TREATMENT OF NON-GAUSSIAN OBSERVATION ERRORS IN THE HIRLAM SYSTEM

P Lönnberg and K Eerola
Finnish Meteorological Institute
Helsinki, Finland

1. INTRODUCTION

The backbone of the observing system in Europe is the radiosonde network, which provides most of the upper-air observations over the continent. Furthermore, several short-range forecasting centres in Europe do not use satellite sounding data from the North Atlantic and rely therefore very heavily on the available radiosonde data. Consequently, the quality of the reports from the radiosonde stations is important to the accuracy of short-range numerical forecasting.

The quality of the temperature/height data from European radiosonde stations is quite variable, while the wind data have rather uniform quality. Several stations have observation errors, which significantly exceed the errors of the background field, usually a six hour forecast. The quality of the European radiosondes in discussed in Section 2. The response of the analysis system and the normal mode initialisation scheme to noisy observations is shown in Section 3. Several approaches have been used to deal with data having a high probability of gross errors. Section 4 proposes a grading of the rejection limits in the first-guess check according to the quality of the station. Forecasts experiments with a station quality list are discussed in Section 5. Some conclusions are presented in Section 6.

2. EUROPEAN RADIOSONDE NETWORK

2.1 Quality of radiosonde observations

The temperature measurements in the stratosphere and upper troposphere have serious errors at several stations as compared to the errors of the background field. Fig. 1 shows the root-mean-square (rms) difference between observed values and short-range forecasts for height at 200 hPa. Several stations with very large rms values are found in Eastern Europe, while the Western European stations generally have very small errors. Even the stations in the North Atlantic area have smaller rms values than the Eastern European stations.

The statistical properties of the differences between observation and first-guess are defined by the distributions of the observation and of the forecast errors. The statistics of the first-guess errors vary slowly in space, particularly over continental Europe and away from the lateral boundaries of the forecast domain. The inter-station variability is mainly caused by different observation error characteristics of the stations.

An upper bound of the error in the background field is given by the lowest rms values found in Europe

(Fig. 1). The rms values are close to or even under 10 m at several stations. Then both the forecast and observation errors are less than 10 m, assuming they are uncorrelated. This suggests that the rms observation error at the worst stations is almost an order of magnitude larger compared to the best stations.

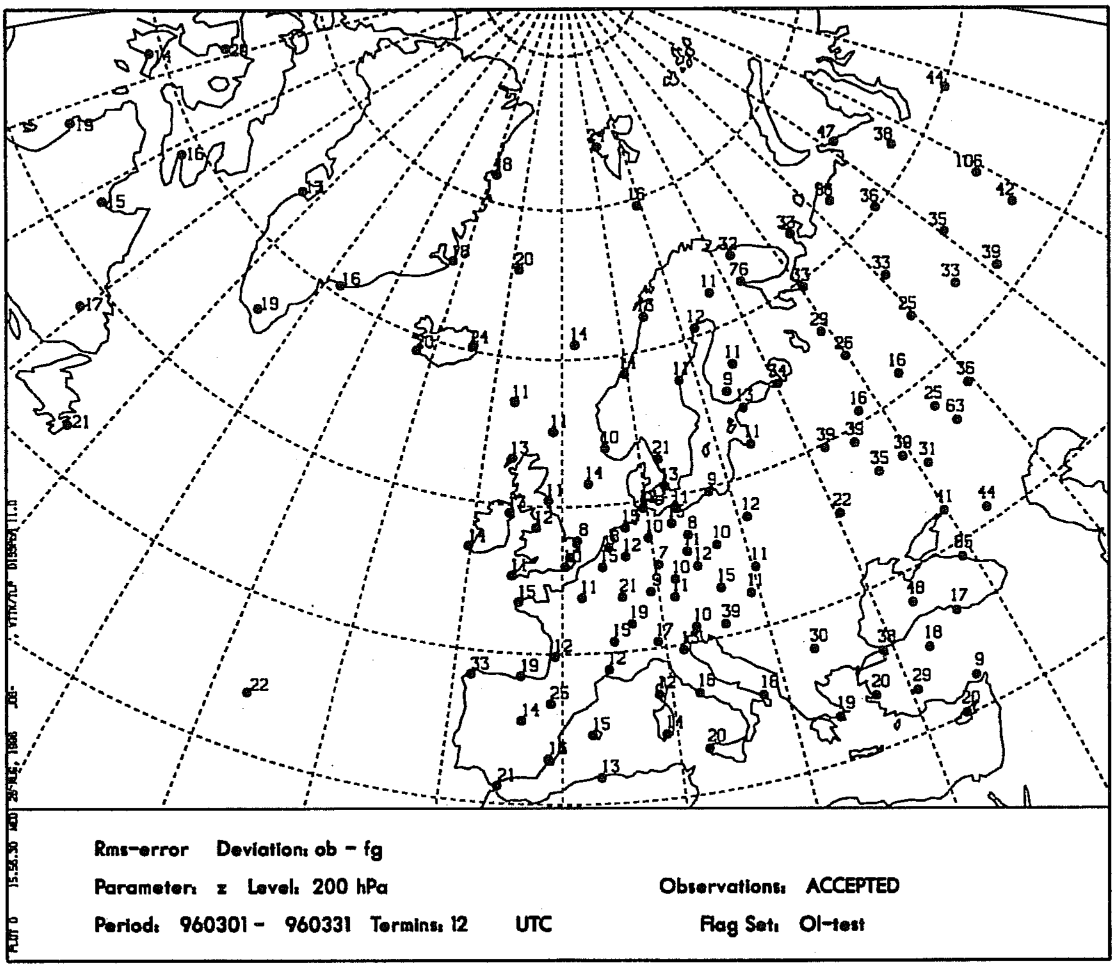

Fig. 1 Root-mean-square values of differences between reported value and 6 hour forecast for 200 hPa height field over Europe for March 1996 at 12 UTC. Unit: m.

The clustering of low quality stations reduces the effectiveness of identifying corrupt observations by using data from neighbouring stations. Only the check against the background field remains as an efficient tool to identify and reject bad observations. Furthermore, the low quality stations have similar biases (not shown) which makes it difficult for the analysis scheme to filter these errors.

Investigation of the time series of the error characteristics and reporting activity shows that there are several types of stations in Europe. Most of the stations stations have small biases and stable magnitudes of random errors. An extreme example of a poor quality station is shown in Fig. 2. The monthly mean difference between observed and first-guess values fluctuates strongly making it even difficult to apply any kind of bias correction. Fig. 2 also shows discontinuities in the station activity.

The investigation of the time series also reveals that considerable improvements have been found at several eastern European stations during the 90's.

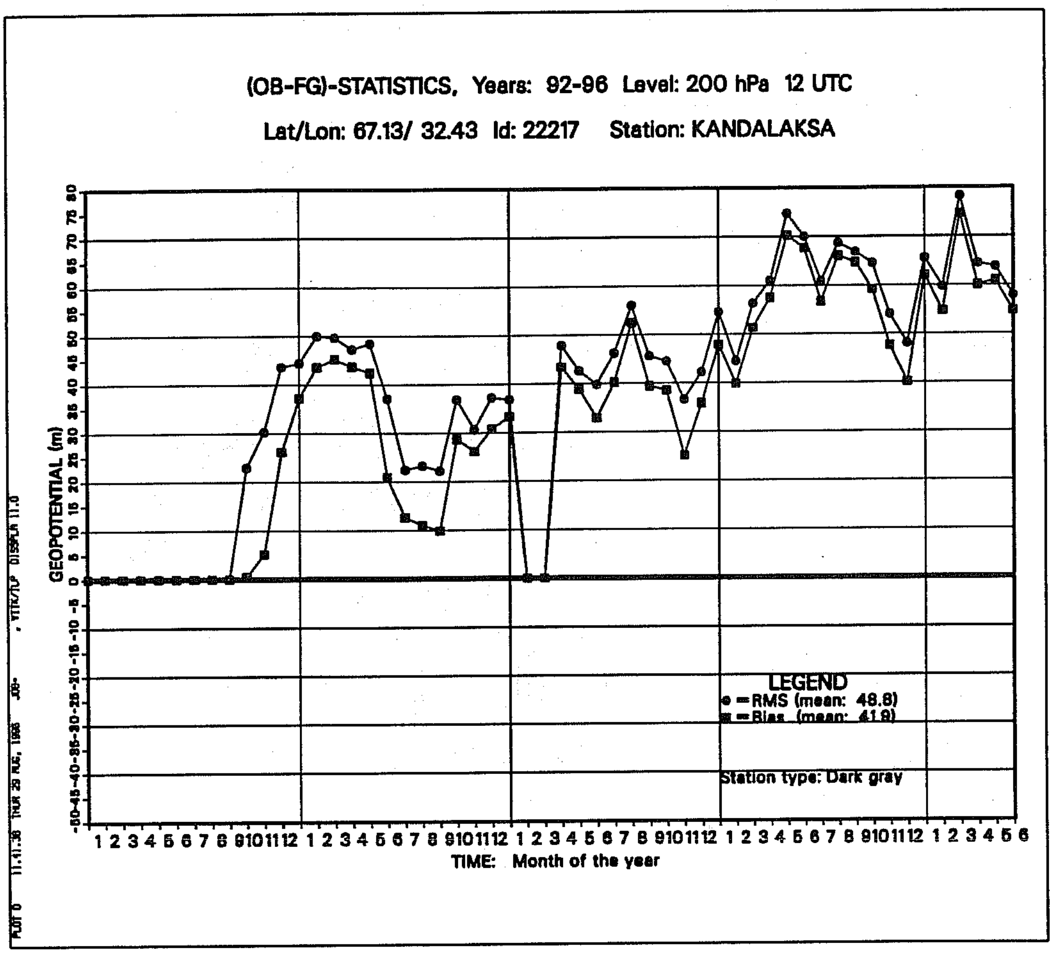

Fig. 2 Time series of the monthly means of root-mean-square and bias values of differences between observed and 6 hour forecast at 200 hPa for station Kandalaksa (67.37 N , 26.65 E) for January 92 - June 96 at 12 UTC. Unit : m .

2.2 Investigation of error distributions

The distribution of the differences between observed values and first-guess values provides an estimate of the likelihood of corrupt observations. Most stations have narrow distributions with a small number of outliers. These distributions are approximately Gaussian.

The HIRLAM analysis scheme assumes that all radiosonde data are unbiased and of equal quality. The assumed observation error depends only on pressure. The validity of these assumptions is tested at each station. Three months of 12 UTC data have been used for the statistics presented in Fig. 3. The whole population has an rms difference of 18 m from the first-guess. The probability of finding normally distributed data more than 2 standard deviations away from the mean is 4.55 %. Fig. 3 shows the percentage of data deviating by more than 36 m from the first-guess. The distinction between high and low quality stations is striking.

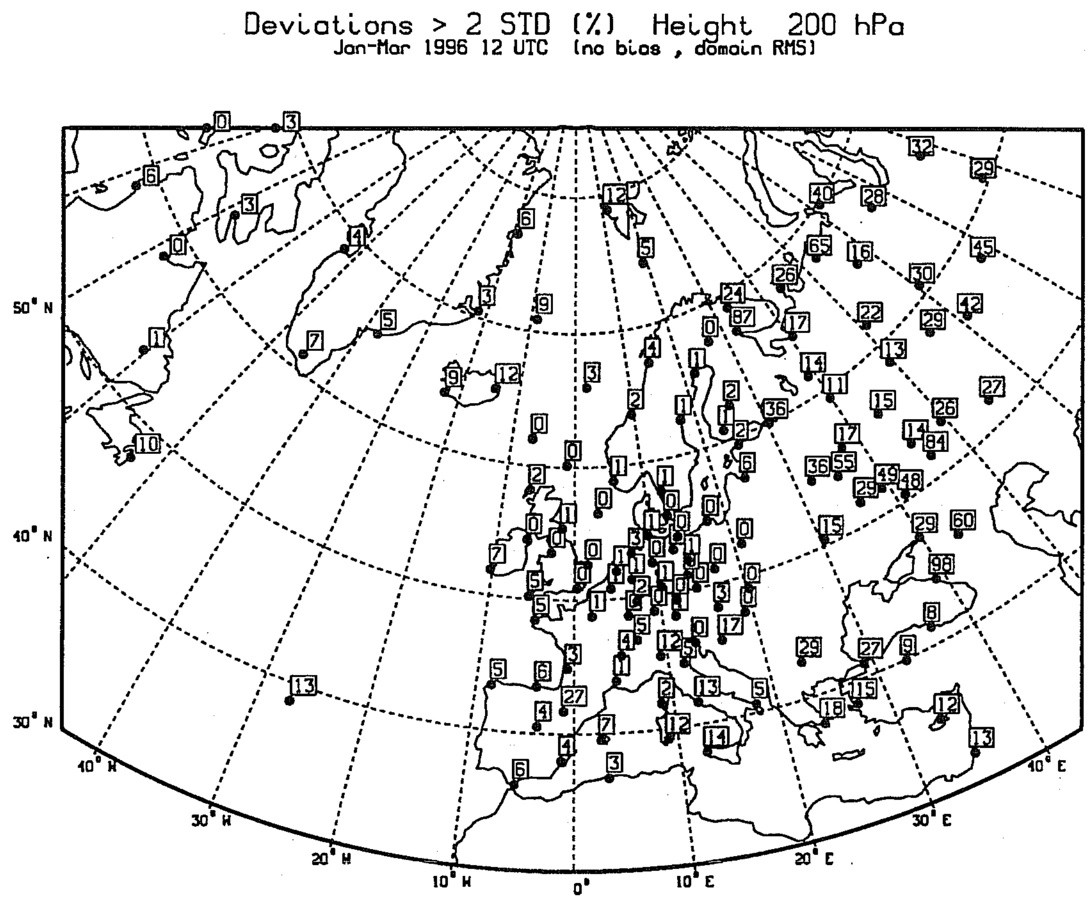

Fig. 3 Percentage of reports deviating by more than 36 m from the 6 hour forecast for January-March 1996 at 12 UTC for 200 hPa height.

3. ANALYSIS IMPACT

The analysis response depends critically on the filtering properties of the analysis system. The horizontal scale length of the background error covariances has been tuned to respond to scales of the same order as the distances between radiosonde stations over Northern Europe. The mean analysis increments are plotted in Fig. 4. The analysis draws to the systematic differences between stations in this region. The mean increments at Sodankylä (67.37 N, 26.65 E) and Kandalaksa (67.17 N, 32.43 E) are very different. Against the first-guess, the biases are -5 m at Sodankylä and 73 m at Kandalaksa (not shown). The analysis draws to 26 % of the first-guess difference. The distance between the stations is only 250 km. The difference in the mean increments between those two stations is 20 m. This corresponds to a mean geostrophic wind increment of 5.8 m/s. A sharp boundary between unbiased stations and biased stations divides Europe (Fig. 4).

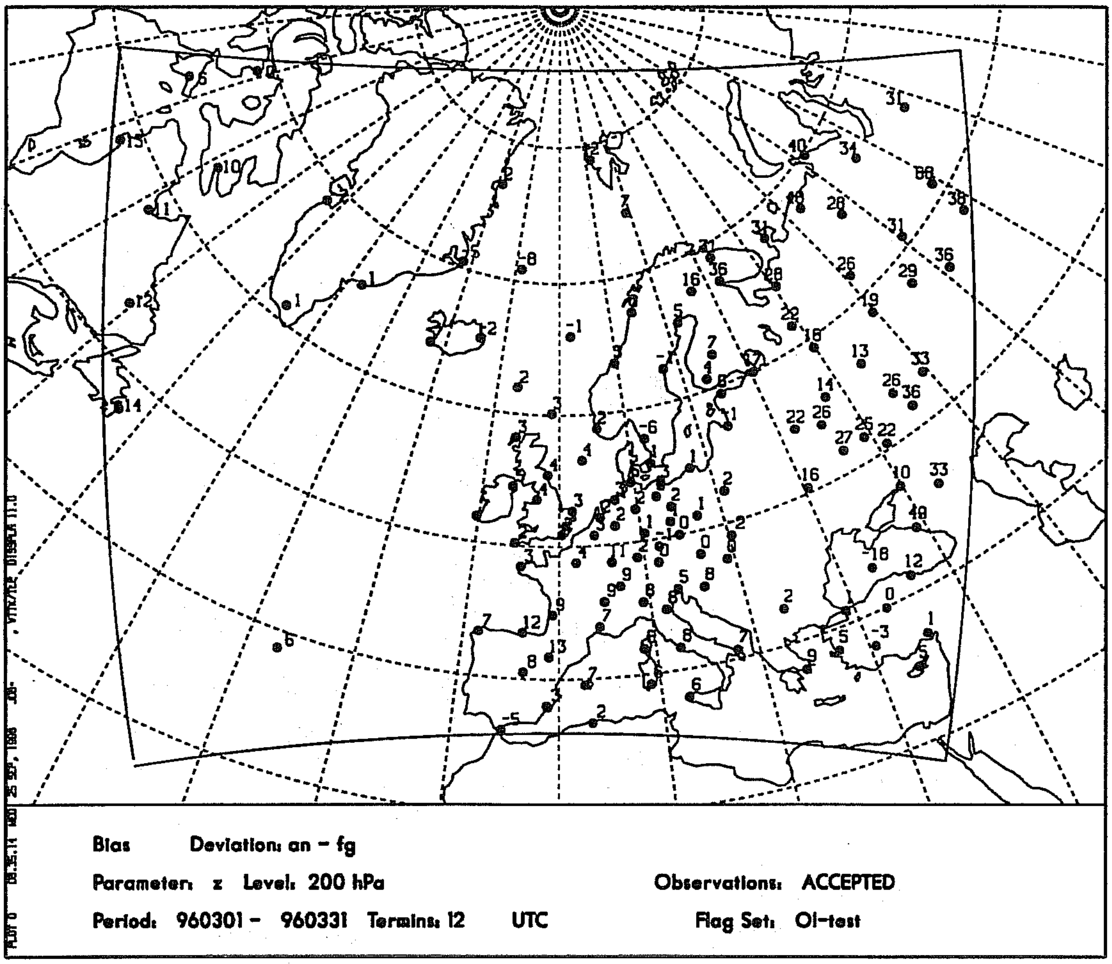

Fig. 4 Mean values of differences between analysis and 6 hour forecast 200 hPa height field over Europe for March 1996 at 12 UTC. Unit: m.

The initialisation removes some of the unbalanced mass field, but spreads the differences on larger scales (Fig. 5). There is still a 7 m mean difference in the increments after initialisation between Sodankylä and Kandalaksa, corresponding to a 2 m/s geostrophic wind.

4. FORMULATION OF A GREY LIST

An analysis scheme supposes unbiased observations with estimates of the observation error characteristics. The large variability of the radiosonde station quality requires attention to be given to each station individually. Ideally, we should correct the data for systematic errors and provide an estimate of the random error. This is a very laborious task, which must be performed on a monthly basis. However, there are significant fluctuations in the monthly values of the bias and a correction based on data from the preceding month(s) might not always be very useful. Especially, the Eastern European radiosonde network has changed dramatically over the past years. Rapid changes in data quality and in the station network makes the maintenance of a blacklist quite demanding.

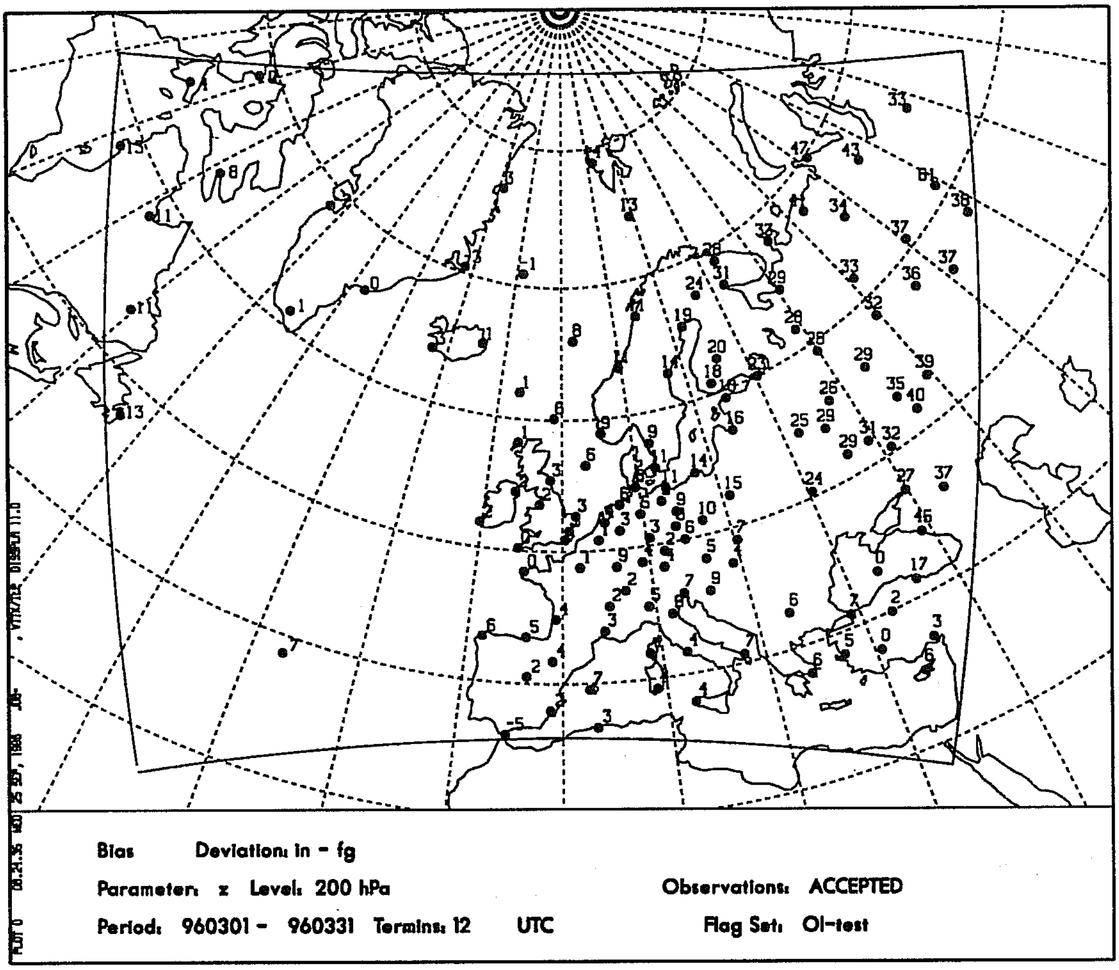

Fig. 5 Mean values of differences between initialised analysis and 6 hour forecast 200 hPa height field over Europe for March 1996 at 12 UTC. Unit: m.

A different approach has been taken in the HIRLAM analysis system, which is a limited area version of the ECMWF OI scheme. Instead of completely excluding data according to a pre-defined list ("blacklist"), data are rejected with different rejection criteria. A grey list is composed of stations with a high probability of corrupt reports. Stations on the grey list have tighter rejection limits and are more readily rejected by the first-guess check (*Lorenc and Hammon, 1988).* Fig. 6 shows the grey list of stations as compiled for the HIRLAM system in 1991.

Detailed investigations of European radiosonde show that there is a very wide range of errors involved. This justifies a division of the grey list into several categories of observations. A list with two shades of "grey" has been recently compiled to correspond to the present situation in the European sounding station network (Fig. 7). The "dark grey" stations have very large systematic and/or random errors. The "light grey" data are clearly worse than those from high quality stations, but better than the "dark grey" ones. The rejection limit are very narrow for the "dark grey" report and intermediate for the

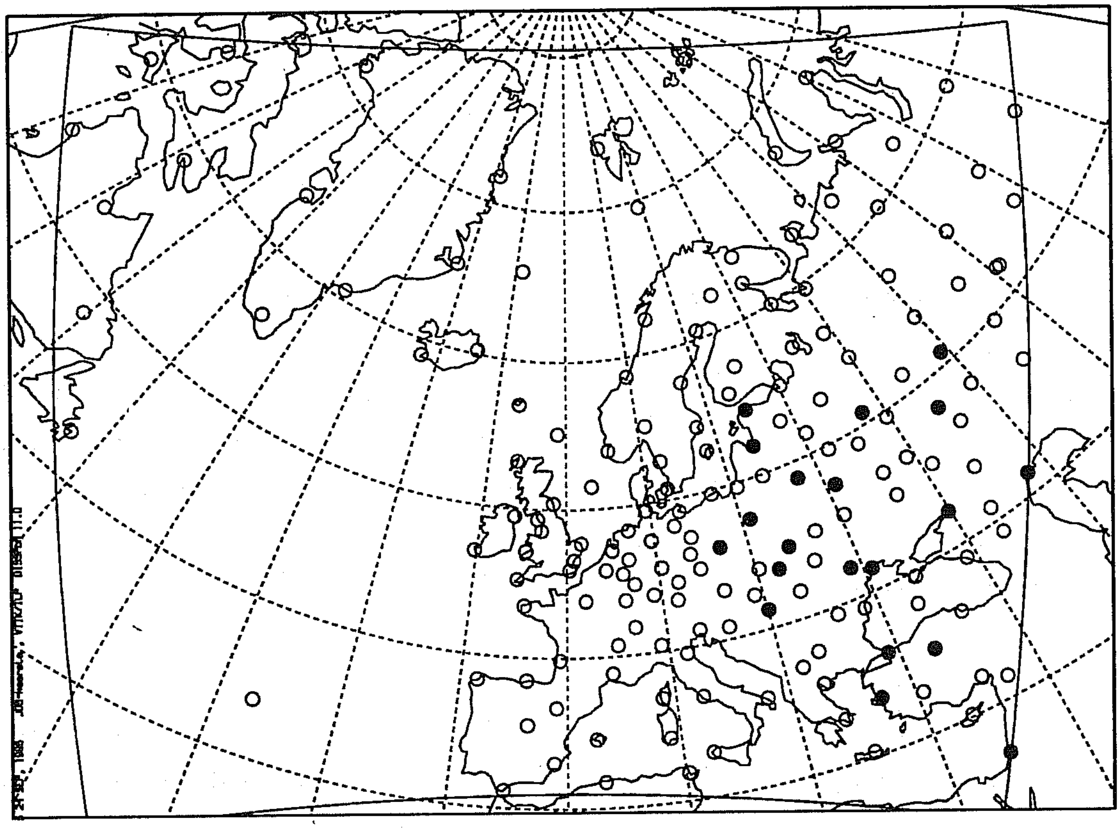

Fig. 6 The HIRLAM "grey" list as compiled in 1991. "Grey" list stations are indicated by filled circles.

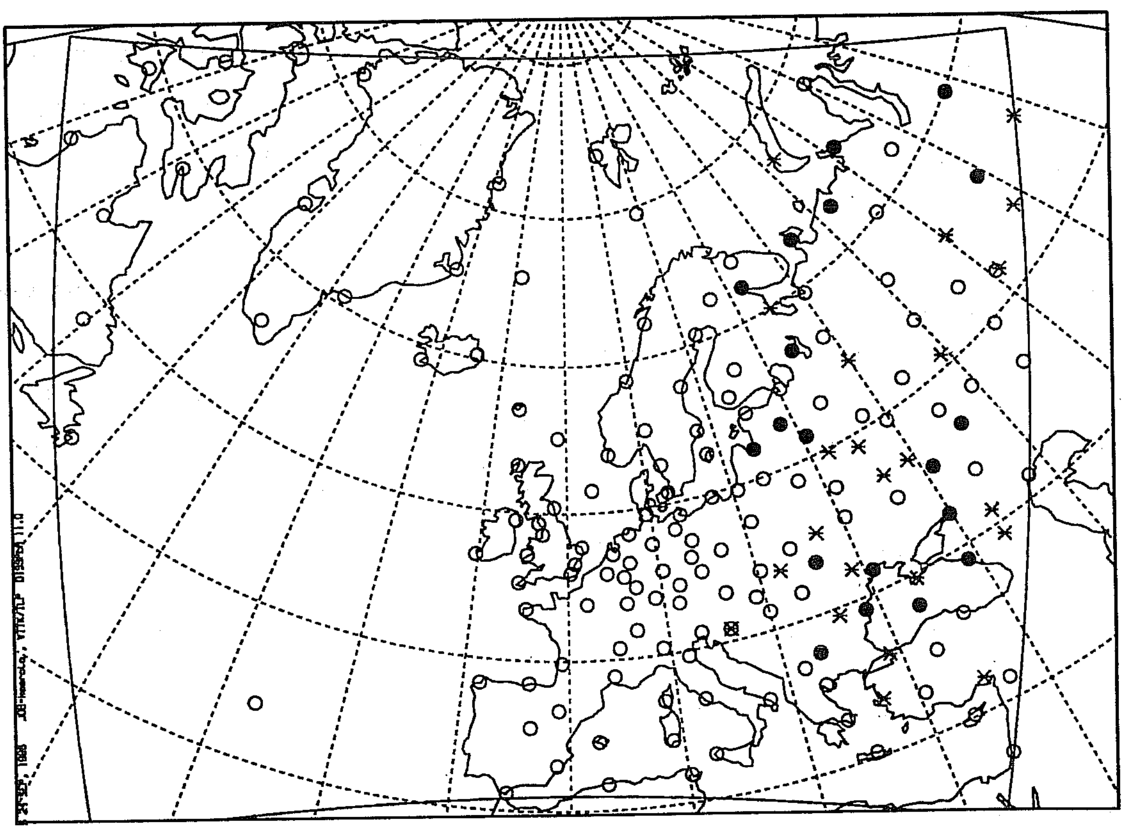

Fig. 7 The 1996 HIRLAM "grey" list with two categories of "grey" stations. "Dark grey" stations are indicated by filled circles and "light grey" stations by the symbol ⁎.

"light grey". The limits for the good data are very high in order to reject only the rare mistakes. The ratio of the rejection limits are 1 : 2 : 4 for "dark grey" : "light grey" : good data. At 200 hPa, typical rejection limits are 40 m : 80 m : 160 m. The "grey" list is used only in the first-guess check for height data. The new "grey" list contains 43 stations with 24 "light grey" and 19 "dark grey" stations.

5. FORECAST EXPERIMENTS

Several short data assimilation experiments with two day forecasts have been run with the limited area HIRLAM system on the FMI forecasting area. All these cases are from summer 1996 which had a strong anticyclone over Northern Europe. In these cases the impact of the "grey" list is small. Typically, the differences are largest in the analysis. The initialisation reduces the peak values of the analysis differences and spreads the differences over larger areas. In the forecasts the differences travel downstream with decreasing amplitude. The analysis differences are confined to the upper troposphere and the stratosphere, but propagate vertically to the surface during a two day forecast.

6. CONCLUSIONS

The quality of European radiosonde temperature (and height) reports is very variable. The problem stations have serious systematic errors and large random errors and the distribution of the differences between observation and short-range forecast is clearly non-Gaussian. Furthermore, the error characteristics are non-stationary, which makes it difficult to apply bias corrections to these stations.

The analysis scheme has been tuned to respond to scales, which are resolved by the radiosonde network in Northern Europe. Efficient quality control methods are therefore required to identify erroneous observations, particularly as the doubtful stations tend to cluster in certain areas. In these cases, the check against the first-guess is the only useful test to identify erroneous reports.

"Blacklisting" of stations is a rough tool, which requires regular supervision. In the HIRLAM system, the rejection limits have been graded according to the quality of the stations. Three groups of stations have been used. Most of the stations are considered good with observation errors of the same magnitude as the error of the background field. Two categories of lower quality data have been added. The list of poor quality stations includes almost a third of all radiosonde stations in the FMI forecasting area. In these two categories the rejection limits in the first guess check have been made smaller.

The forecast impact from the "grey" list was modest in those summer cases which have been run.

REFERENCE

Lorenc, A C and O Hammon, 1988: Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. Q J R Meteorol Soc, 114, 515-543.

BASIC FACTS ABOUT CHAOTIC MOTIONS

Olivier Talagrand
Laboratoire de Météorologie Dynamique, CNRS
Paris, France

Abstract

What makes a deterministic dynamical system chaotic is essentially the fact that finite accuracy in the knowledge of the present state of the system results in total impredictability, after finite time, of the future state of the system. Chaos is normally associated with exponential growth of a small initial error and logarithmic growth of predictability time with respect to the initial accuracy.

Chaos can rather easily be explained, and 'understood', on small dimensional systems: the so-called *baker's transform*, the two-dimensional mapping introduced by Hénon, and the celebrated three-dimensional system of Lorenz. Chaos essentially exists under two forms. In *conservative systems* (of which Hamiltonian systems constitute an important subgroup), volume is conserved in phase space in the course of the temporal evolution of the system. In *dissipative systems*, volume decreases in phase space because of the presence of irreversible processes such as viscosity or heat conduction. For both classes of systems, chaos results from repeated *stretching and folding* of volume elements in phase space. But a major difference is that the orbits of a dissipative system converge to a subset of the phase space, which has zero volume and is called the *attractor* of the system. Because of the repeated stretching and folding, attractors of a chaotic dissipative systems have an infinitely-many-times foliated structure (fractal structure), which is at the origin of the denomination of *strange attractors* given to these objects. Chaos in the Solar System is an example of conservative (and more precisely Hamiltonian) chaos, while turbulence of (even infinitesimally) viscous fluids, such as the atmosphere, is an example of dissipative chaos.

The chaotic character of a dynamical system can be quantified by the Lyapunov exponents of the system, which are cofficients measuring the rate of amplification of small perturbations imposed on that system. For conservative systems, the sum of the Lyapunov exponents is zero, while it is negative for dissipative systems. In either case, the presence of at least one strictly positive Lyapunov exponent is the standard characterization of chaos. The geometrical dimension of the attractor of a dissipative system, which as a rule is not an integer, can be related to the sequence of Lyapunov exponents.

The chaotic character of a dynamical system has major implications for the possibility of assimilating observations performed on that system. Just as finite accuracy on the initial state imposes a finite limit on the predictability of the future state of the system, finite accuracy on the observations imposes a finite limit on what assimilation can achieve. For any dynamical system, whether it is chaotic or not, finite accuracy on observations performed over a finite period of time results in uncertainty on the state of the system at any time. That uncertainty is concentrated in the *unstable* components of the system at the *end* of the period over which the observations have been performed, while it is concentrated in the *stable* components of the system at the *beginning* of the period. As a consequence, past noisy observations, even if they are available over an infinite period of time, cannot define with perfect accuracy the present state of the unstable components of the system. This has very important implications for predictability, since it means that, as a general rule, the uncertainty in the initial condition of a forecast will be

concentrated in those components which are likely to amplify most rapidly in the first stages of the forecast. For a chaotic system, it is clear that assimilation can be useful only if the time interval between successive observations is smaller than the predictability time corresponding to the accuracy of the observations. If such is the case, but the time interval between successive observations is large enough for the tangent linear approximation not to be valid, standard assimilation algorithms such as variational assimilation or extended Kalman filtering will fail.Several approaches have been proposed for coping with such a situation: ensemble Kalman filtering, in which the uncertainty on the state of the system is not represented by a covariance matrix, but rather by a finite sample of possible states (Evensen), Kalman filtering with additional terms representing the non-linear effects not represented by the standard filter (Miller *et al.*), successive variational assimilations performed over progressively longer periods of time (Pires *et al.*). In any case, a major priority is to clearly assess the limitations on assimilation resulting from the chaotic character of the atmospheric and oceanic flows.