VALIDATION OF HIRLAM CLOUD FORECASTS USING SATELLITE-DERIVED CLOUD COVER STATISTICS

Karl-Göran Karlsson
SMHI
Norrköping, Sweden

1. INTRODUCTION

The retrieval of cloud parameters from high resolution satellite imagery is a possibility that is offered by data from many satellite sensors on todays satellite platforms. Several retrieval algorithms have been proposed through the years applied to data from the AVHRR instrument (Advanced Very High Resolution Radiometer, see *Lauritzon et. al.*, 1979) on the polar orbiting NOAA satellites. One example is the algorithm proposed by *Arking and Childs* (1985) and several successors have also been introduced as operational cloud classifications schemes (*Derrien et. al.*, 1993, *Saunders and Kriebel*, 1988, and *Karlsson*, 1989). This paper discusses the possibility of using the cloud information derived from one such operational scheme; the SCANDIA model (*Karlsson and Liljas*, 1990), to verify cloud forecasts from the HIRLAM model (*Gustafsson*, 1991). Cloud cover statistics from SCANDIA for an entire month (August 1994) are compared with cloud forecasts of varying lead times to investigate the HIRLAM models ability to describe general cloud conditions. The quality of both the satellite and the HIRLAM information is discussed in the paper.

2. CLOUD INFORMATION FROM SCANDIA

2.1 The SCANDIA model

The idea of an operational cloud analysis scheme at SMHI, based on digital AVHRR data, was proposed in the early 1980s (Liljas, 1981). Several years of spectral signature studies followed where different cloud and surface types were examined. Typical reflectivity, transmissivity and emissivity characteristics were monitored for each cloud and surface type depending on varying solar elevations, airmass changes and different AVHRR sensors (satellites NOAA-9, NOAA-10, NOAA-11). After defining the main principles for an operational cloud classification method (*Karlsson*, 1989), the final classification scheme, named SCANDIA (the SMHI Cloud ANalysis model using DIgital AVHRR data), was implemented in 1988. Only a brief description of the main characteristics of the cloud classification model is given here. A full description can be found elsewhere (*Karlsson and Liljas*, 1990 or *Karlsson*, 1994).

The classification model makes use of calibrated and geometrically transformed imagery from measurements in five spectral AVHRR bands at maximum horizontal resolution (at nadir 1.1 km). AVHRR scenes are classified by using seven image features derived from the data from all AVHRR channels. Classifications are made in two predefined areas, covering the northern and southern parts of the Nordic area. Each pixel in a scene is classified into one of 23 cloud and surface types. A central image feature is based on AVHRR channel three brightness temperatures related (subtracted) to brightness temperatures in AVHRR channel four. This feature plays a major role in the cloud/no-cloud discrimination, during both day and night. The strength of the feature is to enable discrimination of clouds and snow surfaces during day and to map low and mid-level water clouds during night despite small temperature differences between the surface and the cloud tops.

SCANDIA is a supervised thresholding model, where typical class domains are defined by hyperboxes in a seven-dimensional feature space. A unique set of thresholds is defined for each of several predefined categories related to existing illumination conditions, weather types and satellites. The season (one of the set "Summer, Spring, Winter, Autumn"), the current sun elevation (determined by one of 12 defined sun elevation intervals) and the satellite (at present NOAA-11 or NOAA-12) specify which category and thus which set of thresholds that is used to classify the scene in each of the two areas. In total, 96 different categories are necessary to adapt the cloud classification model to the seasonal and daily changes of classification conditions. A more direct dependence on the current weather situation is furthermore applied by using temperature information at the 700 hPa and 500 hPa levels. This information is taken from meteorological objective analyses and is used for threshold tests of brightness temperatures from AVHRR channel four. In practice, the separation of thick low-, mid- and highlevel clouds are guided by this temperature information. Detection of thin Cirrus clouds is based on other image features.

2.2 Quality of cloud cover information from SCANDIA

A verification of SCANDIA estimations of fractional cloud cover was carried out using hourly cloud observations from 12 SYNOP stations in Sweden during a two-year period between June 1991 and June 1993. Total fractional cloud cover was computed from cloud classification images for pixels at and in the close vicinity to the location of selected SYNOP stations. The used maximum distance from the exact location of the SYNOP station was different depending on the treated cloud types (low-, mid- or highlevel clouds). This was done to compensate for the differences in viewing geometries for the ground observer and the satellite. Furthermore, a maximum time difference of 20 minutes between satellite and SYNOP observation was allowed. More details on the verification experiment are given by *Karlsson* (1993) and *Karlsson* (1994_2).
A way of summarizing the verification results is to divide the material into groups having different absolute errors. We define four such groups: 0 - 2 octas, 2 - 4 octas, 4 - 6 octas and 6 - 8 octas. The first group may be seen as the group representing quite an acceptable error. This can be concluded when considering possible sources of error related to the used comparison method and the inherent errors of the SYNOP observations. The two last groups would then contain the worst cases, i.e. when cloudiness is severely overestimated in almost cloudfree conditions or when cloudiness is severely underestimated in cloudy conditions. Tab. 1 shows

the verification results for the following six situations:

a) Sun elevations above 10°.

b) Sun elevations between -5° and 10°.

c) Night conditions.

d) Summer conditions (15 May - 15 September).

e) Winter conditions (15 November - 15 March).

f) All data (15 May 1991 - 20 June 1993).

Tab. 1 Total distribution of absolute error categories, in % of all studied cases, for conditions **a-f** *(explained in text).*

Absolute error category	a)	b)	c)	d)	e)	f)
0-2 octas	86.0%	78.3%	78.5%	83.6%	78.1%	81.4%
2-4 octas	10.5%	14.0%	13.9%	12.8%	13.0%	12.6%
4-6 octas	3.1%	5.8%	5.8%	3.1%	6.5%	4.7%
6-8 octas	0.4%	1.9%	1.8%	0.5%	2.3%	1.3%
Σ samples	9 352	6 764	7 283	7 164	8 658	23 399

We can see from Tab. 1 that the best agreement between satellite and SYNOP observations is found for sun elevations above 10° and for summer conditions. Here, differences less than two octas can be found in more than 80% of all cases. Results are worse during winter, at night and at sunrise and/or sunset. We may conclude that SCANDIA cloud classifications seem to be excellent during summer with the sun above the horizon while the number of large errors increases noticeably for the rest of the year. However, SCANDIA results are still found quite usable also during winter and dark conditions.

2.3 Monthly cloud frequencies in 1993

The encouraging results from the verification study initiated an experiment trying to use SCANDIA results to compute mean cloudiness month by month for the entire year of 1993. All available satellite passages, received in 1993 at SMHI in Norrköping, with minimum satellite zenith angles below 45° at reception, were selected for this study. Generally, this meant that four satellite scenes were used each day to describe cloud conditions in Sweden and its close surroundings. The chosen satellite scenes, with the present satellites NOAA-11 and NOAA-12, roughly describe cloud conditions during early morning, morning, afternoon and evening. The sun-synchronous orbits of the two satellites guarantee at least one passage with satellite zenith angles below 45° within the approximate time-windows shown in Tab. 2.

Tab. 2 Approximate time-windows valid for the used satellite scenes from NOAA-11 and NOAA-12 in 1993. At least one passage with low satellite zenith angles is guaranteed within a time-window each day.

Time of day	Time-window (UTC)	Satellite
Early morning	03:00 - 05:00	NOAA-11
Morning	06:30 - 08:30	NOAA-12
Afternoon	13:00 - 15:00	NOAA-11
Evening	16:30 - 18:30	NOAA-12

Cloud analyses from SCANDIA were treated in the following way:

a) Each pixel in the high resolution classification image was labelled cloudy or cloudfree depending on the resulting cloud and surface types. Pixels classified as subpixel clouds were treated as cloud-free.

b) A resampling of the result image by a nearest neighbour resampling technique was done to reduce the nominal horizontal resolution from one to four km.

c) Classification results for the two areas (mentioned in Section 2.1) were merged into one result image.

d) Cloud frequencies for an entire month were finally estimated by calculating the fraction of the total available scenes where the pixel was labelled as cloudy. For example, 20% means that the pixel was classified as cloudy in 20% of all used images in a month (normally approximately 120 images).

Almost 90% of all theoretically available satellite scenes were utilized to create analyses of monthly cloud frequencies. The final results, also including analyses of the diurnal variation of cloudiness and comparisons with similar information derived from SYNOP, are published in two papers by *Karlsson* (1994_1 and 1994_2). Here, results are presented for one summer month (May 1993) and one winter month (November 1993) in Figs 1 and 2 to illustrate the conclusions of the experiment. Once again, it was found that satellite-derived cloudiness almost exactly reproduced SYNOP-derived cloudiness during summer while a slight underestimation was found during winter.

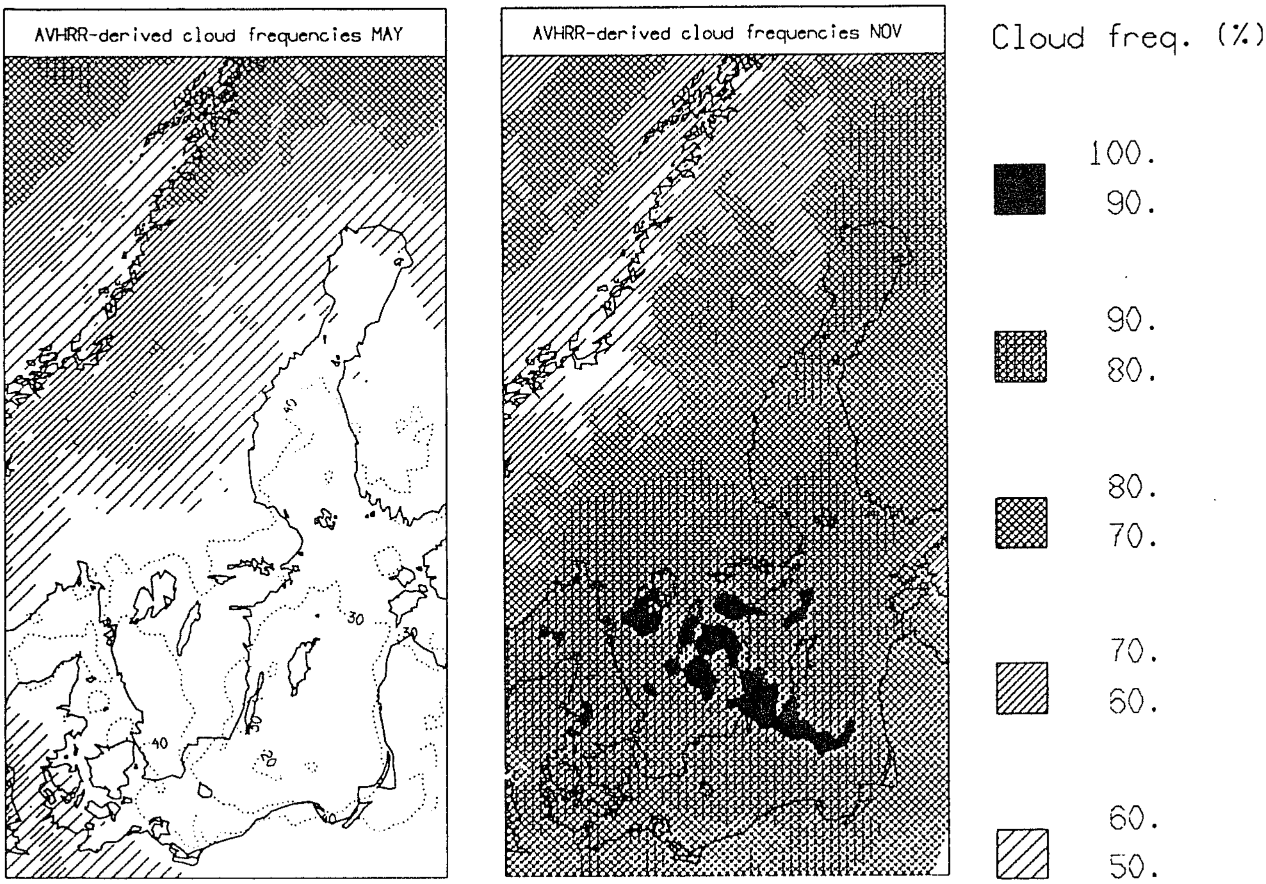

Fig. 1 *Satellite-derived cloud frequencies (%) for May (left) and November (right) in 1993. Results are averaged in a 20 km horizontal resolution.*

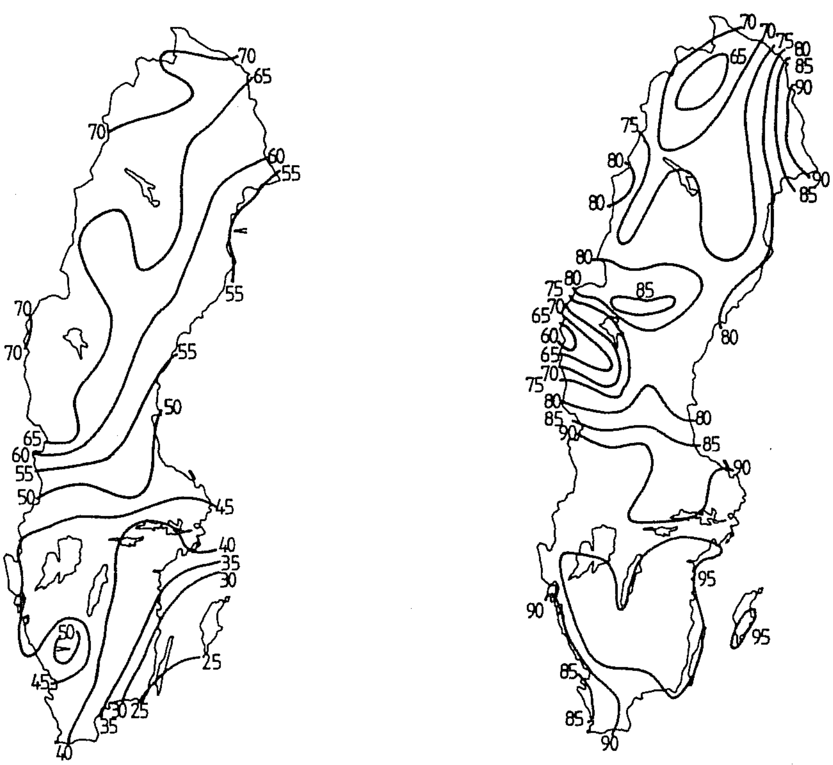

Fig. 2 *SYNOP-derived mean cloudiness for May (left) and November (right) in 1993. Based on SYNOP observations 06 UTC, 12 UTC and 18 UTC.*

3. VERIFICATION OF HIRLAM CLOUD FORECASTS USING SCANDIA

3.1 The set-up of a verification experiment

The cloud cover statistics from SCANDIA (as described in section 2.3) seem suitable for comparison with an ensemble of cloud forecasts from numerical weather prediction models. Such a comparison would indicate if the model is able to reproduce the proper cloud climatology for a certain region. An example will be given where cloud cover statistics are compared with cloud forecasts from the HIRLAM model over northern Europe for the month of August in 1994.

3.2 Cloud forecasts from HIRLAM

The HIRLAM forecasting system (described by *Gustafsson*, 1991) has been used at SMHI since 1992. An explicit prognostic scheme for cloud water (*Sundquist et. al,* 1989) was introduced in the operational HIRLAM version in autumn 1993. This also means that fractional cloud cover is now computed directly, (not diagnosed as in earlier HIRLAM formulations) in the cloud parameterization scheme.

The scheme was initially developed and tested in a version where fields of cloud water and cloud cover from the analysis first guess (a six hour forecast) was used as initial fields at the start of the forecast integration. Unfortunately, a mistake (in fact fully revealed and discovered by the study presented in this very paper!) was made during the operational implementation which meant that the operational HIRLAM since then has been starting the integration with zero cloud water and zero fractional cloud cover. This verification experiment will thus focus on how fast and if the HIRLAM model can generate (or spin up) realistic fields of fractional cloud cover starting initially from a state with no clouds.

3.3 The satellite data set for August 1994

In order to enlarge the area for which the satellite and HIRLAM data could be compared, some modifications to the SCANDIA model were introduced. The most important modification was related to the large variation in sun elevations that had to be accounted for if applying SCANDIA on larger areas. A SCANDIA version was developed to be executed over an area covering the distances from Ireland in the west to Russia in the east and from northern Germany in the south to the Barent Sea in the north. This area was segmented into the different sun elevation categories (mentioned in section 2.1) and the final classification result was enabled by combining several cloud classifications, each of them valid for only one of the segments. Furthermore, the temperature information guiding the separation of thick low- mid- and high-level clouds was now defined as interpolated temperature images based on short HIRLAM-forecasts of the temperature in the 700 and 500 hPa levels. By these actions, the cloud classification for the new area was in effect treating sun elevations and temperature information more accurately than the operational scheme described in section 2.1. A completely new feature was furthermore added to the SCANDIA model, namely the use of short forecasts

(9 or 12 hour forecasts) of surface temperatures. The purpose here was to improve the cloud/no cloud separation at sunrise or sunset when brightness temperatures in AVHRR channel 3 (otherwise effectively used for cloud separation) are sometimes insufficient for conducting an accurate cloud analysis.

The same technique (as described in section 2.3) to compute monthly cloud frequencies was now applied to data over the new area with the following exceptions and modifications:

- All satellite scenes over the area were used instead of only the four best (with lowest satellite zenith angles). This was necessary to obtain as good coverage as possible for the particular area.

- The basic AVHRR images over the area were defined with a 4 km resolution initially. Therefore no special resampling of the classification result was done.

- The monthly cloud frequency was computed by summing the cloud occurrences for all classifications in August and dividing them by the number of available scenes for each location (each 4 km pixel) in the area.

- In order to separately study the contributions to the derived cloud frequencies from the three groups low-, mid- and highlevel clouds, the cloud frequencies for these categories were computed from the classification images.

In all, 206 satellite scenes were used for the studied month. Due to some unfortunate problems with the satellite data reception, the satellite scenes for three days in August (11th, 12th and 21st) were not available. The derived cloud frequencies in the selected area were then based on a minimum of 116 scenes (four scenes per day in the south-east corner of the area) and a maximum of 204 scenes (seven scenes per day in the central northern part).

The computation of cloud frequencies was unfortunately not completely possible without any manual intervention. Manual editing (masking of selected parts of classification images) was conducted for the removal of bands in some of the images caused by the loss of a few scan lines during reception. Furtermore, seven of the used classifications (in the first few days in August) suffered from mis-classifications of very strong sunglints in the southern parts of the North and the Baltic Seas. Very strong sunglints can be misclassified as Stratus cloudiness and, in order to raise the quality of the satellite data set used for this verification experiment, these very obvious defects where removed (by masking out the southern parts of the North and the Baltic Seas) in the scenes. With the exception of these seven scenes and the occurrence of reception loss bands, the cloud classifications otherwise remained untouched.

3.4 The HIRLAM data set for August 1994

The mean of HIRLAM forecasted total fractional cloud cover in August was computed by using HIRLAM forecasts valid every 6th hour. The daily mean of total fractional cloud cover was then computed from four forecasts. All computations were conducted for the same area as for the satellite analysis. Due to the loss of satellite data for the 11th, 12th and 21st of August, the forecasts valid for these days were excluded. Three differing forecast lead times were investigated: 6 hours, 12 hours and 24 hours.

A special study was also devoted to investigate if any of the cloud categories defined as low-, mid- and high-level clouds was in any way mistreated in the HIRLAM forecasts. Cloudiness in vertical layers 1-8 defined here the contribution from high-level clouds, layers 9-11 from mid-level clouds and 12-16 from low-level clouds. The following equation is used in HIRLAM for the computation of total fractional cloud cover ***totcov*** from the fractional cloud cover $\boldsymbol{b_k}$ in each vertical layer $\boldsymbol{k}$:

$$totcov = 1 - cov2d \qquad \textbf{(1)}$$

where

$$cov2d = \Pi \frac{1 - \max(b_{k-1}, b_k)}{1 - \min(b_{k-1}, b_k)} \qquad (2)$$

The principle of maximum overlapping is used here when clouds are occuring in both of two adjacent layers, otherwise random overlapping is applied. The contributions to the total fractional cloud cover (computed downwards from the top layer) from the three cloud categories were computed using equations 1 and 2. These quantities should in theory be the same as the contributions from corresponding cloud categories in the cloud classifications (mentioned in the earlier section).

4. RESULTS

4.1 Results for total fractional cloud cover

Two ways of evalutating the results were applied. The first was to transfer (interpolate) the HIRLAM forecasts to high resolution imagery in the SMHI image processing system for satellite data. The forecasts were averaged over the whole month using the resulting pixel information and then visualized in the same way as the computed cloud frequencies from satellite measurements. This visualization is presented at the workshop. Another way of evaluting the results is to transfer the satellite information to HIRLAM fields

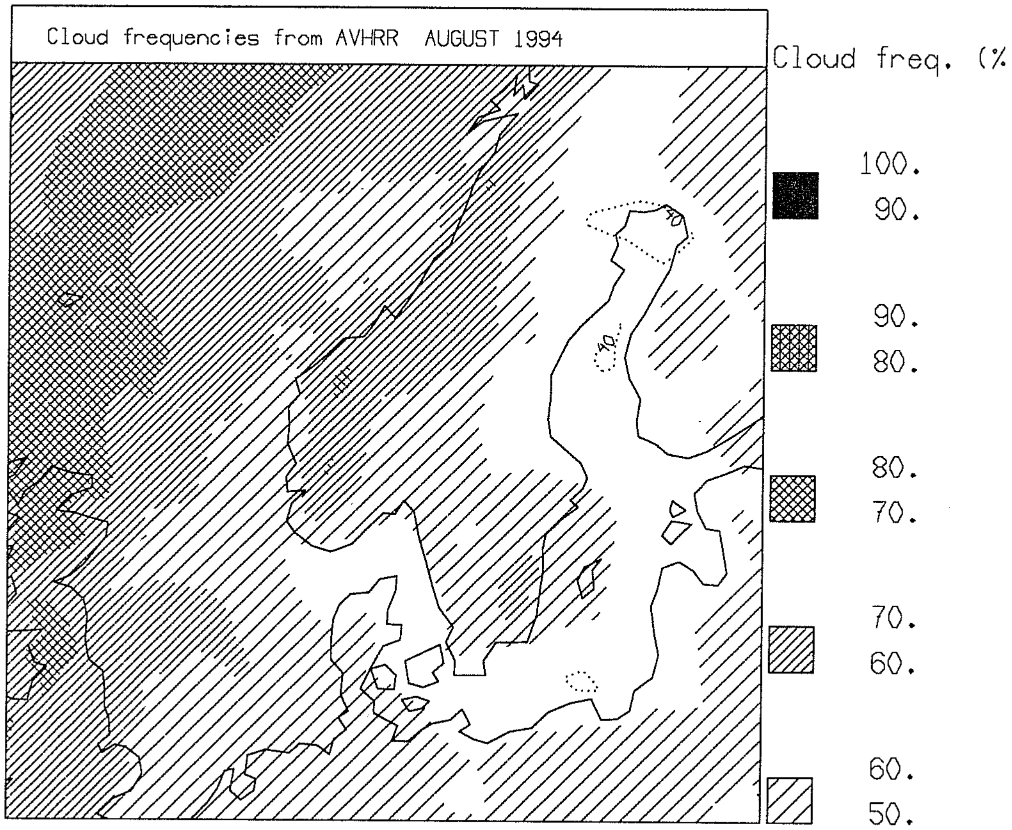

Fig. 3 *Cloud frequencies (%) from 206 AVHRR scenes in August 1994. Averaged in a horizontal resolution of 50 km.*

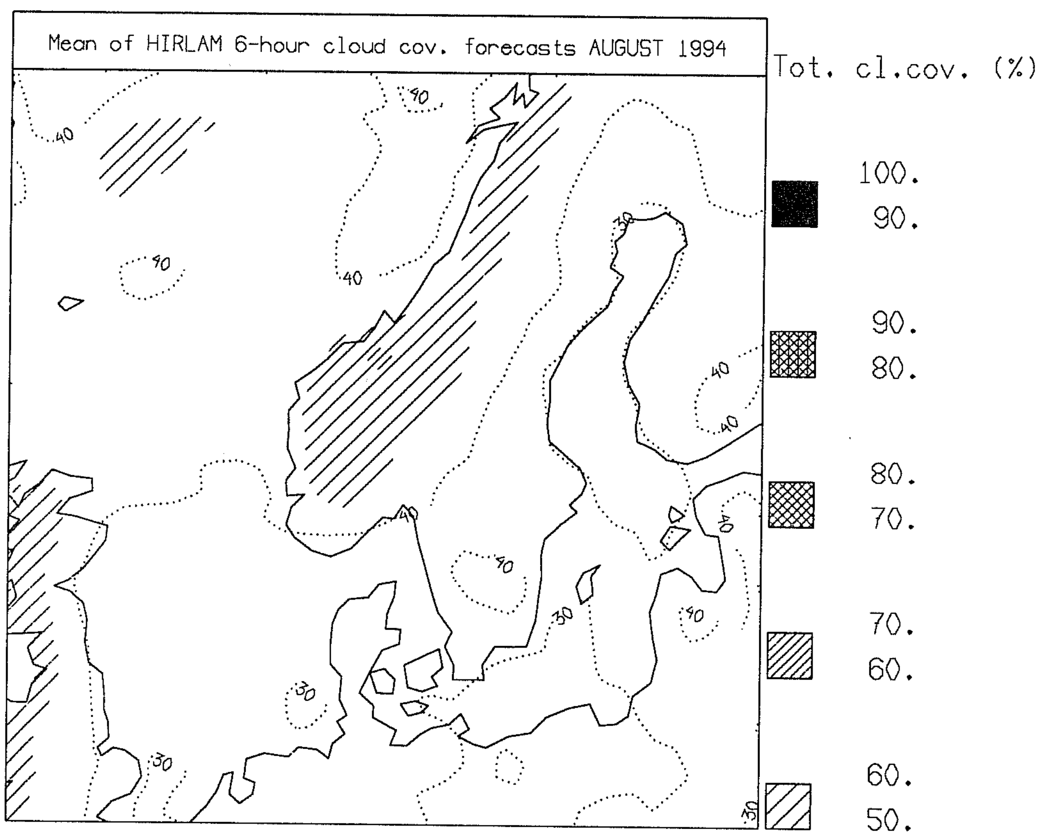

Fig. 4 *Mean of 6-hour HIRLAM forecasts of total fractional cloud cover in August 1994. Composed of four forecasts each day. Horizontal resolution 0.5° (appr. 50 km).*

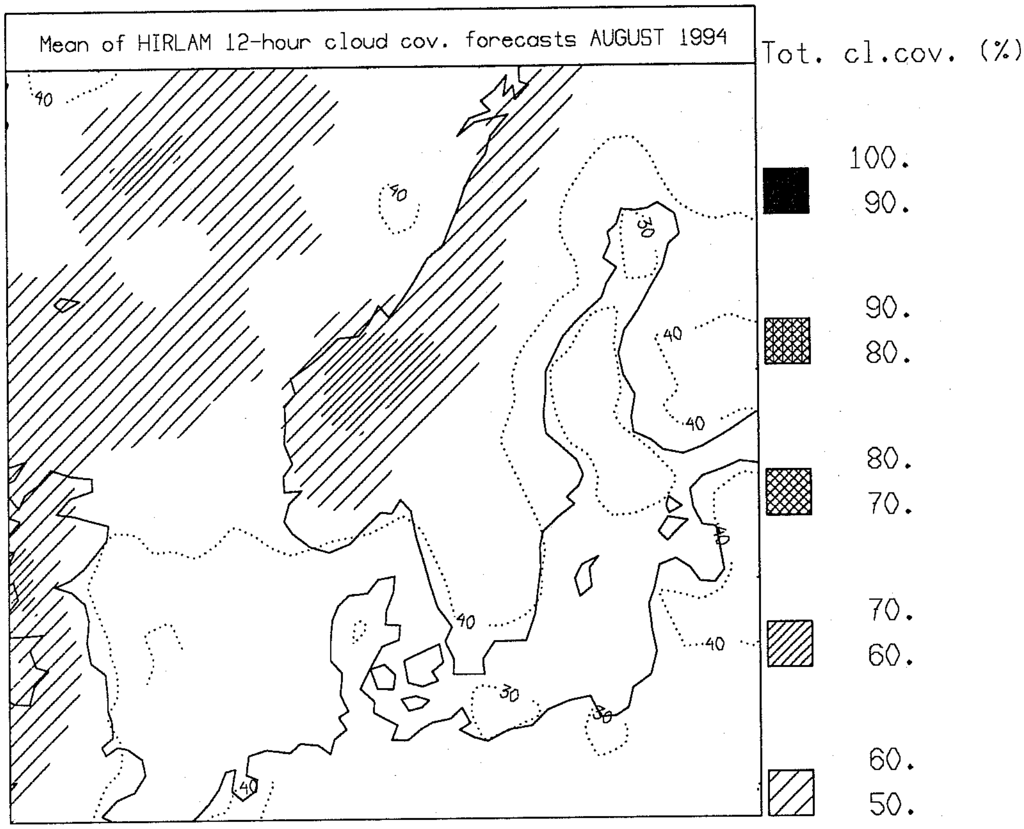

Fig. 5 *Mean of 12-hour HIRLAM forecasts of total fractional cloud cover in August 1994. Composed of four forecasts each day. Horizontal resolution 0.5° (appr. 50 km).*

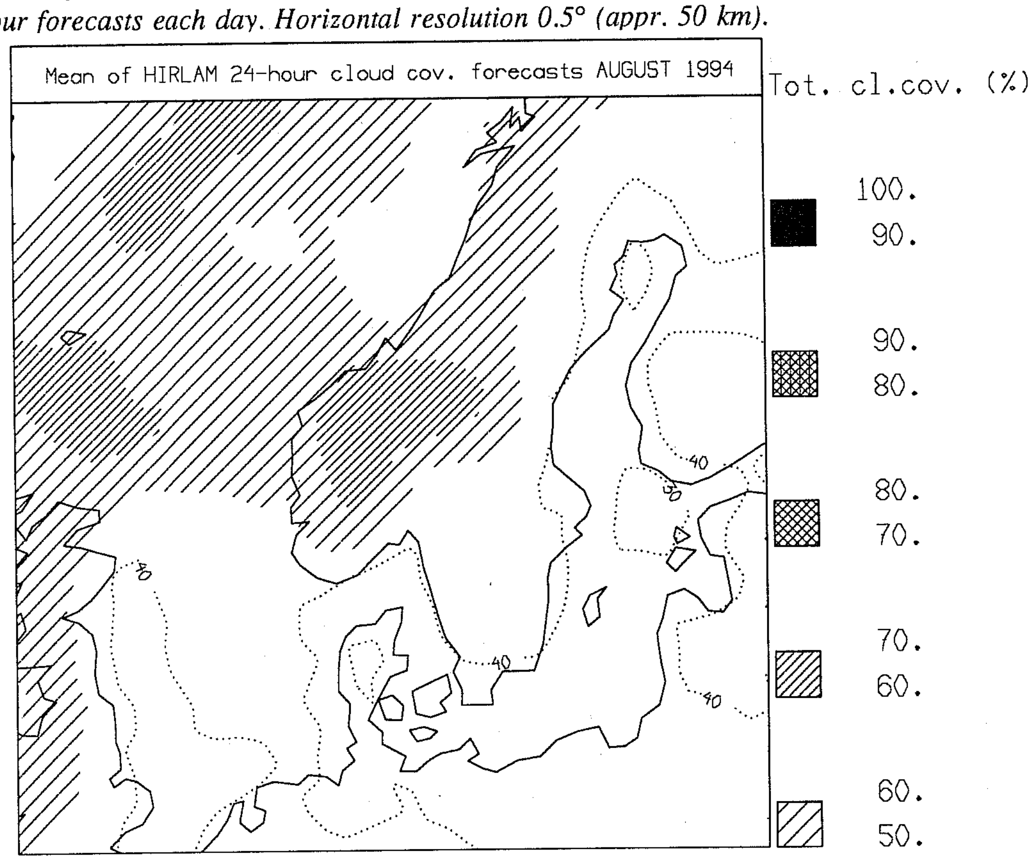

Fig. 6 *Mean of 24-hour HIRLAM forecasts of total fractional cloud cover in August 1994. Composed of four forecasts each day. Horizontal resolution 0.5° (appr. 50 km).*

(GRIB) and then compute the desired statistics describing areal means of cloud coverage and possible biases. Results based on the latter method are presented in this paper.

Fig. 3 shows the satellite-derived cloud frequencies for August 1994 in the same resolution (approx. 50 km) as the HIRLAM grid. Generally, the month was rather cloudy in this region, especially in the western part where cloud frequencies above 70% can be found. Much lower cloud frequencies can be found over the Baltic Sea and the nearby coast areas. Cloud frequencies are also low near Denmark. A local minimum of approximately 35% can be seen in the Bothnian Bay and a secondary maximum of approximately 65% is noticed in the southeastern part of Sweden.

The corresponding HIRLAM forecast information is presented in Figs 4-6, ordered by increasing lead times (6, 12 and 24 hours). We immediately notice that cloud amounts are in general too low for all forecast lead times. Although increasing with longer lead times, cloud amounts have still not reached the satellite-estimated levels for the 24-hour forecasts. However, the general horizontal distribution of clouds is found quite realistic and correlates well with the satellite analysis. Individual maximas and minimas are not captured in the exact places but they are nevertheless indicated near the right locations. Areal means of total cloud cover and biases compared with the satellite data are presented in Tab. 3. The bias for first guess fields (6-hour forecasts) is -14.8% and it improves to -5.9% for the 24-hour forecasts.

Tab. 3 *Areal means of total fractional cloud cover from satellite and from HIRLAM with varying forecast lead times.*

	Satellite	*HIRLAM +06*	*HIRLAM +12*	*HIRLAM +24*
Mean cloudiness (%)	*54.2*	*39.4*	*45.4*	*48.3*
Bias	-	*-14.8*	*-8.8*	*-5.9*

Results have also been compared with the results from the available SYNOP information for August which is also used for the routine verification of various other parameters from HIRLAM forecasts. The same trend for the spin-up of cloudiness can be seen but biases are even larger, especially when using mid-day SYNOP observations. It is suspected that this larger bias is due to the fact that cloud observations from SYNOP often overestimate cloud cover in summer for situations with Cumulus cloudiness dominating. Towering Cumulus clouds hide clear areas between cloud elements from the observer when viewing at larger zenith angles, causing an overestimation of cloudiness.

4.2 Results for the study of low-, mid- and high-level cloudiness

The HIRLAM underestimation of cloud cover could be seen for all the three cloud categories low-, mid- and high-level clouds for all forecast lead times. However, the underestimation was most pronounced for mid- and high-level clouds. The satellite analysis of the contribution from highlevel clouds is presented in Fig. 7 and the corresponding information from the 24-hour forecasts from HIRLAM is shown in Fig. 8. Notice in Fig. 8 that the complete HIRLAM data set, consisting of data from all 16 layers, was only available for the Nordic area in this study. It has earlier been found in several case studies that the HIRLAM generation of Stratus and Stratocumulus cloud decks is inadequate. However, this verification experiment for the summer month of August indicates instead that we have a general underestimation of clouds in all layers during summer.

5. DISCUSSION

In this paper, it has been shown that information on cloud cover for an entire month can be derived from the operational SCANDIA model and that the data can be used to validate cloud forecasts from numerical weather prediction models. The quality of the SCANDIA information is found to be quite comparable to SYNOP information and it is concluded that the satellite information should therefore be a valuable source of data for validating cloud forecasts for large areas, especially oceanic areas.

The verification of the HIRLAM cloud forecasts in August 1994 revealed that the spin-up time to create realistic cloud amounts is of the order 18-24 hours with the present HIRLAM version. Underestimation of cloud amounts could be seen in all vertical levels but the horizontal distribution of clouds appeared quite realistic. Interesting future studies could be devoted to investigating the effect of introducing various cloud initializations and the ability to describe the seasonal variability of cloudiness. The problems with low-level clouds (mentioned in section 4.2) in HIRLAM forecasts may be typical for other seasons which should be investigated further

A study of the diagnosed cloud information from operational ECMWF forecasts during August 1994 was also conducted but results were not completed in time for the preparation of this paper. However, the ECMWF results are presented at the workshop.

Future studies may also address the question of whether AVHRR data can be used to estimate other important cloud parameters. Attempts have recently been made to estimate vertically integrated liquid water content (e.g. *Raustein et. al.*, 1991) from visible AVHRR data. However, such methods suffer from the fact that several assumptions have to be made on typical cloud droplet and ice-crystal distributions and cloud thicknesses. A possibility to improve the results could be to incorporate cloud type information derived from AVHRR data

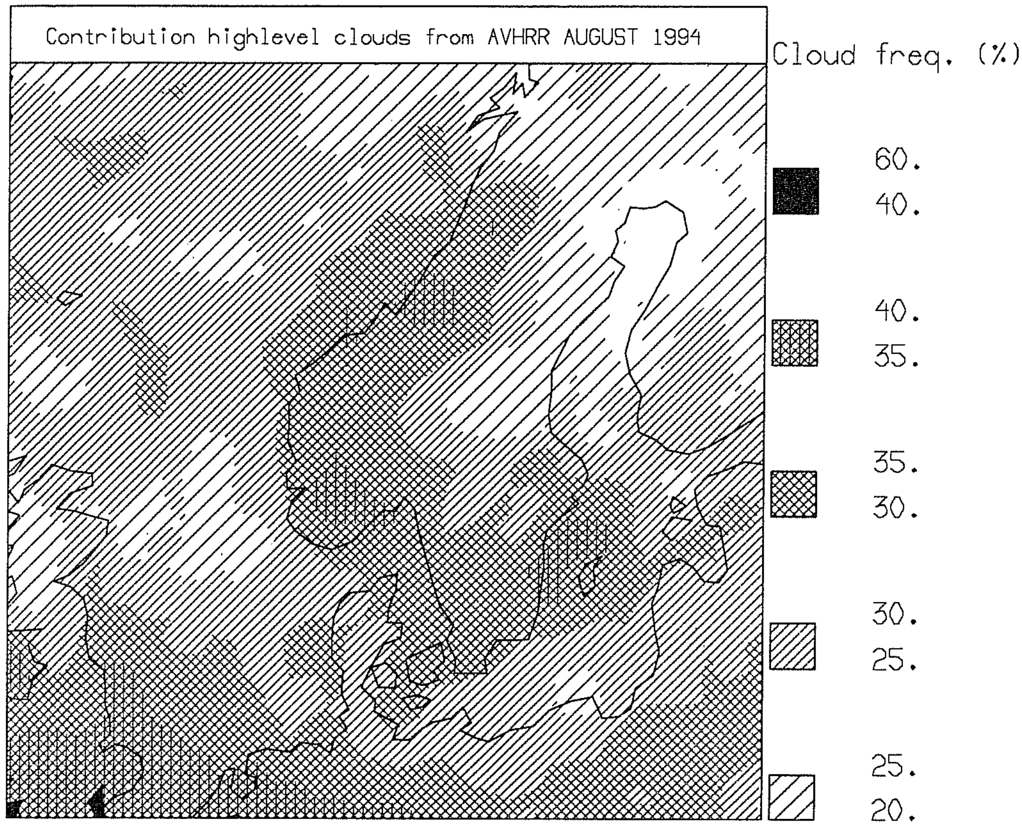

Fig. 7 *The contribution to the satellite-derived monthly cloud frequencies in August 1994 from high-level clouds. Horizontal resolution 0.5° (appr. 50 km).*

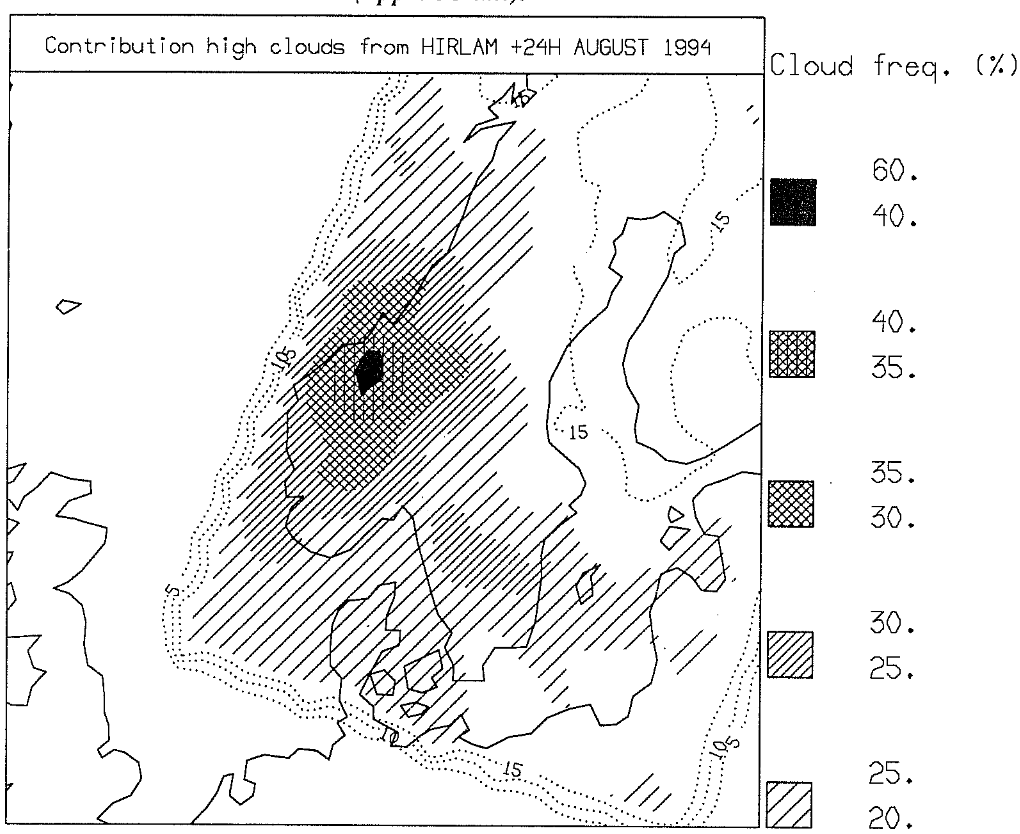

Fig. 8 *Mean of high-level total fractional cloud cover from 24-hour HIRLAM forecasts in August 1994. Composed of four forecasts each day. Horizontal resolution 0.5° (appr. 50 km).*

and to include also other data sources (e.g. SSM/I-data).

REFERENCES

Arking, A. and Childs, J. D. 1985: Extraction of cloud cover parameters from multispectral satellite measurements. J. Climate Appl. Meteor. 24, 322-333.

Derrien, M., Farki, B., Harang, L., LeGléau, H., Noyalet, A., Pochic, D. and Sairouni, A., 1993: Automatic cloud detection applied to NOAA-11/AVHRR imagery. Rem. Sens. of Env., 46, 246-267.

Gustafsson, N., 1991: The HIRLAM model. In: *Proceedings of Seminar on Numerical Methods in Atmospheric Models, ECMWF, Reading, UK, 9-13 September 1991.*

Karlsson, K-G. 1989: Development of an operational cloud classification model. Int. J. Remote. Sens., Vol. 10, Nos 4-5. 687-693.

Karlsson, K-G. 1993: Comparison of operational AVHRR-based cloud analyses with surface observations, In: *Proc. 6th European AVHRR Data Users' Meeting,Belgirate, Italy, 28 June - 3 July, 1993*, EUMETSAT, 223-230.

Karlsson, K-G. 1994_1: Cloudiness in the Nordic summer season estimated from multispectral satellite measurements. Accepted for publication in AMBIO.

Karlsson, K-G., 1994_2: Satellite-estimated cloudiness from NOAA AVHRR data in the Nordic area during 1993. SMHI Reports Meteorology and Climatology, RMK 66.

Karlsson, K-G. and Liljas, E. 1990: The SMHI model for cloud and precipitation analysis from multispectral AVHRR data. PROMIS Report 10, SMHI.

Lauritson, L., Nelson, G. J. and Porto, F. W. 1979: Data extraction and calibration of TIROS-N/NOAA radiometers, NOAA Technical Memo NESS 107, NOAA, Washington D. C.

Liljas, E. 1981: Analysis of clouds and precipitation through automated classification of AVHRR data. SMHI Reports Meteorology and Climatology, RMK 32.

Raustein, E., Sundqvist, H. and Katsaros, K. B. 1991: Quantitative comparison between simulated cloudiness and clouds objectively derived from satellite data. Tellus 43A, 306-320.

Saunders, R. W. and Kriebel, K. T. 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. Int. J. Remote Sens., Vol. 9, No 1, 123-150.

Sundquist, 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.