A comparison of data coverages of conventional observations in different operational HIRLAM routines

Kalle Eerola
Finnish Meteorological Institute

1 Introduction

In any NWP-system, it is essential to utilize all available observations. Eerola (2003) studied one aspect of this problem, the arrival times of conventional observations at Finnish Meteorological Institute (FMI). That study gave useful information when designing the time table and the length of the cut-off time of the RCR (Regular Cycle with the Reference) system at FMI. However, it was based on the observations, which were available at FMI, and therefore did not answer the question: **does FMI get all the conventional observations, which are available and can be used?** Here the word "conventional" refers to those observation types that are set as "active" in the reference HIRLAM system and therefore are used in the RCR system.

To get information about the possibly missing observations or observation types, the author suggested that inside the HIRLAM community we should organize a collection of observation files of one day from the operational HIRLAM systems to see, if the data coverage are similar in different countries. The suggestion received a positive reaction, and the author promised to organize the exchange and to make an investigation of the data coverages in different operational HIRLAM routines, based on the collected data. This report describes the results from that investigation.

The layout of this report is the following. After this introduction, section two describes the data of this study. In section three we discuss the methods used in this study and also the problems in analysing the operational observation files. The results are presented and discussed in section four. Finally, section five summaries the study.

2 Data

In the HIRLAM 3DVAR data assimilations system, the observations, when extracting from BUFR-code, are put into the so-called CMA-files (Central Memory Array-files). The data in this study contain these files from the operational HIRLAM runs of different HIRLAM countries for one day, 17 June 2004. After the data assimilation step the CMA-file (now named as ACMA-file) contain, in addition to observations, also the feedback information from the analysis. So it is possible to pick-up and count exactly those observations that have been used in the analysis.

The ACMA-files for the requested date were received from FMI, SMHI, met.no and DMI. Met Éireann does not save the ACMA-files, so it was not possible to get data for the requested date. Instead, the ACMA-files for the period 2004070709 - 2004071806 were received. INM has problems with the file format: the binary files from their computer are not compatible with other machines. Therefore we could not use their data. From KNMI no data was received.

In addition, HIRLAM beta-version 6.3.3 was run at ECMWF with the default settings to get an impression of the amount of observations available at ECMWF. They use a long cut-off time and therefore they should have a very complete data base.

3 Method

It is not easy to compare the data of this kind, because the inside HIRLAM community the operational HIRLAM configurations vary from country to country. The main difficulties are due to the following differences:

- Cut-off times differ from country to country. However, a commonly used cut-off time is around two hours. Comparing the data coverages from runs with different cut-off time can even give some hints how to define the cut-off time properly.
- Geographical domains are different in different countries.
- The data assimilation cycle is either three or six hours. This affects the length of the data window, i.e. the length of the period from which observations are used. This is important especially when looking at the asynoptic observations.

The method in this study is to compute the number of all used and accepted observations in 5° × 5° horizontal boxes and print them on a map. The term "used and accepted observation" here means that the "report status" `pcma(ncmrst)` is active and "reports event, part 1" `pcma(ncmrev1)` is zero, i.e. no problems have been reported. For more information about the structure of the CMA-file and it's flags, see White (2000).

For some reason, probably due to an older version of HIRVDA package, met.no data reported all SYNOP reports active. Therefore an extra check was done for the SYNOP data that data from any station is not used several times in the same analysis.

Because the data assimilation cycle times and cut-off times differ from country to country, the direct comparison of data coverage is not easy. Two set of maps were produced

1. one containing the number of all accepted and used observations from the 12 UTC run
2. another containing the number of all accepted and used observations from all daily cycles together

In both cases there are aspects, which must be kept in mind while looking at the results. In the first case the data window is different depending on the length of the data assimilation cycle: six hours in case of six hour and three hours in case of three hour cycling. This affects mainly on the number of SHIP, DRIBU and AIREP observations, which don't report at fixed locations. In case of SYNOP and TEMP data, only the observation closest to the analysis time is used from every station.

In this report, only one experiment from every national HIRLAM configuration will be discussed. A larger report, with mainly figures, will be made available in $H_{\mathrm{E}}X_{\mathrm{NET}}$.

4 Results

4.1 Data coverage at 12UTC runs

First we investigate the data coverages at the 12 UTC runs. We don't show here results from all different configurations in different countries, but take only one configuration from every country.

Table 1. shows the cut-off times and cycle intervals for those 12 UTC runs investigated here. Here "ECMWF obs" means the Hirlam 6.3.3 run at ECMWF using the ECMWF observations in the way the HIRLAM reference system does. As can be seen, RCR at FMI, ME at Met Éireann and ECM (HIRLAM reference system using ECMWF observations) use 3 hour data cycling. The other routines use six hour cycling. The cut-off times are mainly of the order two hours, except for DMI. The short cut-off run from DMI was selected to see possible differences caused by this, because they also have a longer cut-off run and direct comparison to it is easy. According to Eerola (2003) a two hour cut-off is a satisfactory for most observation types, but for soundings even a few minutes difference may make a difference in the number of received observations. Note that for Met Éireann the date of observations is different from the others.

Table 1: *The relevant information about the experiments discussed.*

Country	EXP	Cycle interval	cut-off
SMHI	C22	6 h	120 min
met.no	H20	6 h	135 min
DMI	T15	6 h	99 min
FMI	RCR	3 h	120 min
Met Éireann	ME	3 h	120 min
ECMWF obs	ECM	3 h	? min

The number of used accepted SYNOP observations in the selected 12 UTC runs in different countries can be seen in Figures 1 ... 6. The first impression is that the data coverage in different runs are rather similar, taking into account the different areas. A few comments can be made. First of all, it is obvious that all countries has access to such local observations that are not available internationally. For instance, DMI has over Denmark over 100 observations, while other centers have about 60 ... 70 observations on that area. It is noticeable that ECMWF database gives only 38 observations, FMI has 66, although both have a three hour cycling, so a three hour data window. Another note is that FMI(RCR) seems to have slightly less observations from the north-eastern part of USA and from Canada than the other routines.

The distribution and number of SHIP and DRIBU observations are very similar in every country, even if the length of the data windows vary from three to six hours and the cut-off times are different. So the results are not shown here.

Next we look at sounding data. The data coverage maps for TEMP data can be seen in Figs. 7 ... 12. The most striking feature here is that the soundings from Russia are almost missing in the DMI T15 run, probably due to a short, 99 min., cut-off time. This can be confirmed by looking at the same run with long cut-off time, where the missing soundings are present. Otherwise the distributions are very similar, showing that all centers get all the essential sounding information. There are small differences in Central Europe, but that is in the area of good data coverage. So, as already pointed out by Eerola (2003), the length of the cut-off time around two hours is just long enough, but it cannot be much shorter, if a good data coverage is the aim.

The number of aircraft (AIREP and AMDAR) data in the 5° x 5° boxes are shown in Figs. 13 ... 18. Note that aircraft observations are non-synoptic and are not made at fixed locations. So the width of the data window and the cut-off time may play here a bigger role than in case of synoptic observations made at fixed stations.

We can immediately see that both ECMWF and DMI data bases contain much more such aircraft data from North America, which are not present in met.no and FMI runs. Met Éireann has more observations than FMI and met.no in that area, but less than DMI and "ECMWF obs". Especially FMI has problems in getting all aircraft data available. "ECMWF obs" has the longest cut-off time and the best data coverage. However, this does not explain the lack ob observations, because DMI has also a good data coverage, but the shortest cut-off time. On the other hand "ECMWF obs" uses a three-hour cycling, .ie three-hour data window, and has much data over North America. So there must be data, which do not arrive at FMI, SMHI and met.no.

If we look at the 15 UTC run (not shown), where FMI uses a longer cut-off time (over four hours), we see that FMI gets some aircraft data from American area, but not so many as expected. From the statistics shown by Eerola (2003) is is clear that over 90% of AIREP and AMDAR data, which arrives FMI, comes within two hours. The reason for more data in this cycle is probably due to more active aircraft traffic at that time of the day.

The Figs. 13 ... 18 also reveal that FMI(RCR) has less aircraft observations from central and southern Europe compared to SMHI. This can be seen clearly seen, when comparing, for instance, Figs. 12 and 15. On the other hand, the data coverage of DMI (Fig. 13 is very similar to that of FMI. The reason to this is unclear, because the GTS data to FMI comes via SMHI and all data arriving at SMHI should also arrive at FMI. This can be due to the three-hour cycling. This is discussed more in the next chapter.

The different number of aircraft observations is so far the most important result from this survey.

4.2 Data coverage in all termins together

It is more difficult to make any conclusions of the observation coverage, if the number of observations in all the daily cycles are computed together because of different number (four or eight) of data assimilation cycles in a day. The following comments are very practical, giving weight to the importance of the observations in question. Broadly speaking the same comments are valid here as was discussed in the previous chapter for 12 UTC:

- TEMP sounding data is critical for cut-off time: DMI misses many Russian sounding in their short cut-off run.
- DMI and ECMWF data bases have much more aircraft data over the American continent than the other institutes.

Coming back to the queston of aircraft data over Europe, it seems that the difference between FMI and SMHI data coverages is due to the different data assimilation cycle time (three hours at FMI and six hours at SMHI). When looking at the total number of used and accepted aircraft observations during the 24-hour period, we see the the number of observations are comparable to each other (not shown). So SHMI and FMI use about the same number of observations during a 24-hour period, but in case of FMI they are used in eight runs, while in case of SMHI the number of runs is four. However, the difference to the amount of data in ECMWF data base is large (not shown).

5 Summary

The following conclusions can be made:

- Every country has more such domestic SYNOP data available, which is not distributed internationally. This is not very critical at the moment with the current grid resolution and analysis method, but may become more important in the future when using four dimensional data assimilation.
- TEMP soundings are time critical: two hours' cut-off time is needed at the moment to get also all the Russian soundings.
- Only ECMWF, DMI and Met Éireann receives a lot of aircraft information from the American continent. Other centers have only few aircraft data from that area. Generally the amount of aircraft data available varies from country to country. Especially, more aircraft data should be get for the RCR runs at FMI.
- ECMWF data base contains more aircraft observations all over the area compared to the other institutes.
- SMHI and FMI use approximately the same amount of aircraft observations in a 24-hour period, but at SMHI it is used in four runs, while at FMI in eight runs.

The results with more figures will also be published in $H_{\mathrm{E}}X_{\mathrm{NET}}$.

Acknowledgements

I want to thank my colleagues in other HIRLAM countries for providing the data to this study: Ole Vignes (met.no), Lars Meuller (SMHI), Bjarne Amstrup (DMI) and Ray McGrath (Met Éireann).

References

Eerola, K., 2003: Statistics of the arrival times of conventional observations for HIRLAM at FMI. *Hirlam Newsletter*, **43**, 20–32.

P. W. White, Ed., 2000: *IFS Documentation Part I: Observation Processing.* ECMWF, Research Department.

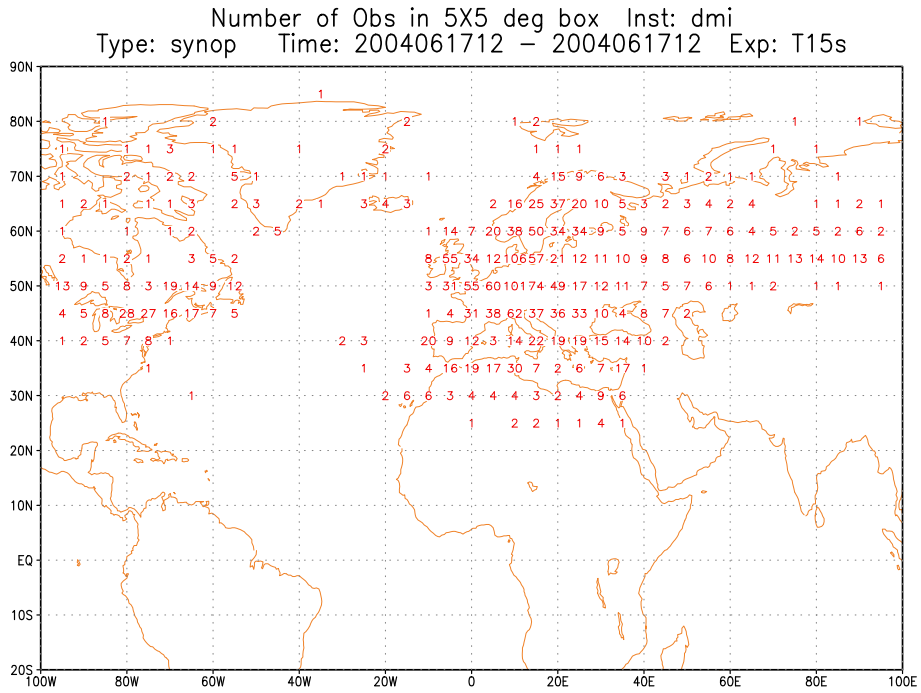

Figure 1: *Data coverage of DMI Hirlam version T15 with 99 min. cut-off for SYNOP observations.*

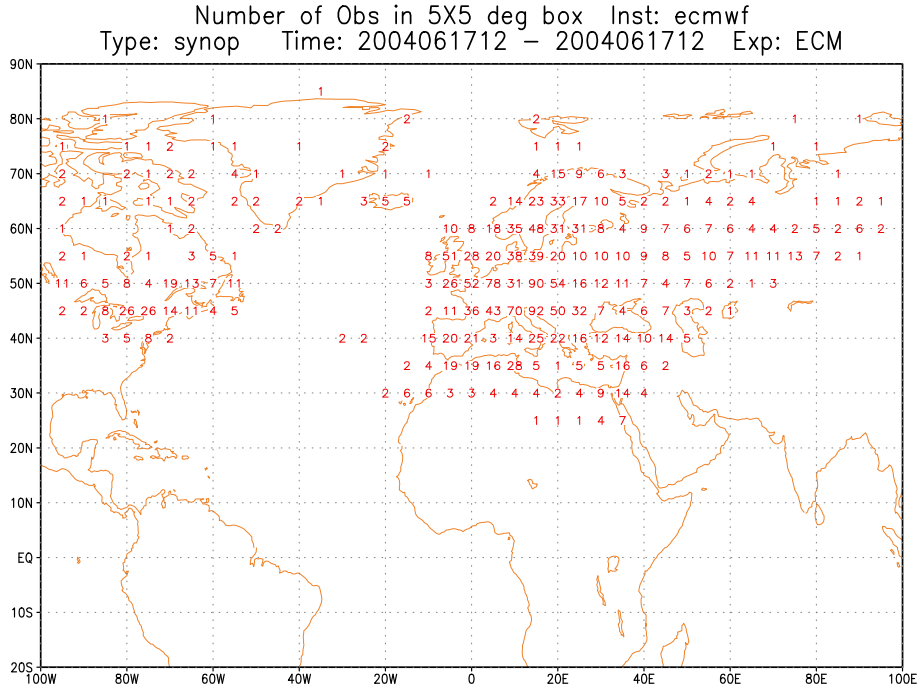

Figure 2: *Data coverage of Hirlam reference at ECMWF for SYNOP observations.*

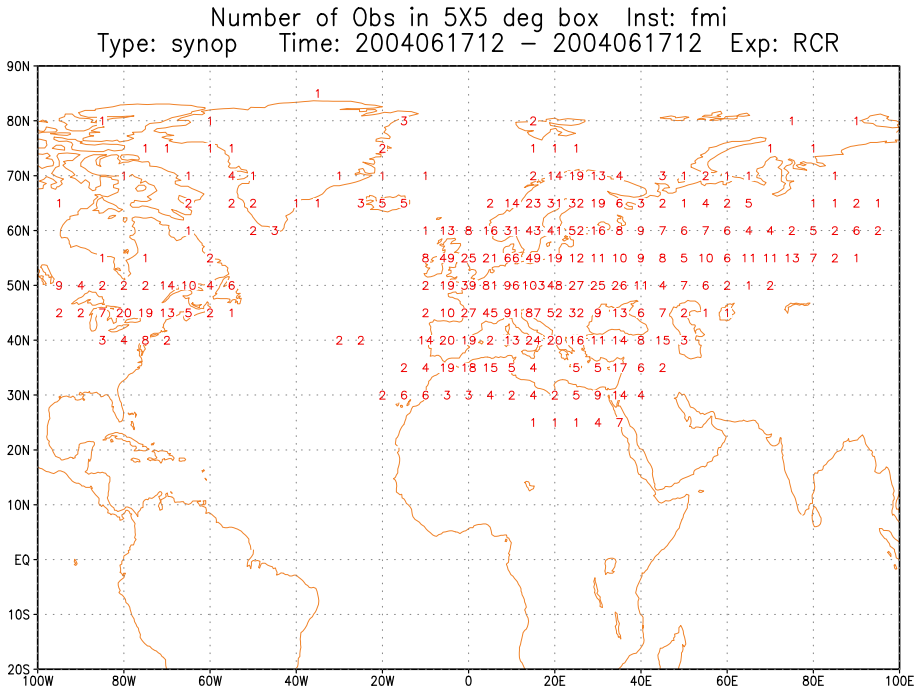

Figure 3: *Data coverage of FMI Hirlam version RCR with 120 min. cut-off for SYNOP observations.*

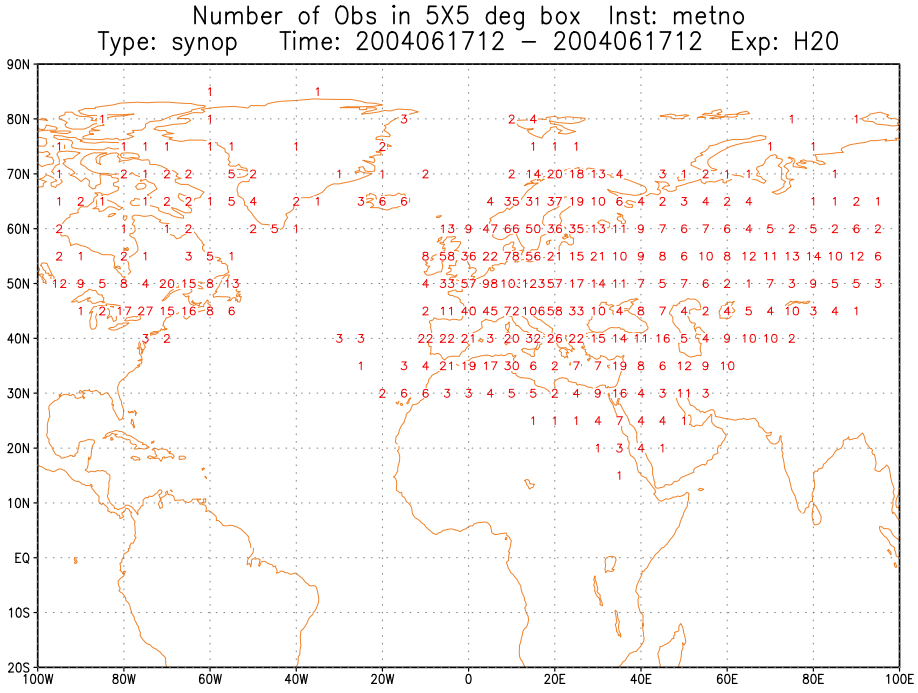

Figure 4: *Data coverage of met.no Hirlam version H20 with 135 min. cut-off for SYNOP observations.*

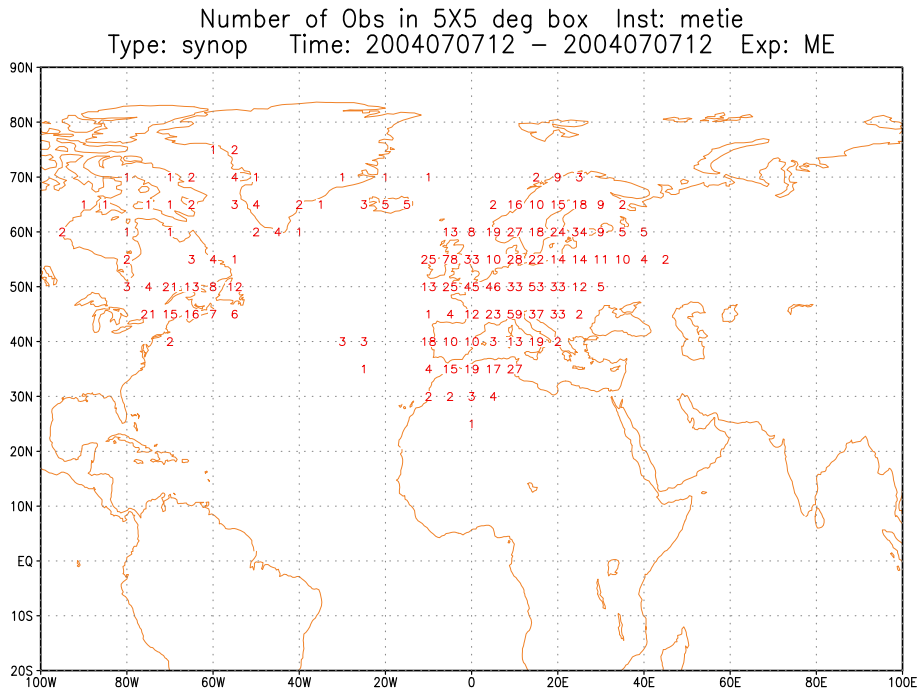

Figure 5: *Data coverage of Met Éireann Hirlam version ME with 120 min. cut-off for SYNOP observation at 07 July 12 UTC.*

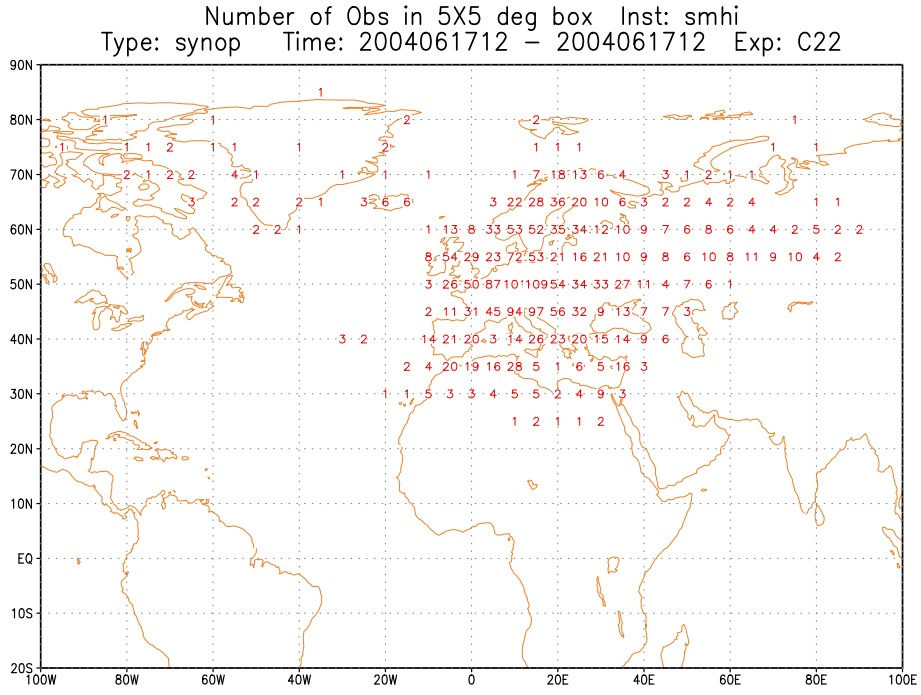

Figure 6: *Data coverage of SMHI Hirlam version C22 with 120 min. cut-off for SYNOP observations.*

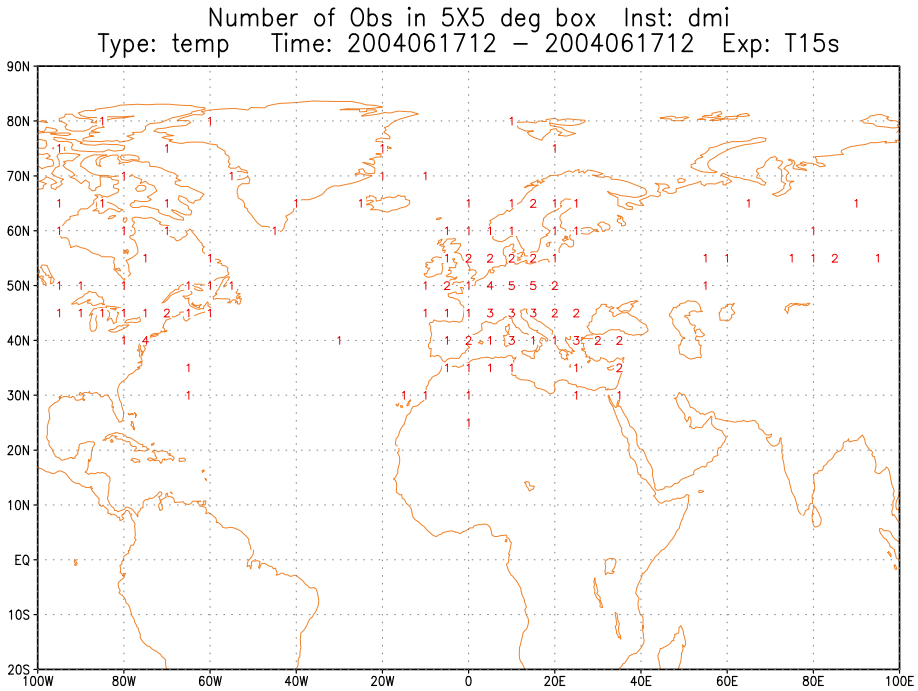

Figure 7: *Data coverage of DMI Hirlam version T15 with 99 min. cut-off for TEMP soundings.*

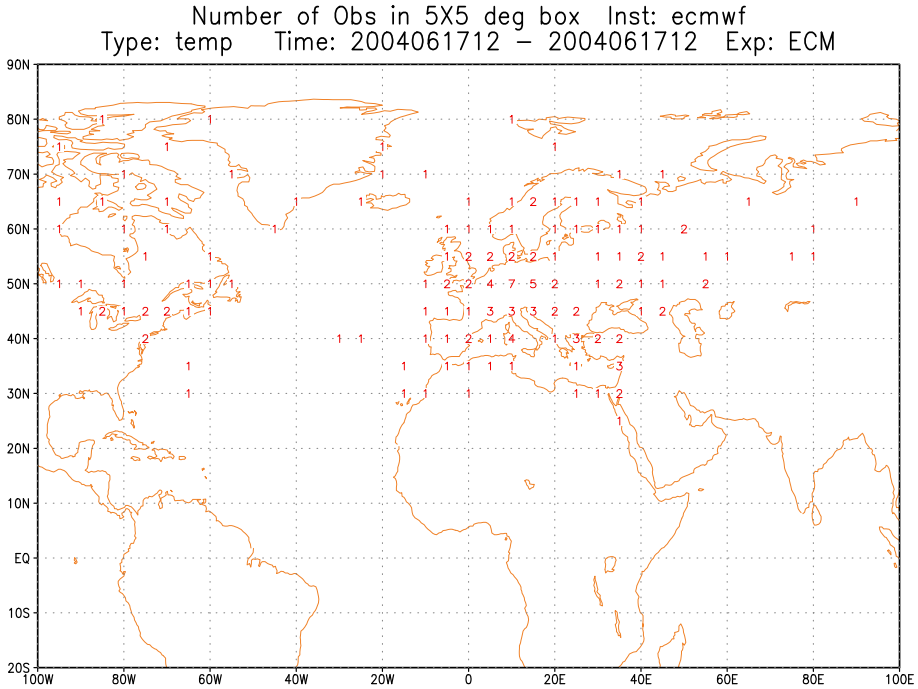

Figure 8: *Data coverage of Hirlam reference at ECMWF for TEMP soundings.*

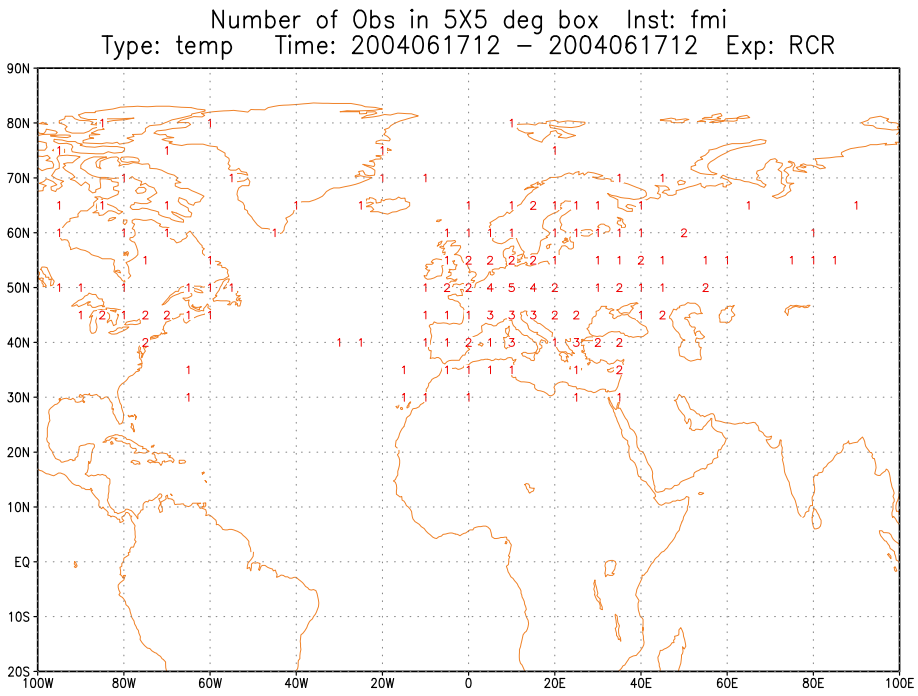

Figure 9: *Data coverage of FMI Hirlam version RCR with 120 min. cut-off for TEMP soundings.*

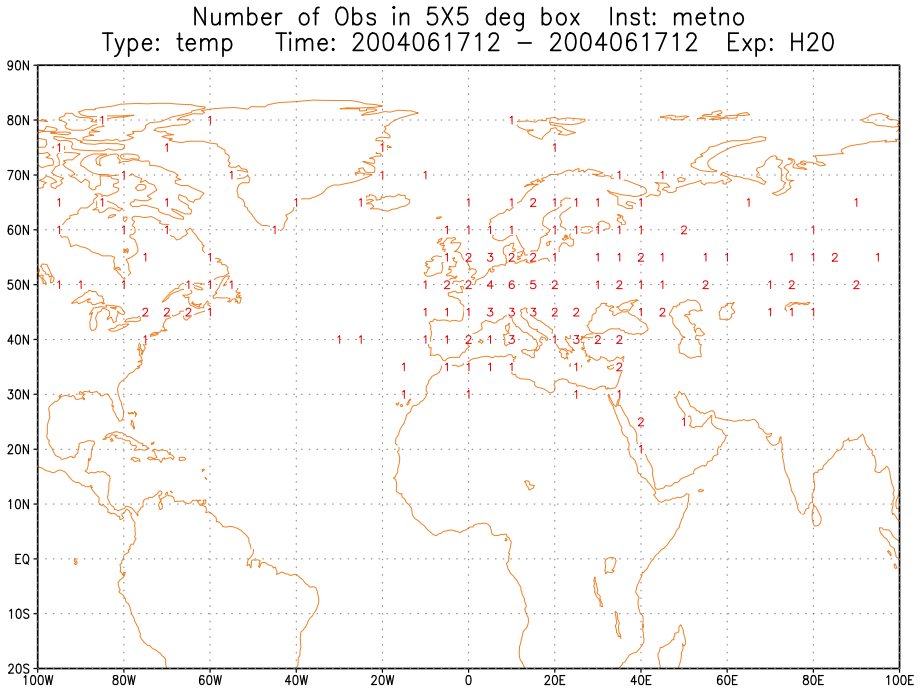

Figure 10: *Data coverage of met.no Hirlam version H20 with 135 min. cut-off for TEMP soundings.*

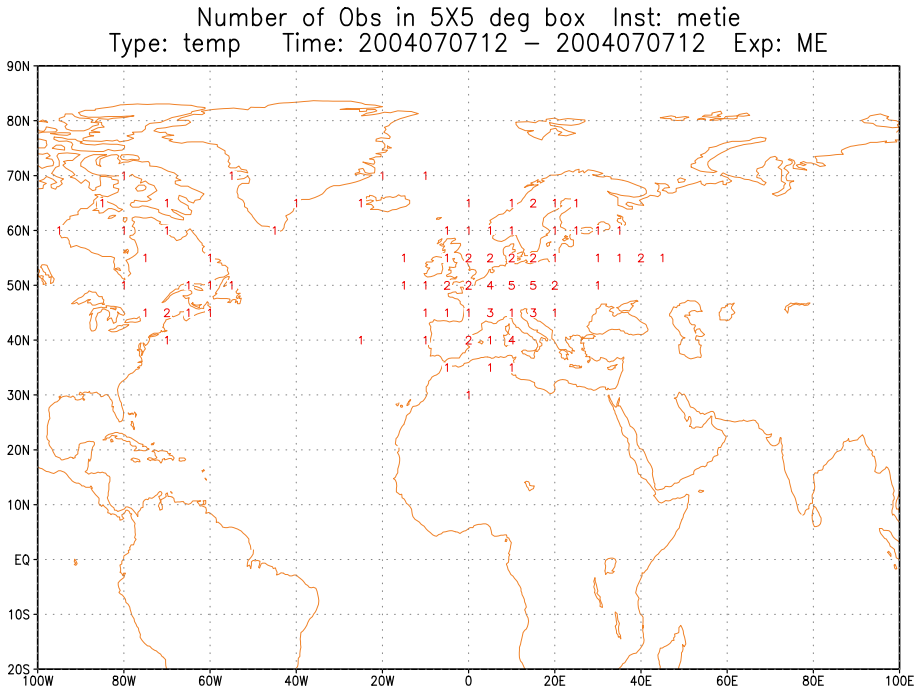

Figure 11: *Data coverage of Met Éireann Hirlam version ME with 120 min. cut-off for TEMP soundings.*

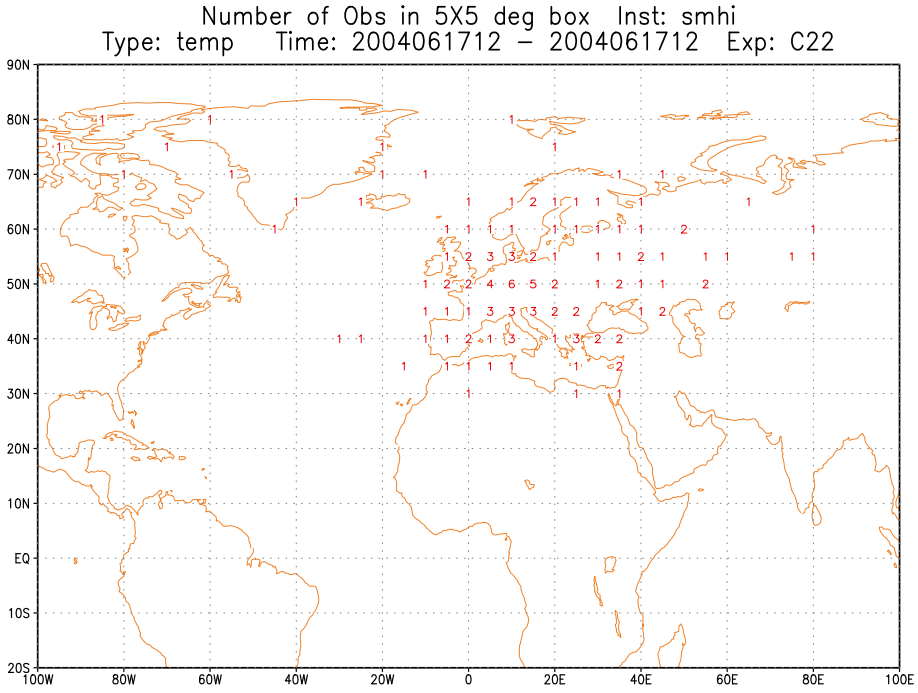

Figure 12: *Data coverage of SMHI Hirlam version C22 with 120 min. cut-off for TEMP soundings.*

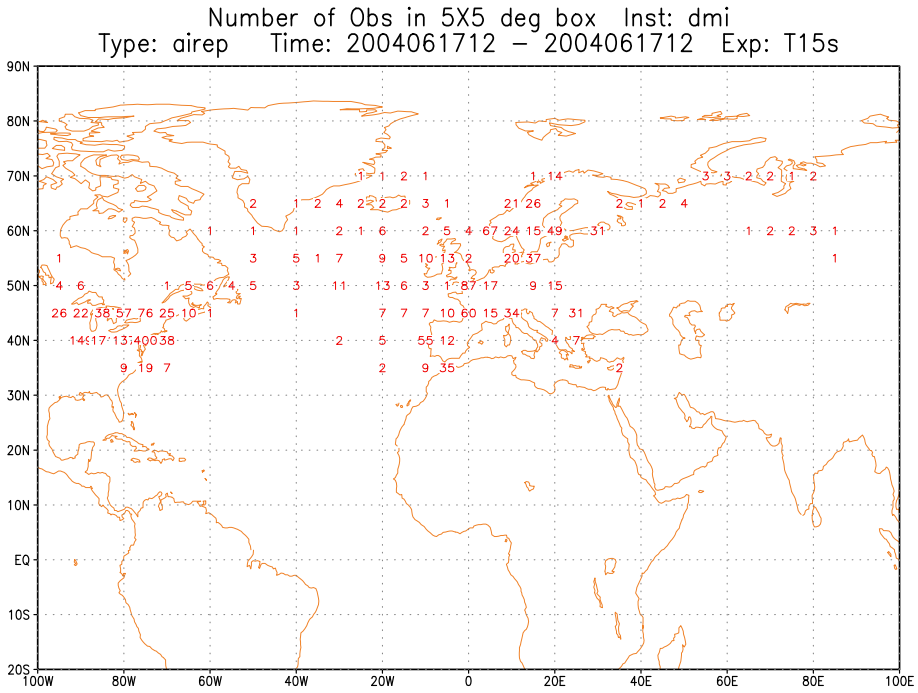

Figure 13: *Data coverage of DMI Hirlam version T15 99 min. cut-off for aircraft data.*

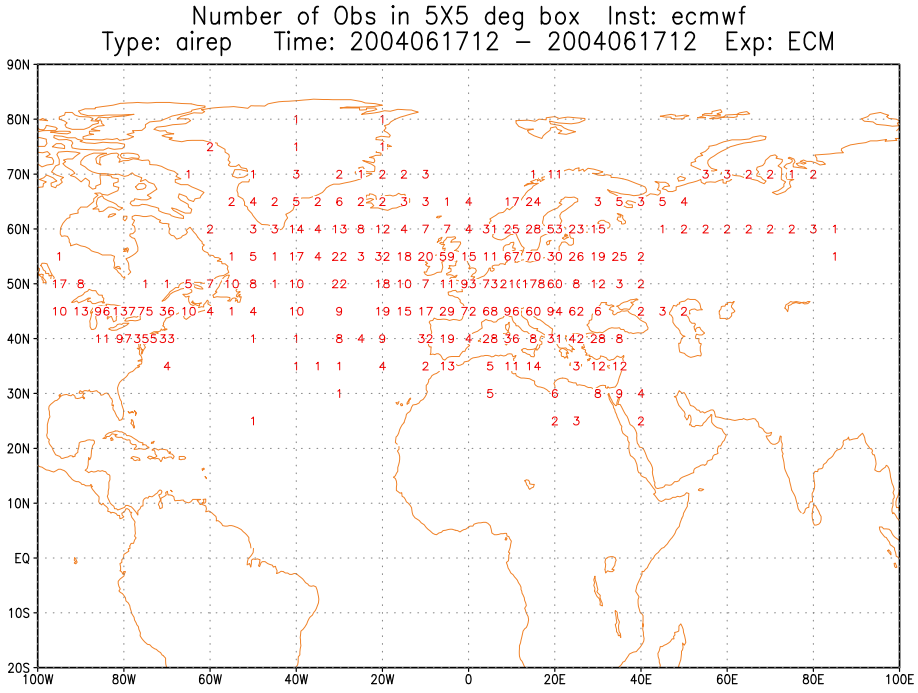

Figure 14: *Data coverage of Hirlam reference version at ECMWF for aircraft data.*

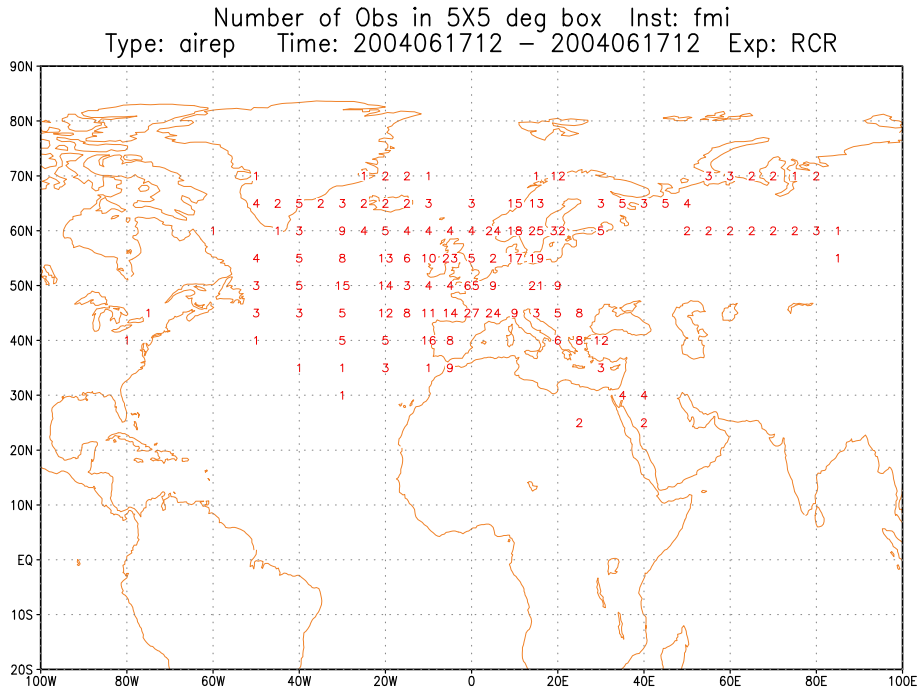

Figure 15: *Data coverage of FMI Hirlam version RCR with 120 min. cut-off for aircraft data.*

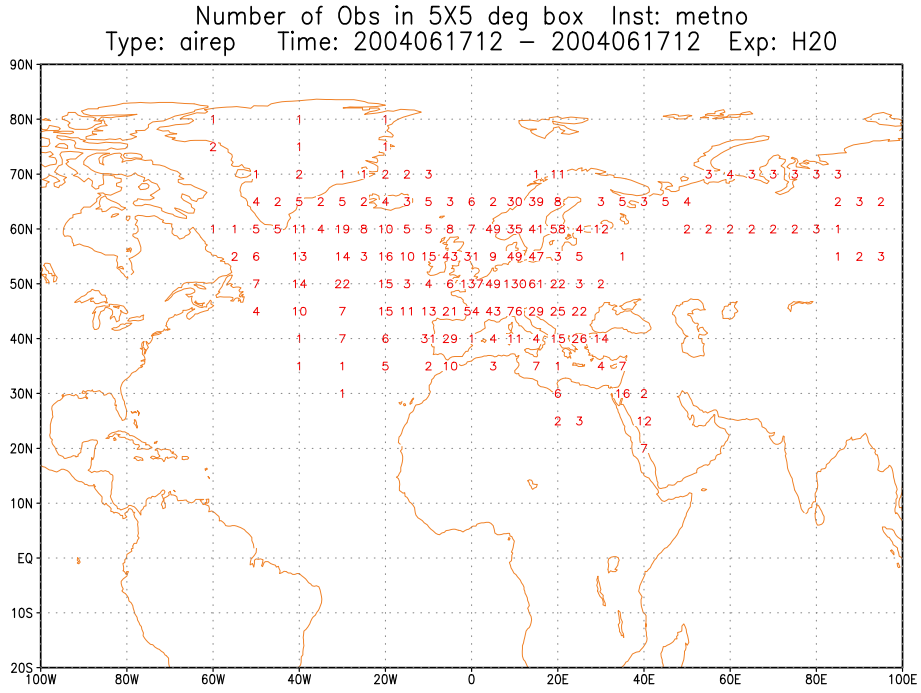

Figure 16: *Data coverage of met.no Hirlam version H20 with 120 min. cut-off for aircraft data.*

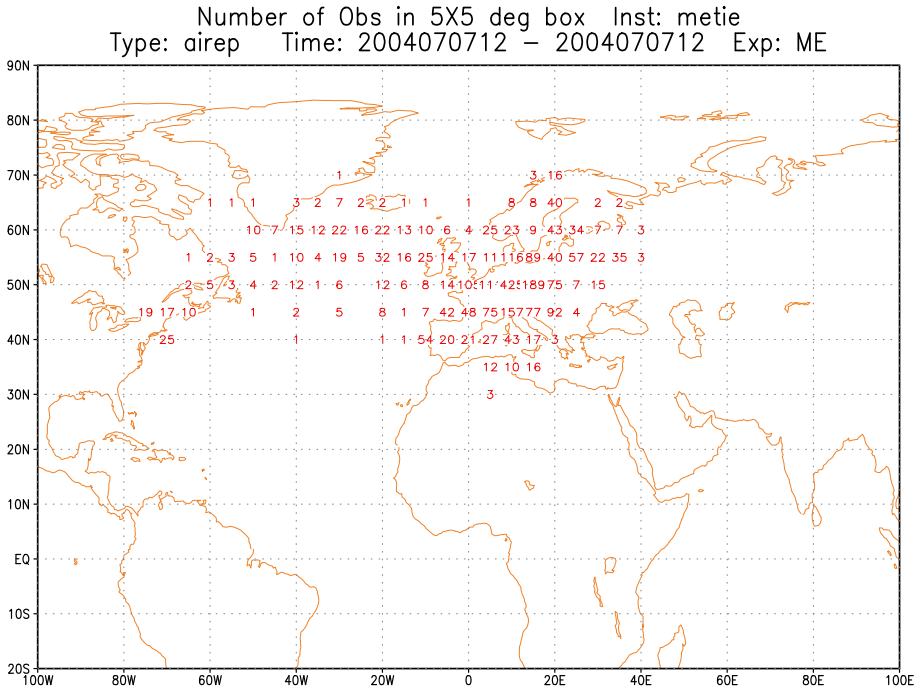

Figure 17: *Data coverage of Met Éireann Hirlam version ME with 120 min. cut-off for aircraft data.*

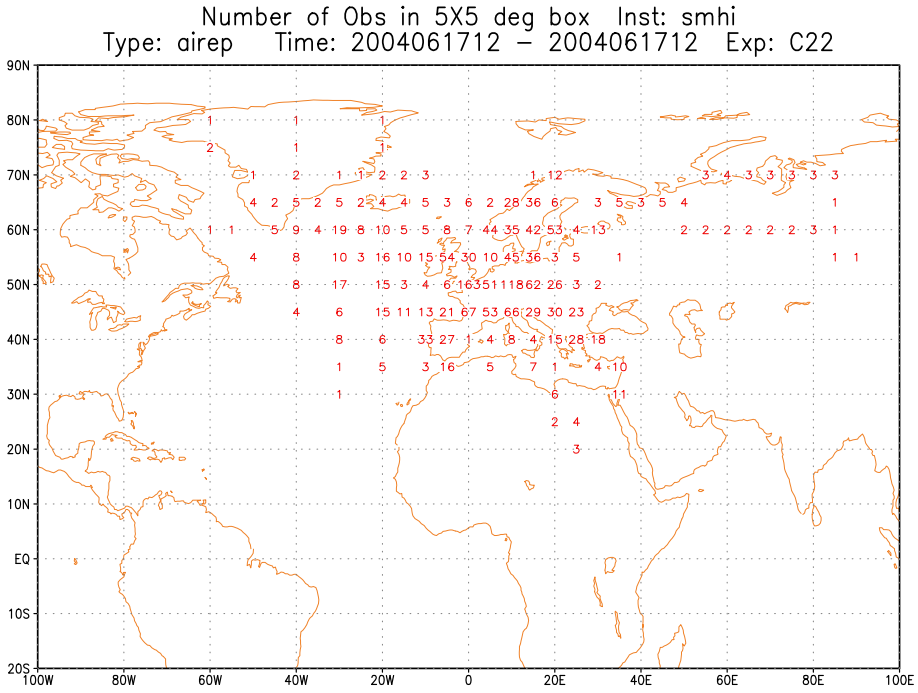

Figure 18: *Data coverage of SMHI Hirlam version C22 with 120 min. cut-off for aircraft data.*