LOCAL WEATHER ELEMENT GUIDANCE FROM THE ECMWF FORECASTING SYSTEM IN THE MEDIUM RANGE: A VERIFICATION STUDY

Horst Böttger
ECMWF

1. Introduction

It has been shown recently (Akesson et al, 1982, Pümpel, 1982, Johannessen, 1982) that the ECMWF model with its comprehensive physical parameterisation scheme, can give forecast guidance to the users of such fields as surface temperature, precipitation and cloud amount, which are directly related to weather forecasts. There is therefore a need to assess the model's accuracy in predicting these parameters. Furthermore, the traditional ways of presenting short range weather forecasts in terms of extremes, and attempts to give exact timing of events, might not be the best form and mode to be applied for the medium range, especially as long as the errors increase rapidly around day 5 of the forecast, this having a detrimental effect on the day to day reliability of forecasts. Verification of weather forecasts, either direct model output or after interpretation, can be used as a guide to the best formats for the presentation of medium range weather forecasts.

A comprehensive software package has been developed and implemented at ECMWF which allows the user, with a minimum of effort, to:

i. access a European archive and extract observations, and analysis and forecast fields;

ii. format the data for local verification and interpretation studies in an on-line data base;

iii. verify the direct model output against weather observations;

iv. develop prognostic equations for statistical interpretation of model output;

v. test the prognostic equations on independent data.

The system was designed to give the maximum choice to users both from the Member States and from within the Centre. Predictors and predictands can be selected, they can be averaged in time and space, and derived predictands and predictors can be created. The application of the software is limited to an ECMWF European archive which now contains data from March 1981 onwards. Table 1 summarises the data available from the European archive. The data sets are described in detail in ECMWF Meteorological Bulletin M3.3/1.

The main goal of the following verification studies and the interpretation studies by Klinker (1982) and Wilson (1982) was to demonstrate the feasibility and flexibility of the ECMWF verification/interpretation software package. Only short data samples were available from the European archive at the time these studies were carried out, and a cautious approach is needed when interpreting the results from the single season available. Nevertheless, it will be demonstrated that verification of forecasts of near surface weather parameters serves two purposes. It provides the modeller with valuable information about the model's performance close to the surface. He could use this information to assess the impact of model changes, e.g. changed orography, changes in physical parameterisation of boundary layer fluxes. Secondly, it provides guidance to the user, who might be a forecaster in the weather services or the eventual end user directly, on the model performance at his location and on the choice of the best model predictors for particular forecasts.

In this study, the temperature at 2m above the model surface (T at 2m) at three locations in Europe in summer and in winter is examined. The data sets are described in chapter 2. We consider, in chapter 3, the impact of ground conditions, such as snowcover and land-sea distribution, which are not represented satisfactorily in the model, and the influence of temporal and spatial averaging on the quality of the forecast. In chapter 4, we discuss briefly the best model predictors for wind forecasts at sea.

2. The data

The ECMWF European archive contains the analysis and forecast data in grid-point form. The resolution of the European archive is 1.5 * 1.5 degrees (latitude and longitude). The current model grid is 1.875 * 1.875 degrees, the archived data being interpolated linearly from the model grid. This results in some smoothing, which for certain fields, e.g. precipitation or soilwetness, might be undesirable.

European area: 33°N, 27°W, 73.5°N, 45°E
Latitude-longitude grid 1.5° x 1.5° mesh.

Values scaled and packed: variable number of bits per grid point.

Period covered: 10 March 1981 onwards for analyses and 1 March 1981 onwards for forecasts

Stored on standard pressure levels:

Surface
1000mb
850
700
500
400
300
250
200
150
100
70

Surface parameters
pressure
temperature
soil wetness
snowdepth
large-scale rain
convective rain
snowfall
boundary layer dissipation
sensible heat flux
latent heat flux
surface stress
surface net radiation
net radiation at the top
cloudcover
u,v wind components at 10 metres
temperature at 2 metres
dewpoint temperature at 2 metres

Upper air parameters:
u,v wind components
temperatures - not available for analyses
geopotential height
relative humidity
vertical velocity - both uninitialised and initialised available for analyses

Analysed upper air fields for 00Z, 06Z, 12Z, 18Z

First guess (6 hour forecast) surface fields for 00Z, 06Z, 12Z, 18Z

Forecasts for H+0 (initialised analysis for 12Z)
H+6 ...H + 72 (6 hour intervals)
H+84...H+240 (12 hour intervals)

Table 1 Data in the ECMWF European grid point archive, grid interval is 1.5 degrees latitude and longitude

The strong gradients along the model's coastlines will also be weakened. All the observations in the European archive Reports Data Base are the synoptic data as received via the WMO Global Telecommunications System, coding or other errors remaining unchanged. As a first check, any user of this observational data is advised to carry out a visual inspection of the data, e.g. by producing time plot series.

The verification of temperature T at 2m was carried out at three locations in Europe, Hannover (WMO index number 10338, 52N 10E, elevation 54m), Nancy (WMO index number 07180, 49N 6E, 217m), and Firenze (WMO index number 16170, 44N 11E, 38m). The data sets were stratified into a summer (15 May to 15 September 1981) and a winter (15 November 1981 to 15 March 1982) season. Around the location of each station, the temperature at 2m above the model surface and all standard pressure level temperatures up to 500mb at an array of 4 * 4 gridpoints was extracted. Additionally, the geopotential height up to 500mb was extracted to derive thickness values related to mean temperatures over deep atmospheric layers. The array of 16 gridpoints covers an area of approximately 600km * 450km in north-south and east-west direction at mid-latitudes, and was found to be sufficient for all trials with space averaging.

For verification of the model wind at sea the observations taken at Ocean Weather Ship (OWS) Lima (57N, 20W) were extracted from the archive. As ocean weather ships tend to move around their nominal position, deviations of two degrees from their fixed location in each direction were permitted. Otherwise, the report was discarded. The u- and v-wind component fields at 10m above the model's surface was extracted at an array of 4 * 4 gridpoints, at 1000mb and at 850mb for the period 15 January to 25 May 1982.

The 2m temperature and the 10m wind are neither analysed quantities nor model parameters. They are derived in the model's postprocessing by interpolation between the lowest model level and the surface. For further details see e.g. Louis, 1982.

Apart from the "visual" inspection of the observational data, no further quality control was carried out. It would, of course, be desirable to replace the synoptic observations by "clean" and checked climatological data.

3. Local temperature verification

3.1 The problem

The direct model output of predicted parameters are valid at the original model gridpoints only. They can be considered to be representative of a model grid square (1.875^o * 1.875^o). Arguments can be made against any grid interpolation or interpolation to the observing site, because the direct model information will, as a consequence, be diffused and lose precision. If the direct model value is left untouched, the statistical verification and interpretation technique will implicitly take into account the displacement from the nearest gridpoint, the local climatic effects, deficiencies in the representation of orography and coastlines. On the other hand, it is not, in general, feasible for the operational user of numerical forecasts to work with data at the original model grid. Changes in numerical formulations might result in additional and/or displaced gridpoints and it is therefore preferable to work from a standard grid, which in our case is the grid in the ECMWF European archive. For medium range forecasts, lack of precision arising from interpolation errors will probably be unimportant. Operational field verification of the ECMWF forecasts indicates a decline in the reliability in phase and amplitude of the synoptic scale features between day 4 and day 6 of the forecast. One way to partially overcome these forecast errors is through averaging in time and space. This results in a loss of information with respect to extreme values and timing, but if it increases reliability, it can be regarded as a first step in the improvement of the use of medium range weather forecasts.

Another way to achieve a reduction in the error variability of temperature forecasts is to select predictors from higher levels in the atmosphere which are less affected by the formulation of the physical parameterisation processes in the boundary layer. Both the impact of averaging and the selection of temperature predictors from various model levels will be discussed in the following sections.

3.2 Daily mean temperature

Hannover is situated in homogeneous and flat terrain in the centre of northern Germany, approximately 150km from the coast. Maritime influence can be expected to be small. Before looking at the forecast, it is worthwhile to see how the model T at 2m in the analysis fits the observation. The field T at 2m is not analysed but evolves throughout the data assimilation cycle within the 6-hour forecasts and it is corrected every 6 hours by analysed thickness values for the lowest pressure levels. Fig. 1 shows the observed and analysed temperatures for Hannover in winter 1981/82. The observed temperature is for 18Z only which we found to be reasonably close to the daily mean, but, during the passage of a front significant deviations from this value occur. However, the analysis of T at 2m fits the observations quite well apart from cold spells, when the observed temperature is close to freezing point or below.

Figures 2 to 5 give the corresponding temperature curves for forecast days 1, 3, 5 and 7 in winter 1981/82. The striking features detectable in these figures are a growing positive bias, the decreasing quality and the growing error variability. In the early stage of the forecast, up to day 5, the warm and cold spells are predicted and clearly distinguished in the meteograms. During the cold periods the model forecast does not capture the intensity of the severe weather conditions; note that the last cold spell in February 1982 is missed completely by day 7. This model error becomes apparent as early as on day 1 and even in the analysis. As expected, the errors grow with forecast time, but even at day 7 (Fig. 5), the model still gives a good indication of the cold spell in December, the subsequent warming and the following rapid drop in temperature, although this significant change was only predicted with a phase lag of one day.

It is still not fully clear why the model cannot predict extreme low temperatures. This may be associated with deficiencies in the representation of observed surface conditions in the model. Fig. 6 gives the observed and model snowdepth over Europe for 10 December 1981 and 17 December 1981. This period falls within the first cold winter spell observed at Hannover. The increased snowcover over Europe during these seven days is not simulated in the ECMWF model. Snowdepth is currently not analysed and evolves throughout the data assimilation cycle in the 6 forecasts used to produce the first guess for the analysis. This explains the poor fit at low temperatures (T at 2m) in the analysis in Fig. 1. A forecast starting from erroneous ground conditions will consequently result in further forecast errors. With the ground being snow-free and too warm, snow falling during the forecast will tend to be melted too quickly.

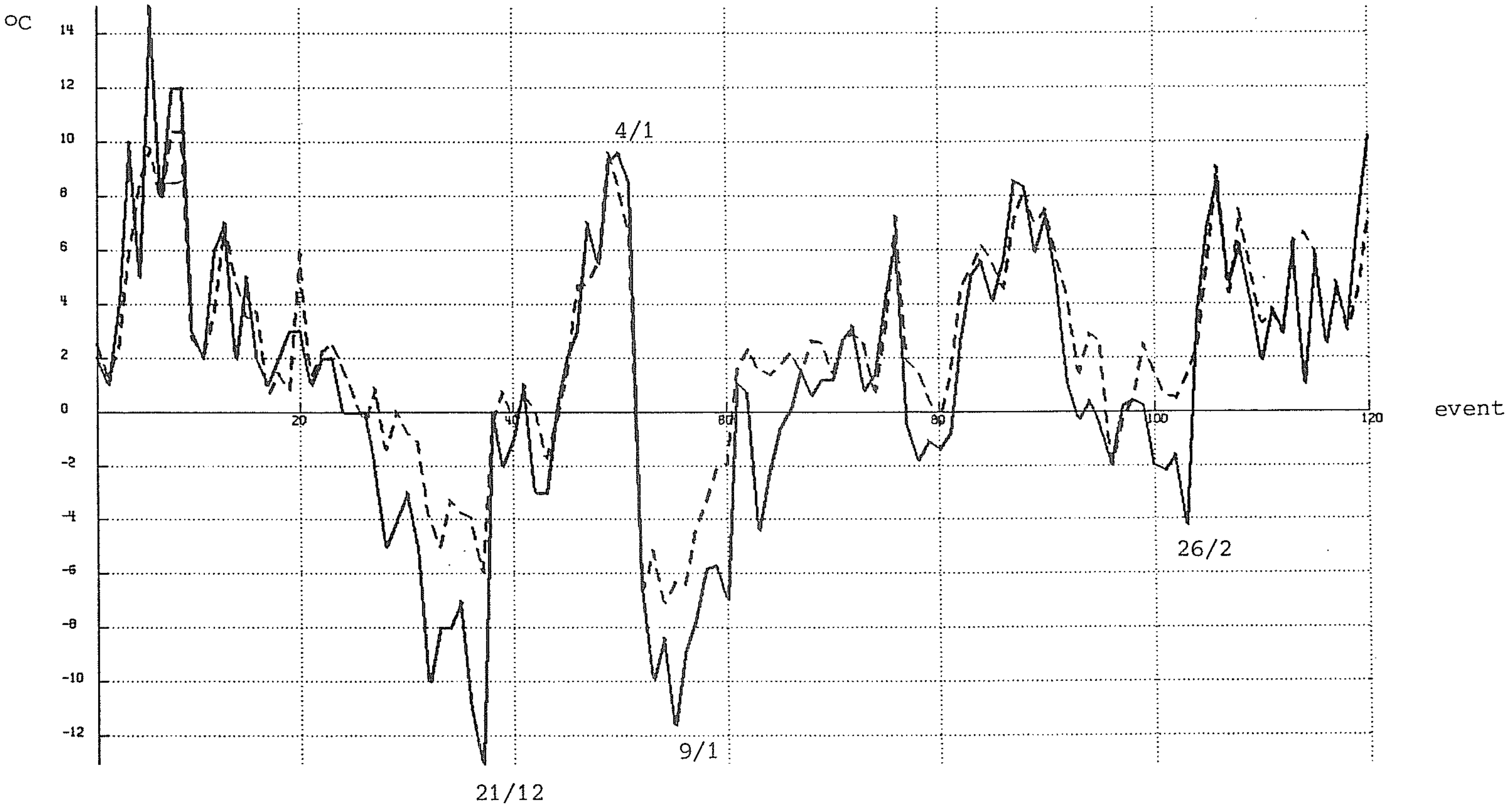

Fig. 1 Temperature at 2m above the surface, observed (full line) and as represented in the ECMWF analysis (broken line) at Hannover, Germany, for each observation (event) received between 15 November 1981 to 15 March 1982. The ECMWF model output is valid at 12Z, the observation at 18Z.

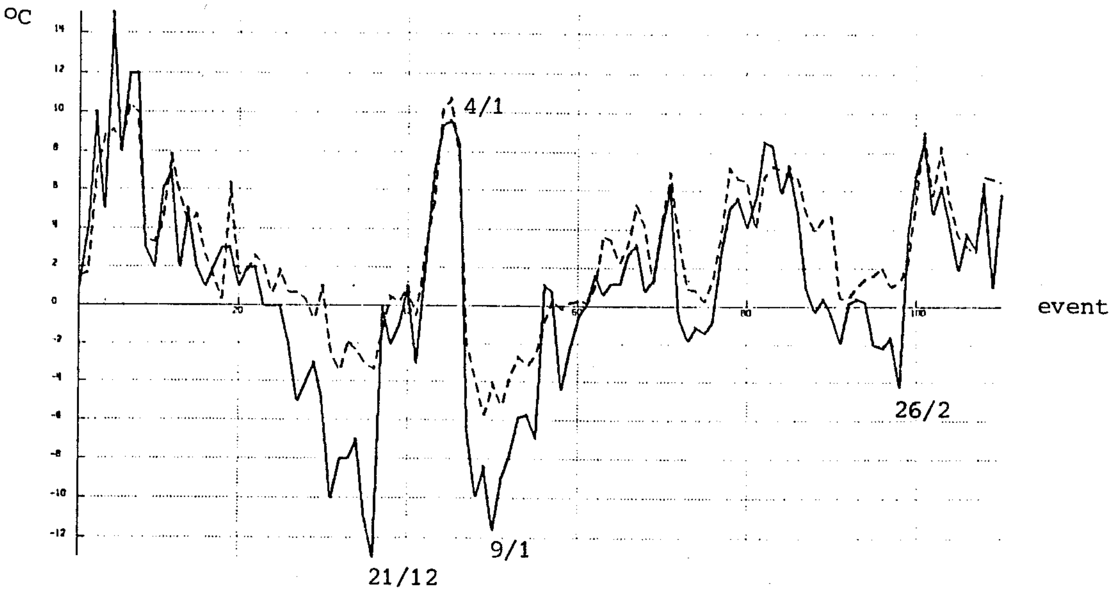

Fig. 2 As Fig. 1 but the broken line is here the 24 hour forecast valid at 12Z.

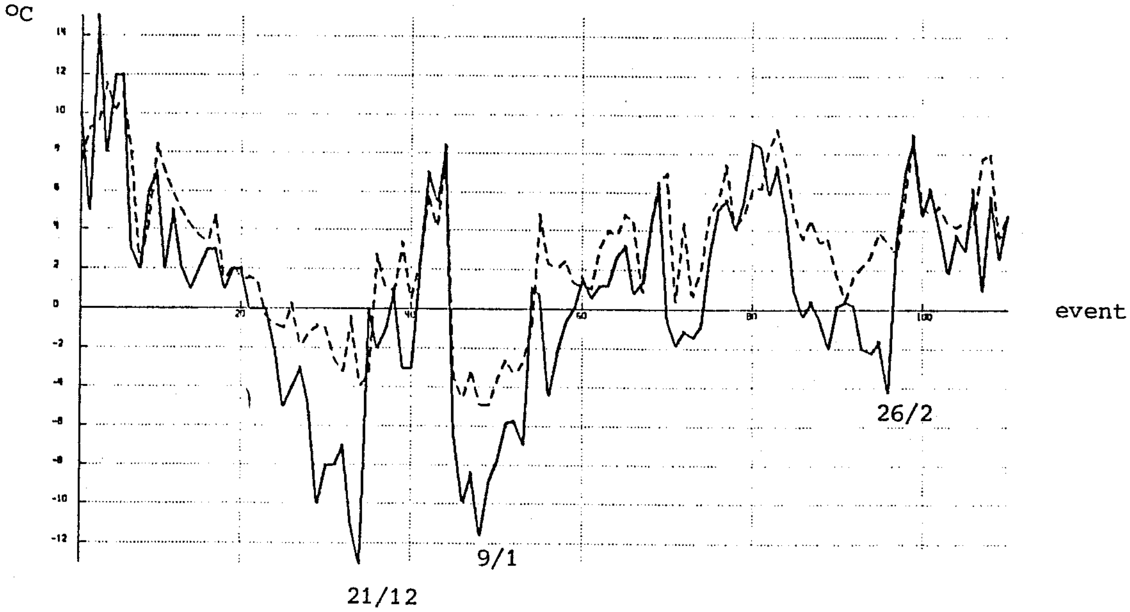

Fig. 3 As Fig. 1 but the broken line is here the 72 hour forecast.

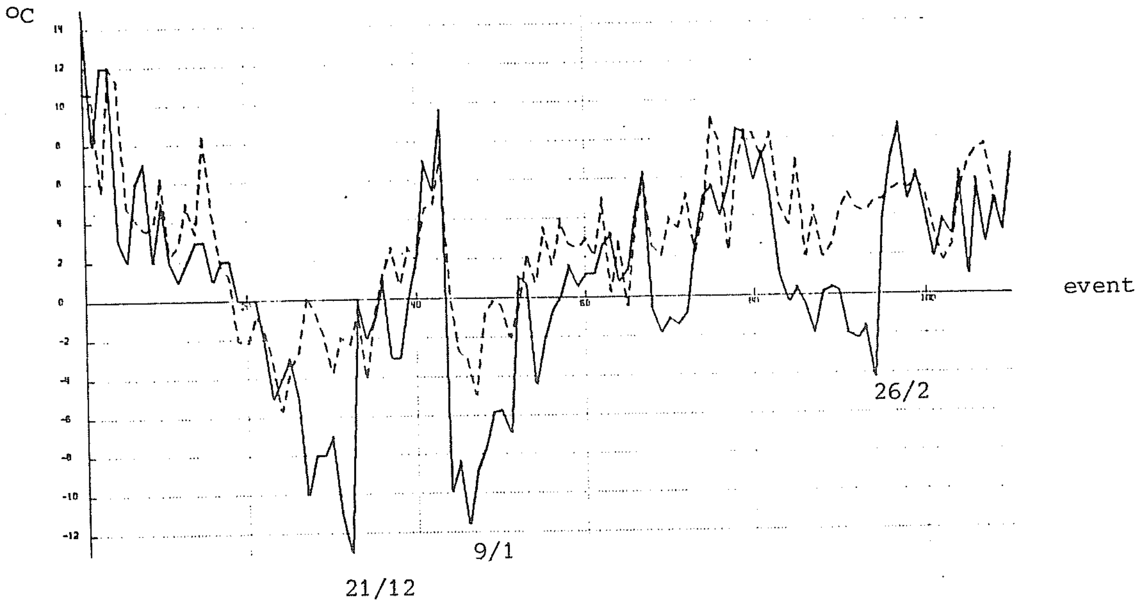

Fig. 4 As Fig. 1 but the broken line is here the 120 hour forecast.

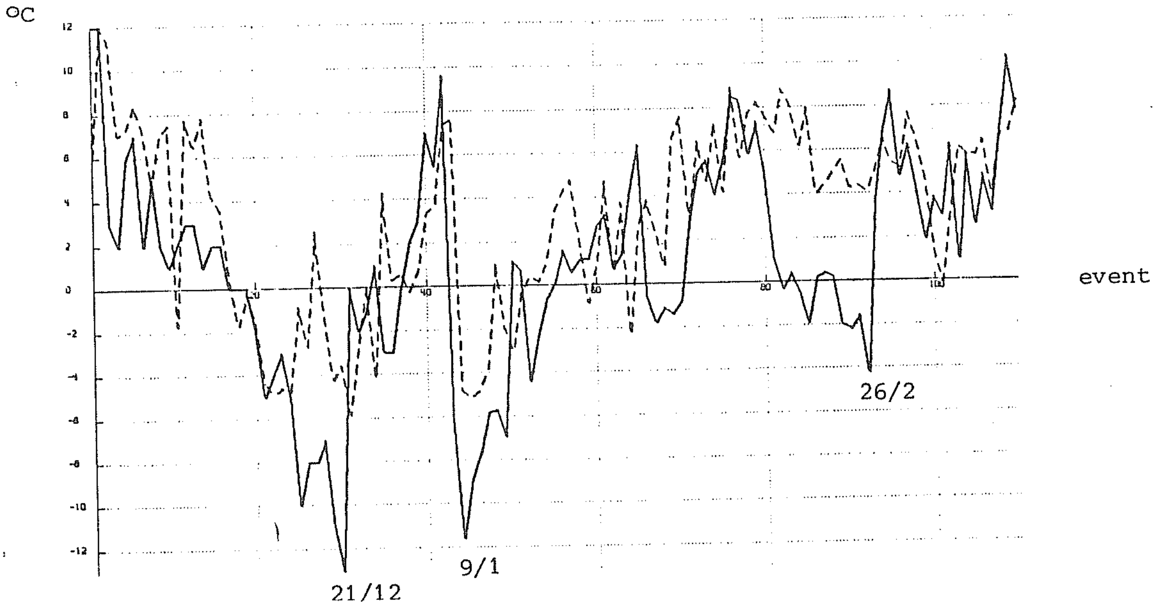

Fig. 5 As Fig. 1 but the broken line is here the 168 hour forecast.

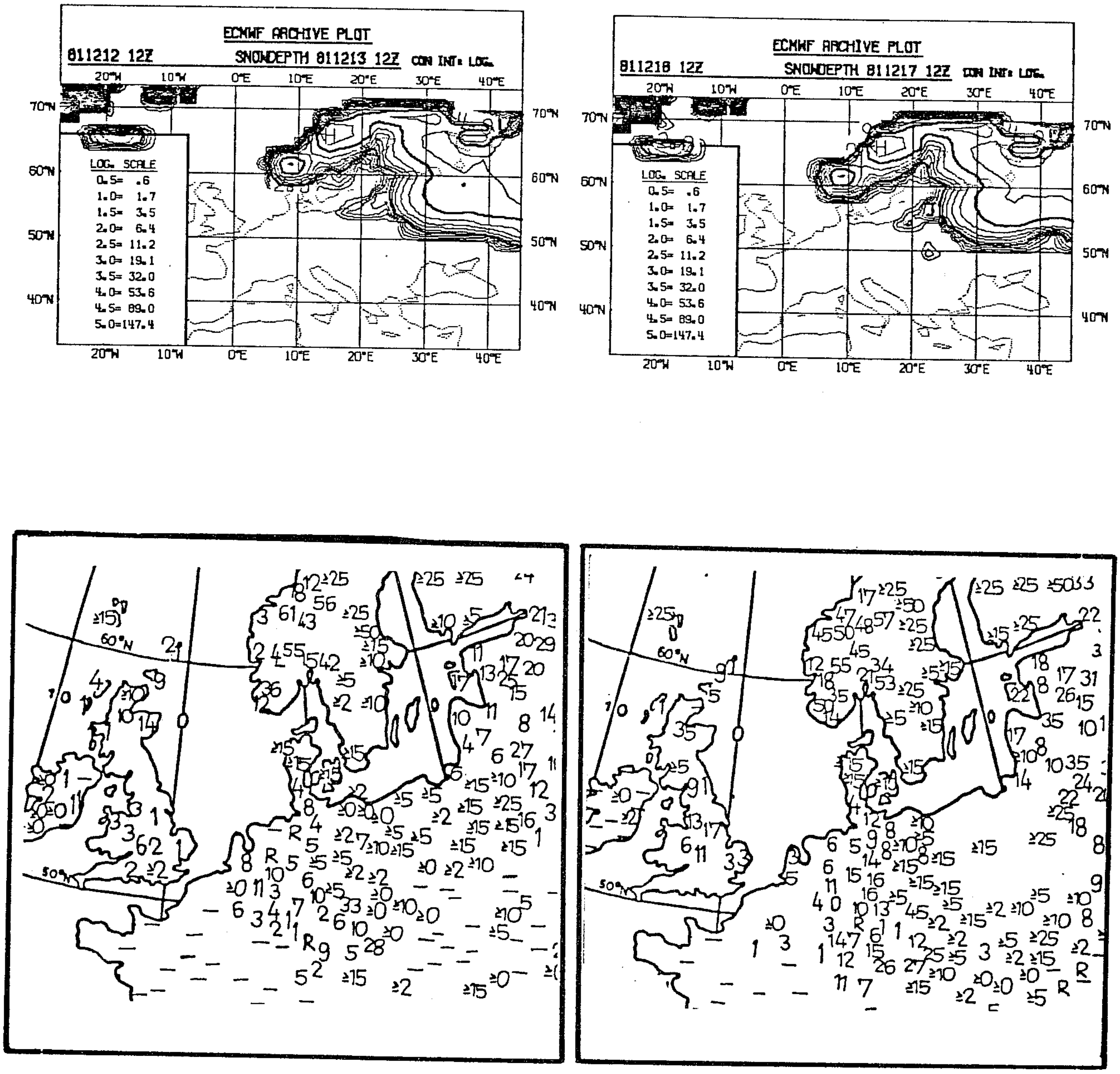

Fig. 6 Snowdepth over Europe as represented in the 24 hour ECMWF forecast, top, and observed values (after Berliner Wetterkarte, Freie Universität Berlin), bottom, for 10 December 1981, left, and 17 December 1981, right. Units of the observations are m^{-1}. The isolines in the forecast field are on a logarithmic scale, the conversion to m^{-1} given in each map.

Temperature forecast errors probably associated with these deficiencies in the representation of the ground conditions in the model become even more obvious at Nancy in France. Fig. 7 shows the two corresponding temperature curves taken from the analysis field and the observations at 18Z for winter 1981/82. As for Hannover, the analysed field fits the observations on most days. The rapid changes between warm and cold days in mid-December are well captured, a change which, however, is not predicted. The low temperatures in the middle of January 1982 were completely missed both in the analyses and forecasts. In the day 5 prediction of T at 2m (Fig. 8) forecast errors exceeding 10°C occurred during this period and the temperature in the late winter cold spell was underestimated, a result very similar to that obtained for Hannover. Fig. 9 demonstrates the complete lack of snow in the ECMWF model on 14 January 1982, whilst the whole of central Europe, including the Nancy area, was covered in deep snow.

It can be expected that these large errors in the representation of ground conditions will not only increase the forecast errors for near surface parameters but will affect the large scale flow pattern and produce systematic errors in the free atmosphere.

3.3 3-day mean temperature

For the same winter periods, 15 November 1981 to 15 March 1982, the surface temperature at Hannover has been averaged over three consecutive days both for the observations and within the forecast. The average over the observations makes use of all available time steps, 00, 06, 12 and 18Z. The verification of the 3-day mean temperature centred on forecast days 3, 5 and 7 is shown in Figures 10, 11 and 12. The graphs may be compared to the corresponding daily temperature forecasts in Figures 3, 4 and 5. One should remember that the x-axis gives the number of events for clean data, e.g. event number 60 of the daily temperature forecast will not necessarily correspond directly to the 3-day mean forecast valid on the same day. Missing observations will cause differences between the two graphs.

Time averaging has had a smoothing effect on the temperature graphs which is more obvious for the observations than for the forecasts. The temperature range, after averaging though, remains almost the same. The reduction in

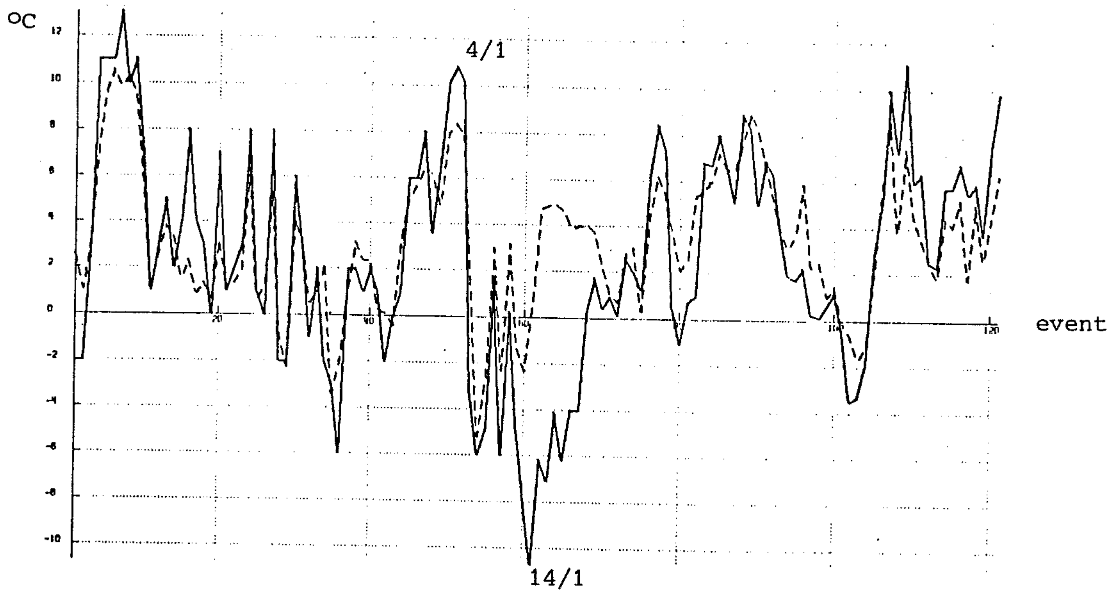

Fig. 7 As Fig. 1 but for Nancy (France)

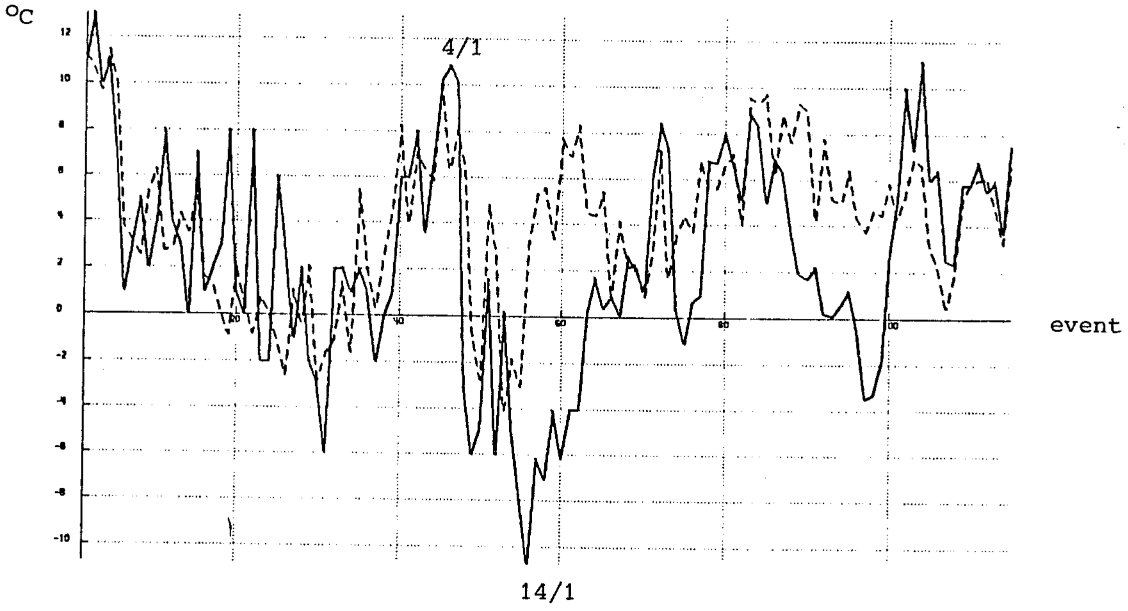

Fig. 8 As Fig. 2 but for Nancy (120 hour forecast)

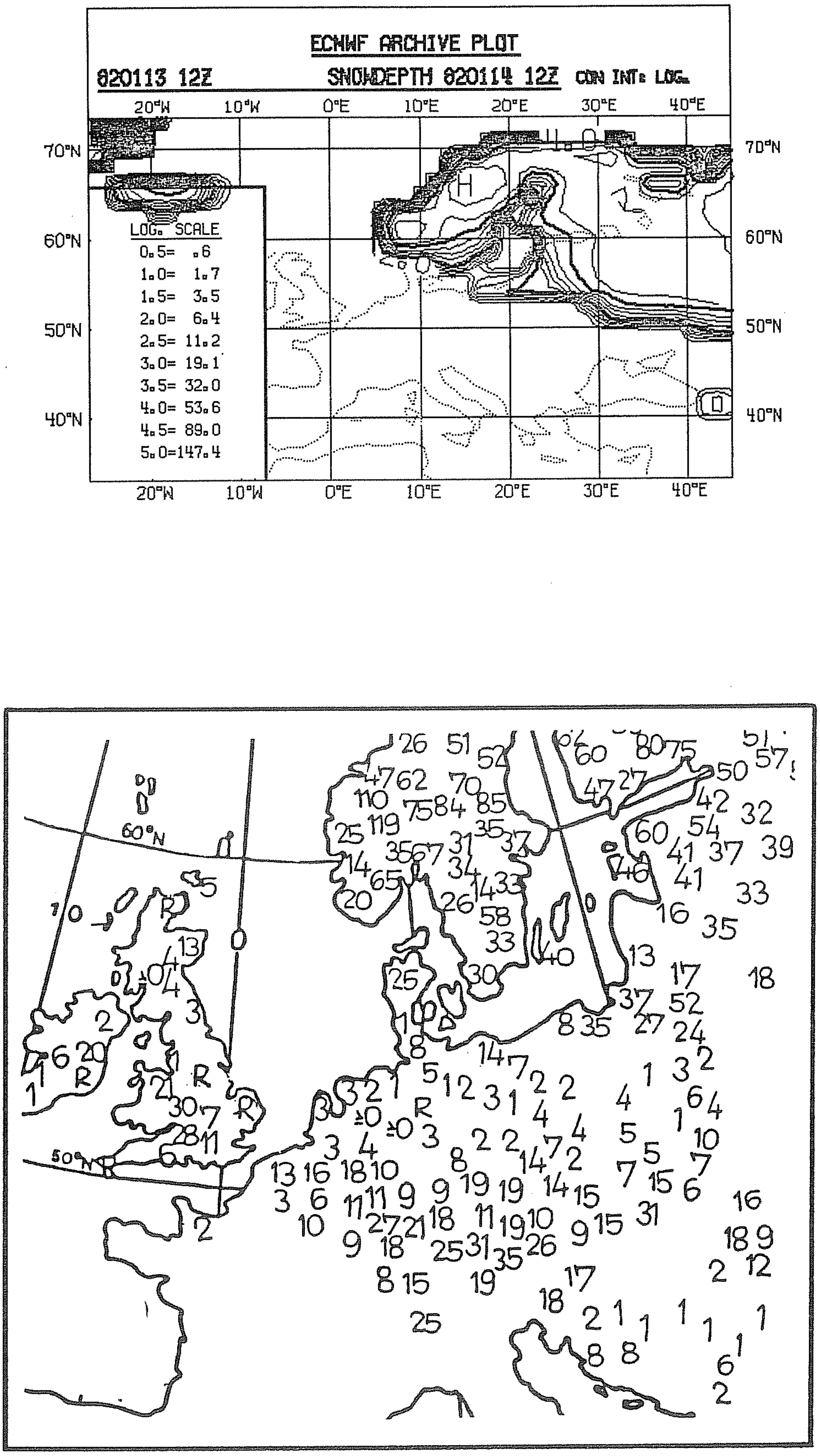

Fig. 9 As Fig. 6 but for 14 January 1982

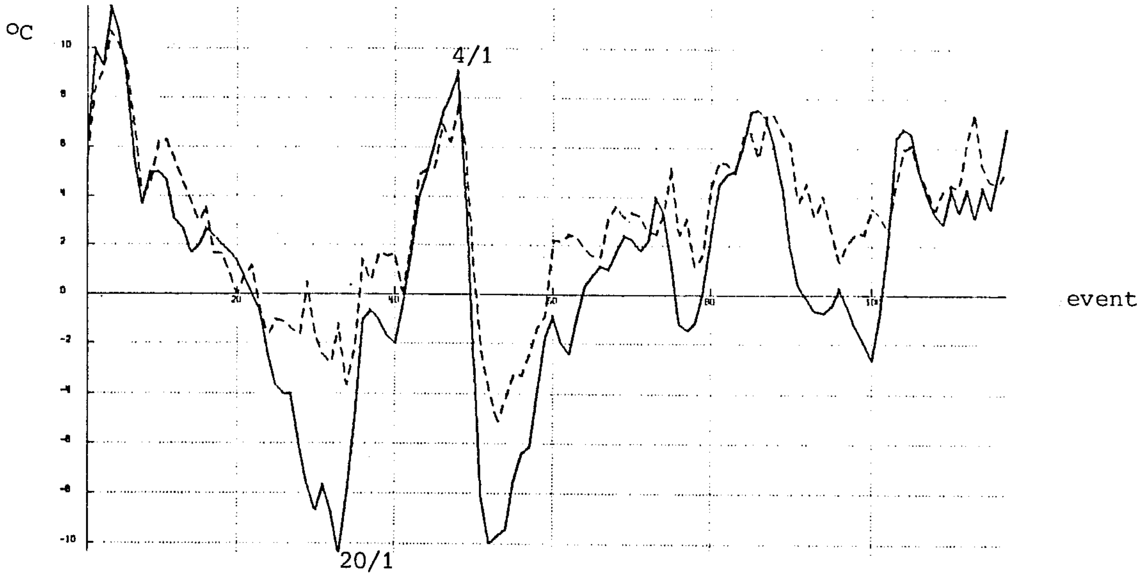

Fig. 10 Mean temperatures at 2m above the surface, averaged over three days, observed (full line) and centred at 72 hour forecast time (broken line) at Hannover (Germany) for each event between 15 November 1981 to 15 March 1982.

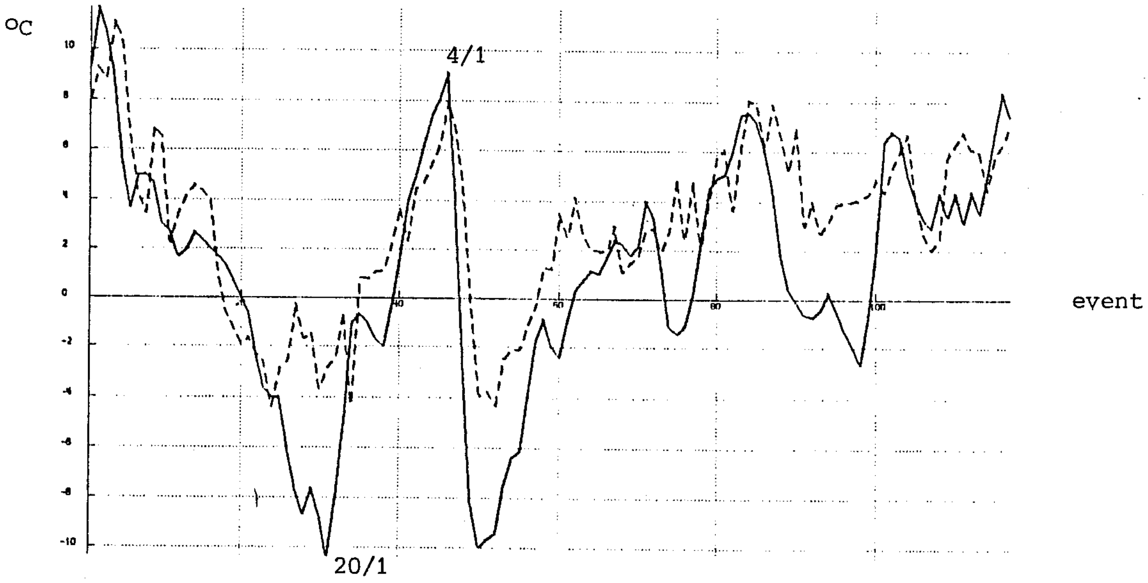

Fig. 11 As Fig. 10 for 120 hour forecast.

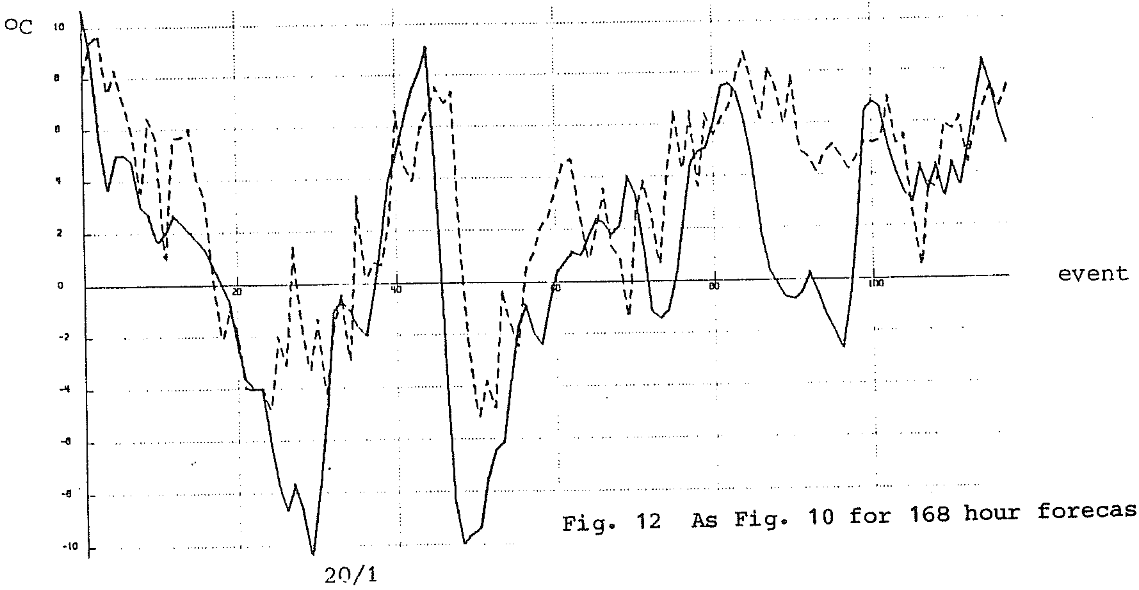

Fig. 12 As Fig. 10 for 168 hour forecast.

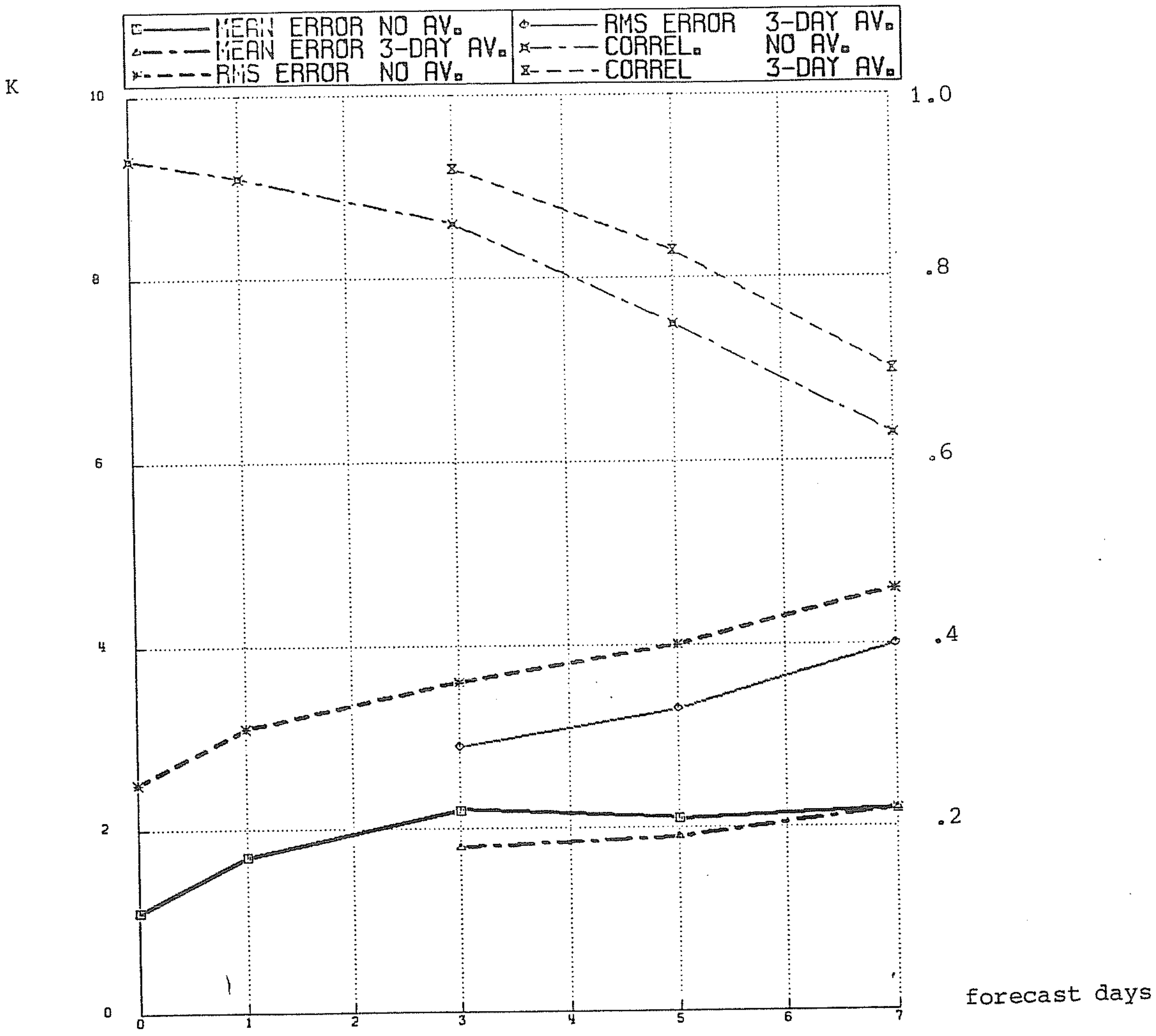

Fig. 13 Correlation, RMS error, and mean error of the 2m temperature forecast for Hannover from the analysis out to day 7, 15 November 1981 to 15 March 1982. The lines covering the full period give the verification results for the daily forecasts while the three day mean scores are indicated from day 3 onwards. The scale for the RMS and the mean error is indicated on the left; the units are Kelvins; the scale of the correlation is given on the right.

amplitude is not significant for the end-user of the forecast. The very cold winter periods interrupted by mild spells are still captured with equal intensity when comparing 1-day with 3-day mean forecasts. The systematic errors described in the previous paragraph obviously remain the same. The phase lag in predicting the drop in temperature after the mild spell in mid-winter becomes more apparent in the 3-day mean forecast.

For many applications of medium range temperature forecasts, e.g. for energy consumption, the prediction of changes in average temperature conditions is of considerable value. The verification scores of correlation, rms-error and mean error for 1-day and 3-day mean temperatures are compared in Figure 13. Time averaging results in a desirable reduction of the errors, improving slightly on the simulation of the observed daily mean and increases the correlation. All this should give the end-user of the forecast more confidence in the quality of the product.

3.4 Comparison of temperature predictors

It has been shown in pilot studies for various locations in Europe (Böttger and Grønaas, 1982, Klinker, 1982) that forecast temperatures close to the surface, e.g. the temperature at 2m or at 1000mb, normally explain the largest part of the observed temperature variance. The boundary layer temperature forecasts might be highly biased with respect to observations, e.g. due to differences between station height and the height of the model's orography, or the observed variability in temperature might be reduced in the model due to an inadequate prescription of the coastlines in the model, but the best correlation is still found between observed and forecast surface temperatures.

Figs. 14 and 15 confirm this result for both the winter and summer months at Hannover. The verification in both figures is valid at day 5, but the explained variance of the observed surface temperature by four different temperature parameters obtained directly from the model is shown for all forecast steps from day 2 to day 6. In winter, the model's 2m temperature and the 1000mb temperature are of equal quality and the day 5 forecast gives the best correlation. Temperature predictors from higher levels of the atmosphere, 850mb temperature, and the thickness, give in comparison a poor explanation of the observed variance. In winter, the boundary layer is often decoupled from the flow above by strong inversions, and temperatures above the inversion will only explain the broad scale conditions in space and time. This is reflected in the flat shape of the correlation graph where neighbouring forecast steps give equal results in predicting the local temperatures.

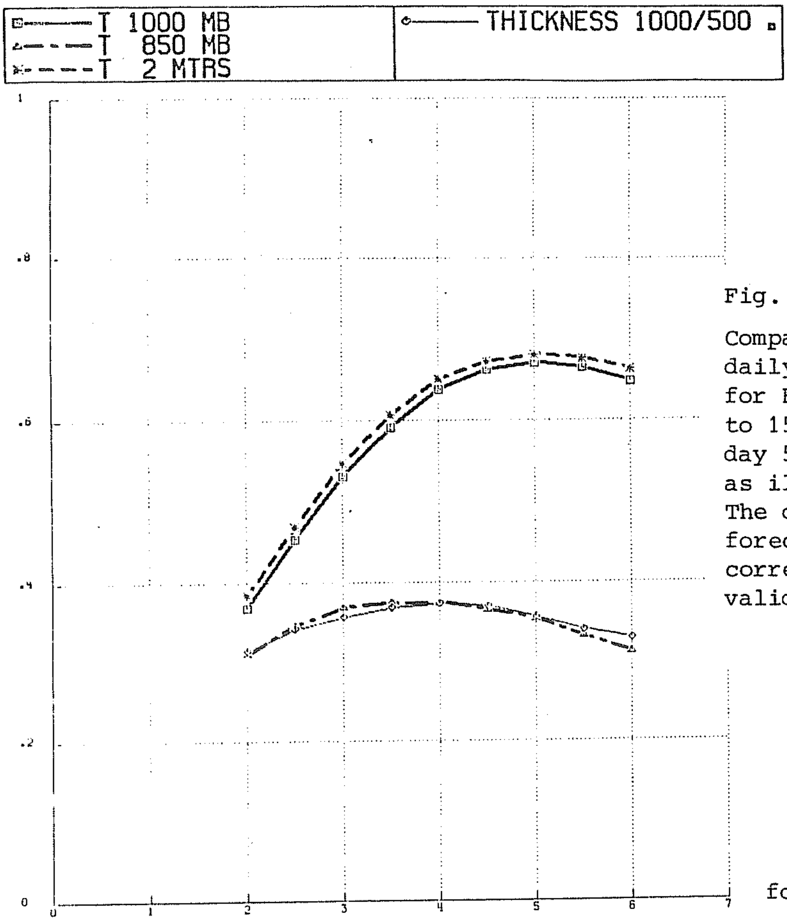

Fig. 14

Comparison of verification of daily temperature forecasts for Hannover, 15 November 1981 to 15 March 1982, valid on day 5, for various predictors as illustrated in the figure. The observations, valid on forecast day 5, are correlated with the forecasts valid on day 2, 2½, 3,....6.

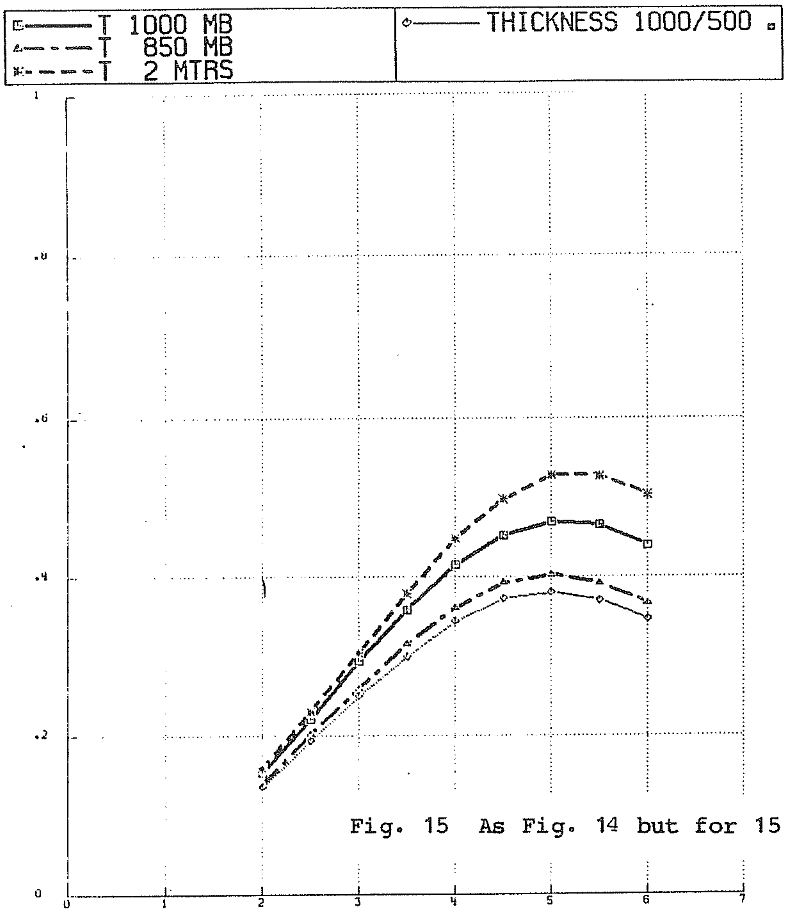

Fig. 15 As Fig. 14 but for 15 May to 15 September 1981.

In summer, the lower atmosphere is frequently well mixed due to enhanced convection, which improves the correlation between the surface temperature and T at 850mb and the thickness, though part of the explained variance remains at the same level as in winter. Near surface model temperatures are still the best predictors but the overall predictability has been reduced compared to the winter time. This is in agreement with hemispheric skill scores (Nieminen, 1983).

3.5 The impact of land-sea contrast

Firenze is situated in central northern Italy, well away from the Mediterranean coasts. However, surrounding model grid points include coastal, sea and land points. Fig. 16 gives the original model grid (1.875 * 1.875 degrees), together with the geographical and the model's coastlines. Large parts of southern Italy and Sicily disappear in the model's Mediterranean, but Firenze is "over land". However, after linear interpolation from the model grid to the archive grid of 1.5 * 1.5 degrees and then further interpolation to the station site, the surrounding gridpoints (some of which are sea points) will affect the value obtained for the location of Firenze. This is confirmed in Figs. 17 and 18, which show the verification of the 2m temperature at Firenze, both for the analysis and for the 5 day forecast.

The observed variability in the surface temperature is captured to a large extent by the analysis, but it is considerably reduced after five days in the forecast. For the forecasts "contributions" from sea grid points reduces the temperature variability. The strong bias, increasing from the analysis until day 5, can be partially ascribed to the lack of a diurnal cycle in the forecast model. In Figs. 17 and 18, the verifying time is 18Z. For comparison, Fig. 19 shows the same forecast verified against the average of maximum and minimum temperature of each day, a value which is closer to the true daily mean and the bias is reduced. Note that the choice of a mean temperature taken from the daily maximum and minimum, observed at 06Z and 18Z, results in fewer events for the verification, indicating that the record of synoptic reports for Firenze is not complete in the ECMWF archive.

The daily mean temperature observed at Firenze exhibits a much larger variability than that predicted by the model. In mid-summer, the climatological sea surface temperature in the northern Mediterranean lies around 22 to 24°C and the model forecast seems to be very close to that range throughout the summer. A statistical adaptation or different choice of gridpoints seems to be necessary to overcome the impact of the sea on the local model forecast.

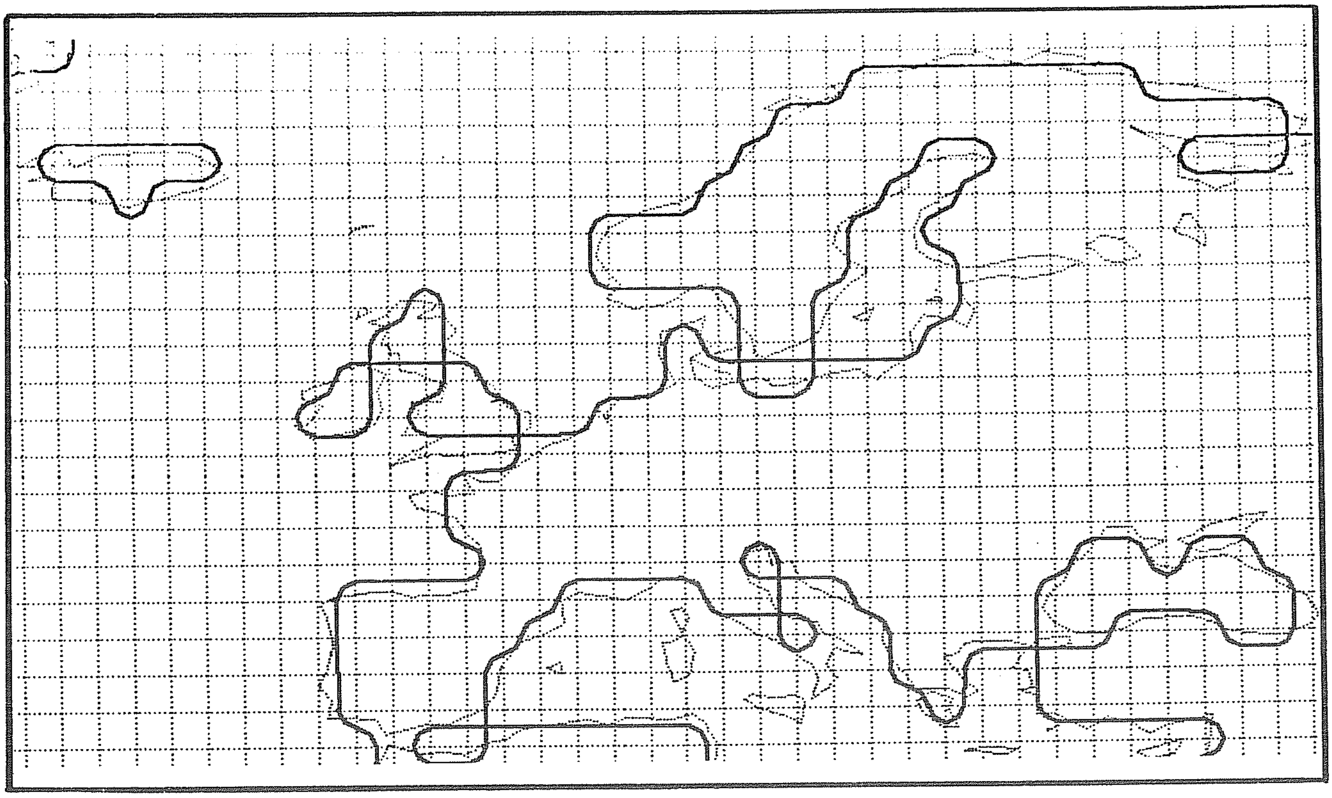

Fig. 16 Horizontal grid of the ECMWF model and the model coastlines.

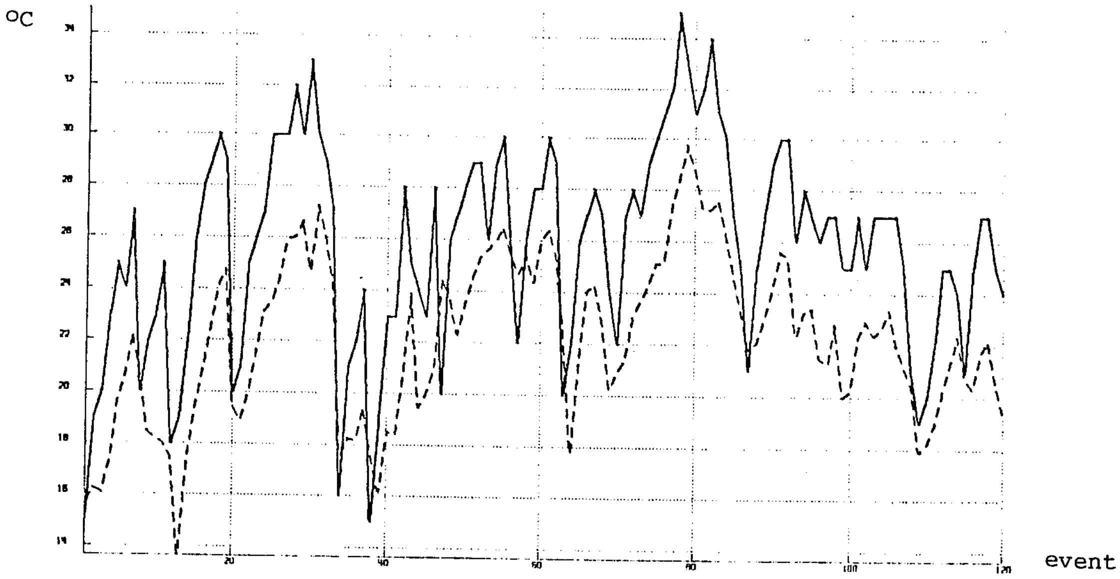

Fig. 17 Temperature at 2m above the surface, observed (full line) and as represented in the ECMWF analysis (broken line) at Firenze (Italy) for each observation (event) received between 15 May and 15 September 1982. The ECMWF model output is valid at 12Z, the observation at 18Z.

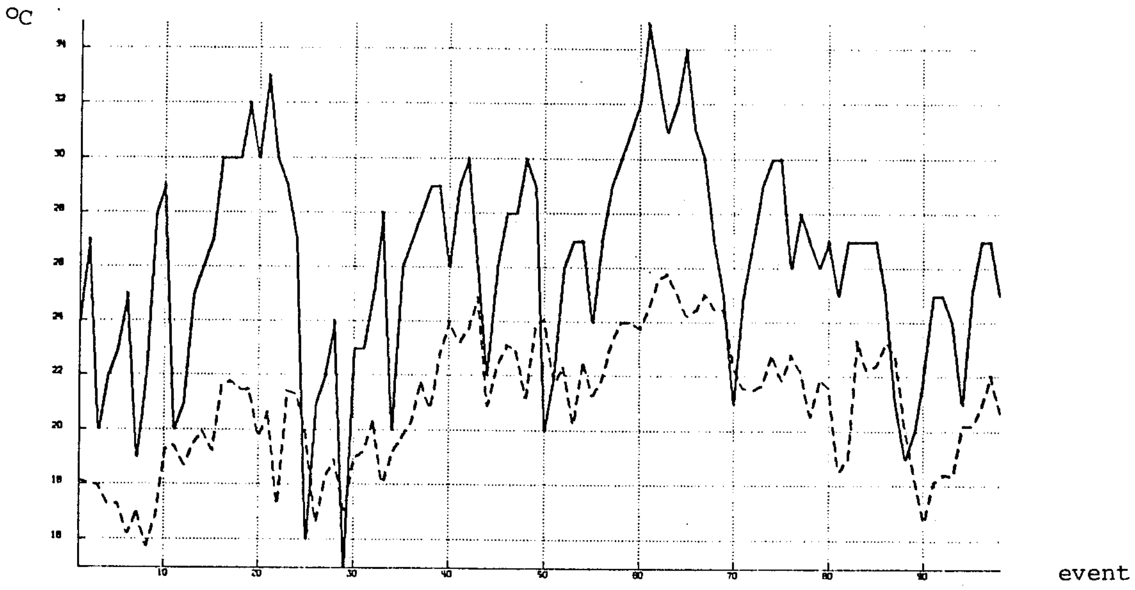

Fig. 18 As Fig. 17 but the broken line is here the 120 hour forecast valid at 12Z.

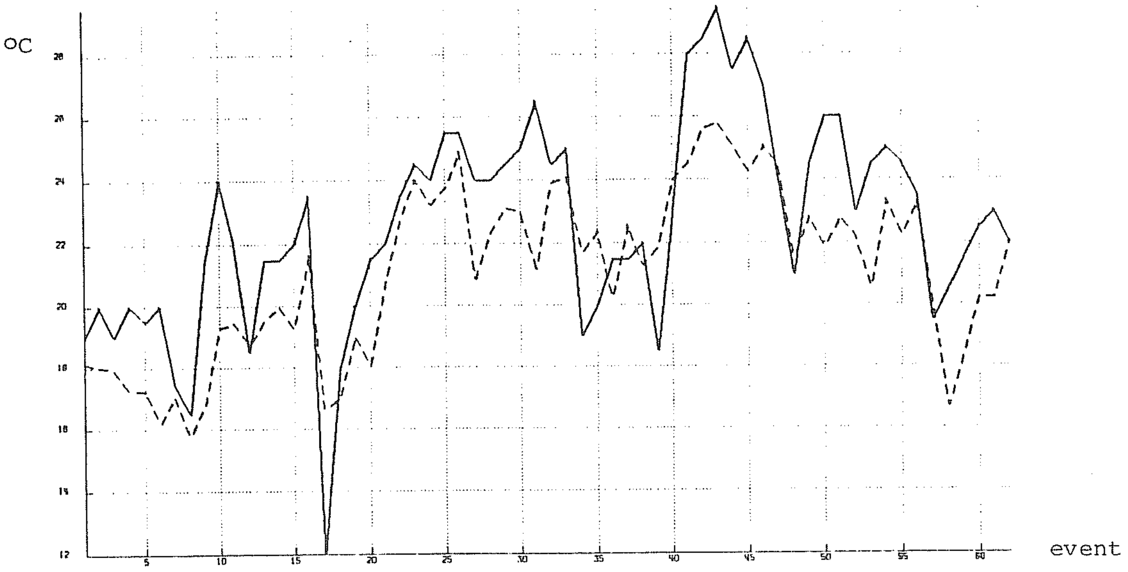

Fig. 19 As Fig. 18, but observed temperature is here the daily mean, i.e. the average of the minimum and the maximum temperature observed at 06 and 18Z respectively (T_{obs} = (Max + Min)/2).

4. Local verification of windspeed

Over land, the model-predicted wind in the boundary layer is unlikely to give satisfactory results when it is compared to local observation, because subgrid scale orographic effects are not represented in the model. There are however several applications of wind forecasts over the coastal areas and over the open sea in the short as well as the medium range forecast period. Examples are ship routing across the Atlantic or ice breaking in the Baltic in winter and spring. In both applications, the prevailing wind direction together with the speed is needed to give guidance to the end user.

Fig. 20 shows the fit of the observed wind speed at OWS Lima in winter and spring 1982 to the ECMWF analysis. It cannot be expected that verification of the forecast will be better than that of the analysis (Pümpel, 1982). The windspeed is slightly under-estimated, giving a negative bias and most of the extremes are smoothed out. The user of the forecast might therefore prefer the model 1000mb wind in the boundary layer (Fig. 21) which is nearly unbiased and gives a much better fit of the extreme wind speeds. The reason for this discrepancy is not quite clear, although the most likely explanation seems to be given by an inadequate interpolation of the 10m wind between the lowest model level (about 32m high) and the model surface.

A summary of the scores is given in Table 2. The correlation of observed and predicted windspeed drops rapidly with the forecast time. The best correlation is obtained for the 850mb wind (not shown). The user who wants to set up a scheme for statistical adaptation of the forecast would be advised to choose a model predicted wind from above the boundary layer, e.g. 850mb. Changes in the operational model will, however, affect the surface wind forecasts and further improvements can be expected. Continuous reassessment of the model predictors will be necessary in the future.

5. Summary and Conclusions

The ECMWF European-area archive and the statistical procedures for local verification and interpretation have been summarised. Verification of model temperature forecast against observations at three locations in Europe has shown that useful guidance can be given by the ECMWF model for predicting average surface temperature conditions in the medium range. Time averaging of the forecast over three days reduces the errors but will increase the reliability of the forecast. As the predicted temperature range will remain almost unchanged, some loss of information due to averaging will not be too serious and the end user of the forecast should gain increased confidence in the quality of the product.

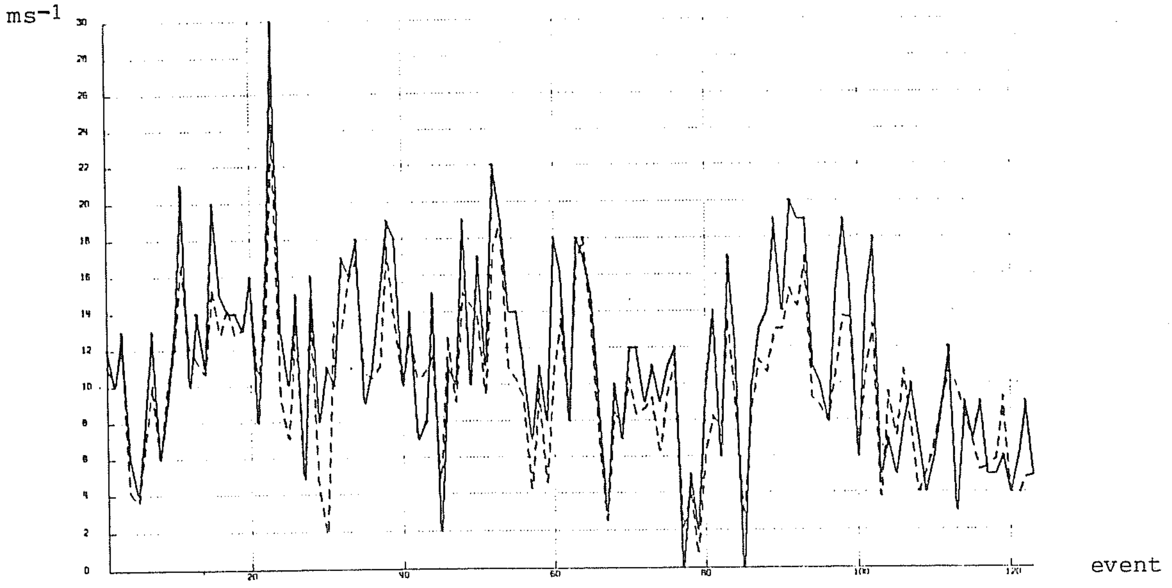

Fig. 20 Windspeed at 10m above the surface, observed (full line) and analysed (broken line) at OWS Lima for each 12Z observation (event) received between 15 January and 25 May 1982.

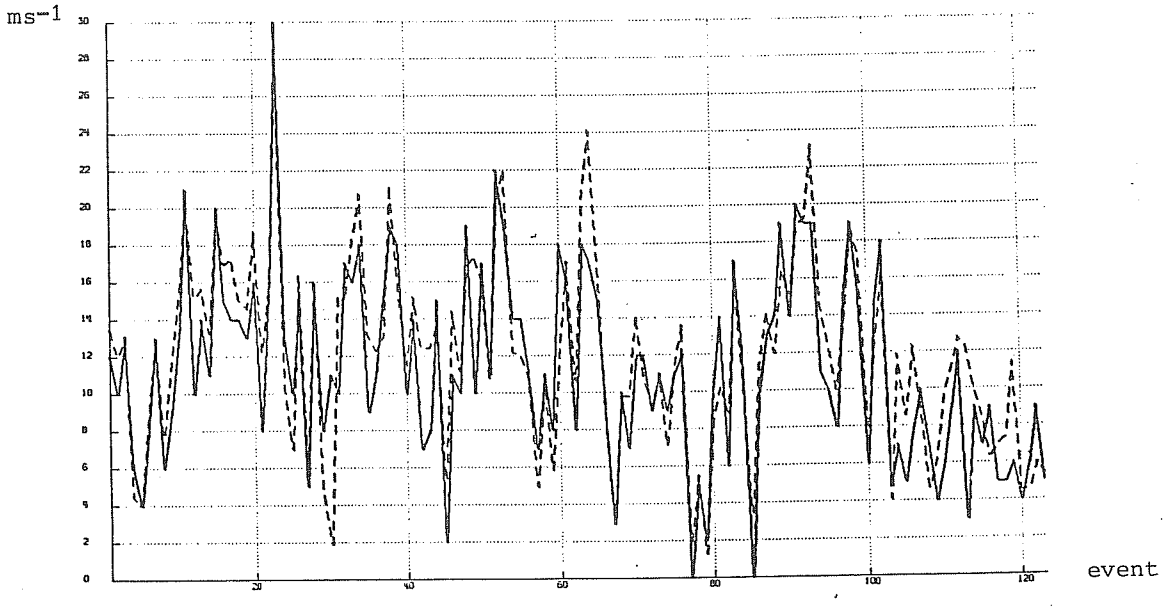

Fig. 21 As Fig. 20, but the dashed line is here the ECMWF model analysis of windspeed at 1000mb (instead of 10m).

Local temperature forecasts obtained from a global model and interpolated to the local site from the surrounding gridpoints are biased, the magnitude of the mean error depending in part on the representation of surface conditions in the model. The user of such forecast products needs to be aware of the characteristics of the model gridpoints from which the forecast is obtained. Inadequacies in the representation of coastlines, ground elevation in the orography and ice and snowcover or soilwetness will influence the bias of local temperature forecasts. Any changes to the surface or near-surface levels of the model are likely to have the largest impact on parameters close to the surface.

Nevertheless, temperature forecasts from the vicinity of the model surface in general give the highest correlation with the observed temperature at 10m. Temperature predictors from above the boundary layer may be less sensitive to model changes but their application in forecasting will require more sophisticated statistical adaptation. The verification of the windspeed at 10m above the model surface exhibits a rapid decrease in correlation within the first three days of the forecast. A somewhat better result is obtained when the observed windspeed is compared to the 850mb model wind.

10m Windspeed

		Forecast time —		
	ANAL	24h	48h	72h
mean error	- 1.3	- 1.3	- 1.4	- 1.7
RMS error	2.8	3.8	4.3	5.0
correlation	.88	.73	.61	.44

1000mb Windspeed

		Forecast time —		
	ANAL	24h	48h	72h
mean error	.8	1.0	.8	.4
RMS error	2.7	4.0	4.5	5.3
correlation	.87	.70	.60	.43

Table 2 Verification of windspeed at OWS Lima, 15 January to 25 May 1982. The observed surface windspeed is compared to the ECMWF model wind at 10m above the surface (top) and at 1000mb (bottom). Units for mean error and RMS error are ms^{-1}.

References

Åkesson, O., Böttger, H., Pümpel, H., 1982: First results of direct model output verification of near-surface weather parameters at 17 locations in Europe, ECMWF, Operations Department Technical Memorandum No. 47.

Böttger, H. and Grønaas, S., 1982: Optimal verification of ECMWF surface temperature forecasts at two locations in Europe. ECMWF, Operations Department Technical Memorandum No. 64.

Johannessen, K.R., 1982: Verification of ECMWF quantitative precipitation forecasts over Europe January 1980 to April 1981, ECMWF, Research Department Technical Memorandum No. 51.

Klinker, E., 1982: Local weather element guidance from the ECMWF forecasting system in the medium range: Interpretation experiments, ECMWF Seminar Proceedings, 1982.

Louis, J-F., 1982: Parameterisation of sub-grid scale processes, ECMWF Seminar Proceedings, 1982.

Nieminen, R., 1983: Operational verification of ECMWF forecast fields and results for 1980-81, ECMWF Technical Report No. 36.

Pümpel, H., 1982: Regional and local verification of near surface weather parameters, ECMWF Seminar Proceedings 1982.

Wilson, L.J., 1982: MOS Probability of precipitation forecasting experiments in the medium range, ECMWF Seminar Proceedings 1982.