

Mesan, an operational mesoscale analysis system

By LARS HÄGGMARK¹, KARL-IVAR IVARSSON¹, STEFAN GOLLVIK¹ and PER-OLOF OLOFSSON², ¹SMHI, S-601 76 Norrköping, Sweden; ²Swedish Military Weather Service, Box 420, S-746 29 Bålsta, Sweden

(Manuscript received 3 August 1998; in final form 1 July 1999)

ABSTRACT

A system for mesoscale analyses of selected variables has been developed. The analysed parameters are of general interest in operational weather forecasting, but normally not available from NWP systems, or available, but with a significantly lower quality than achieved by the mesoscale analysis system. A supplementary objective is to produce initial information to be used for now-casting techniques. Examples of parameters are precipitation, temperature, humidity, visibility, wind and clouds. The basis of the analysis system is the optimal interpolation technique (OI). The use of observations from automatic stations, radars and satellites have been investigated. The investigation indicates that a dense network of ordinary precipitation gauge measurements can produce more accurate analyses than more elaborate systems like radar that suffers from anomalous echoes and other errors.

1. Introduction

The modern systems of numerical weather prediction (NWP) generally use sophisticated methods for data assimilation. These methods could be based on, e.g., optimal interpolation (OI) or variational techniques. The latter could also be used in a 4-dimensional version, using the adjoint model technique. The purpose of all these schemes is to provide initial data for NWP models, that are as accurate as possible.

It is, however, also of interest to analyse other variables, that are not used in the initial state of the model. Examples of these variables could be the temperature at 2-m level, precipitation over some period and fog. It is essential for now-casting techniques to have these variables as gridded data shortly after observation time. In a system developed at the UK Met. Office, Nimrod (Golding, 1998), an automatic analysis/forecasting system, also analyses of such variables are done.

The Nimrod system also contains various methods for now-casting. Another system, producing analyses on these small scales is the LAPS-model (Albers et al., 1996).

At the Swedish Meteorological and Hydrological Institute (SMHI), work is undertaken on rationalisation of the meteorological service, aiming at a high degree of automation. This production is based on a database with gridded information on all relevant parameters. One part of that work is to produce that gridded information. Another part has been to replace the majority of the manual observations by automatic stations. The higher temporal resolution of these automatic observations, together with utilisation of remote sensing data from radars and satellites, give new possibilities for frequent mesoscale analyses. In addition to the ordinary automatic stations, the automatic station network is supplemented by a dense network of automatic stations supplied by the Swedish Road Authority (VViS).

To meet these new requirements, the Mesan

* Corresponding author.
e-mail: lars.haggmark@smhi.se.

system has been developed. This paper describes the essential parts of this work.

2. Method used

The optimal interpolation (OI) technique, has been widely used in meteorological applications, especially in NWP. A key element of the OI technique is the so called structure functions or background error correlation functions. Some effort has been devoted to model these functions for different parameters described below by using historical data. In OI, observations are normally used together with a background field, often referred to as the first guess field. Here, this field is a three or six hour forecast from the Swedish operational model, Hirlam (*H*igh *R*esolution *L*imited *A*rea *M*odel, for reference, see Källén (1996)). The version of Hirlam used during most of the development of Mesan, had a horizontal resolution of 55 km and 16 vertical levels, but in the current version we use 44 km and 31 levels. The integration area covers the area from Greenland in the north-west to the Middle East in the opposite corner. The analysis area is smaller and covers northern Europe. We will not go into details here about the analysis method, but refer to the literature (Daley, 1991).

In the operational setup of Mesan, the analysis is performed using a horizontally smoothed previous analysis as a first guess in cases of missing Hirlam data.

The first step in a numerical analysis is normally a quality control, in which observations with large errors are rejected. The method used here is a standard procedure (Lorenc, 1981), but in cases where the observations are correlated, the method does not work. This is the case for erroneous but consistent data, e.g., radar data in cases of anomalous propagation and the same problem can arise from misclassified satellite images during cold winter days. This is a difficult problem which has been approached by checking correlated data only to uncorrelated observations (e.g., synop). For visibility, a consistency control is done by comparing with the humidity, such as in cases of reported fog.

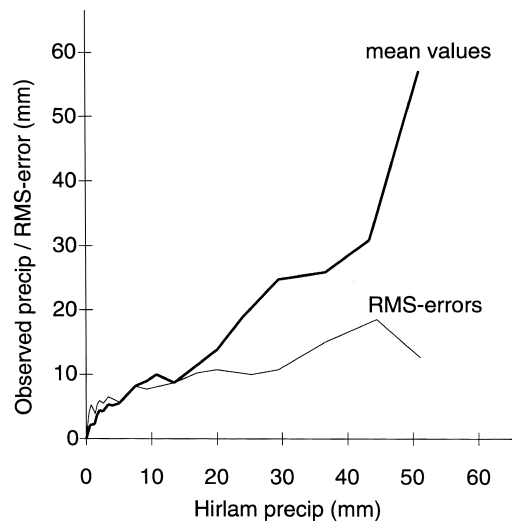


Fig. 1. Observed 24-h precipitation and RMS-errors as functions of the first guess (Hirlam) precipitation values. Heavy line = means, thin line = RMS-errors.

3. Precipitation analysis

Different observation sources for the precipitation analysis will be discussed, such as radar, automatic stations, synoptic stations and automatic stations from the Swedish Road Authorities (VViS). The possibilities of analysis over different accumulation periods will be presented. Fig. 1 shows the observed mean values (24-h precipitation) and the corresponding RMS-errors as a function of the first guess precipitation. The figure shows no large systematic errors in the first guess fields from Hirlam, and that the RMS-errors increase slightly with precipitation rate. This indicates that Hirlam is useful as a first guess.

3.1. Radar

Radar information is not straight-forward to utilise for the analysis of precipitation (T. Andersson, SMHI, personal communication, Riedl, 1995). The advantage is the superior resolution in time and space, and the two main disadvantages are as follows:

- Erroneous information due to non-representative echoes (bright bands, anomalous propagation).
- The curvature of the earth is different from

the curvature of the radar beam. Normally, the echoes used as precipitation information is measured increasingly higher up in the atmosphere, with increasing horizontal distance from the radar, i.e., a lot of precipitation is produced or modified below the echoes.

The problem of bright bands is treated by utilising temperature information from the Hirlam-model. If the reflectivity for the pixels at the height near the zero degree isotherm are higher than a specified value, it is assumed that this is coming from bright bands. In these cases, the reflectivity is reduced using climatological information from a lot of reflectivity profiles of bright bands.

The problem of anomalous propagation is more complicated, since such erroneous echoes can be quite strong and exist together with real echoes. By computing the refraction utilising the stratification and humidity information from Hirlam data, it has been possible to separate false echoes from real. It works well in some situations, but sometimes it fails, since the time and space resolution as well as the accuracy of the Hirlam information is not always sufficient. The approach adopted here is the following.

- If radial winds are present, omit pixels with no or unrealistic radial winds (from Doppler mode). Otherwise, check against other radars (overlapping areas) and if not consistent omit non-zero pixels.
- If many pixels are rejected, take away all pixels which have echoes.
- Create a composite picture, where, when possible, the omitted pixels can be filled with pixels from another radar.

The second problem, when the radar is getting its information too high in the atmosphere for large distances, has been approached by 3 simple corrections.

- If no echoes are seen on distances larger than 150 km (100 km during winter), this information is ignored (treated as no observations). Thus, precipitation observations originating from low level clouds are not suppressed by the radar information.
- A general correction above the zero-degree height, which increases the precipitation at larger distances.
- For the transformation of radar echoes to

estimated precipitation, a correction is added, which increases the precipitation rate in areas of orographic enhancement. Thus the vertical velocity due to orographical lifting $kV \cdot \nabla z$ is used as a correction factor. Here the proportionality factor k is a function of temperature, i.e., $k \propto e_s(T)$, where e_s is the saturation vapour pressure.

3.2. Automatic stations

Two different types of automatic weather stations are used. The first type is the new stations that have been replacing most of the synoptic stations. Those measurements are based on collecting the precipitation in a container, and the change of weight is transformed into precipitation rate. The second type is stations from the Swedish Road Authorities (VViS), which measure the precipitation optically. The VViS-Station network is quite dense, with about 400 stations measuring precipitation. The different observation networks are shown in Fig. 2.

There are very small differences between precipitation records from our automatic stations and traditional manual observations. Fig. 3 shows comparisons between measurements from automatic station and from nearby situated climate stations (mean distance is 11 km, the climate station network is described in Subsection 3.4). The correlation is 0.87. This is expected as natural variations with respect to the structure functions. Comparisons between VViS-measurements and traditional measurements at the same locations (the mean distance is approximately the same) show a correlation of 0.84. A complication is that VViS stations measure precipitation as mm of snow, and not in mm of water. Thus a translation between snow and rain amounts has been included (Fig. 4). The curve is based on both a similar curve from the literature (Gray, 1970) and modified by use of observation statistics from VViS, synoptic and climate station data. The reason for using the wet bulb temperature (T_{iw}) is that it discriminates snow from rain better than temperature.

3.3. Synop

A method of using the weather code (ww) to estimate the precipitation amount has been developed. This is done by estimating a relation-



Fig. 2. The observation networks. Upper left: manual synoptic stations; upper right: SMHI automatic stations; lower left: VViS stations; lower right: radar network coverage.

ship between the weather code and precipitation amount using statistics. Synop information is available every 3 h and precipitation amounts every 12 h, which means that four values comprise the predictor for a 12-h period.

First, a coarse estimate of the accumulated precipitation for each weather code was derived. Each time the weather code was reported, the 12-h precipitation was added to the sum of precipitation for that code. The sums were then divided by the number of observation of the different

weather codes. Using these mean values, the different weather codes were divided into 11 classes with similar values.

Secondly, a regression based method was used. For each 3-h term within the 12-h period, the predictor for a given class was increased by one if the class occurred. The predictand was the accumulated precipitation for the station. Linear regression was used to retrieve coefficients for each class. In solving the regression, the coefficients for a given class are the equivalent of the mean

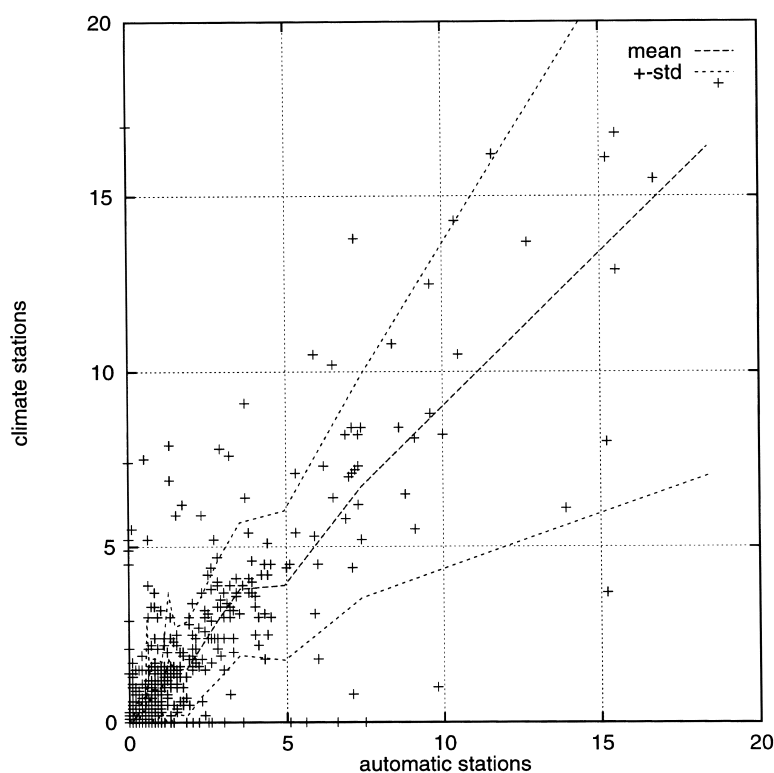


Fig. 3. Comparison between measurements (mm/24 h) from automatic stations and climate stations shown as a scatter plot. In the figure is also shown a running mean of automatic station data plotted against the corresponding values for the climate stations. The dispersion is indicated by lines for standard deviations.

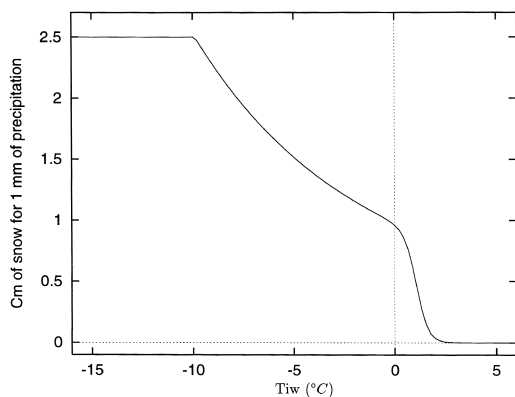


Fig. 4. Assumed amount of snow in cm for 1 mm of liquid precipitation as a function of wet bulb temperature.

precipitation for that class. The large number of data included in the regression, leads to estimates of mean values with a large significance. But since the RMS-error is about 60% of the estimated precipitation, the uncertainty for an individual case is rather large.

Table 1 shows the mean values of 3 h precipitation for different weather codes (ww in the WMO-code). For example, code number 61 (light persistent rain) is in the mean equivalent to 1.7 mm 3-h accumulated precipitation.

3.4. Structure functions

Since the accumulated precipitation is often composed from migrating weather systems, the area covered by precipitation increases with the time over which the accumulated precipitation is considered. The horizontal scale of the structure

Table 1. Mean values of 3-h accumulated precipitation (mm) for different WMO weather codes

Code number decade	Code number within decade									
	0	1	2	3	4	5	6	7	8	9
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	–	–
10	0.0	0.0	0.0	0.5	0.0	0.2	0.5	0.9	0.2	–
20	0.2	0.9	0.2	0.9	0.2	0.9	0.2	0.5	0.0	2.4
30	–	–	–	–	–	–	0.2	–	0.9	–
40	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.0	0.0
50	0.2	0.5	0.5	0.9	0.9	2.4	0.2	0.5	0.9	1.7
60	0.9	1.7	1.7	3.8	5.5	8.0	0.5	0.2	0.9	3.8
70	0.2	0.5	0.5	1.7	0.5	3.8	0.2	0.2	0.2	0.2
80	0.9	2.9	3.8	0.9	1.7	0.2	0.5	0.5	2.4	2.4
90	2.4	1.7	8.0	7.2	7.2	3.8	1.7	7.2	–	5.5

functions are thus dependent on the integration time. The estimation of the correlation function is based on observed auto correlations of first guess errors, followed by curve fitting. The following expression has been used:

$$\text{Corr}(r) = 0.5 \left[e^{-r/R} + \left(1 + \frac{2r}{R} \right) e^{-2r/R} \right].$$

The typical horizontal scale is given by R (110 km, 180 km and 270 km for 3, 12 and 24-h integration time). The function is plotted for different accumulation periods (Fig. 5).

3.5. Use of climatological information

SMHI has a very dense network of climatological stations, which provide 24-h precipitation amounts reported monthly, and thus they are not available in real time. Analysing these observa-

tions, it is easily seen that precipitation is very unevenly distributed, especially in convective situations but also because of orographic effects. At least the latter effects are in principle described by numerical models. Therefore this information could be included in the first guess, but only where precipitation is present in the model field.

It is not unusual that the largest precipitation amounts fall in mountainous areas where synoptic observations are totally or partly missing. If a first guess is not available or gives no precipitation in the area, an analysis based on the sparse observations will not reflect the climate of the area.

If the annual precipitation amounts are normalised by the standard deviation of the observed daily precipitation values at the stations, it can be seen that the derived values are more or less constant, independent of station. That means that an analysis carried out in the normalised values can be done assuming isotropy.

The uneven spatial climatological distribution of the standard deviations of precipitation has been used in the analysis in the following way:

- normalise both the first guess and the observations;
- perform the analysis in this normalised variable;
- do the inverse of the normalisation.

The advantage of the above described approach is that observations from places with little precipitation will not reduce the result in areas with large climatological precipitation.

To use this method we need to know the standard deviation of the precipitation both at observation sites and at grid points. We have

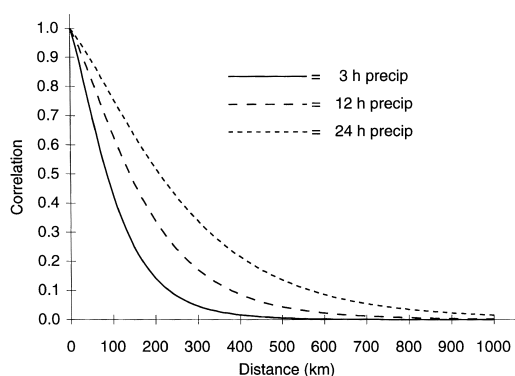


Fig. 5. Correlation functions for precipitation for 3, 12 and 24 h accumulation periods.

chosen to estimate those values by means of statistics. The predictand has been the observed standard deviation of daily precipitation values at the climate stations in Sweden.

Two of the 3 predictors have been chosen from physical considerations, topographical forcing and forcing from variations in friction.

- The frequency of wind directions multiplied by the corresponding upslope gradient of topography. To somewhat limit this effect for steep orography the predictor has been normalised by a logit function (an s-shaped function of the form $e^{ax}/(1 + e^{ax})$).

- The component of the gradient in roughness length perpendicular to the wind direction.

The latter predictor reflects the fact that the crossisobaric flow is enhanced by friction, and can create convergence, mainly in coastal regions, which enhances the precipitation.

The 3rd predictor has been latitude. A linear regression is done and about 50% of the variance of the predictand can be described by these 3 predictors.

Using the wind statistics in cases of precipitation (Alexandersson and Andersson, 1995), the result of the regression for 1994 is shown in Fig. 6 (left panel). Utilising this field as a first guess and the observed standard deviations as observations an analysis can be done (Fig. 6 right panel). Large climatological standard deviations are marked with dark shading and are associated with large climatological precipitation.

This climatological standard deviations can be used in two different ways, dependent on whether wind information is available or not. If wind direction is known (e.g., from the forecast model), the regression equation can be used to create a field with standard deviations which is meant to represent the climate of the current weather situation. In case of no wind information, it is possible to utilise the climatological field of Fig. 6 (right). The effect of using the normalisation method is shown in the following example, where two fictitious observations are producing two different analyses, both with zero first guess. In the right panel of Fig. 7, normalisation has been used but not in the left panel. As the figure illustrates, there is a large difference between the 2 analyses. The climatologically enhanced maximum at the west side of the topographically higher southern Sweden is clearly present in the right panel.

The method increases the quality of the analysis if the first guess:

- gives no precipitation,
- is missing.

But no extra skill is achieved when the spatial distribution of precipitation in the first guess is properly described. On the other hand, if the first guess overestimates precipitation amounts on a large horizontal scale, the method can lead to a minimum where a maximum is wanted. The fact that the NWP-model has a coarser resolution, could possibly be addressed by some kind of down-scaling of the first guess field, but this has not been done in the present study. It is however reasonable to assume that, with an increased resolution of the NWP-models, the climatological behaviour on the smaller scales will be better described (i.e., the orographically driven distribution).

3.6. The importance of different sources

In order to validate the quality of the different observational systems (Fig. 2), sequences of analyses with different observational input have been performed for a period of 6 months. The background fields are taken from Hirlam short range forecasts. The evaluation has been performed for 3 h and 12 h accumulated precipitation, respectively.

In the verification, only data from synoptic stations or from the automatic stations of SMHI have been used. In cases where these are used in the analysis we have utilised a cross validation technique, i.e., 95% of the data are used in the analysis, and the verification is done on the remaining 5%, and this procedure is permuted for the whole data set. In the verification, the analysed values have been interpolated to the observation points. Since this method compares area averages (from the analysis) with point measurements, some discrepancies can be expected.

Since cross validation is computationally expensive, we have restricted our study to the days when precipitation is present. Thus, 18 precipitation days during the period October 1995 to March 1996 are used. The verification area is southern Sweden, and cases with anomalous propagation (false radar echoes) are eliminated.

Fig. 8 shows the explained variance for 3-h and

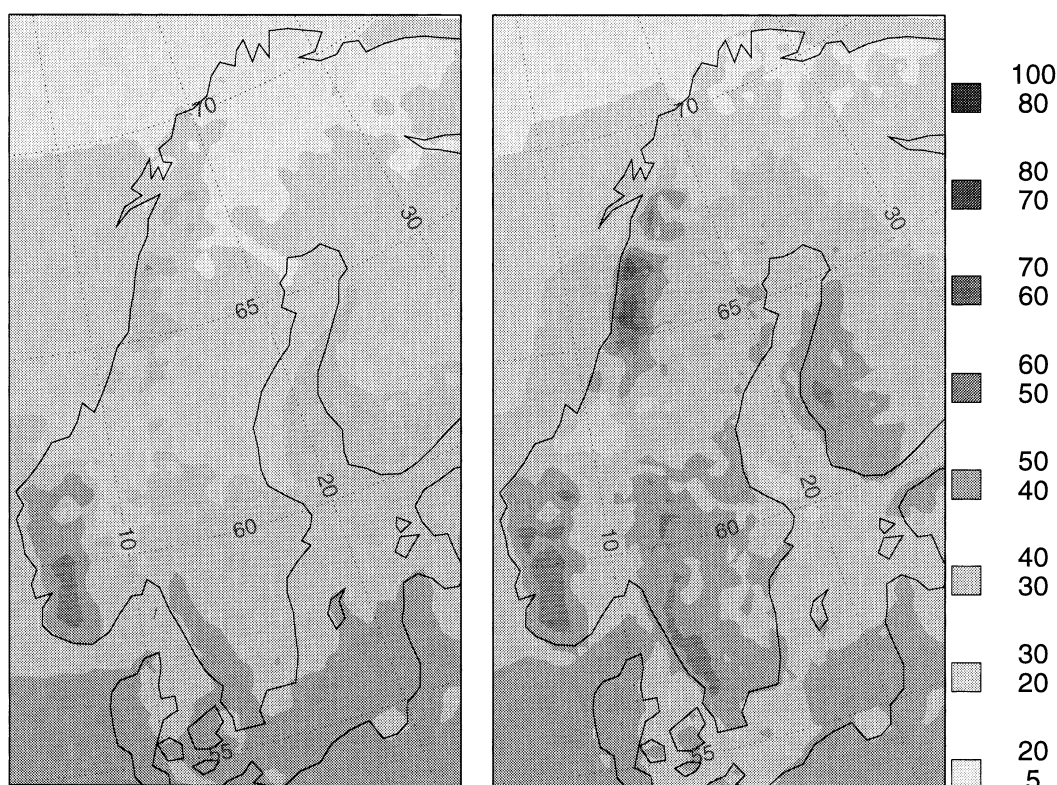


Fig. 6. Standard deviation of 24-h observed precipitation values (unit = tenth of mm). Left panel: from a regression relation; right panel: an analysis, using observed standard deviations and the values from the left panel as the first guess. Dark shadings mean large variations in the precipitation.

12-h accumulated precipitation respectively. The left figure indicates a large improvement for 3-h precipitation when radar information is included together with Hirlam. If the automatic stations are replacing the radar information, almost the same quality is achieved. It is also clear from the figure that a very dense network of measurements on the ground (auto + VViS) have a large positive impact on the result.

For 12-h precipitation (right panel), the quality is almost the same, as long as synop information is included. The improvements from VViS are marginal on a longer time scale. The general conclusion from this study is that radar information improves the result very little, in places where measurements at the ground are present. Over sea, the result could be different, since there are generally very few ground based measurements available. It is also evident that slightly less accu-

rate instruments arranged in a dense network can improve the analysis.

4. Analysis of temperature and humidity at 2 m level

Unlike precipitation measurements, measurement of 2-m temperature not only have the representativity problem between model grid square information and observations, but also an additional complication: The orography of the grid square differs from the real orography where the observations are.

When interpolating the first guess of temperature to the analysis grid, we take into account the vertical variation of temperature from the model state. This procedure has proven to give a small improvement during summer, but not during

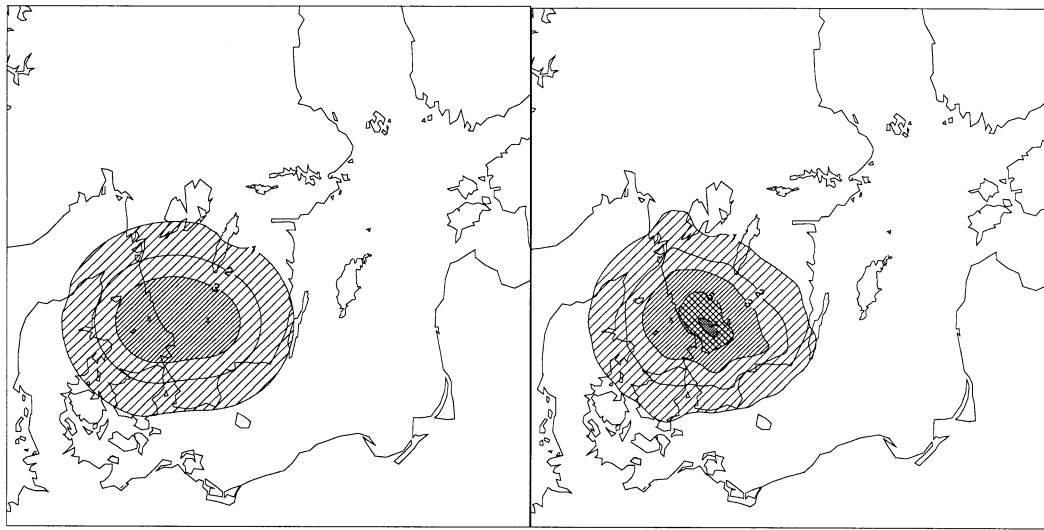


Fig. 7. Two analyses, with zero as first guess, and only two fictitious observations. Left panel: using constant first guess error; right panel: using the spatially varying first guess error from Fig. 6 (right).

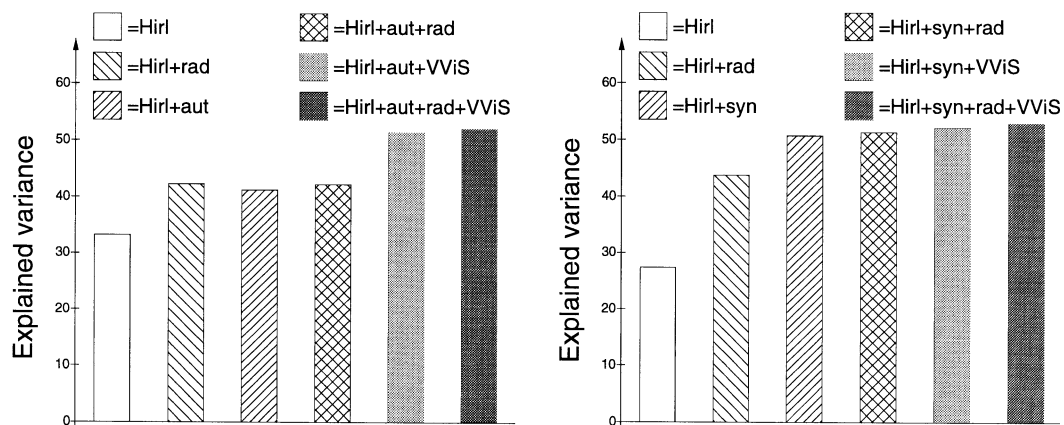


Fig. 8. Explained variance (in percent) for different sets of observations. Left: 3-h accumulated precipitation; right: 12-h accumulated precipitation.

wintertime. It is possible to use the same method when reducing the first guess to observation height above sea level, but practical experiments show no further improvements, mainly due to the problem of estimating low level stratifications.

Using the optimal interpolation technique, it is often assumed that the structure functions are isotropic, i.e., the first guess error correlation is only dependent on distance, regardless of direction. It is clear from statistical investigations that this is not the case for 2-m temperature. The

correlations are dependent on the land–sea contrast, and also on the difference of elevation.

To account for both the effect of different physiography and height, we modify the structure function according to:

$$\text{Corr} = \text{Corr}(r) * F_p(d_p) F_z(d_z),$$

where $\text{Corr}(r)$ takes the same mathematical form as for precipitation (Subsection 3.4), but here R is 190 km. The empirical functions F_p and F_z describe the behaviour due to difference of

land-fraction (d_p) and difference of height (d_z) respectively. Both these functions are linear, and vary from 1, for $d_p = d_z = 0$, until 0.5 for $d_p = 1$ and $d_z \geq 500$ m, respectively.

The impact of the modifications due to fraction of sea and height above sea level is shown in Fig. 9. Using independent observations (VViS-stations), it has been verified that the small scale features of the right panel is not only noise.

The 2-m humidity analysis is done in the variable relative humidity. Thus there is a temperature variation in the variable, and the same correction functions for height and physiography have been used.

The quality of the analysis has been verified using a cross validation technique for the period 10–15 of January 1999 using 22000 observations. For temperature the results shows an improvement over the first guess (Hirlam) from 79.2 to 92.0% explained variance. When physiography is influencing the structure functions there is an addi-

tional but slight improvement to 92.2%. The root mean square error decreases from 3.29 to 2.03 and to 1.98 degrees respectively. For relative humidity the improvement was from 13.6 (first guess) to 46.0 (analysis) and to 47.3 (analysis with physiography) % explained variance and a corresponding decrease in root mean square errors was found.

5. Analysis of wind at 10 m level

The wind analysis may be multivariate by using pressure observations as well as wind. This implies some kind of relationship between wind and mass field information, e.g., geostrophic wind or gradient wind. When a high horizontal resolution is of interest, these relations become doubtful, and moreover, small horizontal pressure differences will be difficult to distinguish from noise. Dynamically, it is also true that on smaller scales

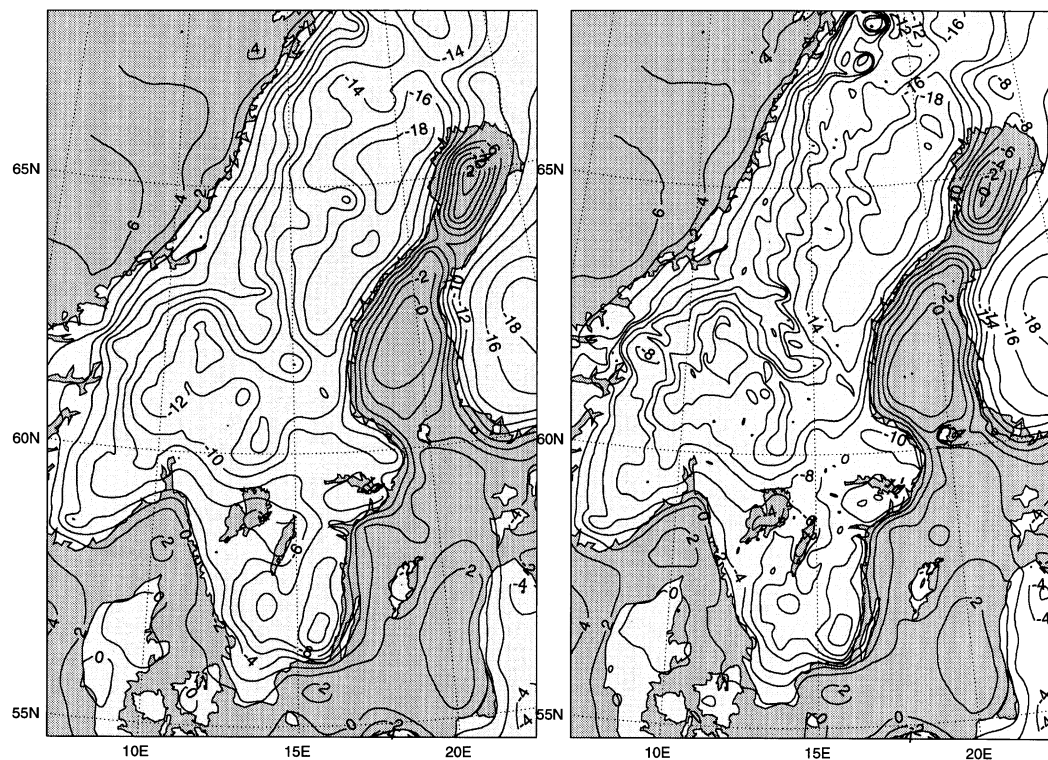


Fig. 9. A 2-m temperature analysis case study. Left: the structure function dependent on distance only; right: as left, but also dependent on altitude and fraction of land.

(mesoscale), the main information is contained in the wind field, and not in the mass field.

We have therefore decided to utilise only wind observations together with the first guess field, which is the 10 m wind from Hirlam. It is (like 2 m temperature) not a model variable, but has to be computed by post processing. The wind measurements from automatic stations and manual observations have been used. The automatic stations are all 10 m high, and observations from lighthouses etc. are properly reduced to 10 m level before they enter the analysis. Manual observations (estimations) are regarded as more uncertain and have been given a larger error.

It is not straight forward to derive the necessary structure functions, since the local effects like roughness and orography plays an important role for the wind at the 10 m level. This is clearly illustrated in Fig. 10 where the correlation (multiplied by 100) of the first guess error of the v -component of the wind, relative to three specific stations (correlation-value = 100) are shown. The data used to derive these structure functions are

3 months of wind measurements over Scandinavia, every third hour. It is notable that for the station in Denmark (left panel), the correlations are fairly high even on large distances, while the two other panels show large differences between land and sea stations.

5.1. Structure functions dependent on roughness

In numerical models, we utilise a roughness field, that is of vital importance for the exchange of momentum, heat and moisture between the earth surface and the atmosphere. The roughness that is used in Hirlam varies from a few mm over sea, to some meters over mountain regions. This so called orographic roughness is needed due to the lack of explicit parameterisation of “gravity wave drag”.

Here we have done a re-normalisation of the roughness in such a way that it is set to 0 for values below 0.2 m and to 1 for values over 6 m and a logarithmic behaviour in between, here called normalised roughness. These values have

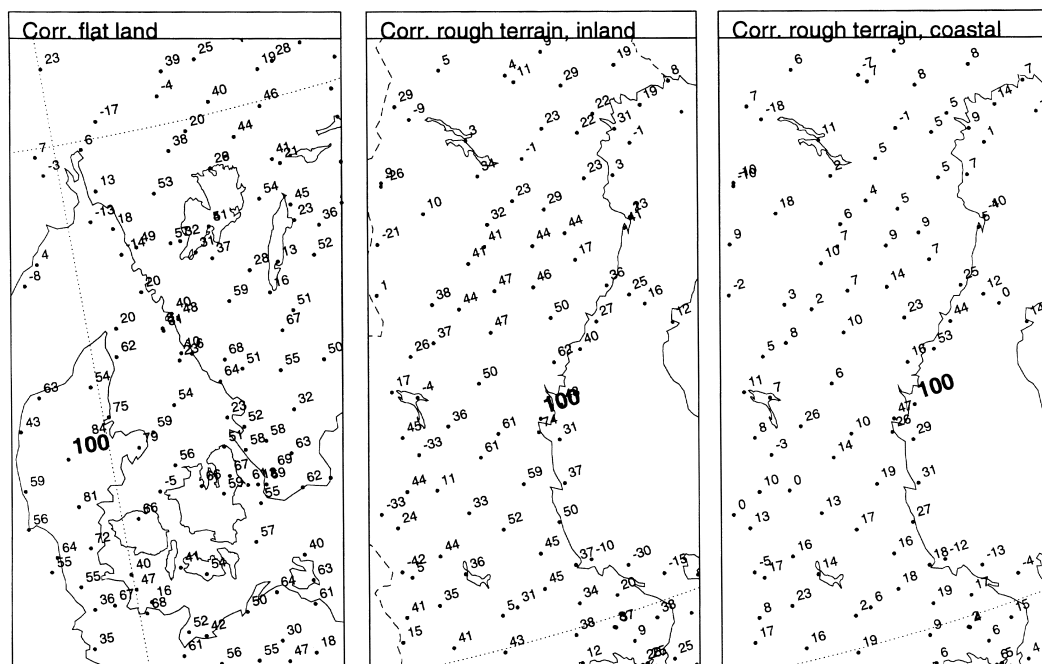


Fig. 10. The correlation of the first guess error for the v -component of the 10-m wind, multiplied by 100, for different reference stations: left: a station in the middle of Jutland; middle: the airport of Sundsvall; right: a coastal station near Sundsvall.

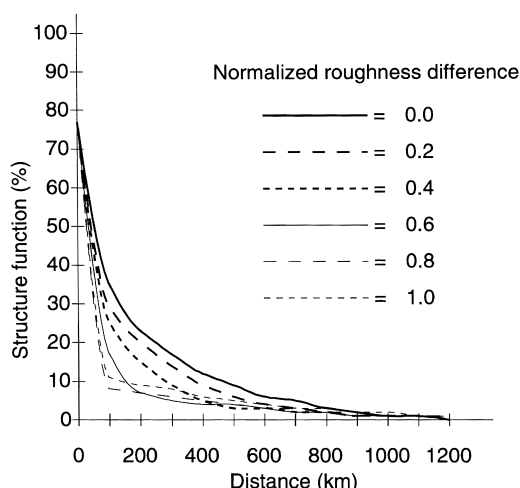


Fig. 11. Empirically derived structure function (units %) of the first guess error of the u -component of the wind, as a function of distance and difference in normalised roughness.

been used, together with a lot of data in a regression to define an empirical structure function (Fig. 11).

5.2. Validation of wind analysis

We have evaluated the wind analyses using cross validation as for the precipitation analysis. Here we have divided the data set into 10 subsets, and thus 90% of the observations are validated against 10% of independent data each time. The evaluation is based on eight randomly chosen cases from January 1997. The first guess was 6 h HIRLAM with 55 km and 22 km grid spacing respectively and the same resolution is used for the analysis as for the first guess. The values in Table 2 are referring to wind velocities. The result is a moderate improvement relative to the HIRLAM

Table 2. The bias, mean absolute error and the RMS-error (m/s) of the first guess and analysis of wind velocity; the number of observations are 1668

Model	Bias	NAE	RMS
HIRLAM 55 km	-0.06	1.76	2.53
MESAN 55 km	-0.50	1.47	2.26
HIRLAM 22 km	0.12	1.81	2.59
MESAN 22 km	-0.42	1.48	2.27

first guess field. The overall result is not significantly different for the high resolution model (22 km) compared to the coarser one (55 km). The reason for the negative bias in the analysis is unknown.

6. Visibility analysis

Analysed visibility can be useful for shipping forecasts and for aviation. Visibility is not a physical variable, but dependent on different phenomena, like aerosol concentration, humidity and precipitation. In Sweden, the humidity can alone describe about 80% of the variations in visibility. This relation is somewhat different for observations from automatic stations than for the manual synoptic ones (Fig. 12). The similarity is larger for small values of the visibility, when the analysis is of interest. For larger values the automatic station observations are modified using the relation between relative humidity and visibility for manual stations.

About 40% of the variations in visibility, that can not be described by the humidity information, can be explained by adding a parameter P_{type} which is 1 for rain, 2 for snowfall and 0 for no precipitation. This simple improvement increases the variance explained to about 90%. By regression, the following formulae for the first guess of visibility (Vis) have been derived.

In cases of no precipitation:

$$Vis = 1.32f(rh) - 14\,361.$$

In cases of precipitation:

$$Vis = 1.11f(rh) - 4970P_{type} - 470R - 1100.$$

Here $f(rh)$ is a function of relative humidity (arrived at by curve fitting) defined as:

$$f(rh) = 1000[7.58 + 122.2(1 - rh) - 100(1 - rh)^2],$$

and P_{type} varies between 1 for rain and 2 for snow. R is the precipitation rate (mm/3h), and the determination of P_{type} is done by utilising the wet bulb temperature T_{iw} :

$$P_{type}(T_{iw}) = 1 + \frac{1 - e^{3.5(T_{iw} - 274.3)}}{1 + e^{3.5(T_{iw} - 274.3)}}.$$

We are using T_{iw} instead of normal temperature, since it is a better discriminator between rain and snow. The formula for P_{type} is derived from

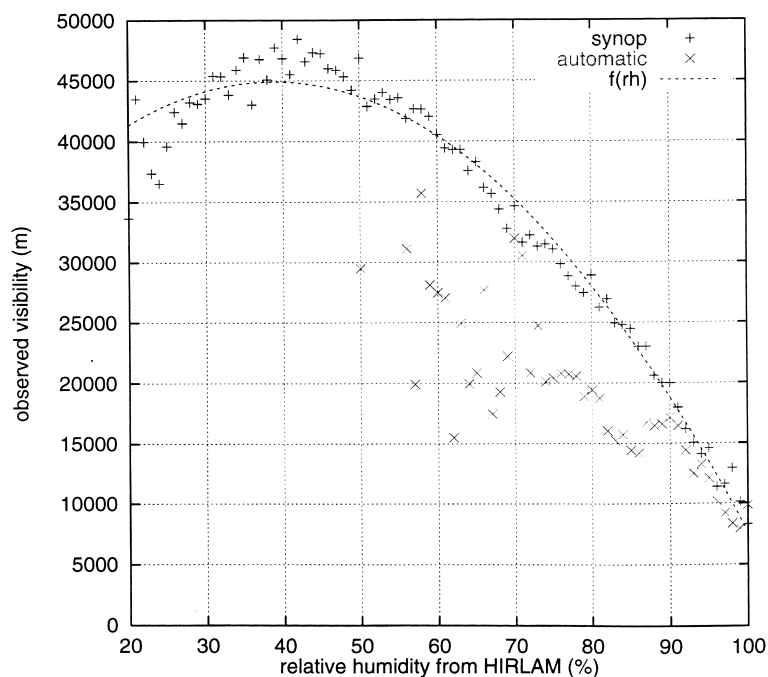


Fig. 12. Mean values of observed visibility (synoptic and automatic stations) as function of forecast relative humidity (%). The dashed line is $f(rh)$, see text.

curve fitting to observed relative distributions of snow and rain. Fig. 13 shows 4 different maps analysed humidity, precipitation, the first guess field using the formulae described above, and finally the corresponding visibility analysis. A comparison between the first and the fourth panel illustrates that most of the information is in the humidity field. By comparing panels 2 and 3, we can see that the precipitation can increase the visibility in cases of light rain.

7. Cloud analyses

For the analysis of clouds, it is obvious that satellite information is a very important source. Here, multi-spectrally classed pictures, based on polar satellites (NOAA) is the most important (Karlsson 1997). Pictures from the METEOSAT geostationary satellite are also used. Since the NOAA satellite pictures have a much higher spatial resolution and more spectral information but coarser temporal resolution than those from METEOSAT, the two sources are used in different

ways. It is worth mentioning that observations that differ more than 60 min from analysis time are not used in the analysis. Other observation sources are synop, metar and automatic stations. The first guess was, as before, the Hirlam model at 55 km resolution, which uses a condensation scheme with explicit cloud water (Sundqvist et al., 1989; Sundqvist, 1993).

The cloud analysis consists of the analyses of cloud base, total cloud cover, amount of low level clouds and top of clouds.

A problem that enters when doing an analysis of cloud base and cloud top is that the variable is not defined everywhere. We have, therefore included a cloud/no cloud analysis.

7.1. The use of METEOSAT data

It is not straight-forward to determine the presence of clouds, when the only information source is the IR-channel in METEOSAT. The characteristics of the METEOSAT IR channel can be found in EUM UG 03, a publication from Eumetsat. Here we have used a statistical relation between

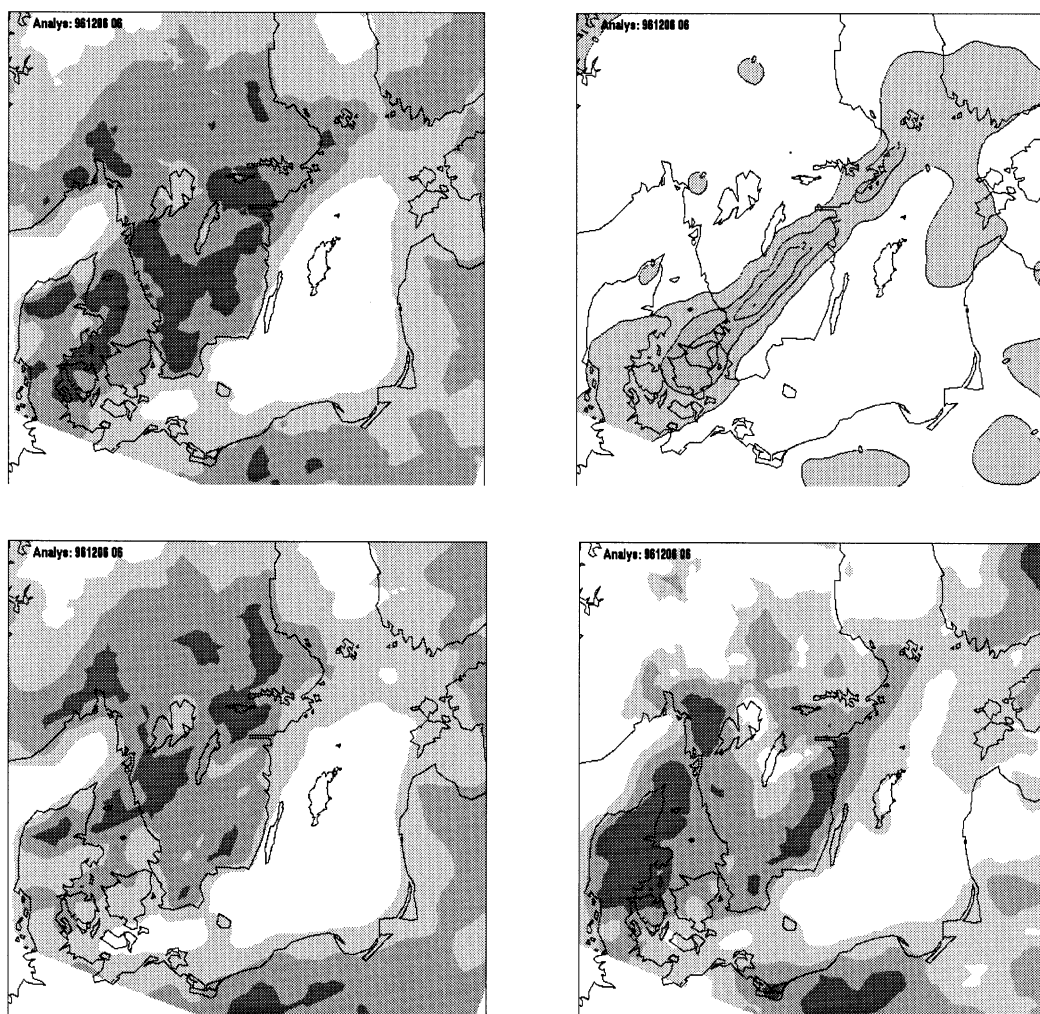


Fig. 13. An example showing different analyses related to visibility; upper left: relative humidity analysis, above 98 (darkest), 95–98, 90–95 and below 90% respectively. Upper right: 3-h accumulated precipitation, intervals 0–1 mm, 1–2mm and above 2 mm (darkest); lower left: first guess visibility; lower right: visibility analysis, intervals below 1 km (darkest), 1–4 km, 4–10 km and above 10 km.

the brightness temperature (the effective radiation temperature of a black body) from the IR-channel in METEOSAT, and the MESAN-analysis 2 m temperature during cloud free conditions. Information from the Hirlam model about the stratification in low levels are also used in some cases.

Fig. 14 shows the brightness temperatures from METEOSAT, during cloud free situations as com-

pared to our analysed 2-m temperatures. A regression line, plus and minus a running average of the standard deviation of the difference is marked with thin solid lines. When the difference between the 2-m temperature and the brightness temperature is more than two standard deviations from the mean difference between the two temperatures, it is assumed that clouds are present. The reverse is not necessarily true, since the temperature of

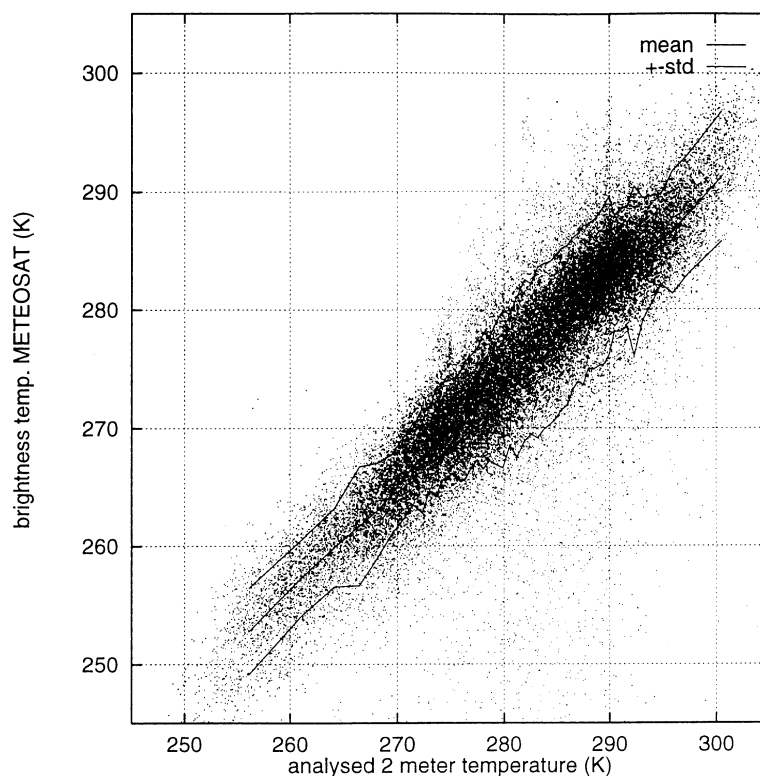


Fig. 14. A scatter plot of the analysed 2 m temperature versus the brightness temperature of METEOSAT. The data are from the period January to November 1996 (45,000 values). Only clear sky condition data are used. The RMS-deviation is 8.5° and the correlation is 0.83.

the clouds may be the same as the 2-m temperature. Therefore we have used information about the stratification from Hirlam in those cases.

In case of a strong inversion near the ground, the temperature near the 2 m level can also be observed far above the inversion, but since clouds normally prevent the creation of inversions, it is regarded as a cloud free case. In neutral or unstable situations, all clouds should be colder than the 2-m temperature, and we assume no clouds. If, on the other hand, the stratification is marginally stable in low levels (a lapse rate of $0.3\text{--}0.7$ degrees/100 m), it is considered as cloudy.

7.2. The use of NOAA data

We have used our operational cloud classification scheme, SCANDIA, (Karlsson, 1996) which is based on NOAA. SCANDIA is a multi-spectral scheme which also utilises the horizontal structure

to distinguish between the different cloud types. Generally this information is of a good quality. If the SCANDIA picture is not available, the IR pictures are used instead, and treated in a similar way as the IR data from METEOSAT.

Since the SCANDIA model has problems for low sun elevations it is not used when the sun angle is between 2 and 6 degrees and in cases with very skew observation angles.

7.3. Total cloud cover

The synop and metar observations of total cloud cover are generally of good quality, but during night time the quality of SCANDIA is better for higher clouds (Karlsson, 1997). Automatic stations can not measure clouds over 3800 m, and thus the use of such stations give an underestimation of total cloudiness in some situations. Fig. 15 shows two cloud analyses, valid at

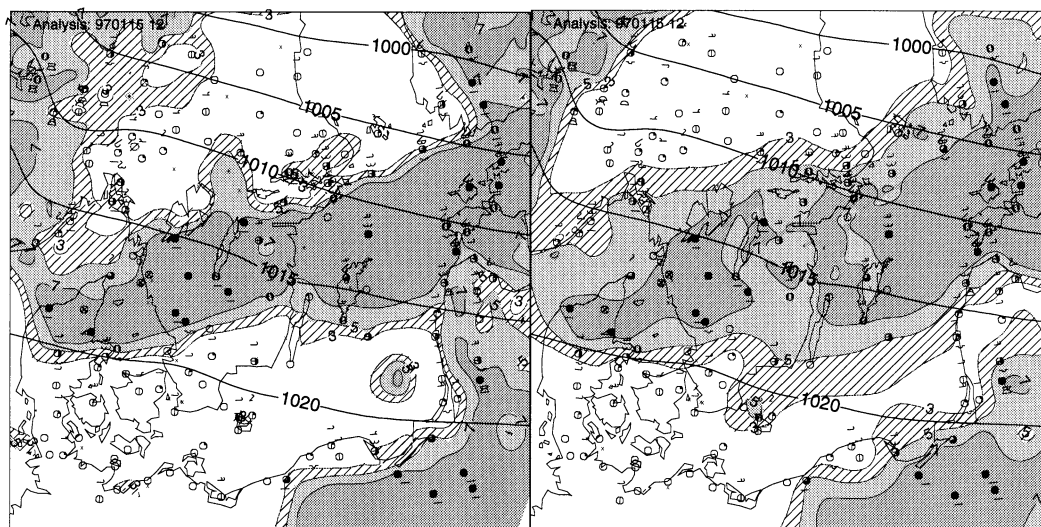


Fig. 15. Example of cloud cover analysis, using synoptic and automatic station observations as well as satellite information from a classified NOAA-image (left) or IR-data from METEOSAT (right). Full lines are isobars in hPa.

the same time, one using NOAA/SCANDIA and the other IR-data from METEOSAT. The smaller scale of the clouds in the NOAA-based analysis is clearly visible in the figure.

7.4. Quality control of cloud observations

We know there are some problems especially with cloud observations from automatic stations and with the classification of clouds from satellite data. Therefore, a quality control of the observational data is necessary. In doing the control, we must take into account that the observation errors from the different types of data can be internally correlated and that we therefore should check observations only against independent data.

Fortunately the two data sets are complementary such that automatic stations only have large problems with high clouds (the instrument does not reach above 3800 m) and the satellite data classification algorithm has difficulties separating low clouds from the surface (especially during winter with not much visible light and strong inversions). To take these characteristics of the observations into account, the quality control is done in the following way.

- Accept manual synop and metar observations as correct. This is done because the quality of the observations are considered acceptable, and the

network of manual stations in Sweden are too sparse to make a spatial consistency control meaningful.

- Accept observations from some classes (Nimbostratus, Cumulonimbus and thick Cirrus) from NOAA/SCANDIA as correct.

- Accept observations from IR satellite images as correctly classified cloudy if brightness temperature is 35° (METEOSAT) or 42° (NOAA) colder than the 2-m temperature.

- Check automatic station data, according to the method mentioned in Section 2, using already accepted observations. Hopefully, erroneous clear sky observations, can be eliminated when actual high or middle high clouds are identified from satellite data.

- Check not yet accepted satellite data using already controlled observations. This step is important to eliminate wrongly classified low clouds using information from ground based observations.

7.5. Significant cloud base

This is a variable, mainly used for aviation purposes, and it is defined as the lowest level, where we have a cloud cover of more than three octas. One complication is that the variable is not continuous (not defined in areas with less than

three octas) and another problem is that low level cloud forecasts from Hirlam are rather poor in many situations. To partly compensate for the so called “spin up” problem (Karlsson, 1996), we lowered the criteria for “significant cloud base” from $3/8$ to $2/8$ of cloud fraction. The “spin up problem” is a tendency of having too little clouds in the Hirlam forecast for short forecast projections. Boundary layer clouds are specially difficult to forecast. Some improvements have been made by utilising the boundary layer humidity to produce clouds. This is done by a method developed for cloud base forecasting (Bergeås, 1985), which mixes the air in the boundary layer. The boundary layer height is computed as in Holtslag and Boville (1993).

Both synop/metar and automatic stations are used. The observation error is different for measured and estimated cloud bases. The limitation of 3800 m in cases of automatic stations is not very

severe, since such high cloud bases often are of less interest. The satellite information is used schematically, since pixels that are classified as stratus or stratocumulus are given a standard value of 300 m. This is used only over data sparse regions like the sea, and the observation error is larger in these cases than those of synop/metar.

The areas where significant clouds are present are determined by a separate analysis procedure.

- Set all observations to one if a significant cloud base is observed and to zero elsewhere.
- Make a first guess field by doing a transformation of the analysis of total cloudiness, i.e., put this field to unity where total cloudiness is exceeding $3/8$ and to zero elsewhere.
- Perform the analysis

Fig. 16 shows the final result, where the areas of non-significant cloud base are unshaded.

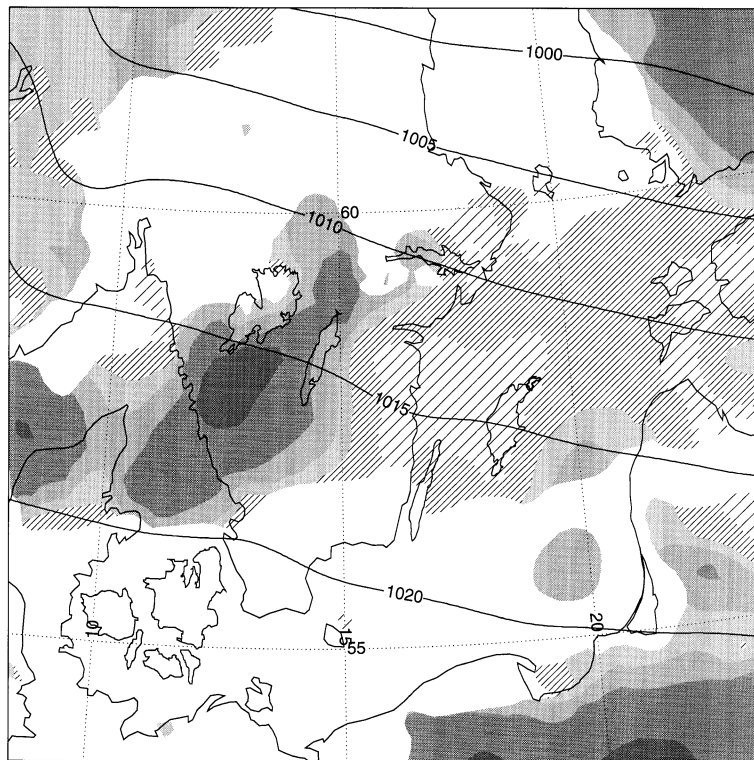


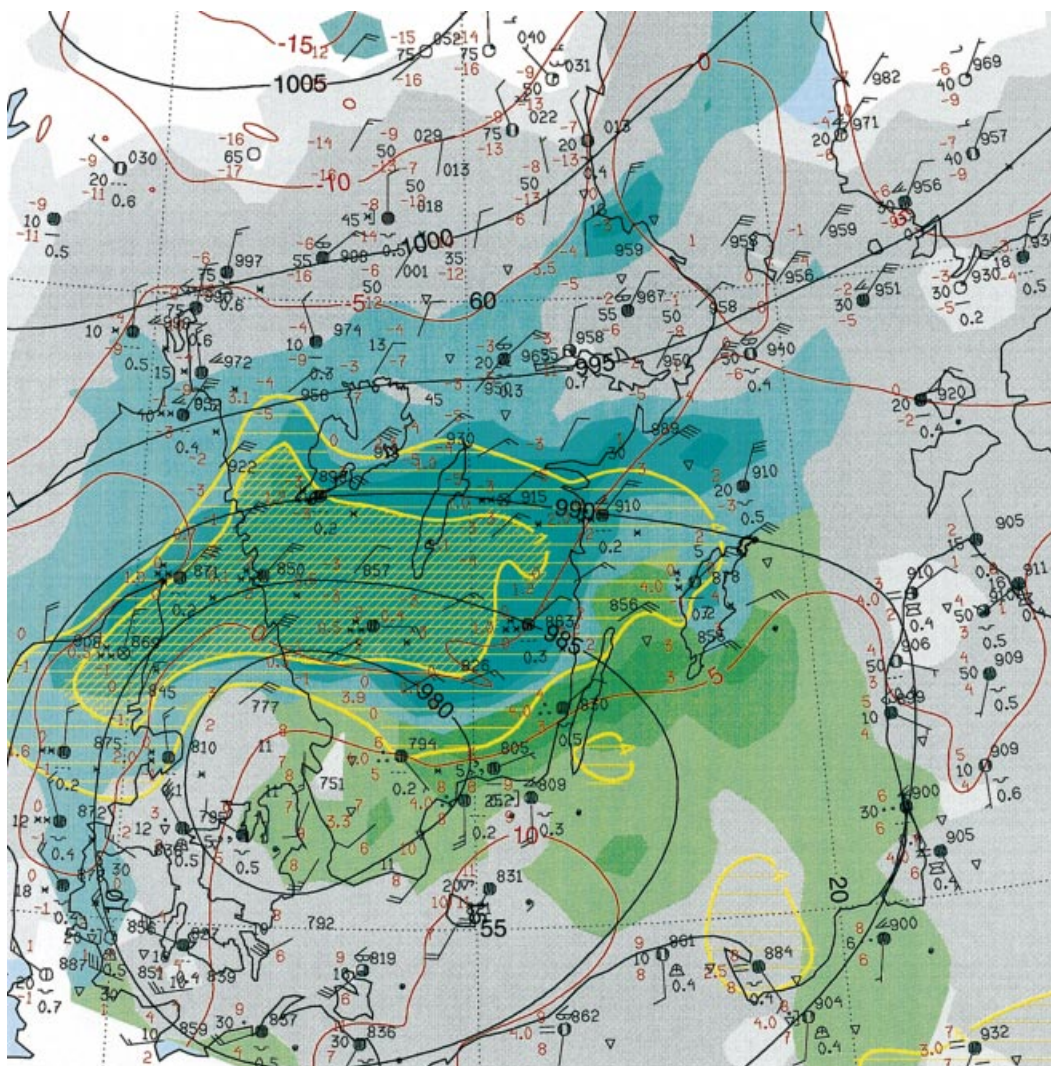
Fig. 16. An analysis of significant cloud base. The shaded areas corresponds to cloud bases less than 100 m (darkest), 100–300 m, 300–600 m, and 600–1000 m. Values 1000–3000 m are indicated with dense diagonal lines and values above 3000 m with coarser lines. Full lines are isobars in hPa.

8. Operational status

We are presently (since October 1996) producing mesoscale analyses every hour, and the results are presented for the operational forecaster as maps with a lot of information on each picture. Therefore it is necessary to use coloured fields, and here we have used a palette similar to that of the old hand analyses, when possible. An example of the operational presentation is shown in Fig. 17.

The case shown is a classical situation (1995-11-17) where there was a severe snow fall in western Sweden. We note the correlation between the visibility and the snow fall pattern.

As mentioned in the introduction these analyses are not directly used in NWP, but besides the maps given directly to the forecaster, of course all fields are stored as gridded information to be used in other applications. An example of such an application is to produce initial information



for now-casting purposes, like a system of 1-dimensional models, in each grid point (Gollvik and Olsson, 1995). Other applications are to create input data to runoff-models (Lindström et al., 1996), and for direct use by the road authorities.

9. Summary and conclusions

An operational mesoscale analysis system has been developed. It is based on optimal interpolation, and most of the work has been devoted to estimate structure functions, and to identify and compensate for erroneous observations. The general idea has been to produce gridded information, also of variables that are not used directly in NWP-models, but also for other purposes like

now-casting. The analysis has been done for precipitation, 2-m temperature and humidity, wind at 10 m level, visibility and clouds. In all analyses, we have tried to use all available information, e.g., for precipitation we have used not only synoptic and automatic stations, but also radar data, a dense network of present weather sensors from the road authorities, and climatological information, where wind direction and orography also are utilised. The Hirlam model has been used as first guess and also for interpreting some remote sensing data. So far, the operational experience from the system is encouraging, and we believe that this is a good and necessary step towards a rational treatment of an increasing amount of high-frequency types of observations.

REFERENCES

- Albers, S. C., McGinley, J. A., Birkenheuer, D. L. and Smart, J. R. 1996. The local analysis and prediction system (LAPS): analysis of clouds, precipitation and temperature. *Weather and Forecasting* **11**, 273–287.
- Alexandersson, H. and Andersson, T. 1995. Precipitation and thunderstorms. In: *Climate, lakes and rivers*. National Atlas of Sweden, edited by Bra Böcker, Höganäs, Sweden.
- Bergeås, L. 1985. *A bulk model for the unstable planetary boundary layer over the sea. A sensitivity investigation*. Report DM 51, Dep. of Meteor., University of Stockholm, Sweden.
- Cramer, H. E. 1967. Turbulent transfer processes for quasi-homogeneous flows within the atmospheric surface layer. In: *Boundary layers and turbulence. Phys. Fluids (Suppl.)*, 240–246.
- Daley, R. 1991. *Atmospheric data analysis*. Cambridge University Press. ISBN 0-521-38215-7.
- EUM UG 03, 1995. *Meteosat high resolution and WEFAX imagery. User guide*. Eumetsat, December 1995. Available from EUMETSAT (www.eumetsat.de).
- Golding, B. W. 1998. Nimrod: a system for generating automated very short range forecasts. *Meteorol. Appl.* **5**, 1–16.
- Gollvik, S. and Olsson, E. 1995. A one-dimensional interpretation model for detailed short-range forecasting. *Meteorol. Appl.* **2**, 209–216.
- Gray, M. 1970. *Handbook on the principles of hydrology*. Water Information Centre, Inc. ISBN 0-912394-07-2.
- Holtzlag, A. A. M. and Boville, B. A. 1993. Local versus nonlocal boundary layer diffusion in a global climate model. *J. Climate* **6**, 1825–1842.
- Häggmark, L., Ivarsson, K. I. and Olofsson, P. O. 1997. *Mesan, Mesoskalig analys, RMK 75*. SMHI, Norrköping, Sweden (in Swedish, available from SMHI, 601 76 Norrköping, Sweden).
- Karlsson, K. G. 1996. Validation of model cloudiness using satellite-estimated cloud climatologies. *Tellus* **48A**, 767–785.
- Karlsson, K. G. 1997. Cloud climate investigations in the Nordic region using NOAA AVHRR data. *Theor. Appl. Climatol.* **57**, 181–195.
- Källén, E. 1996 (ed.). *Hirlam Documentation Manual, system 2.5* (available from SMHI, Norrköping, Sweden).
- Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S. 1996. Development and test of the distributed HBV-96 model. *J. Hydrology* **201**, 272–288.
- Lorenc, A. C. 1981. A global three-dimensional multivariate statistical interpolation scheme. *Month. Wea. Rev.* **109**, 701–721.
- Riedl, J. 1995. Examples of improvements to clutter suppression in current operational weather radar systems. COST 75 Weather Radar Systems, Ed. C. G. Collier. International Seminar, Brussels, Belgium, 20–23 Sep. 1994. EUR 16013 EN. ISBN 92-826-9576-X, 114–124.
- Sundqvist, H., Berge, E. and Kristjansson, J. E. 1989. Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.* **117**, 1641–1657.
- Sundqvist, H. 1993. Parameterization of clouds in large scale numerical models, In: *Aerosol-cloud-climate interactions*, ed. P. V. Hobbs. Academic Press, Inc., 175–203.