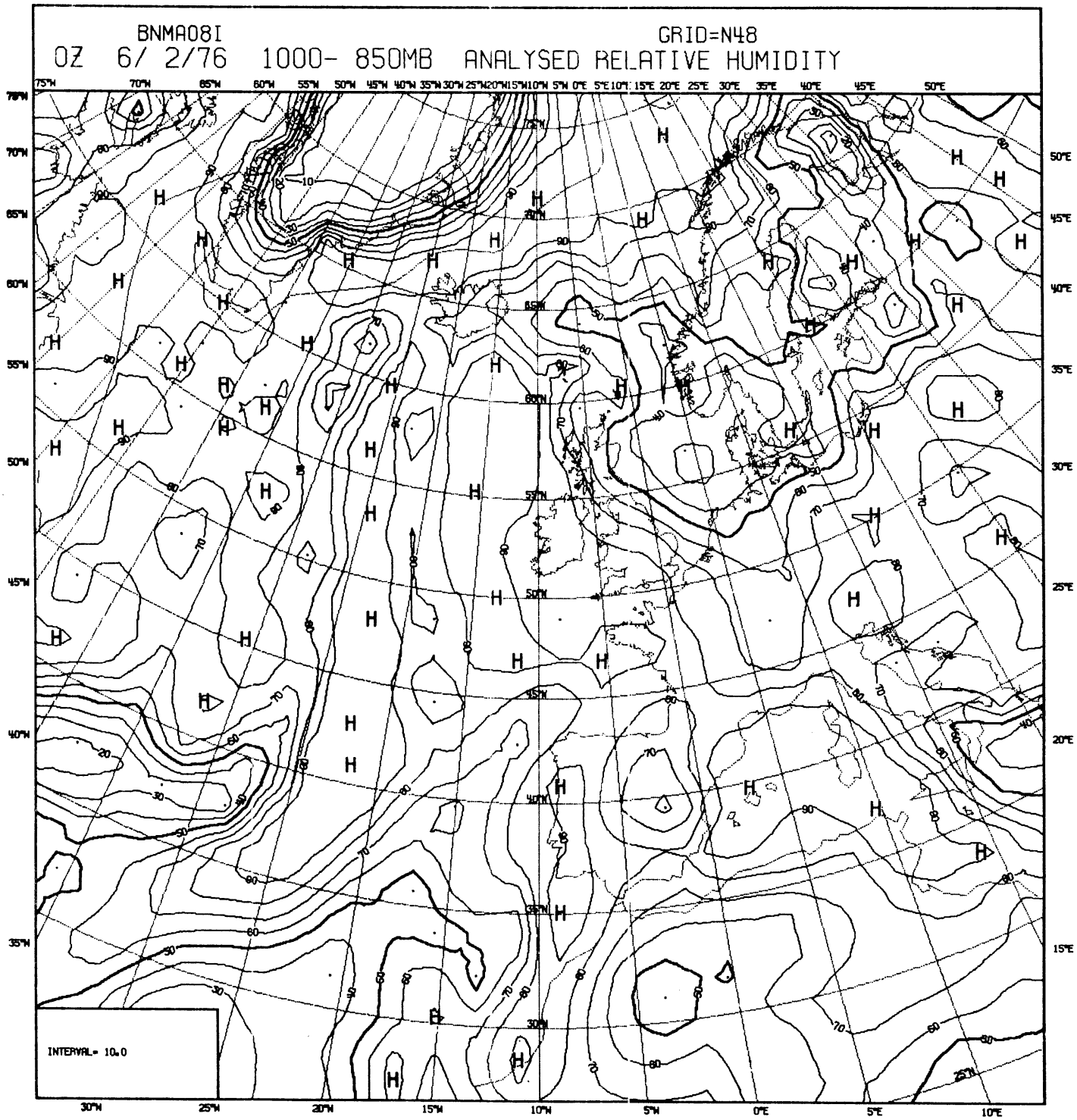


Fig. 3.5 Analysis of relative humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_2$.

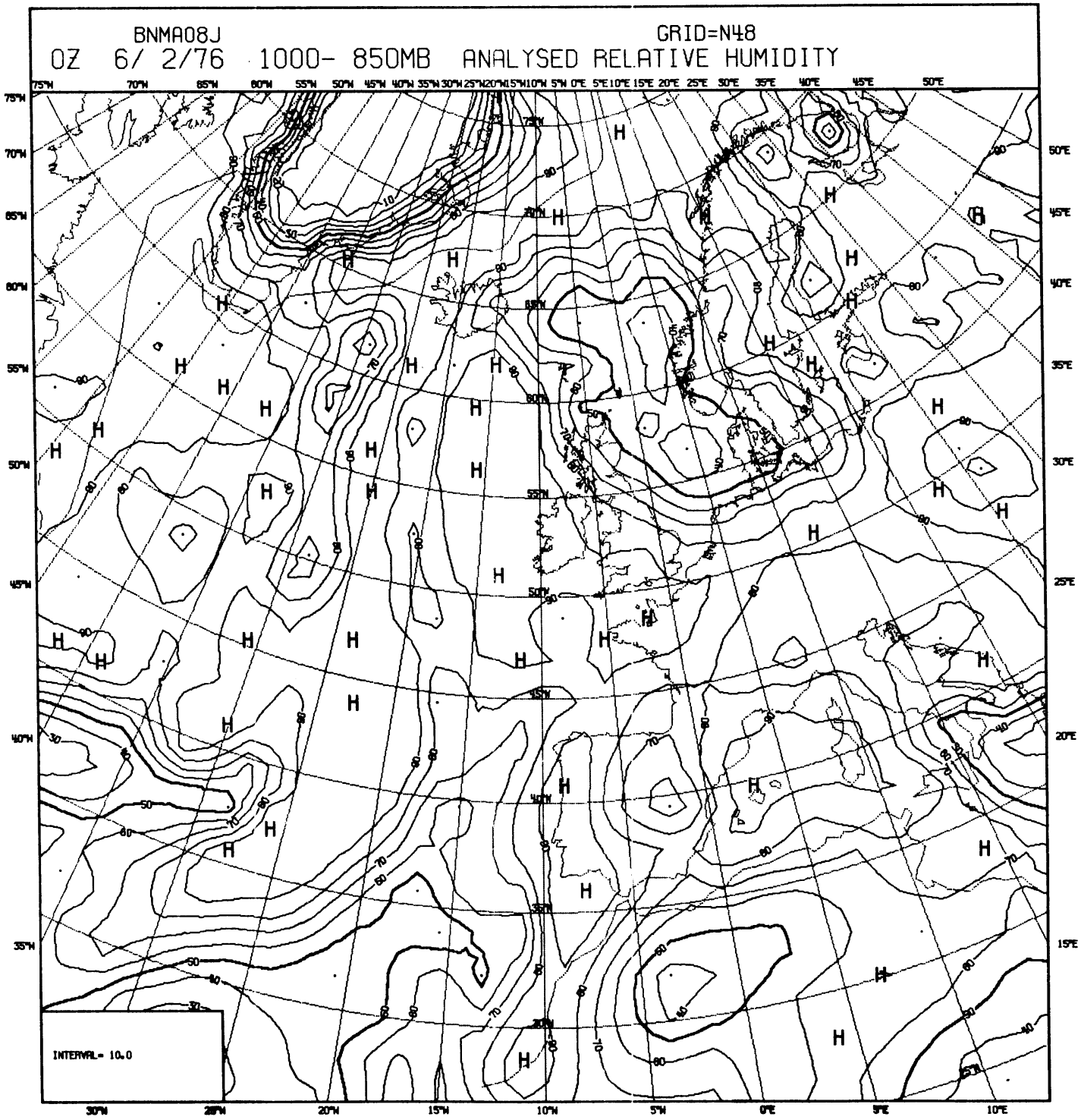


Fig. 3.6 Analysis of relative humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_3$.

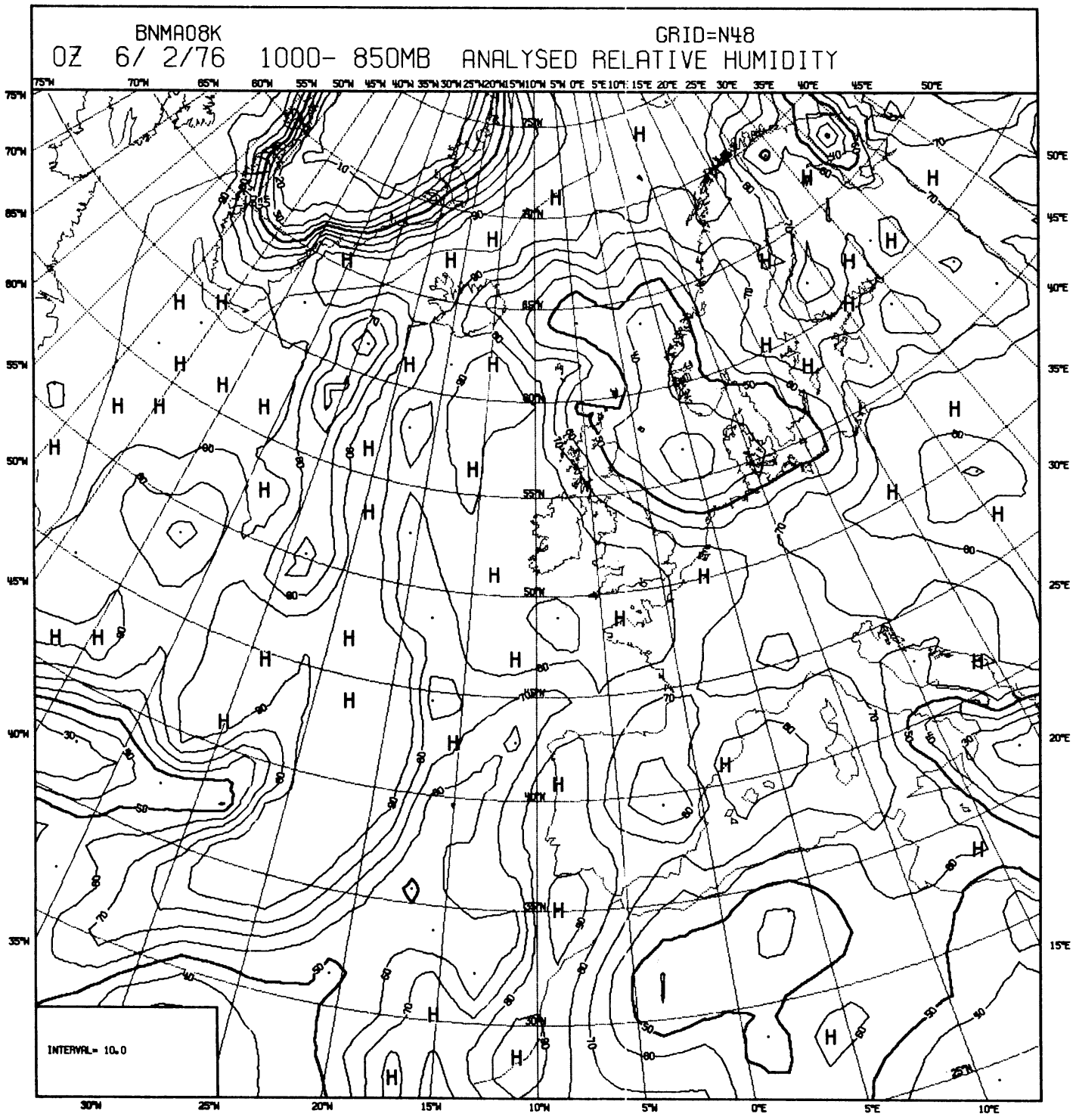


Fig. 3.7 Analysis of relative humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_4$.

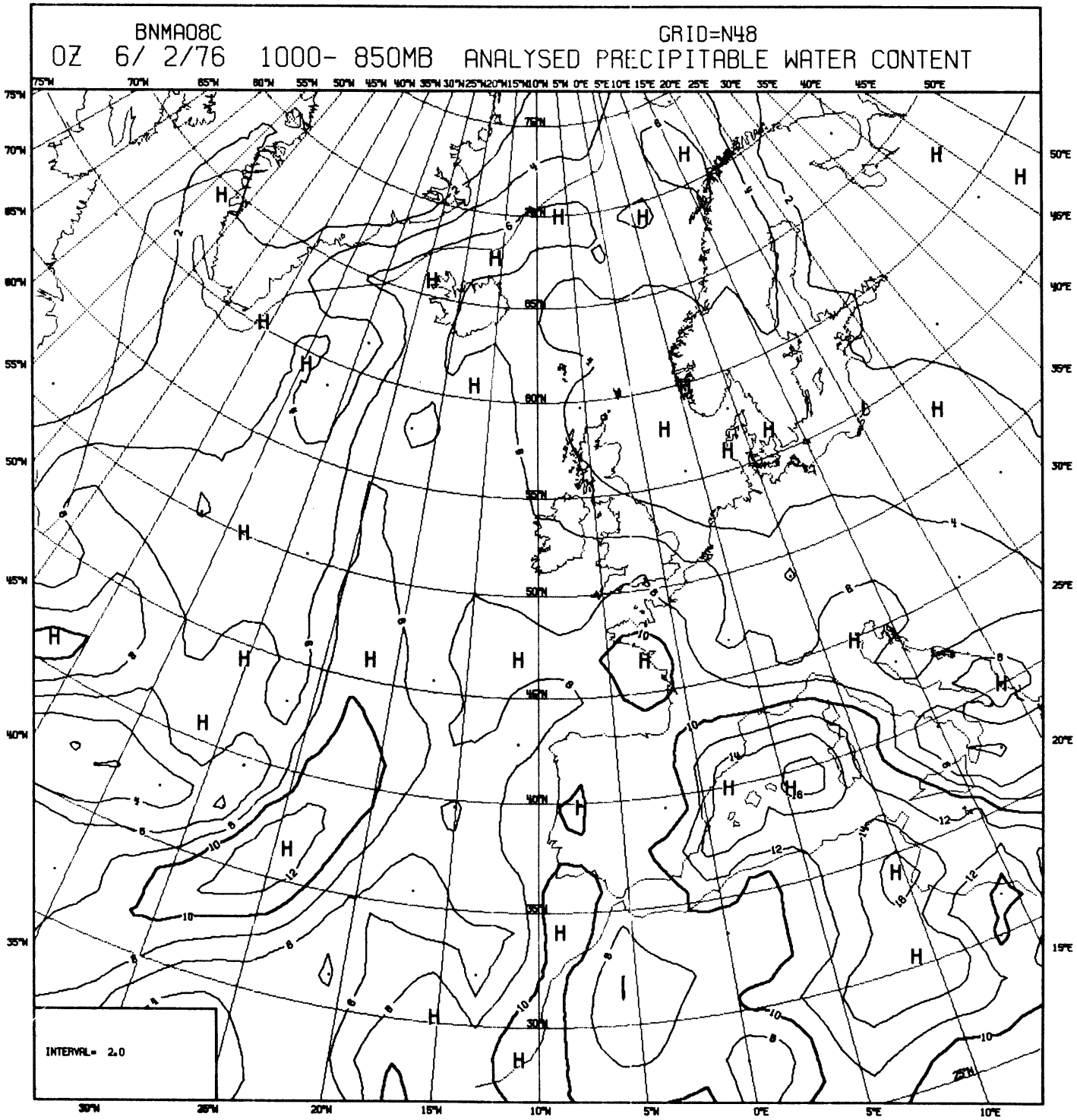


Fig. 3.8 Analysis of absolute humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_1$.

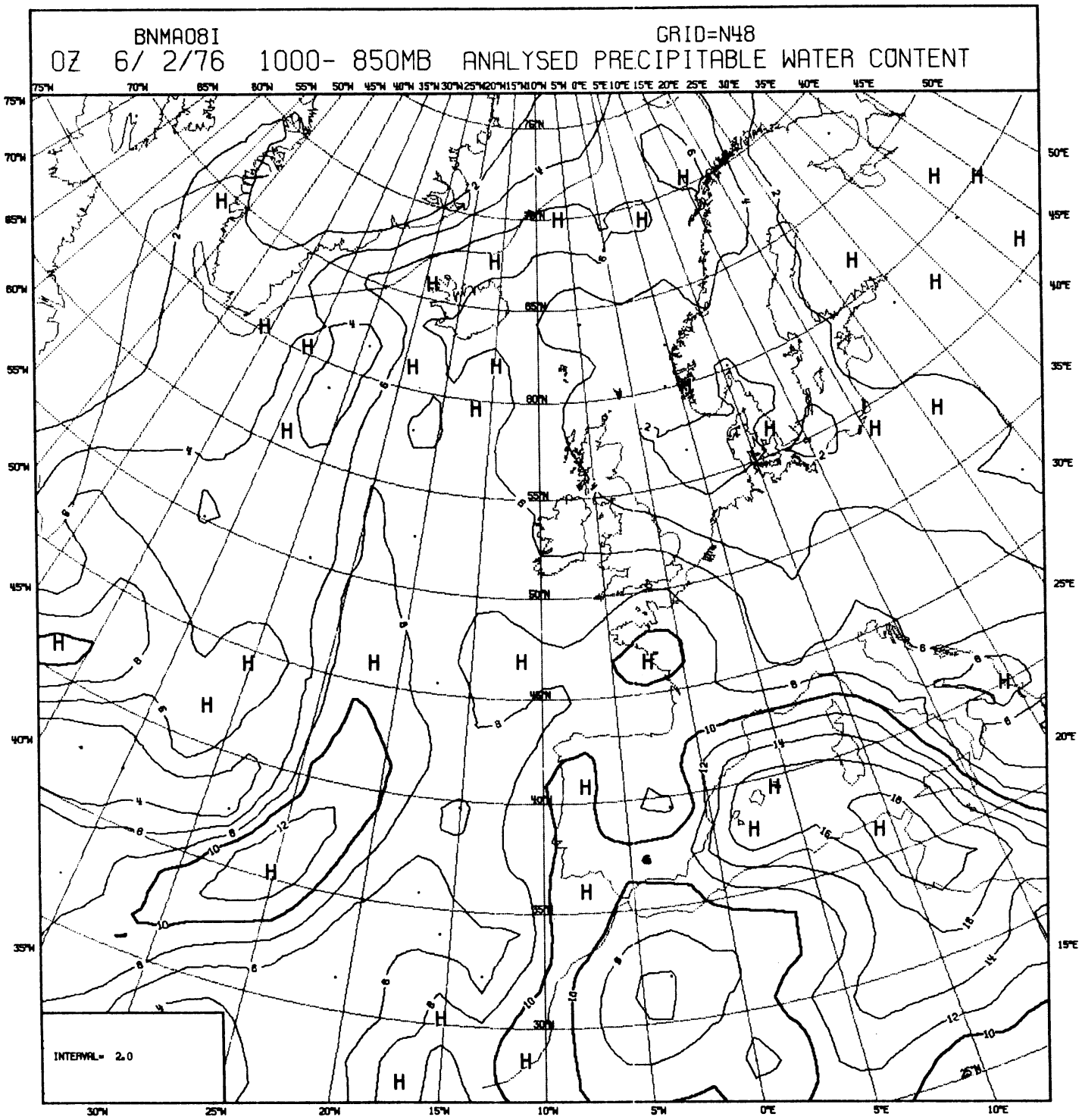


Fig. 3.9 Analysis of absolute humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_2$.

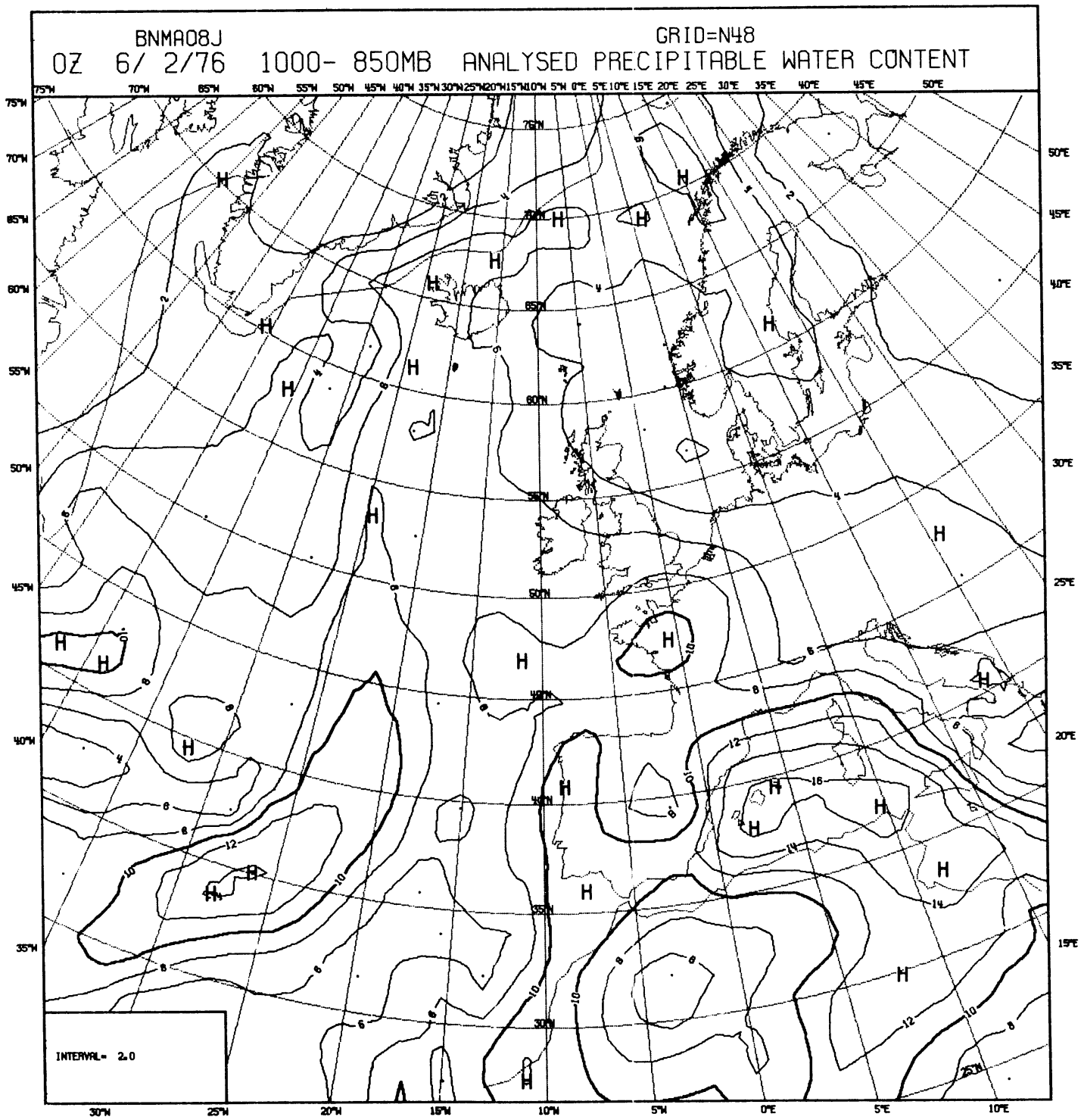


Fig. 3.10 Analysis of absolute humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_3$.

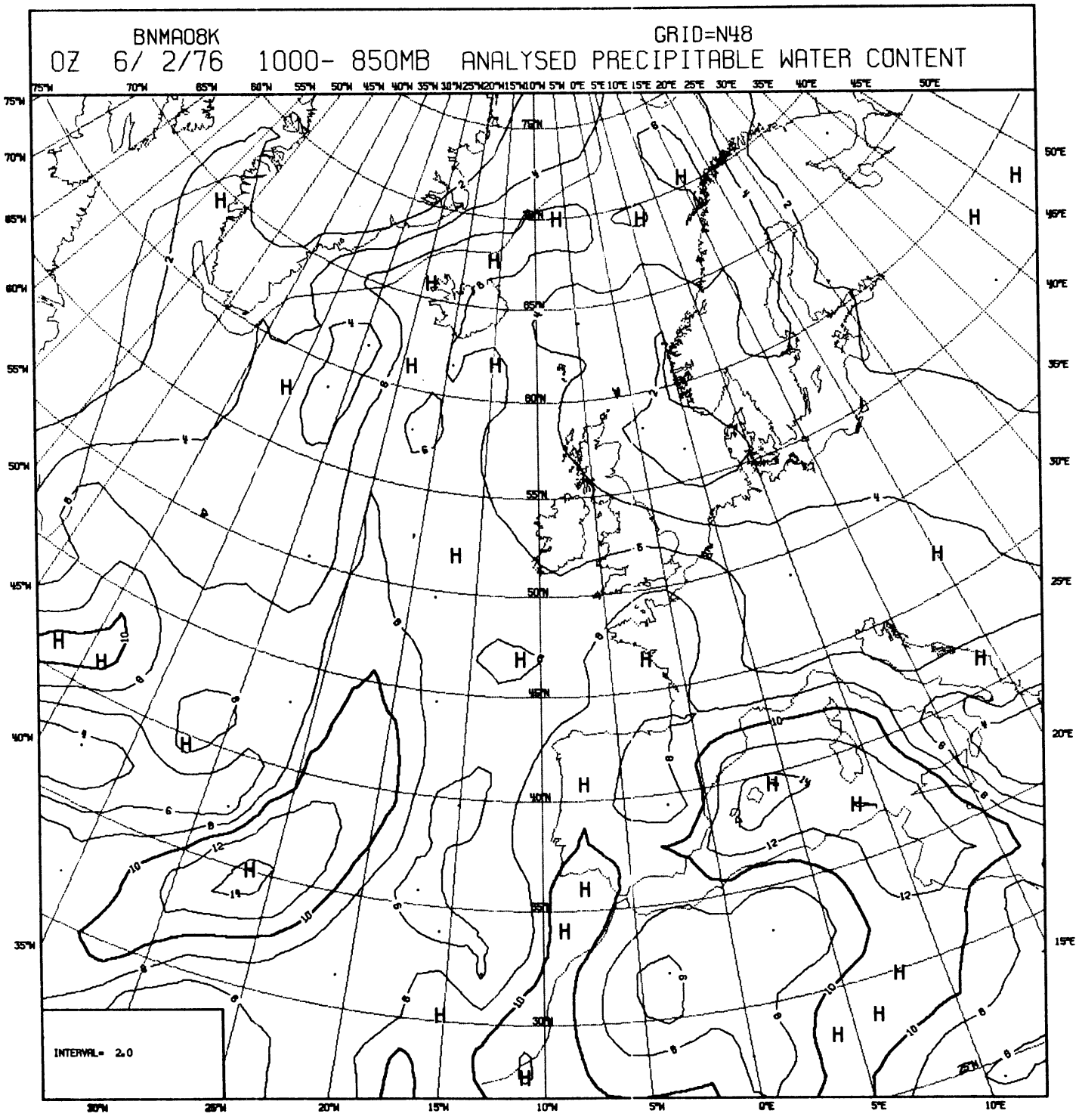


Fig. 3.11 Analysis of absolute humidity of 1976 Feb. 6th, 00 gmt, using weight function $W_4$.

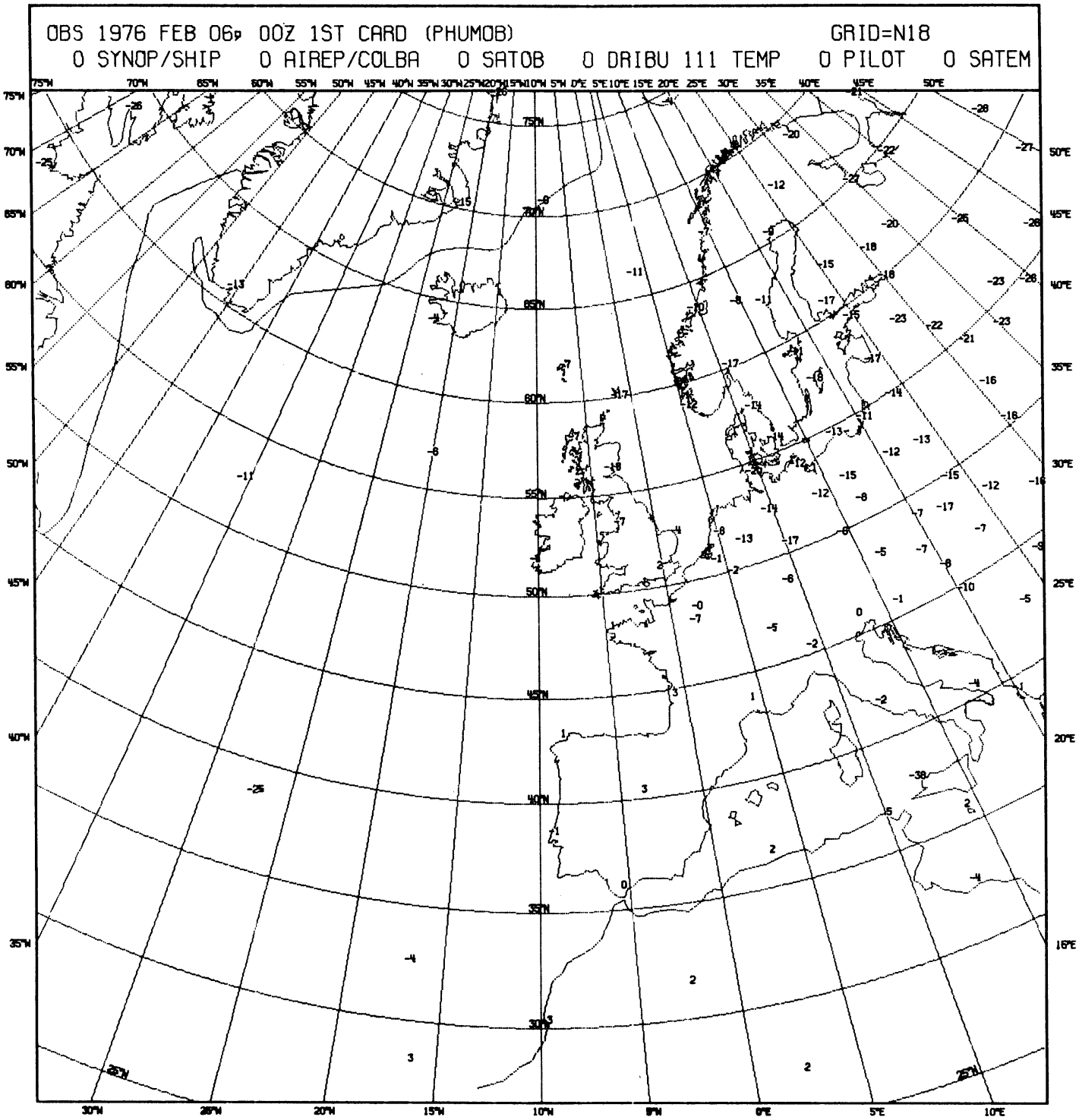


Fig. 3.12 Observations of 850 mbar dewpoint temperature of 1976 Feb. 6th, 00 gmt.

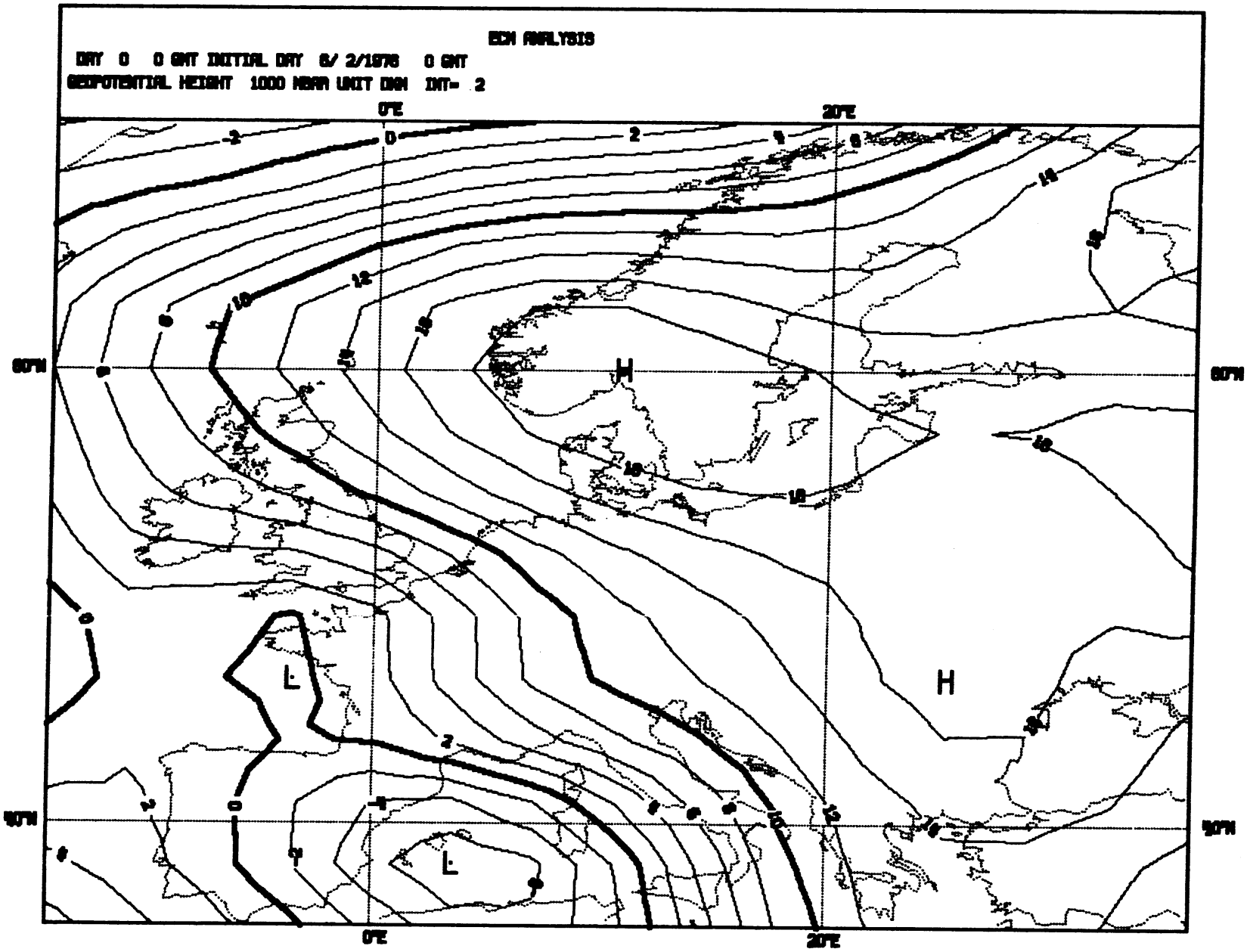


Fig. 4.1 The ECMWF-analysis of the 1000 mbar level.

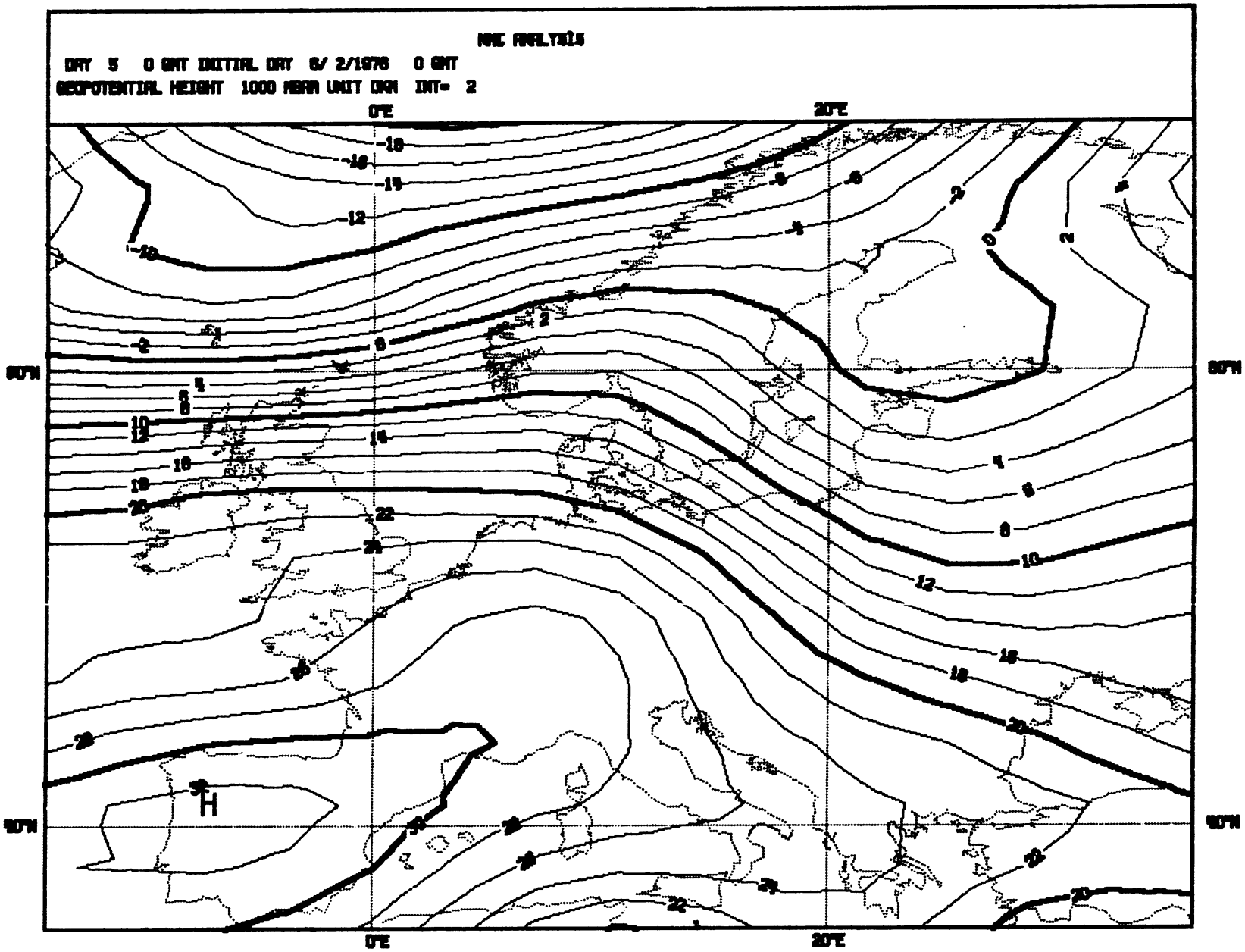


Fig. 4.2 5-day forecast of the 1000 mbar level starting from the NMC-analysis.

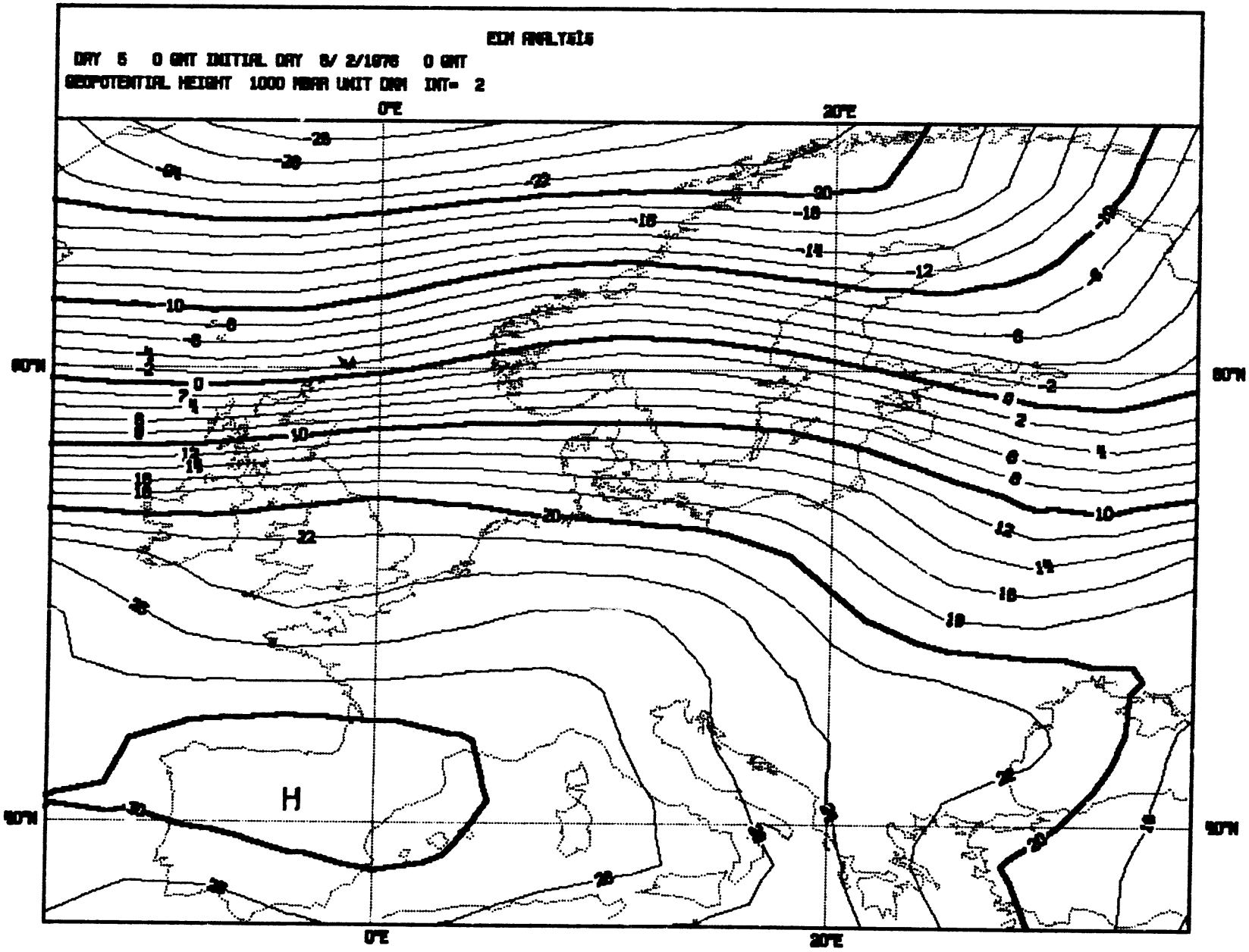


Fig. 4.3 5-day forecast of the 1000 mbar level starting from the ECMWF-analysis.

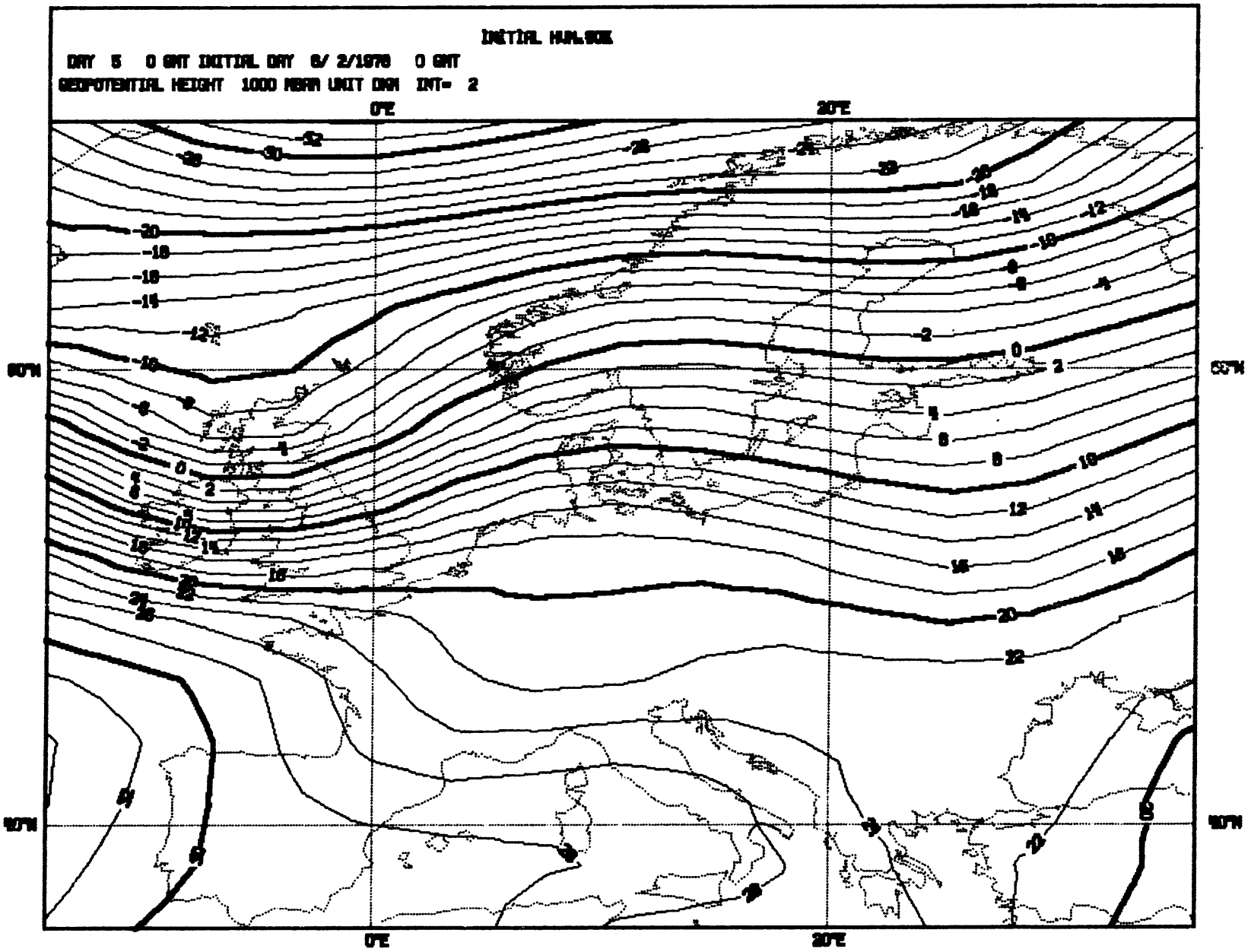


Fig. 4.4 5-day forecast of the 1000 mbar level starting from the ECMWF-analysis with initial humidity changed to 60%.

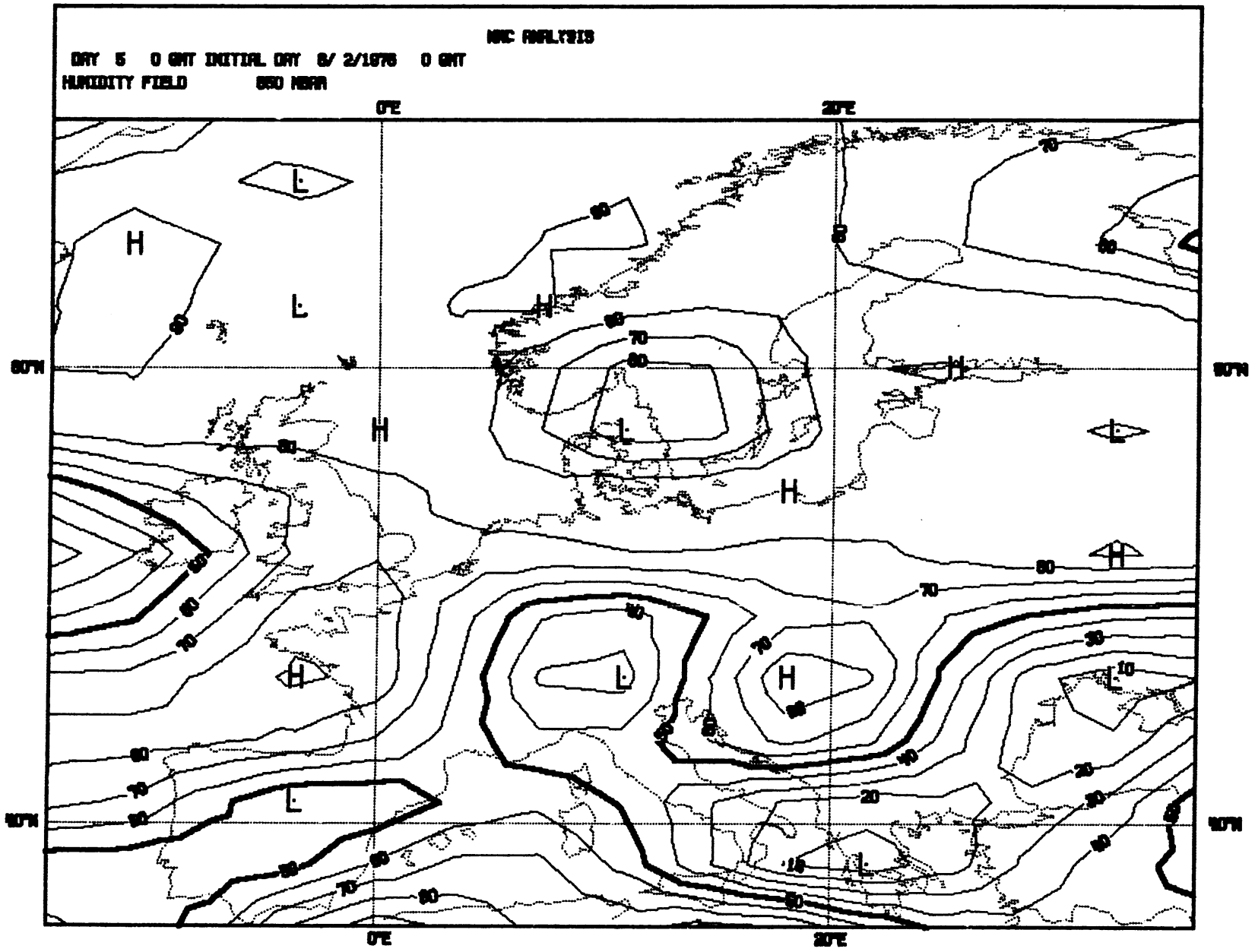


Fig. 4.5 5-day forecast of 850 mbar relative humidity starting from the NMC-analysis.

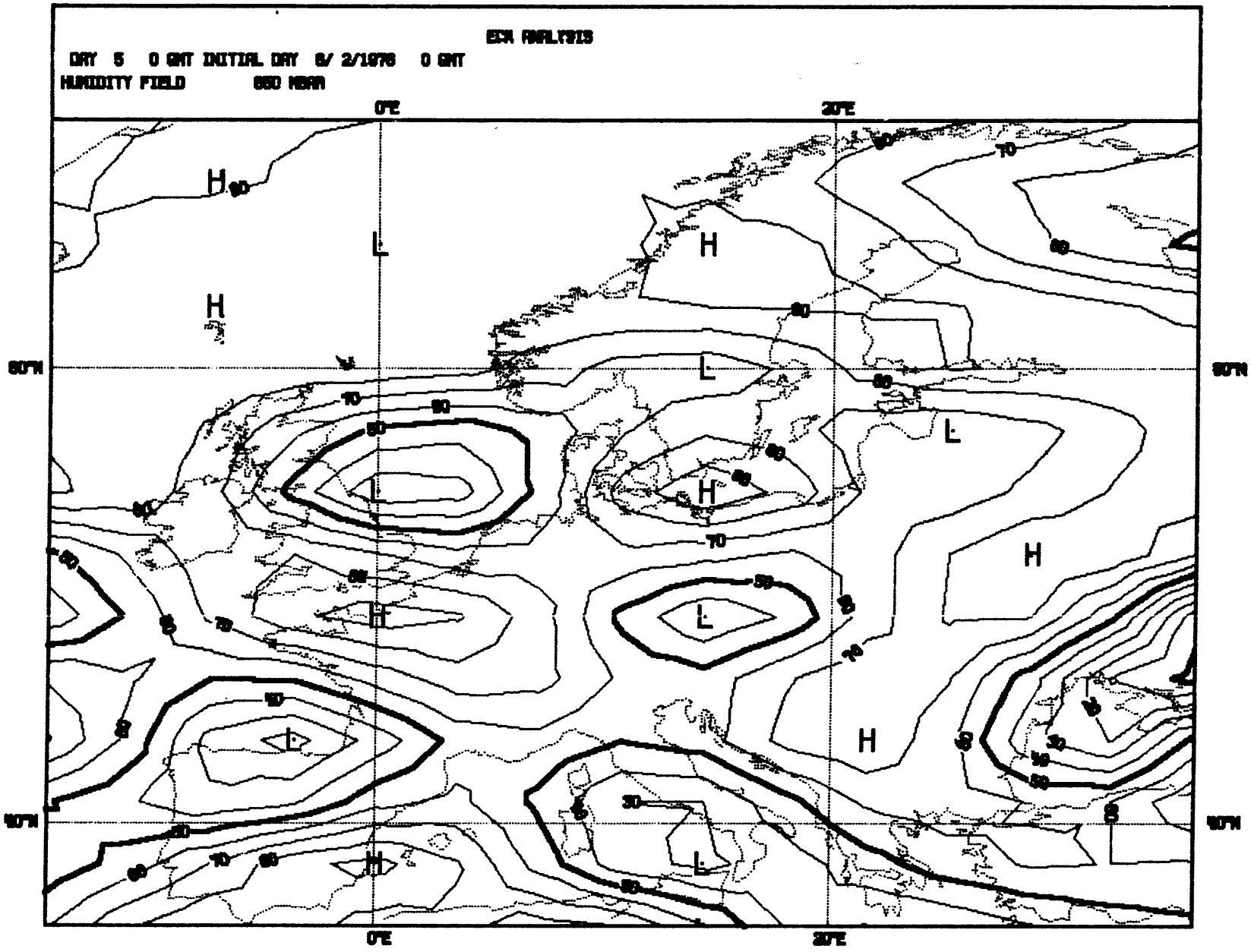


Fig. 4.6 5-day forecast of 850 mbar relative humidity starting from the ECMWF-analysis.

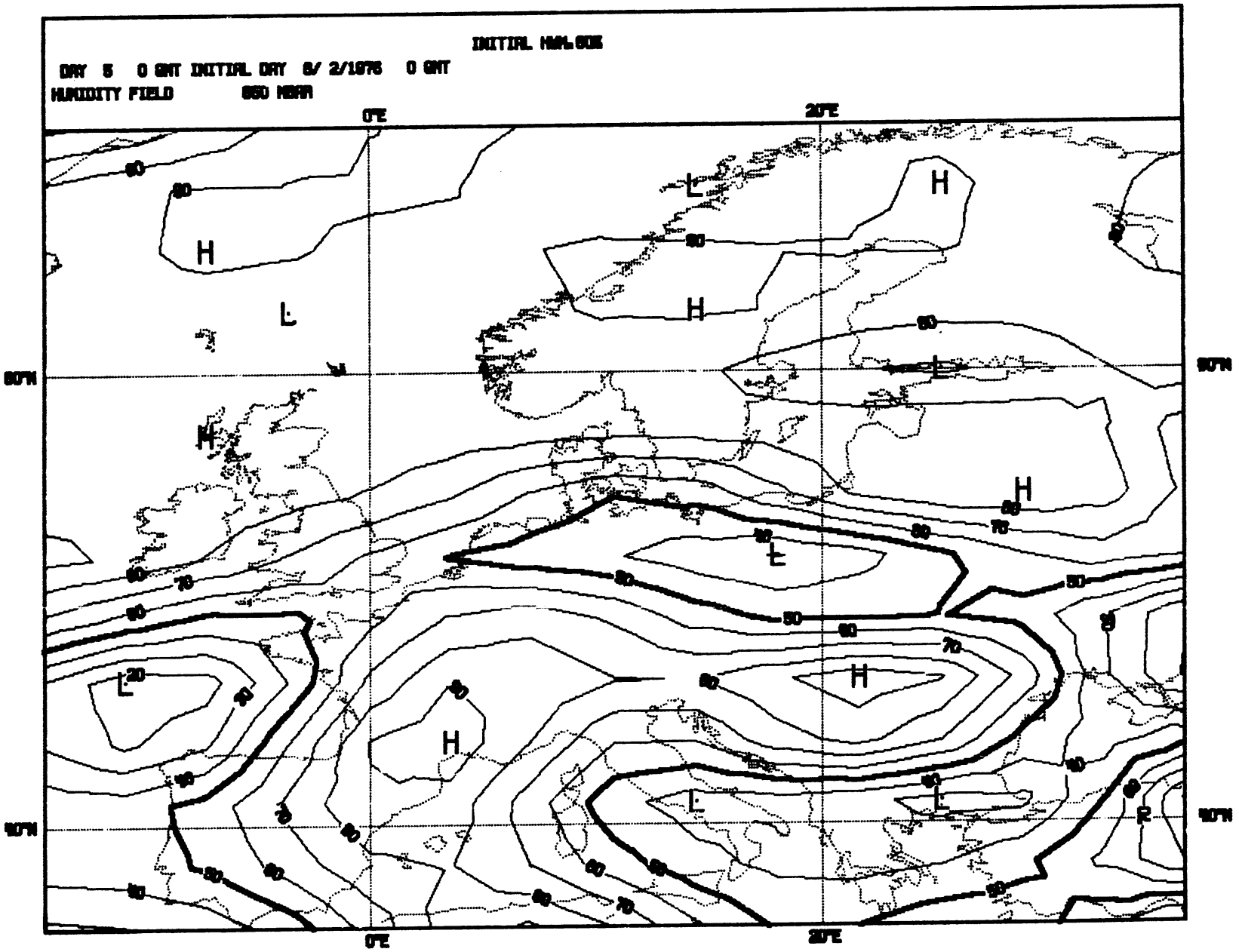


Fig. 4.7 5-day forecast of 850 mbar relative humidity starting from the ECMWF-analysis with initial humidity changed to 60%.

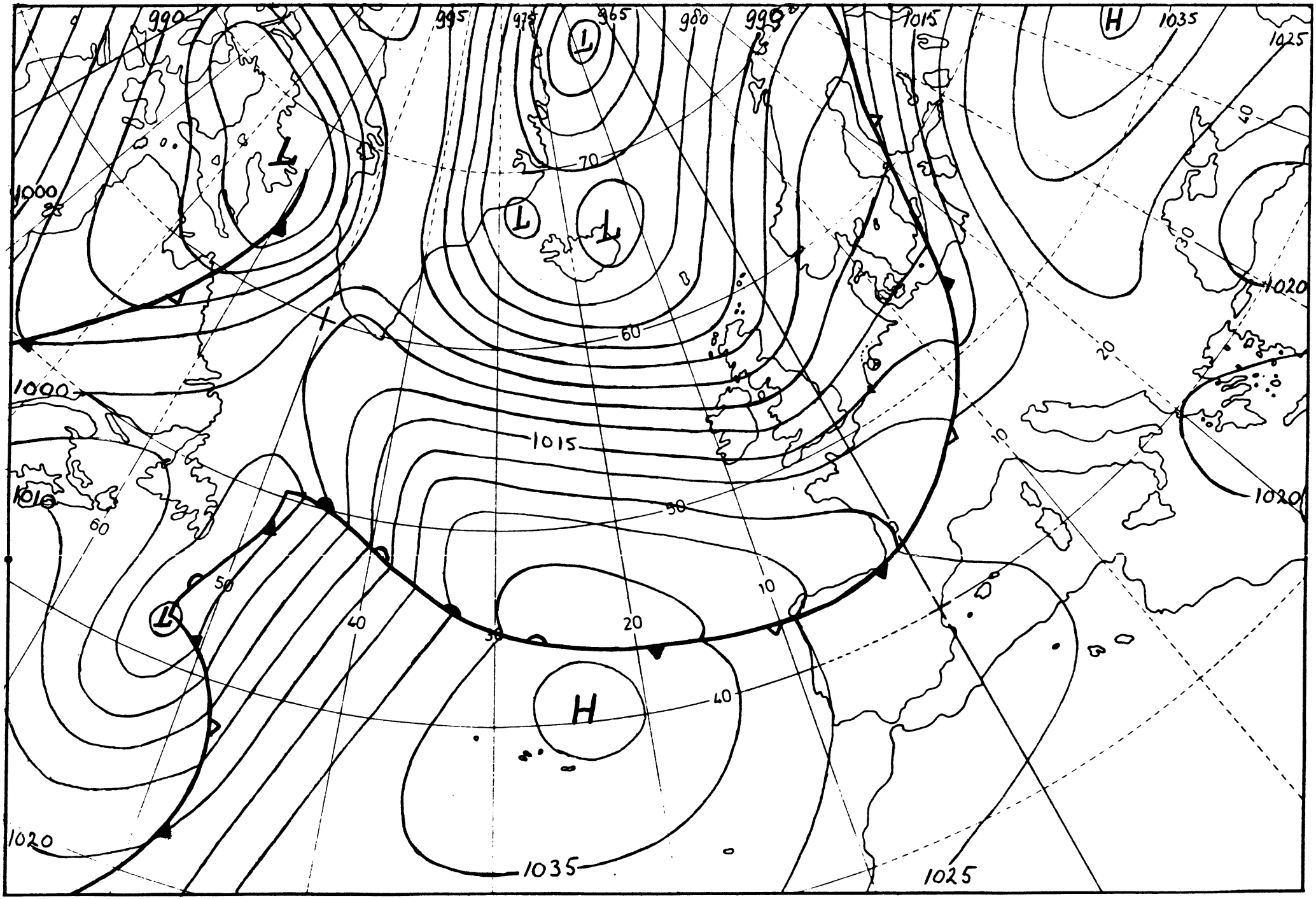


Fig. 4.8 Subjective analysis of sea-level pressure of 1976 Feb. 11th, 00 gmt.