ECMWF 3D Variational Data Assimilation of Conventional Observations

D. Vasiljević, C. Cardinali, P. Undén
ECMWF, Reading, U.K.

ABSTRACT

We describe the main features of the ECMWF 3D variational assimilation of conventional observations. Preliminary results are discussed in terms of comparisons with the current ECMWF OI analysis. We use near surface observations in a manner consistent with the model's surface layer parametrization.

1. Introduction

A variational analysis consists in minimizing a (scalar) cost function $J(x)$ of a (vector) control variable x. The minimization proceeds through an iterative descent algorithm until an appropriate convergence is reached. The cost function measures the distance of a model solution to the observations J_O and to a background field J_b:

$$J = J_o + J_b \tag{1}$$

The technique relies on the notion of an adjoint operator which provides a convenient way of computing the gradient of the cost function with respect to the control variable.

The variational analysis, reproducing theoretically the present OI (optimum interpolation) analysis, can formally be written in terms of J_O and J_b as (*Lorenc*, 1986):

$$J_o = (y - Hx)^t O^{-1} (y - Hx) \tag{2}$$

$$J_b = (x - x_b)^t B^{-1} (x - x_b) \tag{3}$$

where x_b is the background, B is the covariance matrix of background errors, y is the vector of the observed data, O is the covariance matrix of observation errors which also includes the representativeness errors (*Lorenc*, 1986), and H is the ensemble of operators (linear or/and non-linear) transforming the control variable x into the equivalent of each observed quantity.

In the OI analysis the operator H takes a very simple form (basically linear) and as a consequence cannot extract from the observations any quantity non-linearly (even weakly) related to the model variables. This is a strong limitation not only in the case of currently available space based observations (TOVS and scatterometer) but also in the case of conventional (mostly ground based) observations. For example, the most widely available observations of 10m wind, 2m temperature and relative humidity (available from SYNOP/DRIBU reports) are either not used or only used crudely in OI systems.

This paper describes the computation of J_O and its gradient for the conventional observations in the ECMWF 3D variational assimilation (*Pailleux et al.*, 1991). Special attention will be given to the use of near surface observations. Preliminary results of the ECMWF 3D variational mass and wind analysis will be shown. The results will be presented as comparisons with the ECMWF OI analysis.

Heckley et al. (1992) discuss the background term J_b of the cost function used for our experiments, *Andersson et al.* (1992) discuss the use of radiances and *Thépaut et al.* (1992) discuss the use of scatterometer.

2. EVALUATION OF THE COST FUNCTION AND ITS GRADIENT

2.1 Conventional observations

By conventional observations we mean ground based observations whose observation errors are not correlated in the horizontal. In addition, the space based satellite wind measurements (SATOB) are included in this group. The following is the list of the observing systems:

- SYNOP: surface pressure (p_S), 10m wind components (u_{10m}, v_{10m}), 2m temperature (T_{2m}) and 2m relative humidity (RH_{2m})
- AIREP: upper air wind components (u,v) and temperature (T)
- SATOB: upper air wind components (u,v)
- DRIBU: surface pressure (p_S), 10m wind components (u_{10m}, v_{10m}) and 2m temperature (T_{2m})
- TEMP: 10m wind components (u_{10m}, v_{10m}) 2m temperature (T_{2m}), 2m relative humidity (RH_{2m}) geopotential heights (Z), upper air wind components (u,v), upper air relative humidity (RH) and upper air temperature (T)
- PILOT: upper air wind components (u,v)

Most of the observed quantities listed above are used by the current ECMWF OI system with the exception of near surface observations. Specifically in the OI, 10m wind and 2m relative humidity observations are used only over sea, whereas 2m temperature is not used at all (10m wind is also used in tropical land areas with low altitudes). Although not all of these quantities are directly observed, as for example relative humidity, we will continue to treat them as observed in order to stay, at this stage, as close as possible to the OI system.

2.2 Computation of the cost function and its gradient

The J_O term given in equation (2) is a quadratic form comprised of the covariance matrix O of the observation errors and the observation departures from the model $(y\text{-}Hx)$. Specification of the matrix O (or its inverse) requires a specification of the observation errors; specification of the observation departure requires a specification of the operator H.

2.2.1 *Observation errors*

It is reasonable to assume that observation errors associated with different observation types are uncorrelated. This assumption is used in most operational analysis systems. Thus following *Pailleux et al.*(1991), O is a block-diagonal matrix and hence J_O can be split in a number of independent terms for each individual observation type:

$$J_o = J_{o(\text{synop})} + J_{o(\text{airep})} + J_{o(\text{satob})} + \ldots \tag{4}$$

Within the individual observation type, it is reasonable to assume that observation errors of different observation points are not correlated. Then the contributions from different observations points can be computed independently. For example, the radiosondes' contribution to the cost function can be split:

$$J_{o(\text{temp})} = J_{o(\text{temp})1} + J_{o(\text{temp})2} + \ldots\ldots \tag{5}$$

Furthermore, an additional assumption for each individual observation can be made: the observation errors of the different observed quantities are uncorrelated, allowing to split the above term even further,

$$J_{o(\text{temp})i} = J_{o(u,v)i} + J_{o(\text{Z})i} + J_{o(\text{RH})i} + \ldots \tag{6}$$

where each of the terms on the right hand side of (6) is a quadratic form built on the vertical covariance matrix of the observation error. We further assume all observation errors to be uncorrelated in the vertical except for the geopotential height from TEMP observations.

The observation errors decorrelation assumptions made so far are reasonable for the conventional observation types. However, since the observation error contains the representativeness error one can expect the observation error of adjacent SYNOP temperatures to be correlated in certain weather conditions such a winter inversions over cold land. It should also be borne in mind that the representativeness error includes not only the local (sub-grid) variability, but also the error in the operator H. No explicit provisions have yet been made in the ECMWF 3D variational analysis to take account of the observation error correlations induced by the representativeness errors.

The separations made so far are equally applicable to the gradient of J_O with respect to the model variables:

$$\nabla_x J_o = \nabla_x J_{o(\text{synop})} + \nabla_x J_{o(\text{airep})} + \nabla_x J_{o(\text{satob})} + \ldots \tag{7}$$

$$\nabla_x J_{o(\text{temp})} = \nabla_x J_{o(\text{temp1})} + \nabla_x J_{o(\text{temp2})} + \ldots \tag{8}$$

$$\nabla_x J_{o(\text{temp}i)} = \nabla_x J_{o(u,v)i} + \nabla_x J_{o(Z)i} + \nabla_x J_{o(RH)i} + \ldots \tag{9}$$

2.2.2 *Observation operator (H) and its adjoint (H^*)*

In computing the cost function J_O, one of the first steps is to find the observation departure (y-Hx). Since, with rare exceptions, no model variable is directly observed, H plays the role of a "post-processing" operator providing the model equivalents of the observed parameters. Since the ECMWF forecast model is spectral and the control variable is the vector of spectral model variables, the operator H is the product of several distinct operators. These include the inverse spectral transform (H_{it}) which provide the model values on the gaussian grid at all model levels and a horizontal interpolation (H_{hi}) from grid points to observation points for all the model levels. The current version of the ECMWF 3D variational analysis provides two interpolation options: 1) bi-linear (4 points), and 2) bi-cubic (12 points). The horizontal interpolation is common for all the observation types. The next step is to perform the vertical interpolation (H_{vi}) to get the model equivalents of the observed quantities at the observed levels. H_{vi}, unlike H_{hr}, will differ for different parameters. The interpolation techniques and additional mathematical operations employed will be the subject of the next two sub-sections (2.2.2.1 and 2.2.2.2).

Given the above decomposition, the operator H can be written as product of the individual operators:

$$Hx = H_{vi} H_{hi} H_{it}(x) \tag{10}$$

The evaluation of the cost function is straightforward, given the observed departures and the inverse of the matrix O.

The variational analysis procedure requires the gradient of the cost function with respect to the control variable. This computation uses the adjoint relationship:

$$\nabla_x J_o = H^* \left(\nabla_{Hx} J_o \right) \tag{11}$$

or alternatively, using the chain rule for differentiation:

$$\nabla_x J_o = H_{it}^* H_{hi}^* H_{vi}^* \left(\nabla_{Hx} J_o \right) \tag{12}$$

where the term in brackets in the above equation (12) is simply:

$$\nabla_{Hx} J_o = -2O^{-1}(y - Hx) \tag{13}$$

The gradient of J_O given by (12) is used in an iterative descent procedure to perform the minimisation needed for the variational analysis.

2.2.2.1 Vertical upper air observation operators

The following vertical interpolation techniques are employed:

- wind: linear in ln(p) on full model levels but quadratic at the two top levels
- geopotential: as for wind but the interpolated quantity is the departure from ICAO standard atmosphere
- humidity: linear in p on full model levels
- temperature: linear in p on full model levels

A detailed description of the vertical interpolation is given in the appendix A1.

The integration of the hydrostatic equation used for the geopotential operator depends on both temperature and specific humidity.

2.2.2.2 Surface observation operators

The vertical gradient of the model variables varies strongly in the lowest part of PBL, where flow changes are induced on very short time and space scales, due to physical factors such as turbulence and terrain characteristics. In such a situation the operator H_{vi} should take account of the model surface layer. The operators for 10m wind and 2m temperature are

based on Monin-Obukhov (MO) similarity theory, whereas in the case of 2m relative humidity it is a somewhat simpler (still highly non-linear) operator. The similarity functions match the physical parametrization of the forecast model.

The post-processed values of 10m wind $(u_{10m}{}^{pp}, v_{10m}{}^{pp})$ and 2m temperature $(t_{2m}{}^{pp})$ are functions of temperature T_1, specific humidity Q_1 and wind components U_1, V_1 at the lowest model level. The surface pressure (p_s) and temperature (T_s) are also included as well as the roughness length (Z_0), vegetation (C_v), soil wetness of the first model soil layer (W_s), water content of the skin reservoir (W_1) and snow (S_n), or formally written:

$$(u_{10m}{}^{pp}, v_{10m}{}^{pp}, T_{2m}{}^{pp}) = f(T_1, Q_1, U_1, V_1, p_s, T_s, Z_0, C_v, W_s, W_1, S_n) \tag{14}$$

On the assumption that the relative humidity is constant in the model's surface layer, the post-processed 2m relative humidity $(RH_{2m}{}^{pp})$ is assigned the value of the lowest model level which is a function of the lowest model level pressure (p_1), specific humidity (Q_1) and temperature (T_1):

$$RH_{2m}{}^{pp} = f(p_1, Q_1, T_1) \tag{15}$$

Details of post-processing operators are given in the appendix A2.

3. 3D VARIATIONAL MASS AND WIND ANALYSIS: PRELIMINARY RESULTS

We compare the performance of the 3D variational analysis with that of the operational ECMWF OI in the following experiments, described in detail in section 3.1 and 3.2:

a) OI analysis using:
- p_S from SYNOP and DRIBU
- (u_{10m}, v_{10m}) from SYNOP, DRIBU, TEMP and PILOT over sea and low altitude land areas in the tropics
- (u, v) from AIREP, SATOB, TEMP and PILOT
- Z from TEMP

b) 3D VAR analysis using the same set of observations as a)

c) 3D VAR analysis using only 10m wind observations from SYNOP, DRIBU, TEMP and PILOT everywhere except near the coast

d) 3D VAR analysis using only 2m temperature observations from SYNOP, DRIBU and TEMP everywhere but near the coast

e) 3D VAR analysis using the observations of b) plus c) plus d).

No observations of relative humidity were used because, so far, only the "dry" part of the 3D variational analysis has been validated. The resolution used was T63/L19 corresponding to 19 levels in the vertical and spectral truncation of T63 in the horizontal.

The J_b term, used in addition to J_0 is discussed by *Heckley et al.* (1992). J_b is made multivariate via Hough mode separation with a large penalty on the unbalanced part of the departure from the first guess. The horizontal structure functions are prescribed in spectral space using a parametric form which in grid point space closely resembles the form which is currently used in the OI. A length scale of about 500km is used. The prediction errors were kept horizontally constant to the global average of OI values. The vertical correlation matrix is a block diagonal matrix with one block for each 3D state variable. The vertical correlation matrices have been closely modelled on those used by the OI in the southern hemisphere. The control variable includes all the model upper air variables defined in spectral space, normalised by prediction errors (see *Heckley et al.*, 1992). The mass variable is $P = \phi + RT \ln p_s$ variable (see *Heckley et al.*, 1992). Surface fields are not updated in the analysis but kept constant. A quasi-Newton algorithm is used in the minimisation for 3D VAR experiments, with the number of iterations set to 30.

3.1 3D variational experiments - OI comparison

In this section we compare OI and 3D VAR experiments as described in a) and b) above. The OI is used to quality control the observations for the 3D VAR. The analysis experiments were performed for 91/12/30 12Z, and the OI experiment was preceded by 12 hours (two analysis cycles) of assimilation. Both analysis experiments used the same first guess.

Figs. 1-7 show the N.H. analysis increment maps for $\log p_s$, wind and temperature at the lowest model level (about 32 metres above the ground), level 11 (about 500hPa) and the top model level (7.35hPa). The surface pressure increment maps show similar features; the 3D VAR increments have slightly smaller amplitude and a broader spatial scale. Wind increments are very similar in the troposphere, but at the top level the 3D VAR has larger values. The largest increment differences can be seen in temperature, particularly at the top and bottom of the model atmosphere. Repeating the OI experiment with constant prediction errors, used by the 3D VAR, gave essentially the same results (not shown). This indicates that the differences in the increments between 3D VAR and OI cannot be explained by the different prediction errors used. The temperature increments reflect a bias over continental land masses: the scale of the bias is such that, with the current formulation of the horizontal structure functions in 3D VAR, this information is spread almost globally also over sea areas. The easiest way to improve this problem is to "refine" the horizontal structure functions (*Heckley*, personal communication).

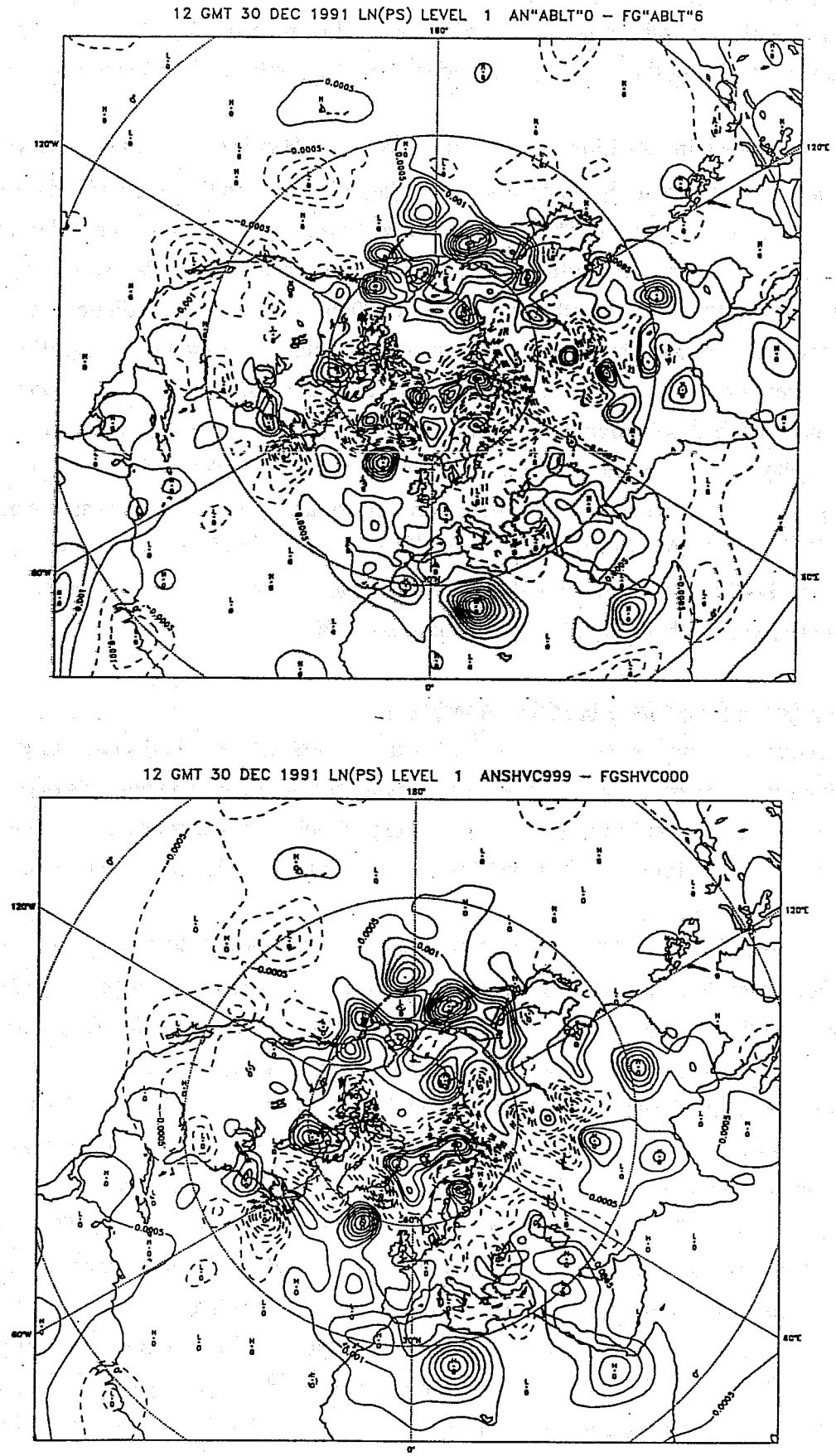


Fig.1: Northern hemisphere surface pressure ($\ln p_s$) increment maps: (upper panel) exp. a (OI) and (lower panel) exp. b (3D VAR); contouring interval is 0.0005.

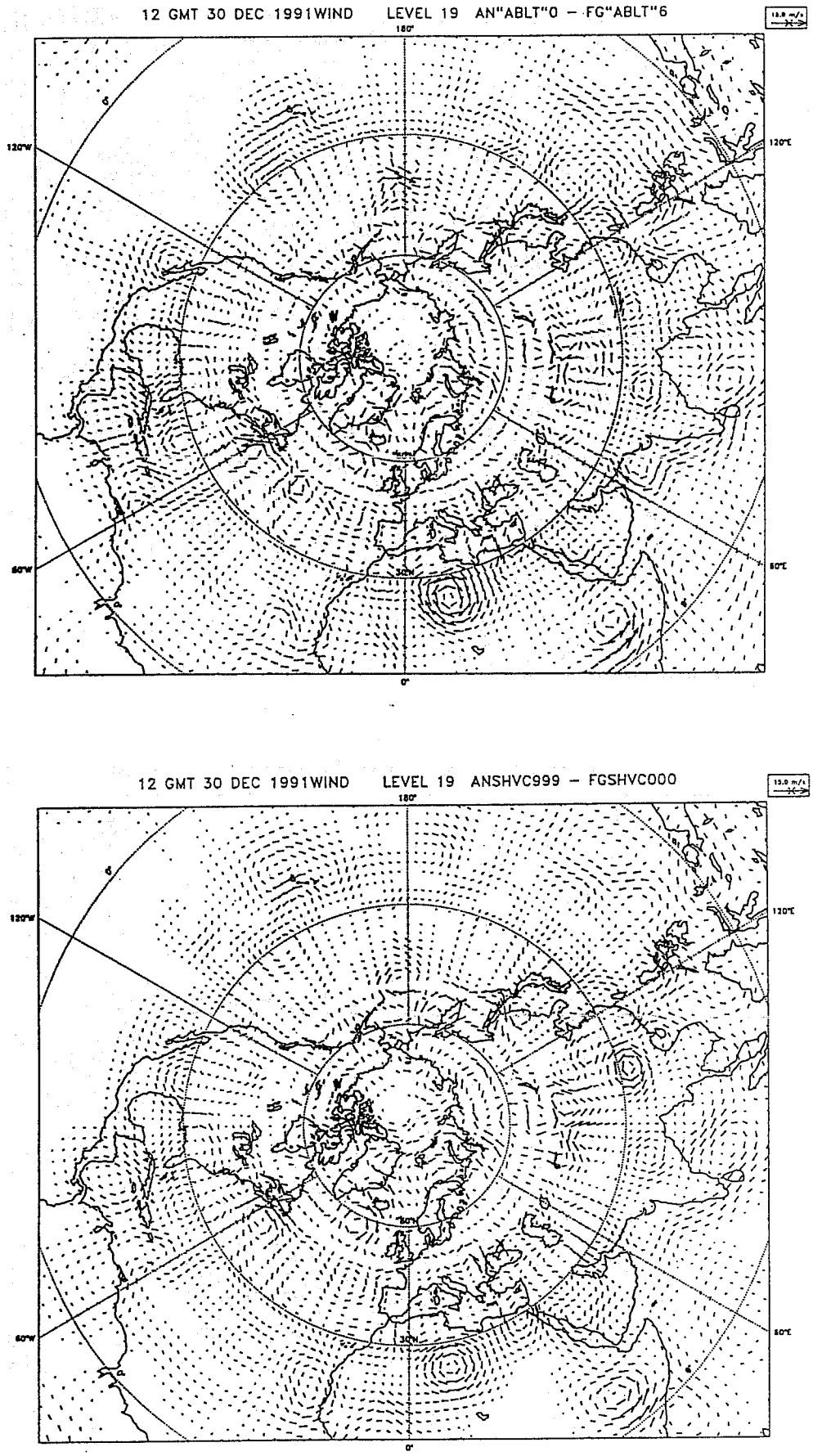


Fig. 2: Northern hemisphere lowest model level wind increment maps: (upper panel) exp. a (OI) and (lower panel) exp. b (3D VAR).

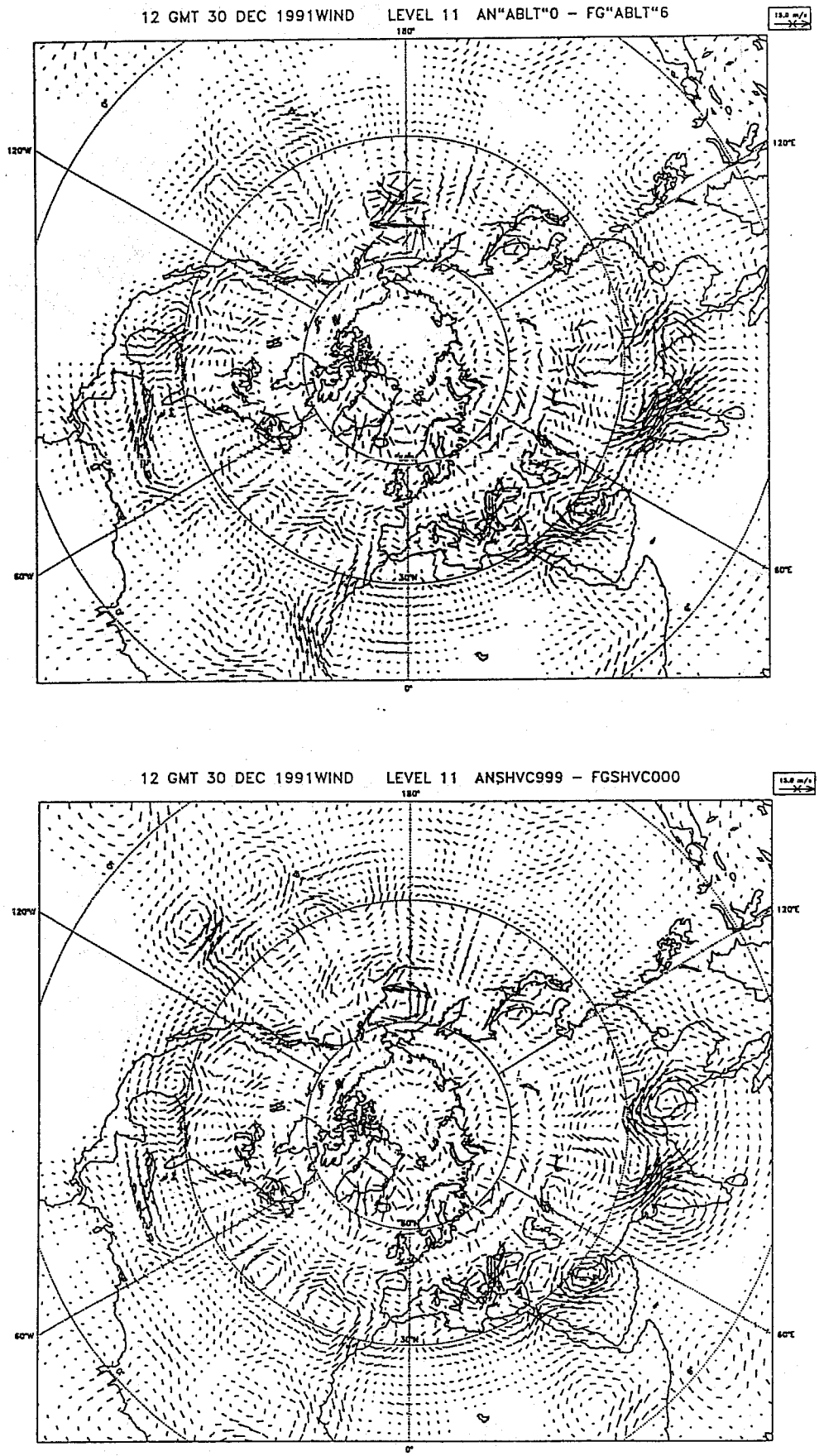


Fig.3: Northern hemisphere 11th model (about 500hPa) level wind increment maps: (upper panel) exp. a (OI) and (lower panel) exp. b (3D VAR).

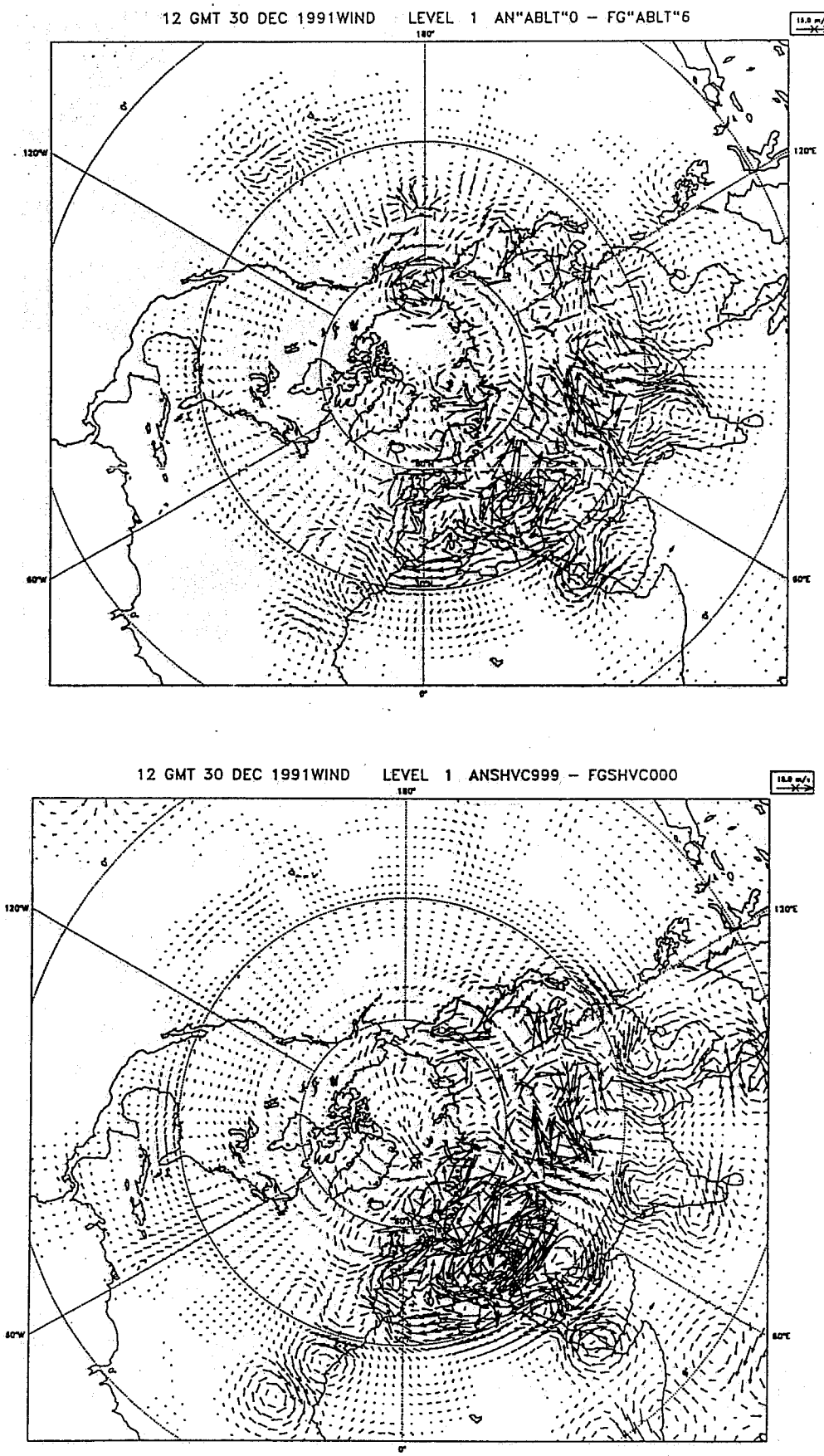


Fig.4: Northern hemisphere top model level wind increment maps: (upper panel) exp. a (OI) and (lower panel) exp. b (3D VAR).

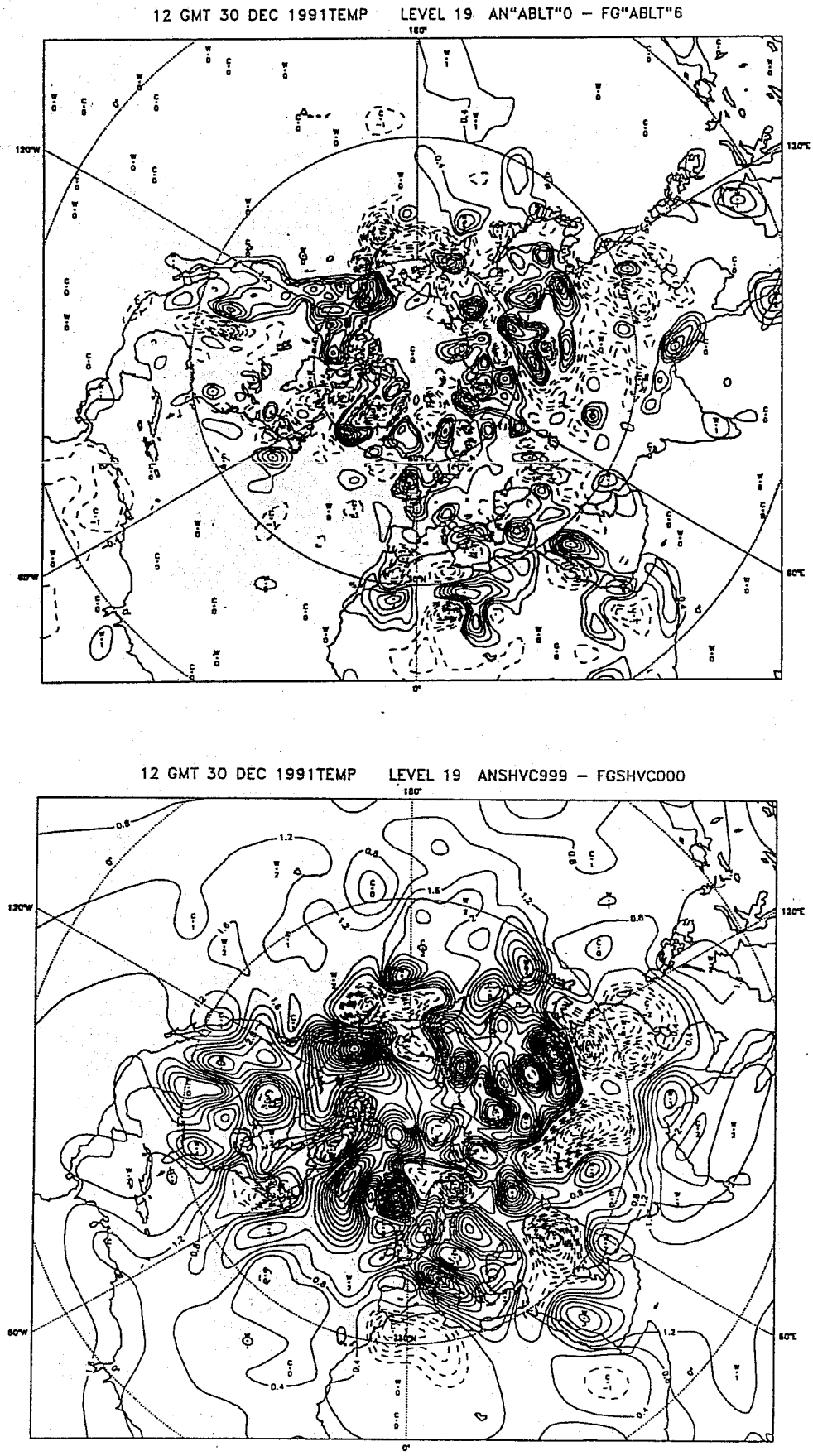


Fig.5: Northern hemisphere lowest model level temperature increment maps: (upper panel) exp. a (OI) and (lower panel) exp. b (3D VAR): contouring interval is 0.4 degrees.

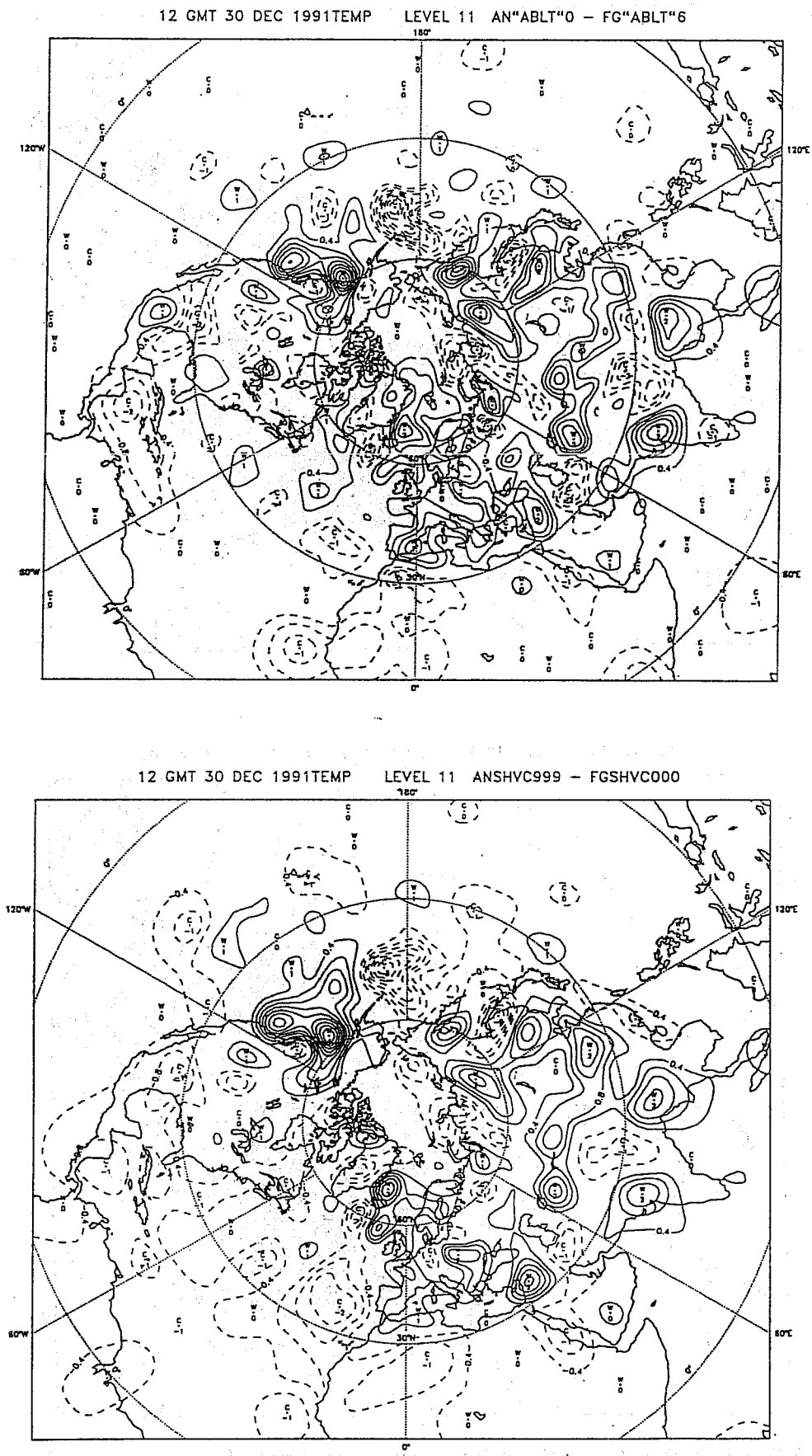


Fig.6: Northern hemisphere 11th model level temperature increment maps: (upper panel) exp a (OI) and (lower panel) exp b (3D VAR); contouring interval is 0.4 degrees.

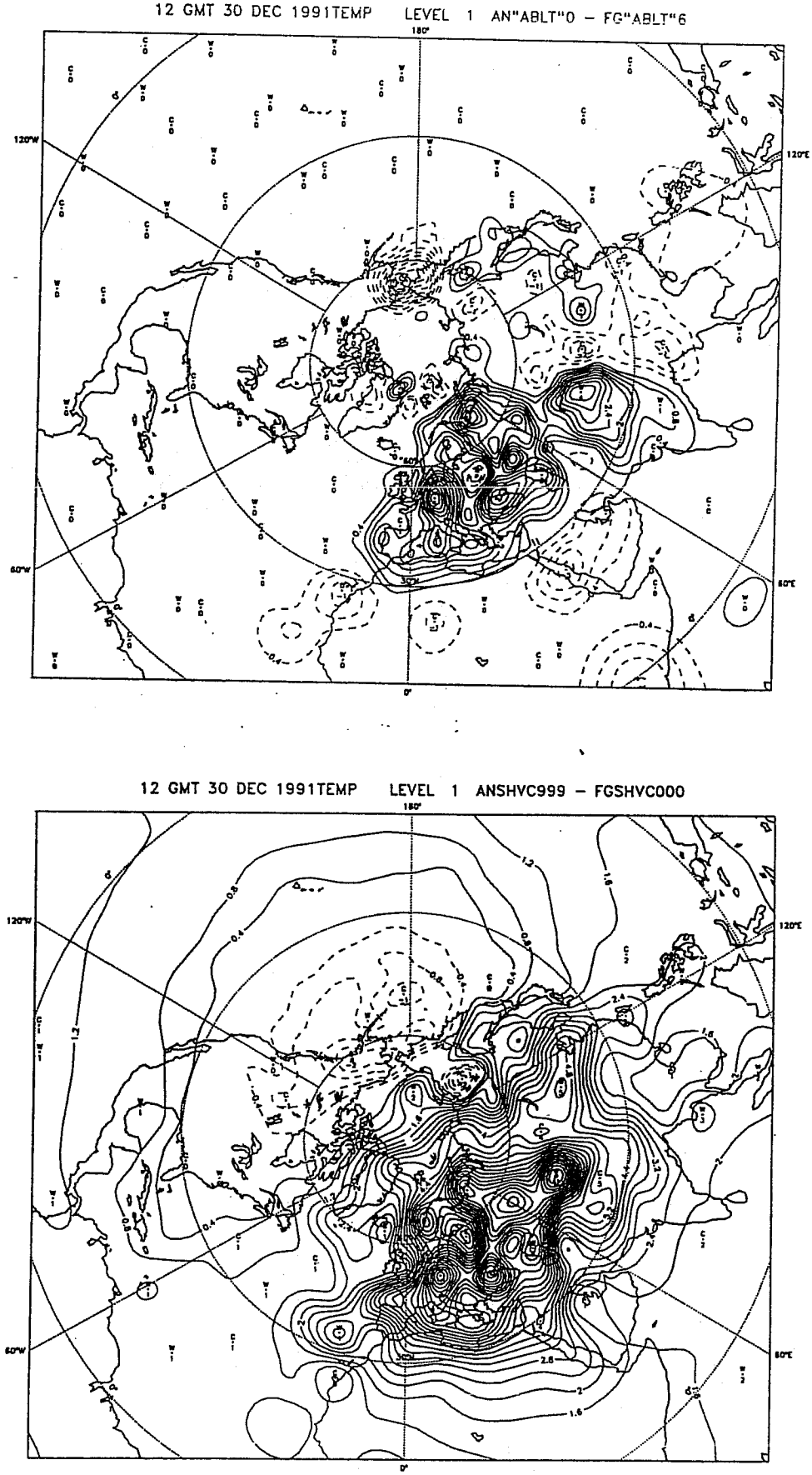


Fig.7: Northern hemisphere top model level temperature increment maps: (upper panel) exp a (OI) and (lower panel) exp b (3D VAR); contouring interval is 0.4 degrees.

Fig. 8 shows the sea level pressure with 850hPa temperature (top panel), and 500hPa height and temperature (bottom panel) for the OI uninitialised analysis. A comparison with the corresponding pictures for 3D VAR (fig. 9) show very small differences at 500hPa and somewhat larger discrepancies near the surface. It is worth noticing the low off the coast of the Greenland, which is deeper in the OI system.

Fig. 10 shows the RMS fit (solid line) and bias (dotted line) to TEMP wind and height data. The thicker lines are the uninitialised OI and the thinner 3D VAR analysis. OI fits better the data in the extratropics with the exception of the tropospheric geopotential height in the southern hemisphere and the stratospheric winds in both hemispheres. Biases for both systems are very similar, 3D VAR being marginally better. In the tropics the better RMS fit to data by OI is even more pronounced. A possible explanation for this lies in the fact that 3D VAR imposes mass and wind balance in the tropics whereas it is absent in OI.

Fig. 11 compares the initialized OI analysis with 3D VAR. 3D VAR outperforms the OI in the case of height at every level and in the case of wind components only in the stratosphere, whereas results seem to favour OI for the tropospheric winds.

Two 24 hour forecast experiments have been performed using the OI and 3D VAR analyses as initial conditions. The model used was the adiabatic version of the ECMWF IFS model with the addition of vertical diffusion. Figs. 12 and 13 show the one day forecasts, and fig. 14 is the verifying OI analysis. The differences between the forecasts (fig.12 - fig.13) are of the same order of magnitude as the differences between the initial fields (fig.8 - fig.9). Note that the low off the Greenland coast is deeper in both 24 hour forecasts than in the verifying analysis.

3.2 3D variational experiments - 10m wind and 2m temperature observations

3D variational experiments were performed to examine the impact of using 10m wind or/and 2m temperature observations. Direct comparison with the OI is not possible since the only surface observations used by the OI are 10m wind over sea and in low altitude land areas in the tropics. Experiments have been carried out by using only either 10m wind or 2m temperature observations at a time. These observations were only used over unambiguous sea or land areas, avoiding possible problems in the vicinity of the coast. In order to avoid using observations in the areas where there is a large mismatch between the model's orography and the station height, we reject observations with larger mismatch than 200m. In addition to this, 2m temperature observations were corrected for the difference between the model's orography and station height assuming a constant lapse rate of $6.5\ 10^{-3}$ K/m.

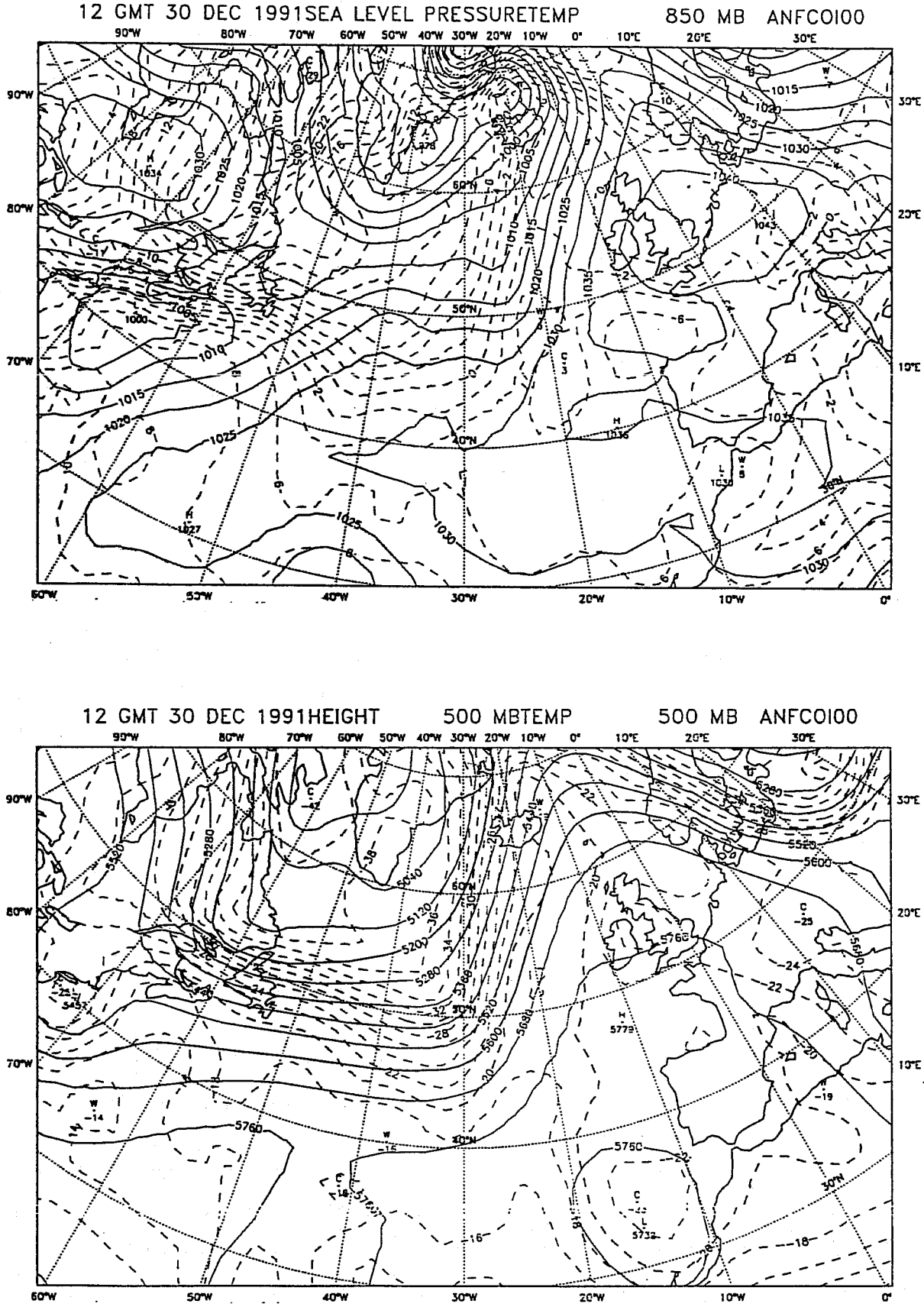


Fig.8: OI analysis (exp. a): (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

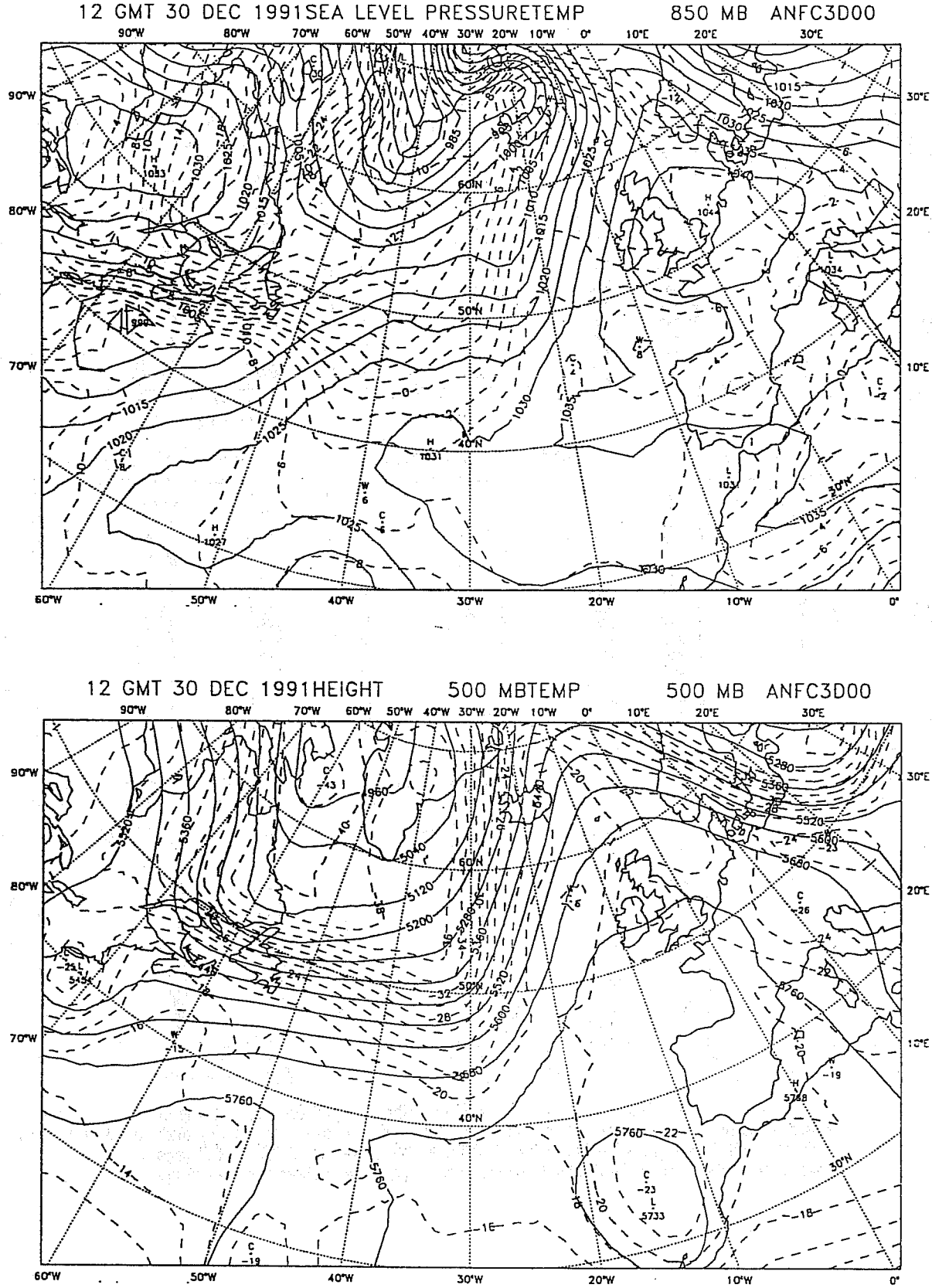


Fig.9: 3D VAR analysis (exp b): (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

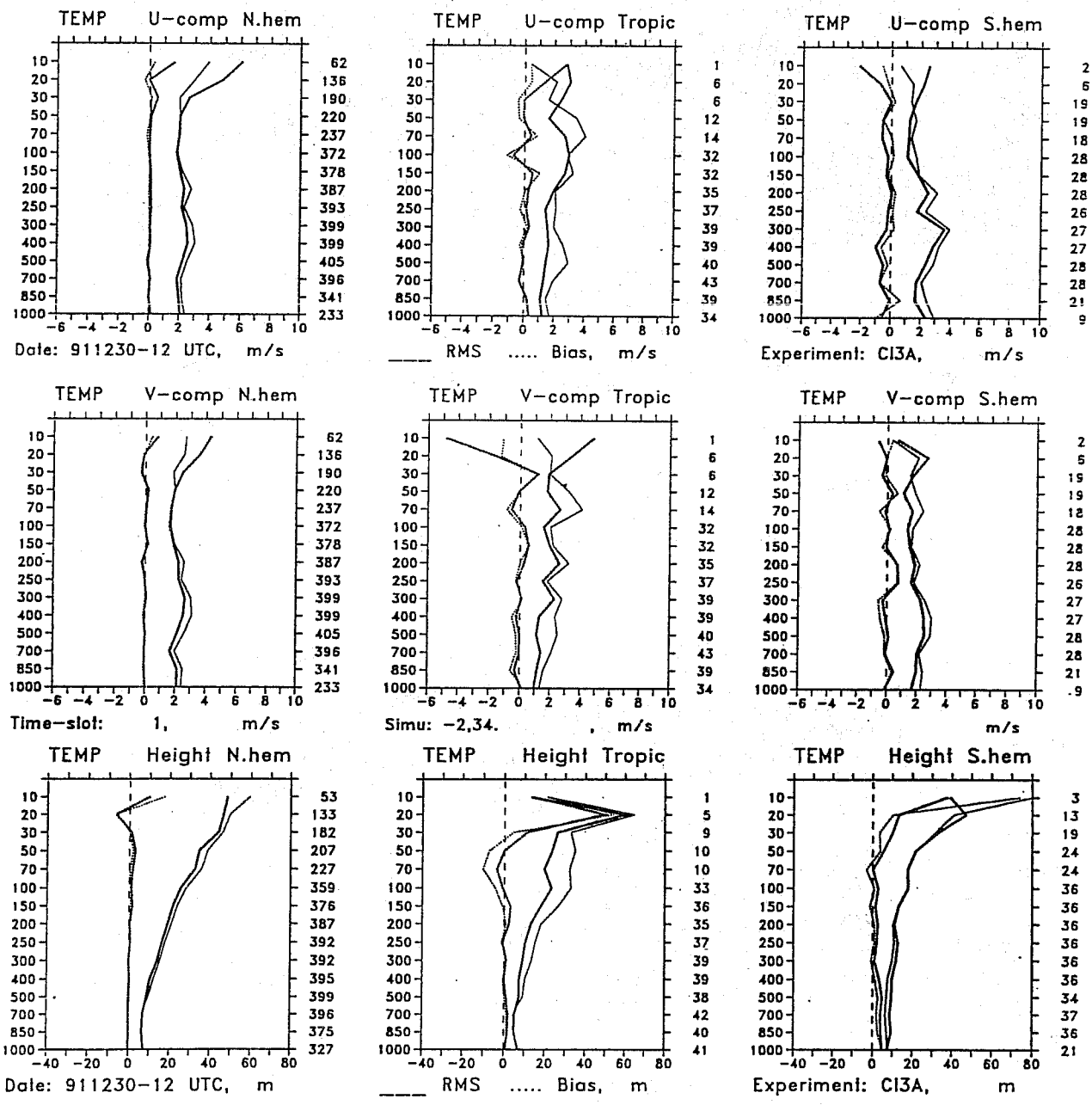


Fig.10: Bias (dotted line) and RMS (solid line) fit to TEMP: (three upper panels) u component, (three middle panels) v component and (three lower panels) height data; thicker lines (solid and dotted) represent uninitialised OI fit (exp. a), whereas thinner (solid and dotted) represent 3D VAR fit (exp. b); each of three panels refer to the northern hemisphere, tropics and southern hemisphere from left to right.

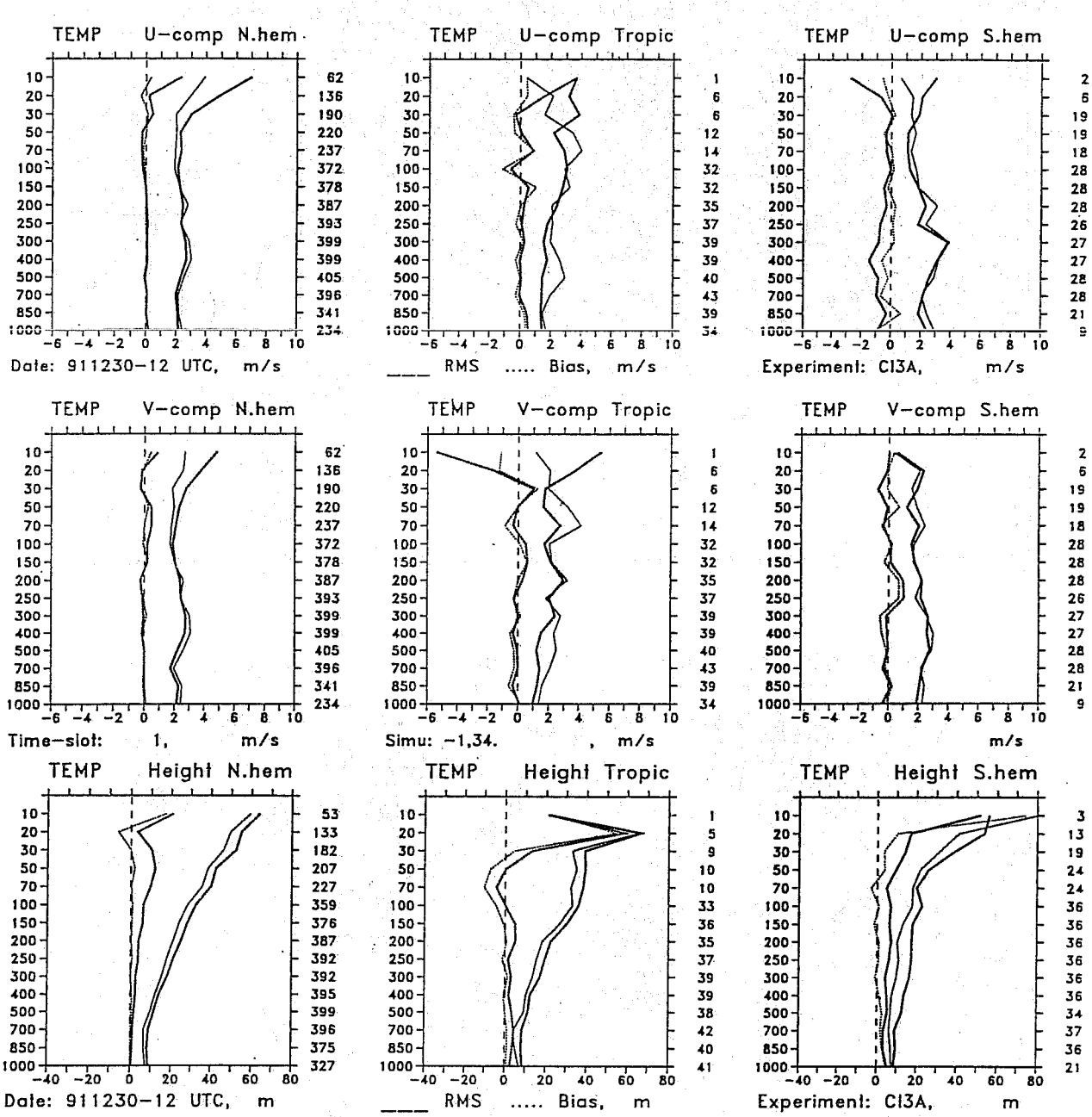


Fig.11: Bias (dotted line) and RMS (solid line) fit to TEMP: (three upper panels) u component, (three middle panels) v component and (three lower panels) height data; thicker lines (solid and dotted) represent initialised OI fit (exp. a), whereas thinner (solid and dotted) represent 3D VAR fit (exp. b); each of three panels refer to the northern hemisphere, tropics and southern hemisphere from left to right.

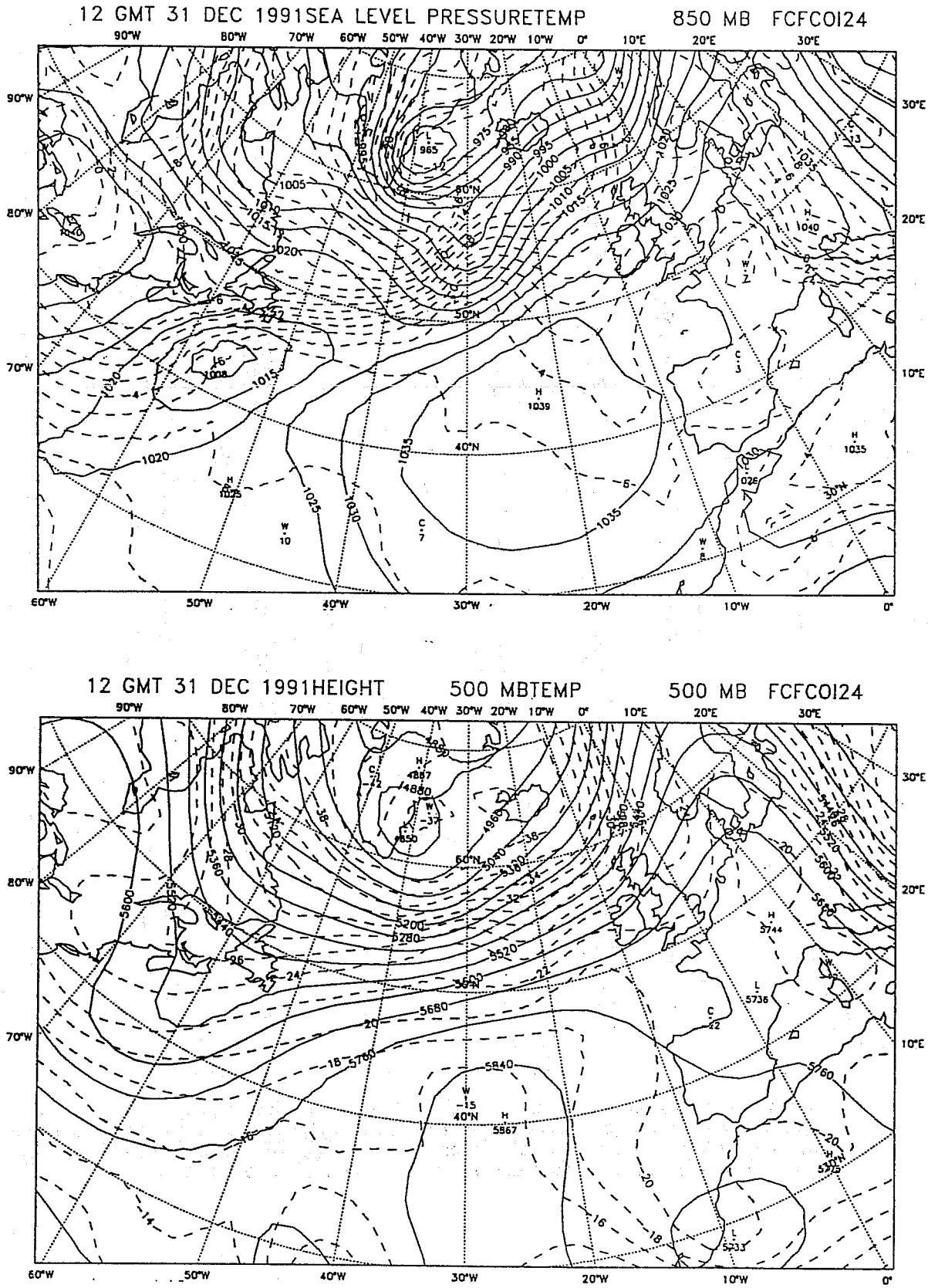


Fig.12: 24hr forecast: (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); initial condition is OI analysis (exp. a; see fig. 8); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

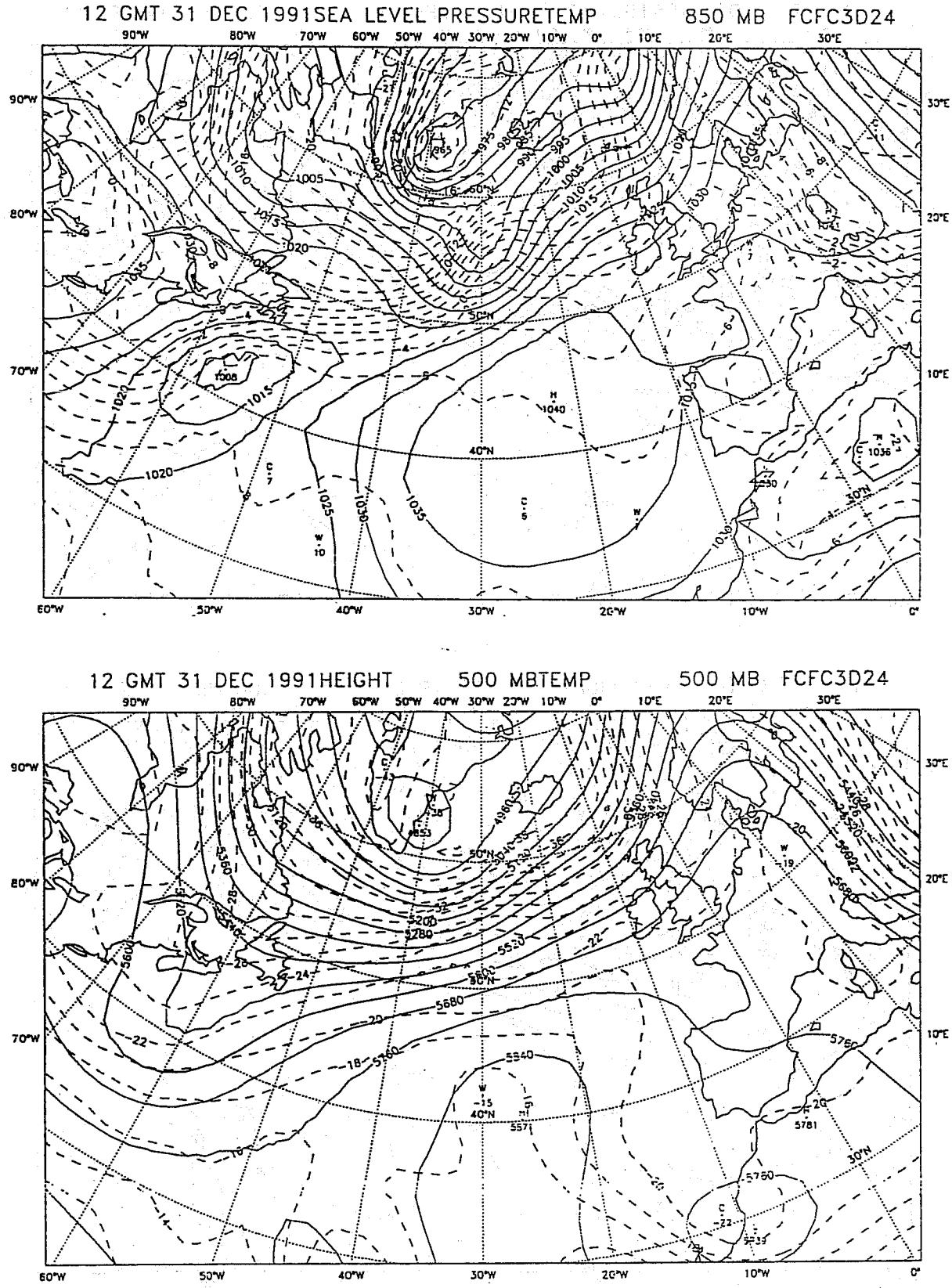


Fig.13: 24hr forecast: (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); initial condition is 3D VAR analysis (exp. b; see fig. 9); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

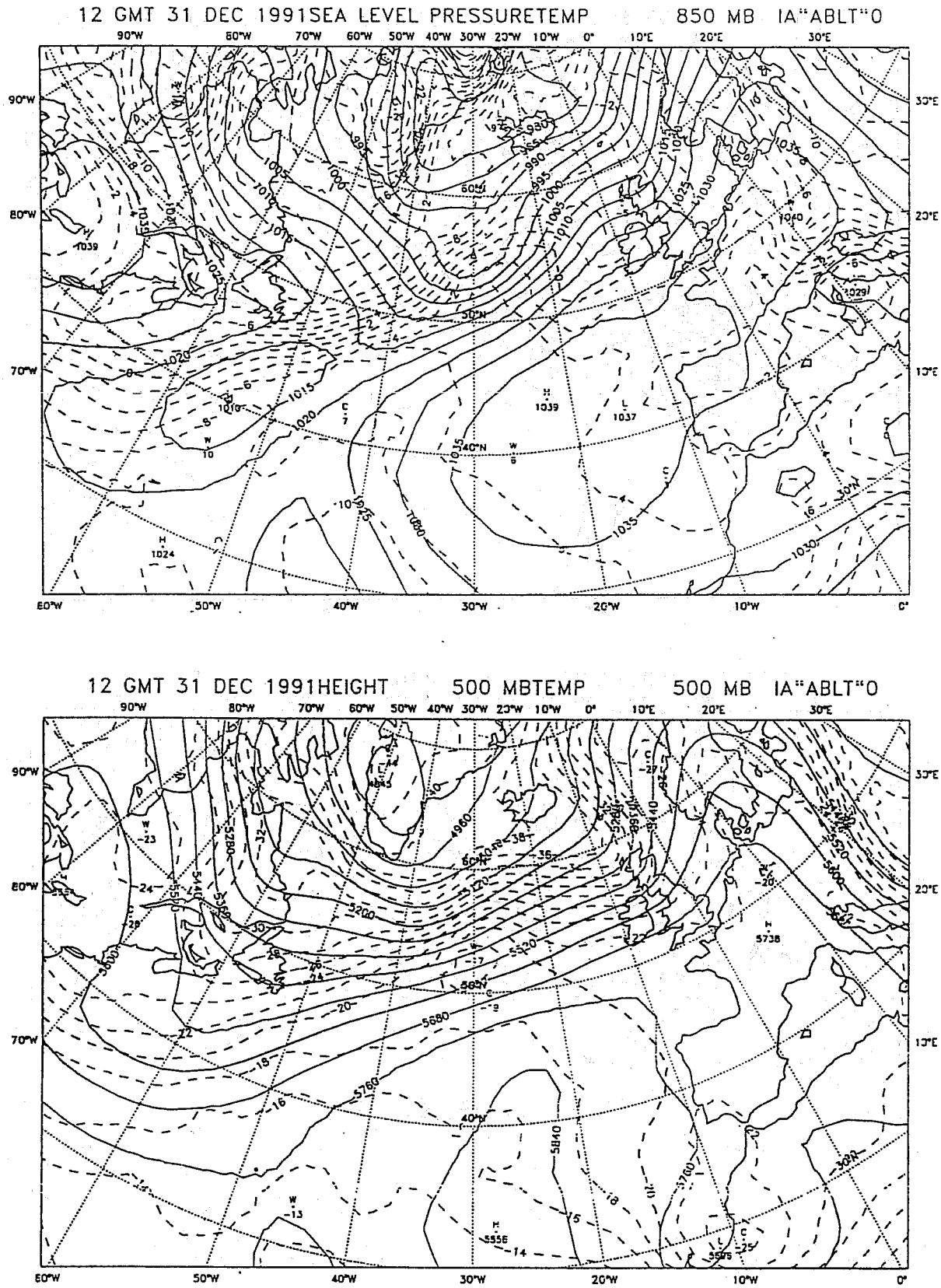


Fig.14: OI (verifying) analysis: (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

Figs. 15-17 show 3D variational analysis increments for surface pressure and lowest model level wind and temperature for the northern hemisphere when only the 10m wind observations were used. Although temperature observations are not used, wind induced temperature increments are created. The impact has also been felt even at the higher levels, through the vertical correlation of J_b.

Figs. 18-20 are equivalent to the maps figs. 15-17 when 2m temperature observations are the only observations used.

Fig. 21 show the analysis for sea-level pressure/850hPa temperature and 500hPa height/temperature when near surface observations are used together with all the upper air observations used previously. This should be compared with fig. 9. Small differences can be observed. Using this as an initial condition, a 24 hour forecast experiment, fig. 22, has been carried out. For forecast impact this should be compared with fig. 13. Differences between forecast fields are of the same order of magnitude as differences between analysed fields.

In order to understand the behaviour of the surface operator a more detailed analysis was carried out. Fig. 23 shows two scatter diagrams for post-processed $\left(T_{2m}{}^{pp}\right)$ against observed $\left(T_{2m}{}^{\text{obs}}\right)$ temperatures at the beginning (initial point) and at the end (final point) of the minimisation for all used observations. The variational analysis reduces the distance between the post-processed and the observed values. Additional insight can be gained by separating the points into stability classes. Fig. 24 depicts the scatter for all the observed points in unstable case. There is a migration of points from stable/neutral to unstable conditions which is also substantiated by fig. 25 showing the bulk Richardson number R_i (final) versus R_i (initial).

Since no model surface variables are included in the control variable the following two figures show how this can affect the overall analysis. Fig. 26 compares $\left(T_1 - T_s\right)$ versus $\left(T_{2m}{}^{pp} - T_{2m}{}^{\text{obs}}\right)$ for the initial field for both sea and land points, whereas fig. 27 is for the final field. The comparisons between the initial and the final situation in two different cases show a similar distribution before the minimisation: points are more spread horizontally than in the vertical. At the end of the analysis the distance between the post-processed and observed temperature has been reduced. This reduction of the horizontal spread has been achieved at the expense of increased vertical spread, much more evident for points over land than over sea. Because of the large sea thermal inertia, the surface temperature is changing slower over sea than over land: therefore having T_S constant through the minimisation is not an important limitation over sea. The next two figures (bar charts) will help in highlighting this problem. Figs. 28 (land points) and 29 (sea points) show the

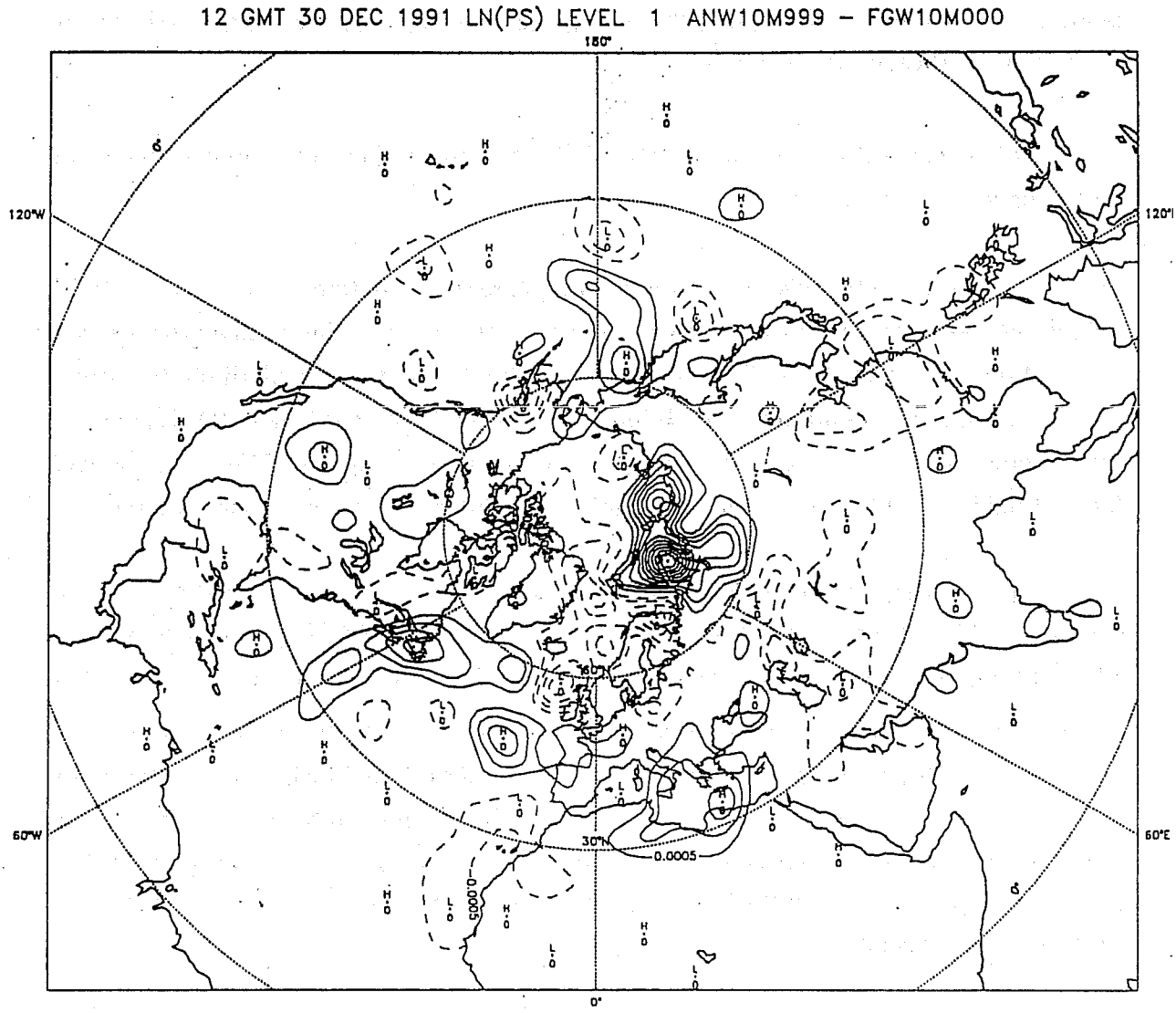


Fig.15: Northern hemisphere 3D VAR analysis (exp. c) surface pressure ($\ln p_s$) increment map; contouring interval is 0.0005.

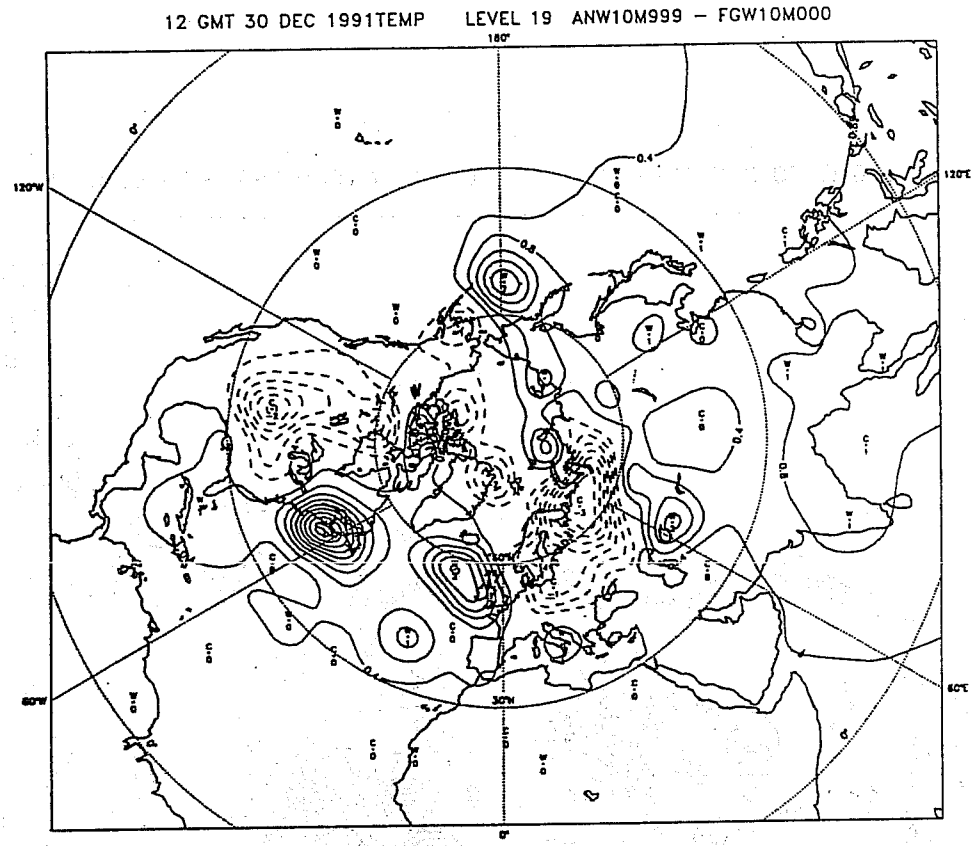


Fig.16: Northern hemisphere 3D VAR analysis (exp. c) lowest model level wind increment map.

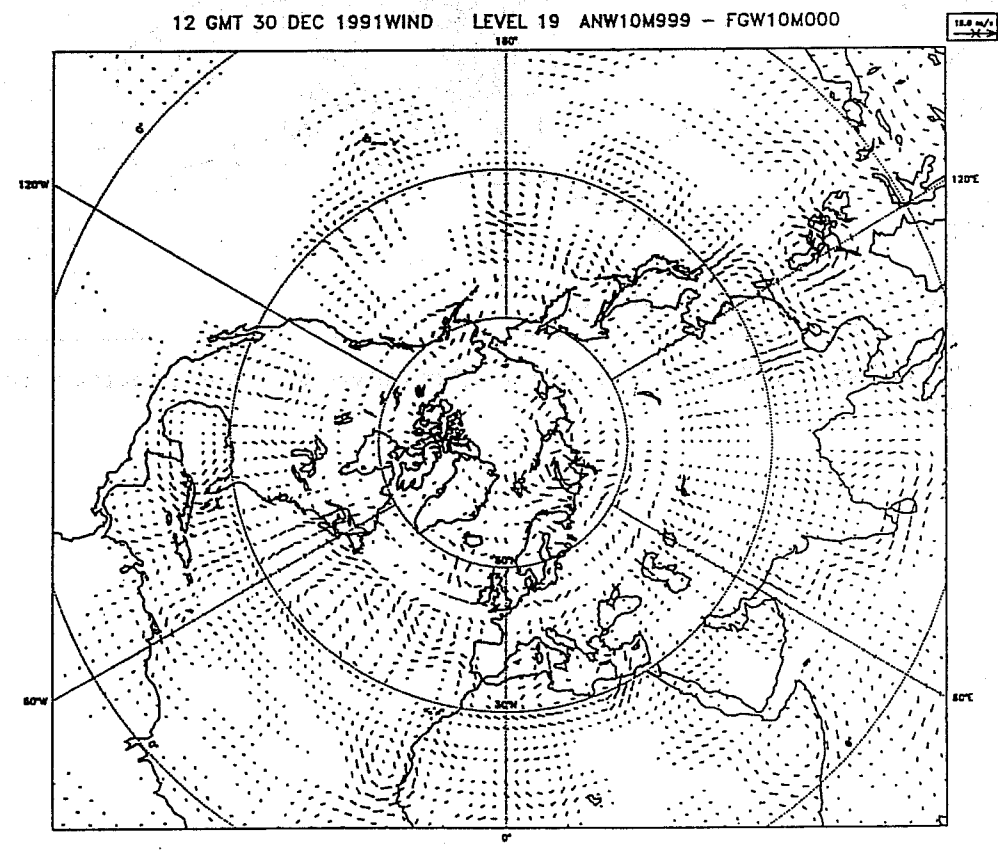


Fig.17: Northern hemisphere 3D VAR analysis (exp. c) lowest model level temperature increment map; contouring interval is 0.4 degrees.

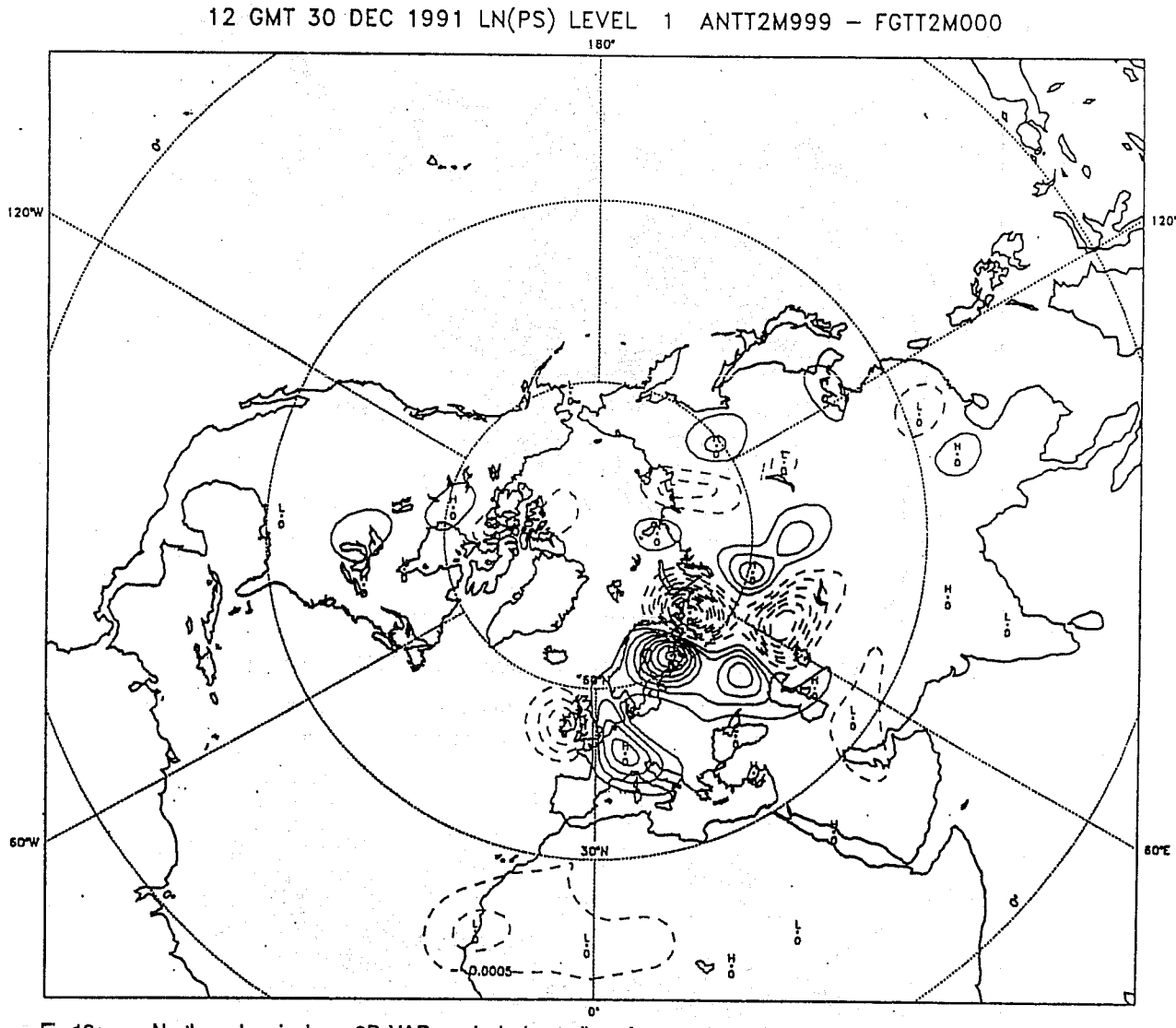


Fig.18: Northern hemisphere 3D VAR analysis (exp. d) surface pressure ($\ln p_s$) increment map; contouring interval is 0.0005.

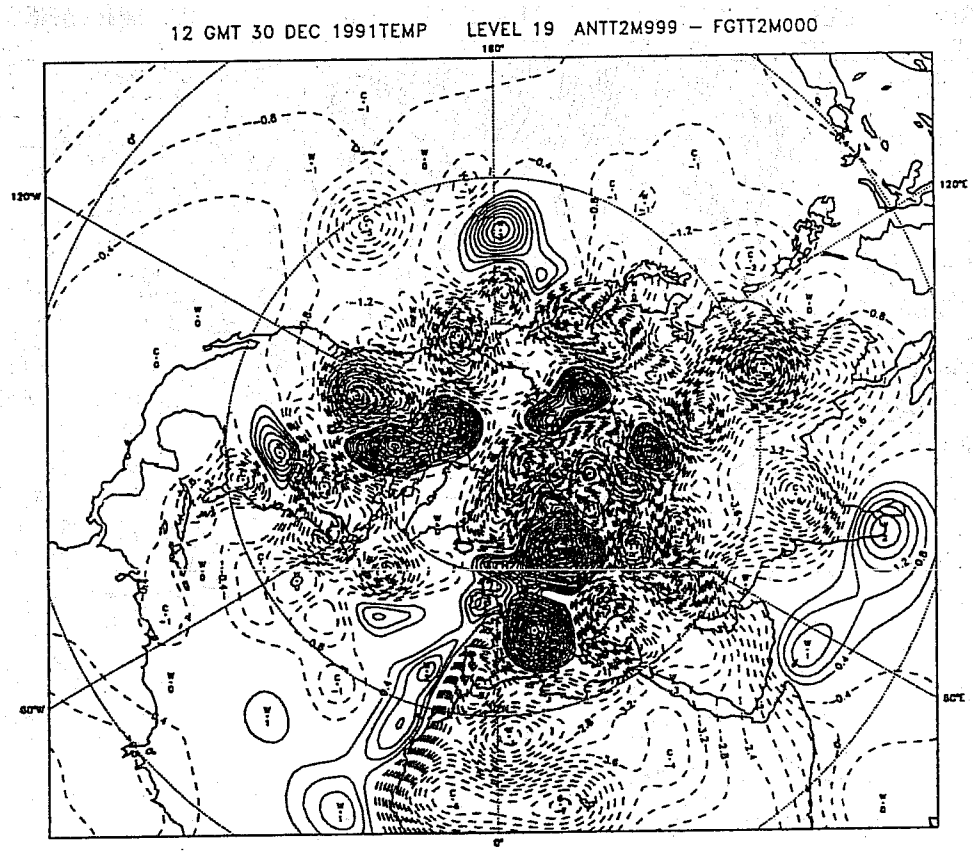


Fig.19: Northern hemisphere 3D VAR analysis (exp. d) lowest model level wind increment map.

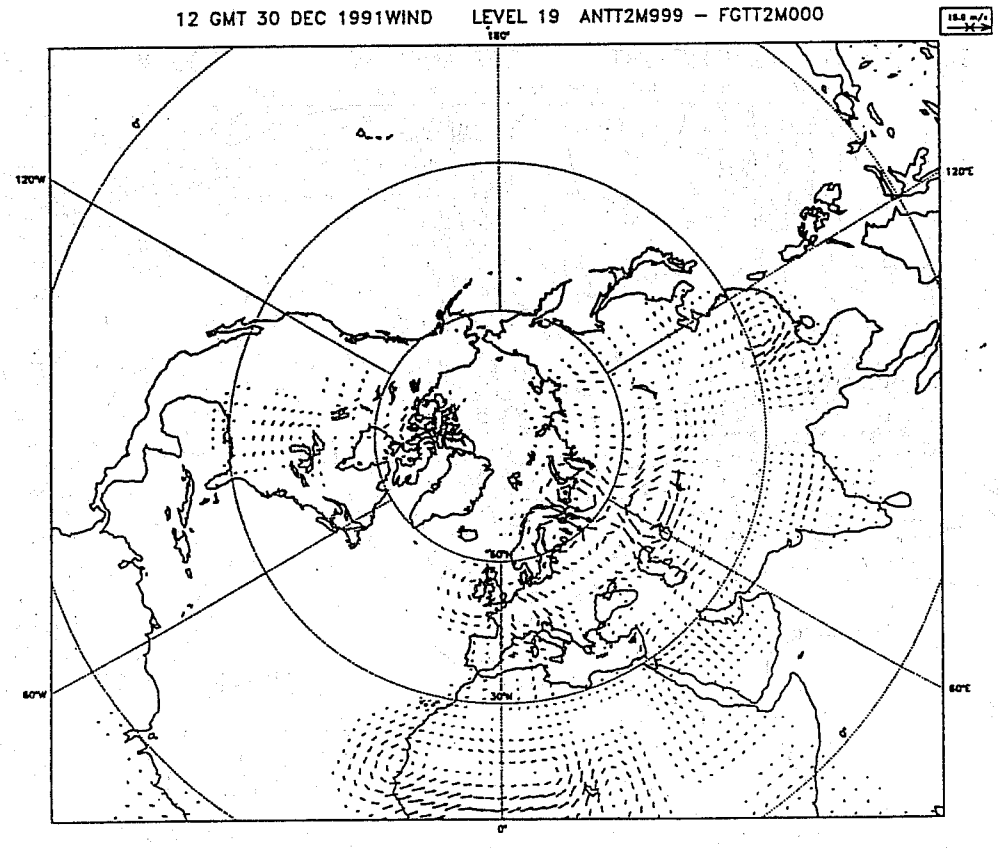
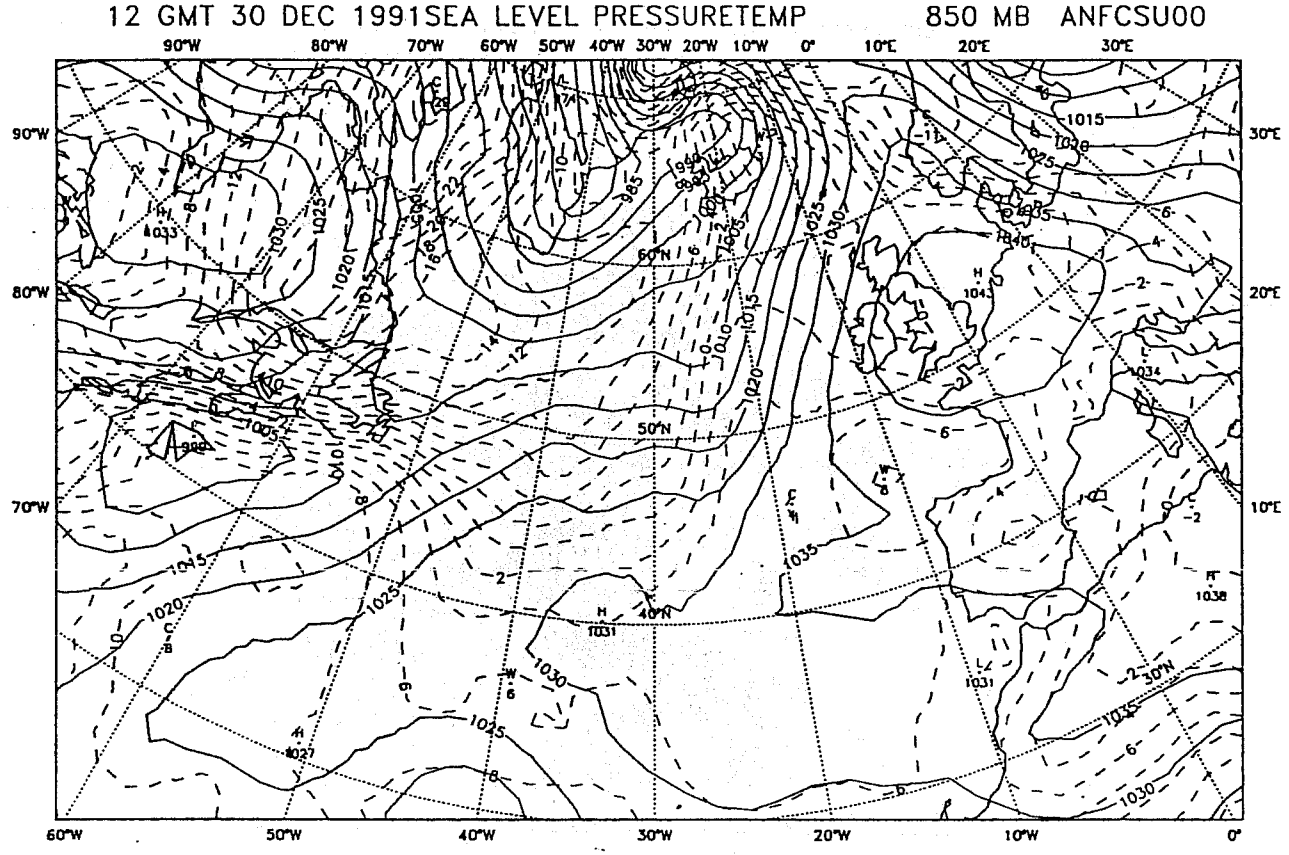


Fig.20: Northern hemisphere 3D VAR analysis (exp. d) lowest model level temperature increment map; contouring interval is 0.4 degrees.



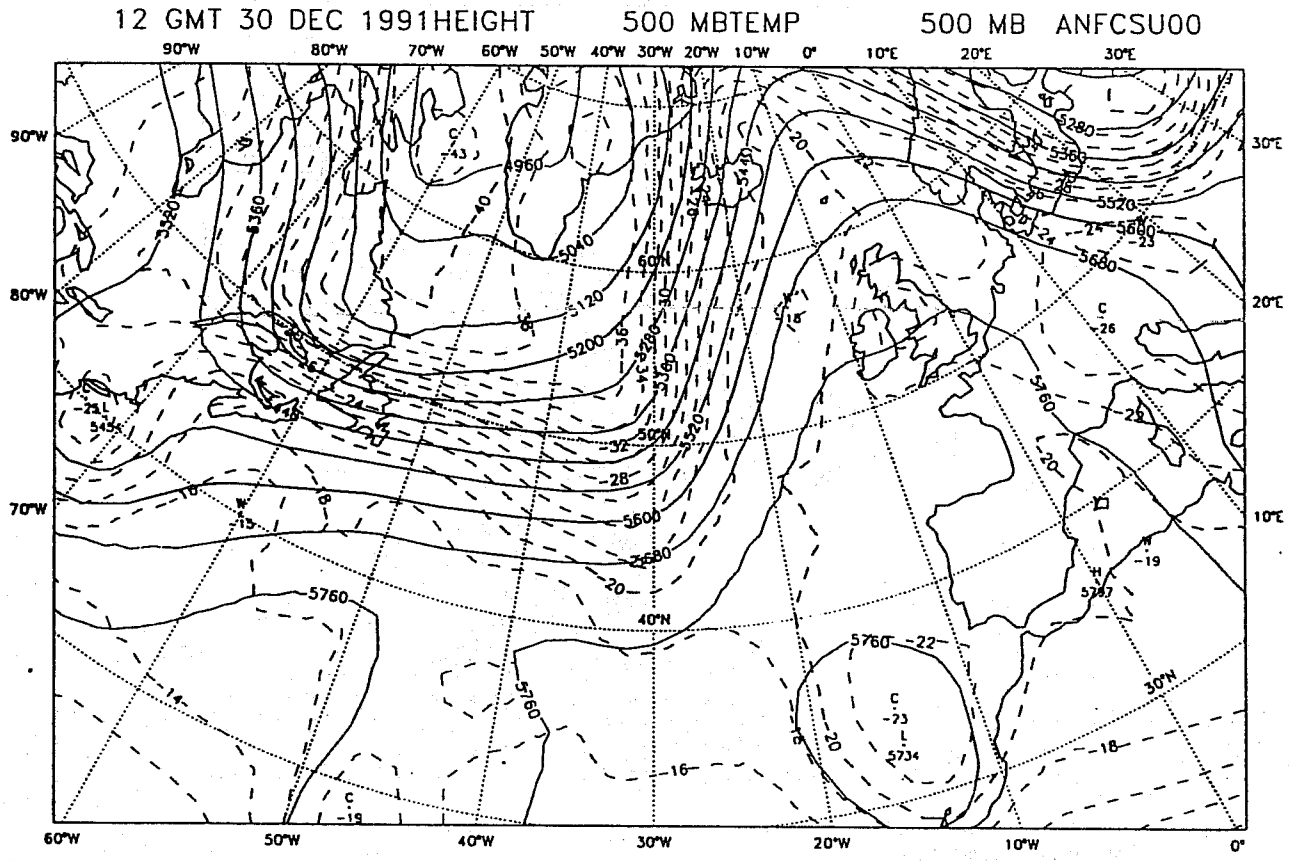


Fig.21: 3D VAR analysis (exp. e): (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

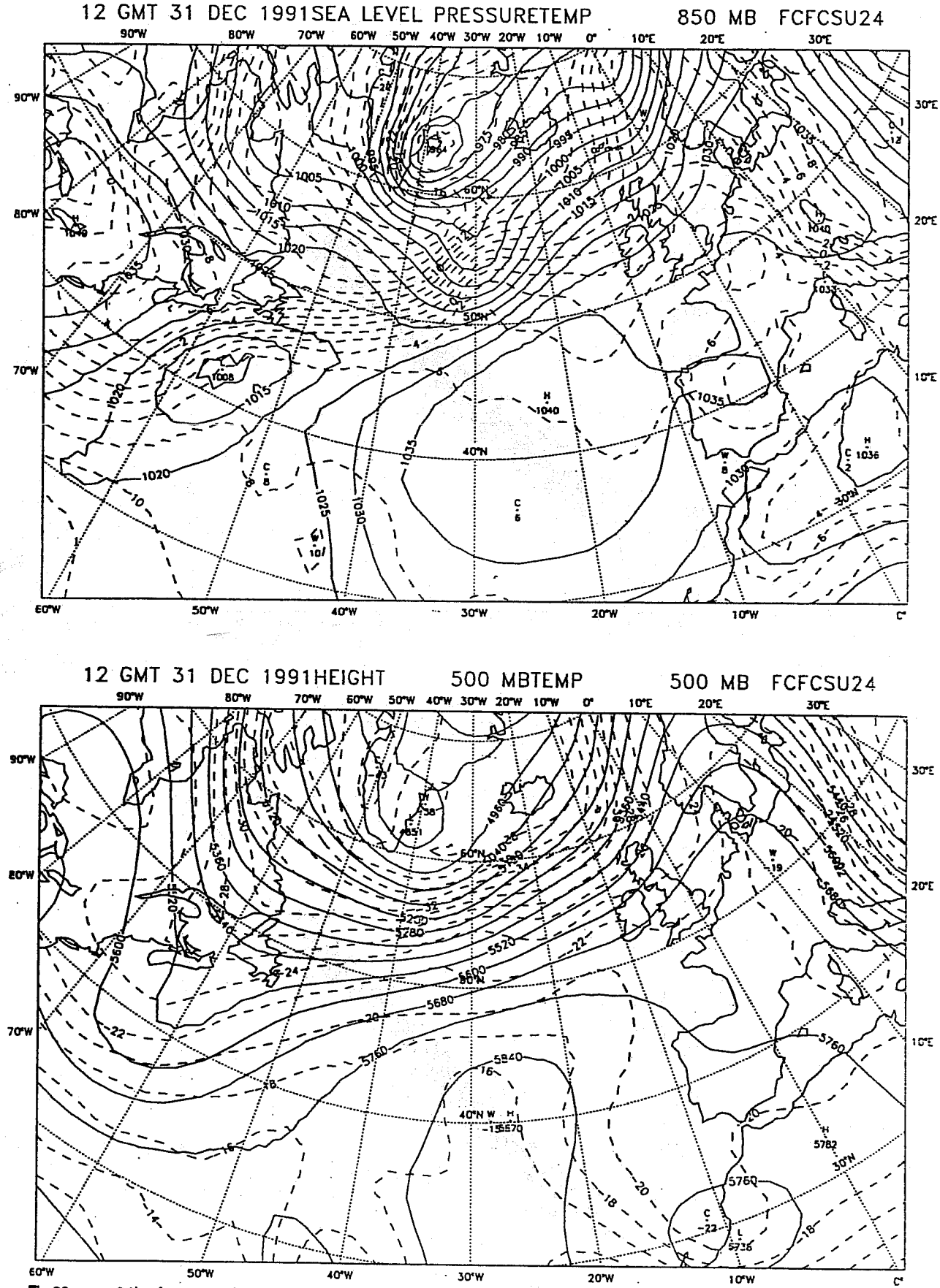


Fig.22: 24hr forecast: (upper panel) sea level pressure (solid line) with 850hPa temperature (dashed line) and (lower panel) 500hPa height (solid line) with temperature (dashed line); initial condition is 3D VAR analysis (exp. d; see fig. 21); contouring intervals are: pressure 5hPa, height 80m and temperature 2 degrees.

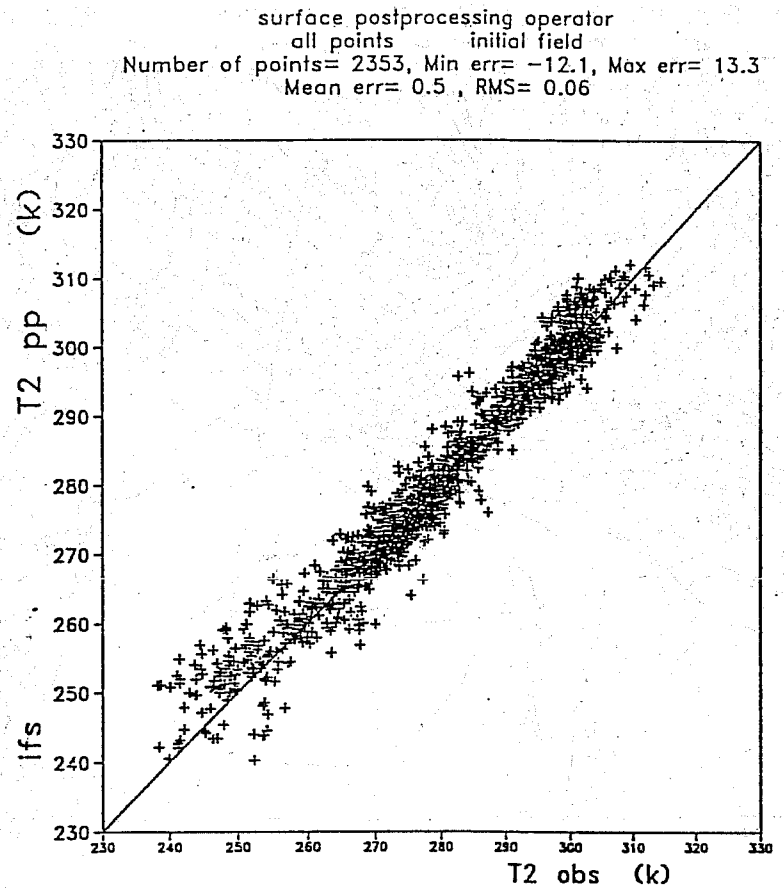
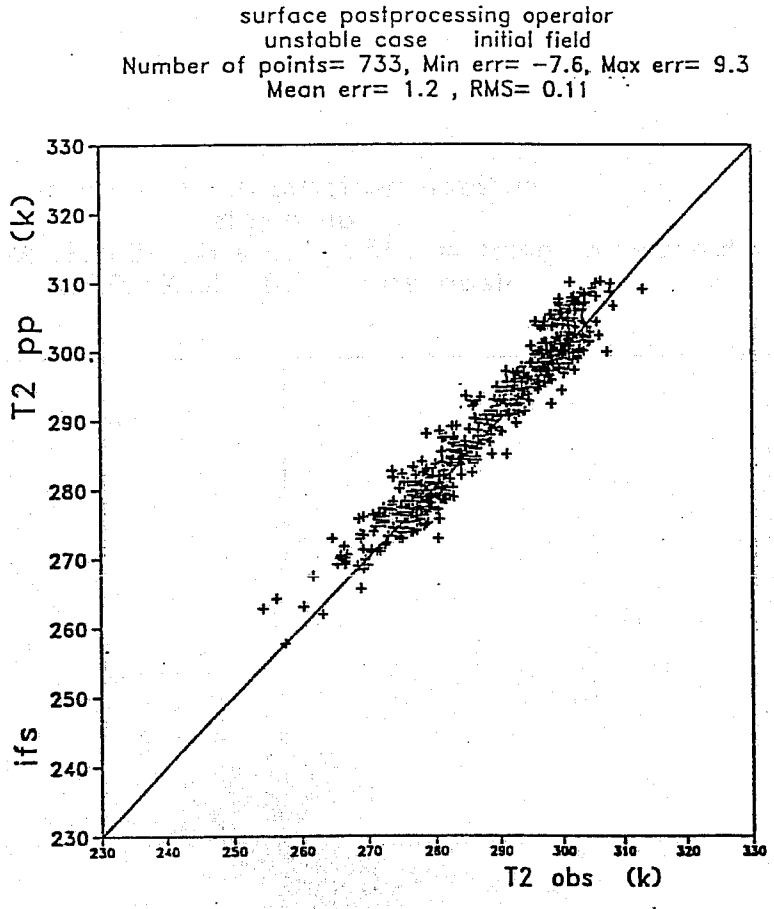


Fig.23: The post-processed 2m temperature ($T2_{pp}$) is plotted against the observed 2m temperature ($T2_{obs}$) for all the observed values at the initial point (top panel) and the final point (bottom panel) of the minimisation.



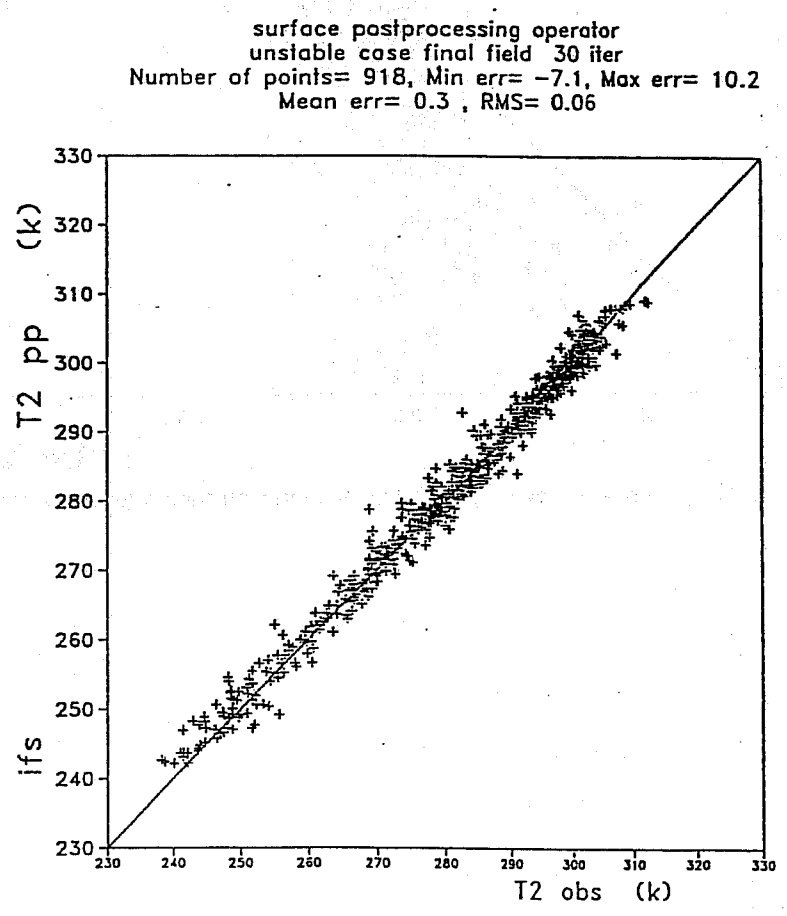


Fig. 24: The post-processed 2m temperature ($T2_{pp}$) is plotted against the observed 2m temperature ($T2_{obs}$) for the observed values in the model unstable conditions at the initial point (top panel) and the final point (bottom panel) of the minimisation.

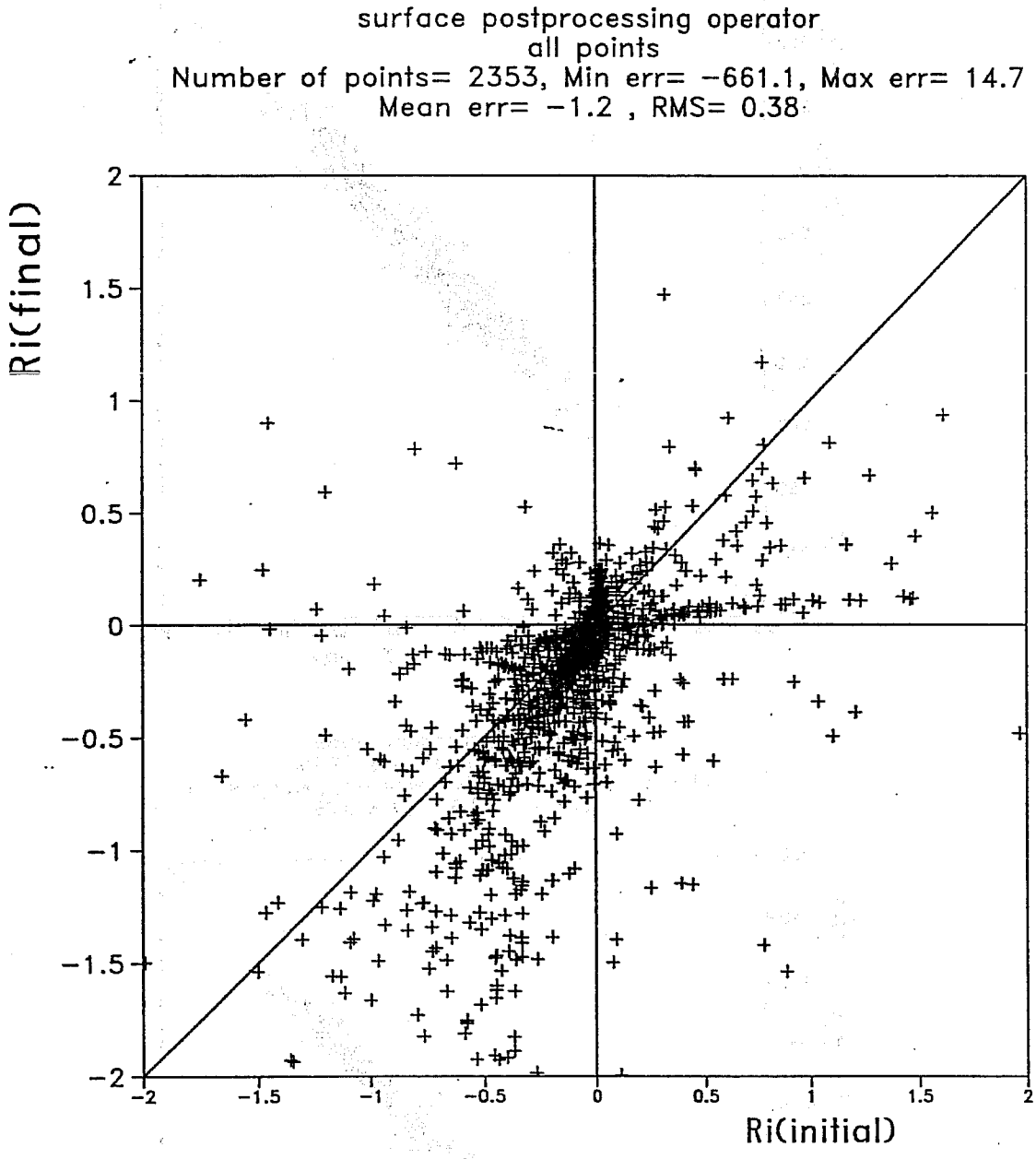
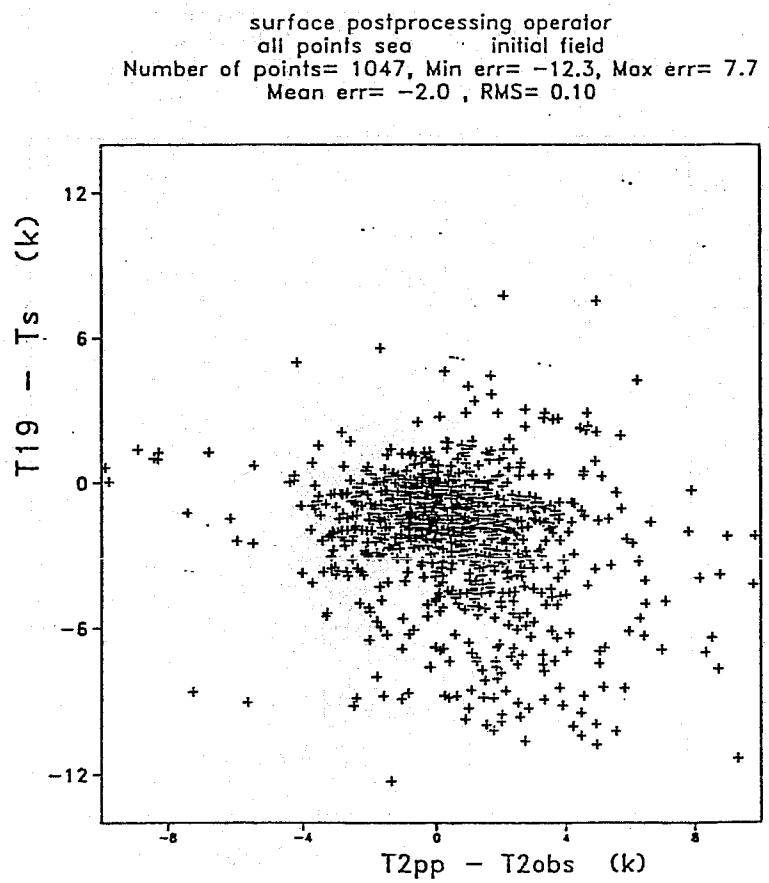


Fig. 25: The final against the initial bulk Richardson number for all the observed values.



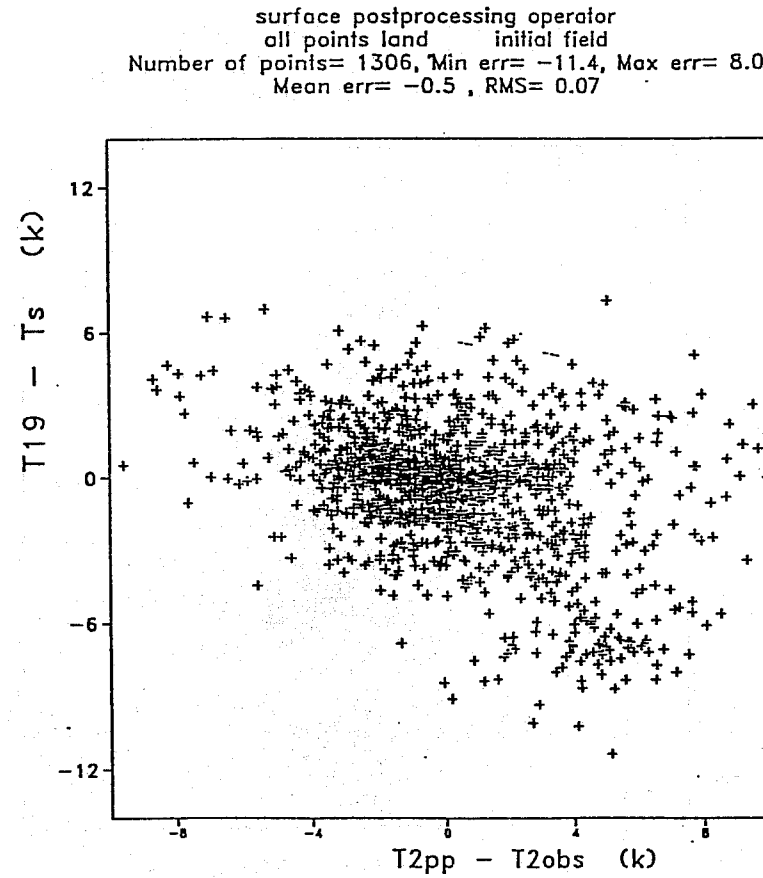
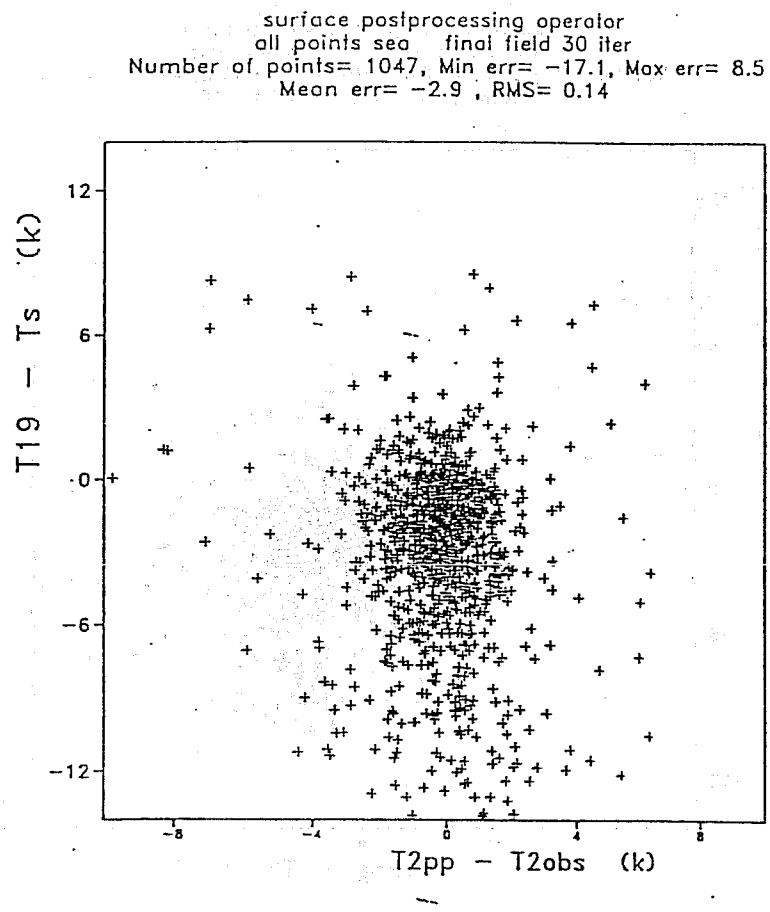


Fig. 26: The differences between the post-processed and the observed temperature ($T2_{pp}$-$T2_{obs}$) are plotted versus the differences between the 19th model level and the surface temperature ($T19$-T_s) for all points over sea (top panel) and all points over land (bottom panel) at the beginning of the minimisation.



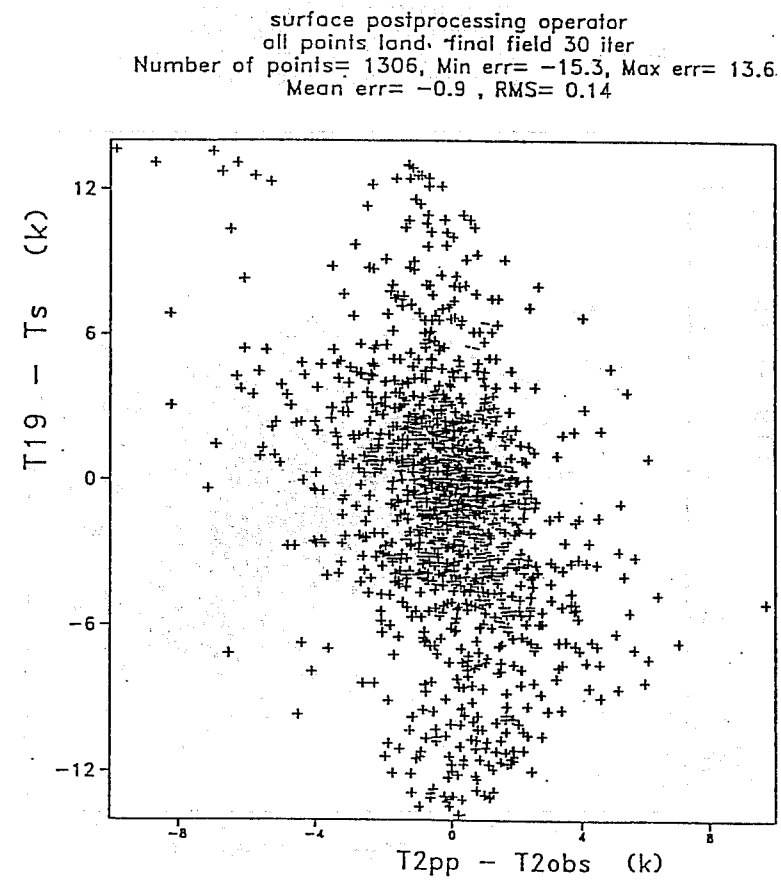


Fig. 27: The differences between the post-processed and the observed temperature ($T2_{pp}$-$T2_{obs}$) are plotted versus the differences between the 19th model level and the surface temperature ($T19$-T_S) for all points over sea (top panel) and all points over land (bottom panel) at the end of the minimisation.

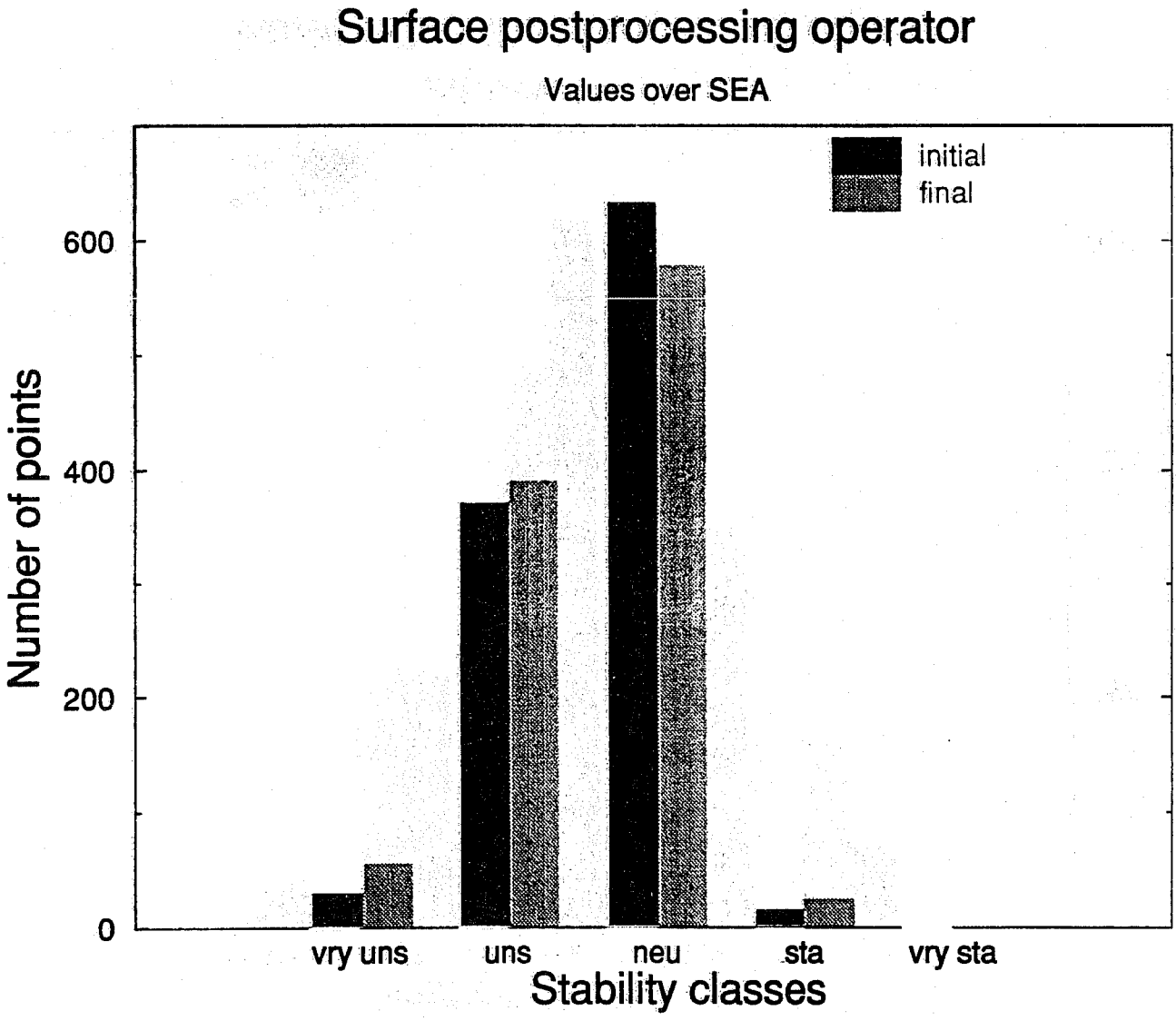


Fig. 28: Sea points distribution across five stability classes before (black) and after (white) the minimisation.

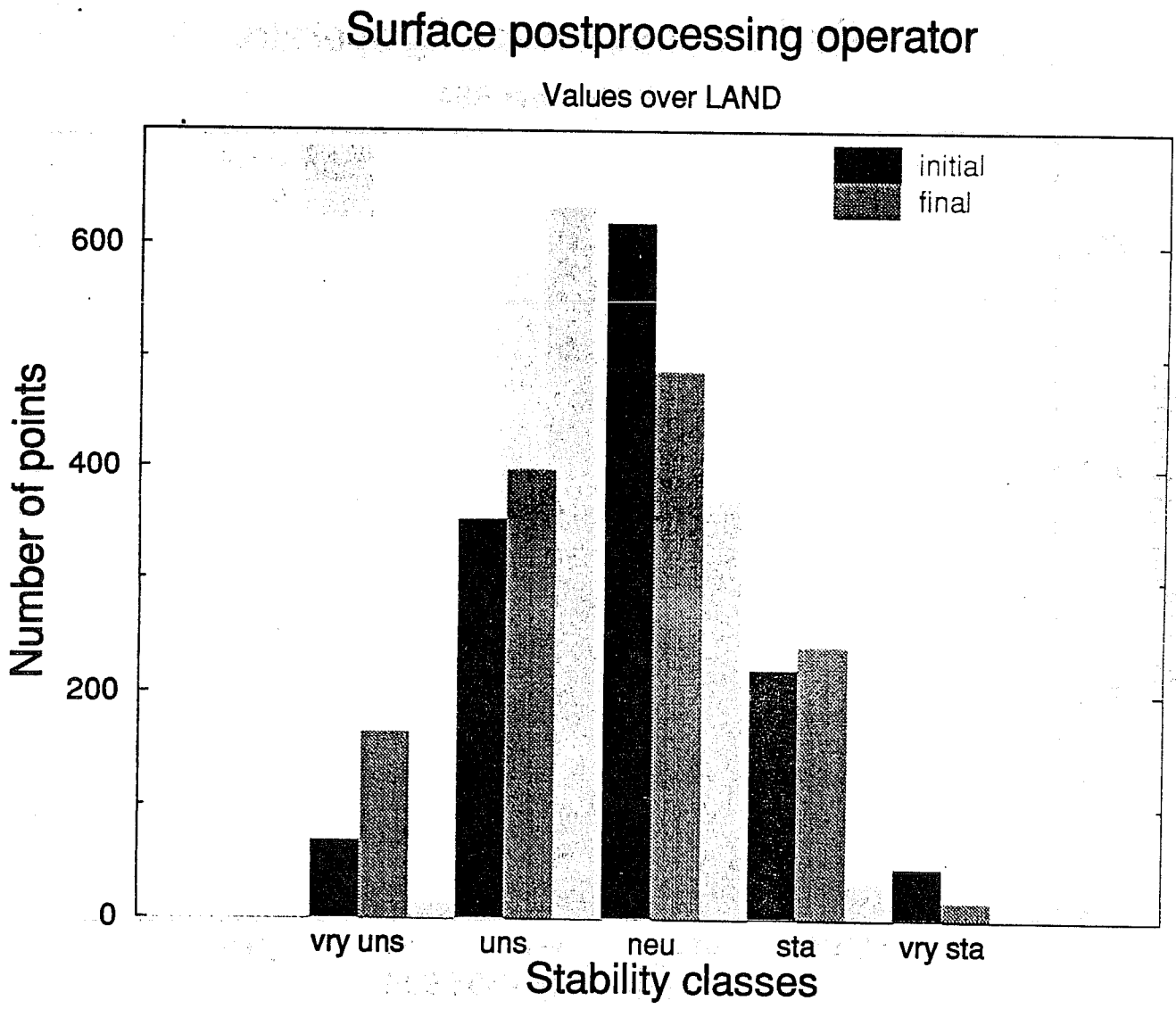


Fig. 29: Land points distribution across five stability classes before (black) and after (white) the minimisation.

distribution across the stability range before and after the minimisation. There is a displacement towards the unstable and very unstable cases more evident over land than over sea, suggesting again the necessity of having T_S included as control variable.

4. CONCLUSION

The ECMWF 3D variational assimilation of conventional observations has been developed and the presented preliminary results show that it is comparable with the current ECMWF OI analysis. In spite of a number of differences noticed, particularly in temperature increment maps, the results are rather encouraging. It is shown that the near surface observations can be used in the 3D variational analysis in the consistent way with the model's surface parametrization definitions. The results have also shown that, in order to make a better use of the near surface observations, it is necessary to extend the control variable by including surface parameters.

5. ACKNOWLEDGEMENTS

We are very grateful to the ECMWF's IFS team as well as P. Viterbo and A. Beljaars for their help and suggestions given. We also wish to thank A. Hollingsworth for his helpful comments.

References

Andersson, E.,J.-N. Thepaut, J. Eyre, A.P. McNally, P. Courtier and J. Pailleux, 1992: Use of radiances in 3D/4D variational assimilation. This volume.

Blondin, C., 1991: Parametrization of land-surface processes in numerical weather prediction. Land surface evaporation: measurement and parametrization, T.J. Schmugge and J.-C. André, Eds., Springer-Verlag, 31-54.

Geleyn, J.F., 1988: Interpolation of wind, temperature and humidity values from the model levels to the height of measurement. Tellus, 40, 347-351.

Heckley, W. A., P. Courtier, J. Pailleux and E. Andersson, 1992: On the use of background information in the variational analysis at ECMWF. This volume.

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

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

Lorenc, A. C., 1988: Optimal non-linear objective analysis. Q.J.R. Meteorol. Soc., 114, 205-240.

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

Pailleux, J., W. Heckley, D. Vasiljević, J.-N. Thépaut, F. Rabier, C. Cardinali and E. Andersson, 1991: Development of a variational assimilation system. Technical Memorandum No. 179, ECMWF.

Thépaut, J.-N., P. Courtier and R. Hoffman, 1992a: Use of dynamical information in four-dimensional variational assimilation. This volume.

APPENDIX

A1. UPPER AIR OBSERVATION OPERATOR

A brief description of the vertical coordinate used is given, and then the vertical interpolation techniques employed for each parameter (height, wind, humidity and temperature) are described separately.

A1.1 Vertical coordinate

The forecast model's vertical coordinate is a hybrid η coordinate which combines a sigma coordinate at lower levels with a pressure coordinate at higher levels. The hybrid coordinate is defined as:

$$p_{k+1/2} = a_{k+1/2} + b_{k+1/2} p_s; \quad 0 \le k \le NLEV$$

and

$$\eta_{k+1/2} = \frac{a_{k+1/2}}{p_0} + b_{k+1/2}$$

where $a_{k+1/2}, b_{k+1/2}, p_0$ are constants and k stands for the full model levels with levels ordered from the top to the bottom. The vertical structure can be schematically represented as:

$\eta_{k-1/2}$ ——————————— $\phi_{k-1/2}$

η_k ..... u_k, v_k, T_k, q_k

$\eta_{k+1/2}$ ——————————— $\phi_{k+1/2}$

The pressure gradient term of the dynamics implies the following definition of the pressure at full model levels (p_k):

$$\ln p_k = \frac{p_{k+1/2} \ln p_{k+1/2} - p_{k-1/2} \ln p_{k-1/2}}{p_{k+1/2} - p_{k-1/2}} - 1$$

A1.2 Height

(a) *above the model's orography*

The heights are computed by integrating the hydrostatic equation exactly using the ICAO temperature profile and interpolating only the difference between the model level heights

and the ICAO heights. The interpolation is a linear in ln(p) up to the second model level and a quadratic interpolation in ln(p) for levels above it.

The ICAO temperature profile is defined as:

$$T_{ICAO} = T_0 - \Lambda z \tag{A1.2.1}$$

where T_0 is 288K, z is the height above 1013.25hPa and Λ is 0.0065K/m. Using this temperature profile and integrating the hydrostatic equation provides T as a function of pressure:

$$T_{ICAO} = T_0 \left(\frac{p}{p_0} \right)^{\alpha} \tag{A1.2.2}$$

where $\alpha = \frac{\Lambda R}{g}$ and p_0 is 1013.25hPa. Combining the last two equations gives the ICAO height profile as a function of pressure:

$$z_{ICAO} = \frac{T_0}{\Lambda} - \frac{T_0}{\Lambda} \left(\frac{p}{p_0} \right)^{\alpha} \tag{A1.2.3}$$

with the ICAO height at the surface:

$$z_{s_{ICAO}} = \frac{T_0}{\Lambda} - \frac{T_0}{\Lambda} \left(\frac{p_s}{p_0} \right)^{\alpha} \tag{A1.2.4}$$

The ICAO heights above sea level using the ICAO temperature profile are:

$$z = z_{ICAO} - z_{s_{ICAO}} + z_s + \Delta z \tag{A1.2.5}$$

where z_S is the height of the model orography. Δz is the interpolated geopotential departure from ICAO to the observed pressure levels. In order to find Δz , Δz_k (geopotential departure at the model's full level) is computed:

$$\Delta z_k = \sum_{j=NLEV}^{k+1} \frac{R_d}{g} \left(T_{v_j - T_{ICAO}} \right) \ln \left(\frac{p_{j+1/2}}{p_{j-1/2}} \right) + \beta_k \frac{R_d}{g} \left(T_{v_k} - T_{ICAO_k} \right) \tag{A1.2.6}$$

where:

$$\beta_k = 1 - \frac{p_{k-1/2}}{p_{k+1/2} - p_{k-1/2}} \ln\left(\frac{p_{k+1/2}}{p_{k-1/2}}\right) \equiv \ln\left(\frac{p_{k+1/2}}{p_k}\right) \qquad \text{(A1.2.7)}$$

for $k>1$ and $\beta_1 = \ln 2$. The first term on the right hand side of equation (A1.2.6) integrates the hydrostatic equation to half levels and the second term adds the contribution form $p_{k+1/2}$ to p_k

The ICAO stratosphere is defined where the temperature has decreased to T_{St}=216.5K and above which it is kept constant. Thus, z_{ICAO} is integrated to the tropopause with a constant temperature gradient and above the tropopause the integration is done using the constant temperature T_{St}.

Above the second full level of the model, the linear interpolation is replaced by a quadratic interpolation in ln(p):

$$z(\ln p) = a + b(\ln p) + c(\ln p)^2 \qquad \text{(A1.2.8)}$$

where a, b and c are constants determined so that the above equation fits the heights at the three top levels (k=1,2,3). The interpolation formula is:

$$z(\ln p) = z_2 + \frac{(z_2 - z_1)(\ln p - \ln p_2)(\ln p - \ln p_3)}{(\ln p_2 - \ln p_1)(\ln p_1 - \ln p_3)} - \frac{(z_2 - z_3)(\ln p - \ln p_1)(\ln p - \ln p_2)}{(\ln p_2 - \ln p_3)(\ln p_1 - \ln p_3)} \qquad \text{(A1.2.9)}$$

where 1,2 and 3 refer to k=1,2 and 3.

(b) *below the model's orography*

The extrapolation of the geopotential below the model's orography is carried out as follows:

i) Find T_* (surface temperature) by assuming constant lapse rate :

$$T_* = T_{NLEV} + \Lambda \frac{R_d}{g} T_{NLEV} \ln\left(\frac{p_s}{P_{NLEV}}\right) \qquad \text{(A1.2.10)}$$

and

$$T_* = \left(T_* + \max\left(T_y, \min\left(T_x, T_*\right)\right)\right) / 2 \qquad \text{(A1.2.11)}$$

ii) Find the temperature at mean sea level T_0:

$$T_0 = T_* + \Lambda \frac{\phi_s}{g} \tag{A1.2.12}$$

and

$$T_0 = \min\left(T_0, \max\left(T_x, T_*\right)\right) \tag{A1.2.13}$$

where T_x is 290.5K and T_y is 255K.

iii) The geopotential is calculated as:

$$\phi = \phi_s - \frac{R_d T_*}{\gamma}\left(\left(\frac{p}{p_s}\right)^{\gamma} - 1\right) \tag{A1.2.14}$$

where $\gamma = \frac{R_d}{\phi_s}\left(T_0 - T_s\right)$.

A1.3 Wind

A linear interpolation in $\ln(p)$ is used to interpolate u and v to the all observed pressure levels up to the second full model level and a quadratic interpolation in $\ln(p)$ is used above it. Below the lowest model level wind components are assumed to be constant and equal to the values of the lowest model level.

A1.4 Humidity

The specific humidity (q), the saturation specific humidity (q_S) and the relative humidity (RH) are linearly interpolated in p.

q_S is given by

$$q_s = \frac{\frac{R_d}{R_v}\frac{e_s(T)}{p}}{1 - \left(\frac{R_v}{R_d} - 1\right)\frac{R_d}{R_v}\frac{e_s(T)}{p}} \tag{A1.4.1}$$

The saturation vapour pressure $e_s(T)$ is defined:

$$e_s(T) = c_1 \exp\left(\frac{c_2(T - T_0)}{T - c_3}\right) \qquad \text{(A1.4.2)}$$

where T_o is 273.16K, c_1=610.78 (17.269 for TT_o), c_2=21.875 (35.86 for TT_o) and c_3=7.66.

The relative humidity RH is first calculated at the model's full levels:

$$RH = \frac{pQ\frac{R_v}{R_d}}{\left\{1 + \left(\frac{R_v}{R_d} - 1\right)Q\right\}e_s(T)} \qquad \text{(A1.4.3)}$$

and then interpolated to a required pressure level.

Below the lowest and above the top model full levels q, q_S and RH are assumed to be constant and equal to the values of the lowest and the top model's full levels, respectively.

A1.5 Temperature

The linear interpolation in p is used to interpolate the temperature to the all pressure levels. Above the highest model level the temperature is kept constant and equal to the value of the highest model level. Below the lowest model level temperature is extrapolated linearly with respect to pressure using temperature values at the lowest and one above model's full levels.

A2. SURFACE POSTPROCESSING OPERATOR

A2.1 Mathematical formulation

The analytical technique (*Geleyn*, 1988) is used to interpolate values between the lowest model level and the surface. When MO theory is applied:

$$\frac{\partial u}{\partial z} = \frac{u_*}{k(z+z_o)}\phi_M\left(\frac{z+z_o}{L}\right) \tag{A2.1.1}$$

$$\frac{\partial s}{\partial z} = \frac{s_*}{k(z+z_o)}\phi_H\left(\frac{z+z_o}{L}\right) \tag{A2.1.2}$$

$$L = \frac{C_p}{g}\frac{T}{k}\frac{u_*^2}{s_*} \tag{A2.1.3}$$

where

u, s:	wind and energy variables
u^*, s^*:	friction values
k = 0.4:	Von Karman's constant

The temperature is linked to the dry static energy s by:

$$s = C_p T + \phi \tag{A2.1.4}$$

$$C_p = C_{pd}\left[1+\left(\frac{C_{pv}}{C_{pd}}-1\right)q\right] \tag{A2.1.5}$$

Defining the neutral surface exchange coefficient at the height z as:

$$C_N = \left[\frac{k}{\ln\left(\frac{z+z_o}{z_o}\right)}\right]^2 \tag{A2.1.6}$$

The drag and heat coefficients as:

$$C_M = \frac{u_*^2}{[u(z)]^2} \tag{A2.1.7}$$

$$C_H = \frac{u_* s_*}{u(z)[s(z) - \tilde{s}]} \tag{A2.1.8}$$

we can set the following quantities:

$$B_N = \frac{k}{\sqrt{C_N}}, B_M = \frac{k}{\sqrt{C_M}}, B_H = \frac{k\sqrt{C_M}}{C_H} \tag{A2.1.9}$$

and considering the stability function in stable conditions as:

$$\phi_{M/H} = 1 + \beta_{M/H}\frac{z}{L} \tag{A2.1.10}$$

we obtain integrating the (A2.1.1-2) from 0 to z_1 (lowest model level):

$$u(z) = \frac{u(z_1)}{B_M}\left[\ln\left(1 + \frac{z}{z_1}\left(e^{B_N} - 1\right)\right) - \frac{z}{z_1}(B_N - B_M)\right] \tag{A2.1.11}$$

$$s(z) = \tilde{s} + \frac{s(z_1) - \tilde{s}}{B_H}\left[\ln\left(1 + \frac{z}{z_1}\left(e^{B_N} - 1\right)\right) - \frac{z}{z_1}(B_N - B_H)\right] \tag{A2.1.12}$$

In unstable conditions the stability function can be expressed as:

$$\phi_{M/H} = \left(1 - \beta_{M/H}\frac{z}{L}\right)^{-1} \tag{A2.1.13}$$

and the vertical profiles for wind and dry static energy are:

$$u(z) = \frac{u(z_1)}{B_M}\left[\ln\left(1 + \frac{z}{z_1}\left(e^{B_N} - 1\right)\right) - \ln\left(1 + \frac{z}{z_1}\left(e^{B_N - B_M} - 1\right)\right)\right] \tag{A2.1.14}$$

$$s(z) = \tilde{s} + \frac{s(z_1) - \tilde{s}}{B_H}\left[\ln\left(1 + \frac{z}{z_1}\left(e^{B_N} - 1\right)\right) - \ln\left(1 + \frac{z}{z_1}\left(e^{B_N - B_M} - 1\right)\right)\right] \tag{A2.1.15}$$

The temperature can then be obtained from s as:

$$T(z) = \frac{s(z) - zg}{C_p} \tag{A2.1.16}$$

When z is set to the observation height, formulae (A2.1.11-12), (A2.1.14-16) give the postprocessed wind and temperature. To solve the problem, we have to compute the dry static energy at the surface $\tilde{s} = \tilde{s}(T_s, q = 0)$ (pa.A2.2) and B_M, B_N and B_H values depending on the drag and heat exchange coefficients (A2.1.9) (par. A2.3).

A2.2 Surface values

To determine the dry static energy at the surface we use (A2.1.4-5) where the humidity at the surface is defined by:

$$\tilde{q} = q(z = o) = h(C_{SN}, C_l, C_v) q_{sat}(T_s, p_s) \tag{A2.2.1}$$

h is given by (*Blondin*, 1991):

$$h = C_{SN} + (1 - C_{SN})\left[C_l + (1 - C_l)\bar{h}\right] \tag{A2.2.2}$$

with

$$\bar{h} = \max\left\{0.5\left(1 - \cos\frac{\pi W_s}{W_{scap}}\right), \min\left(1, \frac{q}{q_{sat}(T_s, p_s)}\right)\right\} \tag{A2.2.3}$$

The formula (A2.2.2) assigns a value of 1 to the "surface relative humidity over the snow covered and wet fraction of the grid box. The snow cover fraction C_{SN} depends on the snow amount S_n: $C_{SN} = \min(1, S_n / S_{ncr}), S_{ncr} = 0.015m$ is a critical value. The wet skin fraction C_l is derived from the skin reservoir water content (W_l): $C_l = \min(1, W_l / W_{lmx})$, where $W_{lmx} = W_{l\max}\{(1 - C_v) + C_v L_t\}$ with $W_{l\max} = 2\ 10^{-4} m$ being the maximum amount of water that can be held on one layer of leaf $(L_t = 4)$ or on bare soil and C_v is the vegetation ratio. W_{scap} is the field capacity (2/7 in volumetric units).

A2.3 The transfer coefficients over land

Comparing the equations (A2.1.1-2) integrated from z_o to $z + z_o$ with (A2.1.6-7), C_M and C_H can be analytically defined:

$$\frac{1}{C_M} = \frac{1}{k^2}\left[\int_{z_o}^{z_o + z} \frac{\phi_M\left(\frac{z'}{L}\right)}{z'} dz'\right]^2 \tag{A2.3.1}$$

$$\frac{1}{C_H} = \frac{1}{k^2}\left[\int_{z_o}^{z+z_o} \frac{\phi_M\left(\frac{z'}{L}\right)}{z'}dz' \int_{z_o}^{z+z_o} \frac{\phi_H\left(\frac{z'}{L}\right)}{z'}dz'\right] \tag{A2.3.1}$$

Because of the complicated form of the stability functions, the former integrals have been approximated by analytical expressions, formally given by:

$$C_M = C_N f_m\left(R_i, \frac{z}{z_o}\right)$$

$$C_H = C_N f_h\left(R_i, \frac{z}{z_o}\right) \tag{A2.3.2}$$

where C_N is given by (A2.1.5). The bulk Richardson number R_i is defined as:

$$R_i = \frac{g\Delta z \Delta\Phi_v}{C_p \Phi_v |\Delta \underline{u}|^2} \tag{A2.3.3}$$

where Φ_v is the virtual potential temperature. The functions f_m and f_h correspond to the model instability functions and have the correct behaviour near neutrality and in the cases of high stability (*Louis*, 1979)

(a) *unstable case* $R_i < 0$

$$f_h = 1 - \frac{3bR_i}{1 + 3bCC_N\sqrt{\left(1+\frac{z}{z_o}\right)(-R_i)}} \tag{A2.3.4}$$

$$f_m = 1 - \frac{2bR_i}{1 + 3bCC_N\sqrt{\left(1+\frac{z}{z_o}\right)(-R_i)}} \tag{A2.3.5}$$

C = 5

(b) *Stable case* $R_i > 0$

$$f_m = \frac{1}{1 + 2bR_i(1+dR_i)^{1/2}} \tag{A2.3.6}$$

$$f_h = \frac{1}{1+3bR_i(1+dR_i)^{1/2}} \qquad \text{(A2.3.7)}$$

d = 5

A2.4 2m Relative humidity

The interpolation formula for the relative humidity is expressed by:

$$RH = \frac{pq\dfrac{R_v}{R_d}}{\left[1+\left(\dfrac{R_v}{R_d}-1\right)q\right]e_s(T)} \qquad \text{(A2.4.1)}$$

where $e_s(T)$ is the saturation pressure (A1.4.2). The relative humidity at 2m depends on the specific humidity, temperature and pressure (q, T and p respectively) at the lowest model level: it is constant in the surface model layer.