Bias Correction Issues In Limited Area Models: Strategy For HIRLAM

Per Dahlgren

SMHI
Swedish Meteorological and Hydrological Institute
Folkborgsvägen 1, 601 76 Norrköping, Sweden

1 Introduction

HIRLAM, HIgh Resolution Limited Area Model, is a regional model developed, maintained and used by Finland, Sweden, Norway, Denmark, Ireland, Iceland, The Netherlands and Spain. Météo-France is also involved in a research cooperation. At the member-states weather-services, HIRLAM is used for operational short-range synoptic forecasts.

When a 3DVAR data assimilation scheme was developed around 2000, direct assimilation of satellite radiances became possible. Today, AMSU-A radiances are used operationally in Denmark, Norway, Sweden and Spain. Operational use with a good data coverage became feasible with EARS, EUMETSAT ATOVS Retransmission Service. A positive impact from AMSU-A on forecast scores, especially temperature and geo-potential, has been seen in most impact studies, figure 1 and also in e.g. [Schyberg et al., 2003].

2 Use Of ATOVS In HIRLAM

The observations operator, that projects the NWP model first guess field into radiance space, used in the HIRLAM 3DVAR is RTTOV. At present, the most used version is RTTOV-7, but RTTOV-8 is being tested at DMI, Denmark. In the operational implementations, cloud-cleared radiances over sea are used. The use of high peaking channels over all surfaces is being tested at met.no (Norway) and DMI (Denmark). A limitation for the use of AMSU-A in HIRLAM is the low model top , 10hPa. This means that channels 11-14 can not be used since they peak above the model top. Another complicating factor is that the RTTOV model need profiles up to 0.1 hPa. Therefore seasonal varying climate profiles are used between 10-0.1hPa. The channels used are thus 5-10, channel 4 may be used over sea in the future with a better treatment of the surface skin temperature T_{skin}. Development of this is ongoing.

3 Bias Correction Of AMSU-A

The distribution of a large sample of observation minus first guess, $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$, statistics should ideally be Gaussian distributed around zero. This is not the case, however, and therefore a bias correction procedure is needed, figure 3 on the left. In the HIRLAM community we have not put an extensive effort in finding the sources of the biases and how to correct them. We have instead adopted an already existing method, the Harris/Kelly scheme, [Harris and Kelly, 2001]. This method adds an empirical correction to the observations:

$$y_{corr} = y_{raw} + b(\mathbf{x}_b, y_{raw}) \tag{1}$$

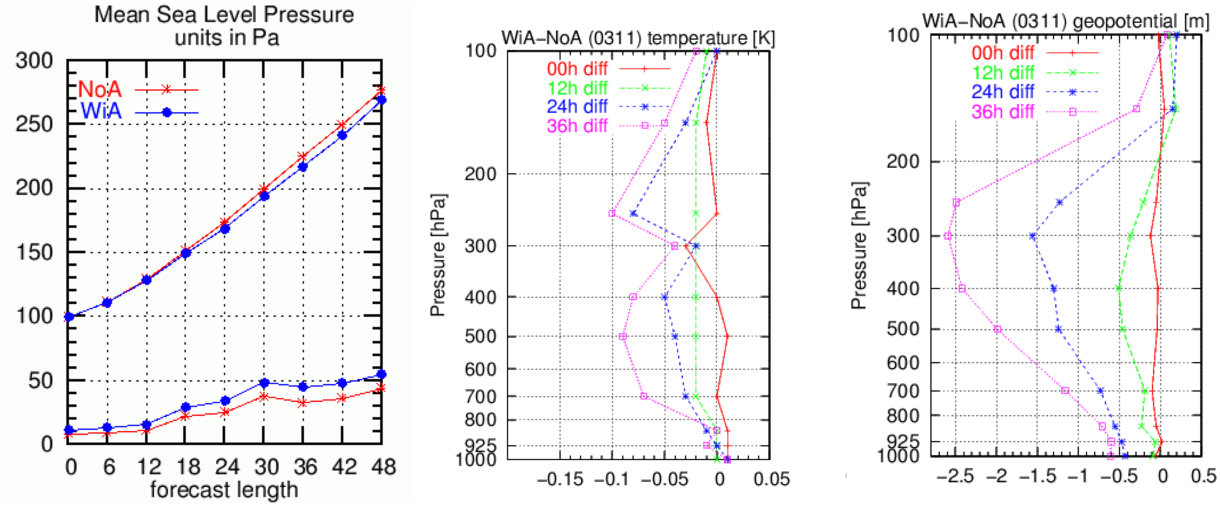

Figure 1: Verification scores from DMI, Denmark, Nov 2003. WiA means that AMSU-A from NOAA15 and 16 has been used. The left plot shows RMSE and also the bias. The middle and right plot show the difference in RMSE between NoA and WiA. Since those curves are negative, NoA (No AMSU used) has a larger RMSE. From [Amstrup, 2004]

Where b is given by a regression formula:

$$b(\mathbf{x}_b, y_{raw}) = \sum_{j=1}^{N} c_j P_j(\mathbf{x}_b, y_{raw}) \tag{2}$$

c_j are regression coefficients determined from a reference dataset. P_j are a set of predictor variables. In the present implementation these variables are, [Schyberg et al., 2003]:

1. Constant shift
2. Mean temperature between 1000hPa and 300hPa
3. Mean temperature between 200 hPa and 50hPa
4. Surface temperature
5. Integrated water vapor
6. Square of the observation scan angle
7. The observation scan angle (with sign)

As mentioned, the correction is applied to the observation even though it also corrects for biases originating from the radiative transfer model **H**. The number of predictors has lately been questioned. In the future it may be important to revise the number of predictors that we use. For example, when testing AMSU-B (a humidity sounder) at SMHI, predictors 4 and 5 were not used. In Denmark, when testing AMSU-A over ice and land, predictors 4 and 5 were also removed.

	Period	Number of observations in dataset	Latitude bands
Sweden	Jan-June 2004	500.000	00-60N, 60-90N
Denmark	Jan-May 2005	1.000.000	00-45N, 45-60N, 60-90N
Norway	Nov 2003 - Feb 2004	300.000	25-65N, 65-90N

Table 1: Examples of how the reference data-set for calculations the regression coefficients has been chosen in some HIRLAM countries. This example shows NOAA16 data.

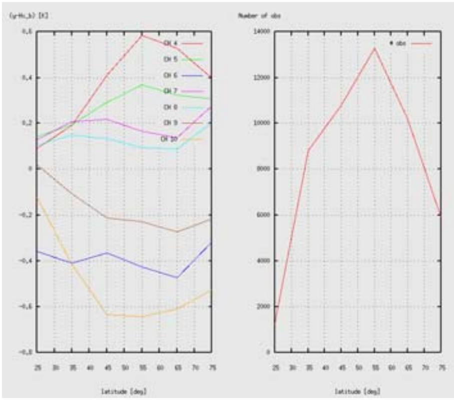

Figure 2: Bias as a function of latitude for AMSU-A channels 4-10, NOAA16. The data is from March 2005, SMHI operational suite. The number of observations are on the right.

3.1 Calculating The Coefficients

The most common way in the HIRLAM countries has been to collect data from a couple of months. This gives data-samples ranging between $3*10^5$ and 10^6 depending on model area and satellite, table 1. The observations used have all passed the quality checks of the assimilation-system. This means for AMSU-A, cloud contamination and first guess check.

The biases also have a latitude dependence, figure 2, so it is a good idea to have different coefficients for different latitude bands. However, in a regional model the latitude bands must be chosen so that it is possible get a large data-sample in each band. In practice, the selection of latitude bands is a compromise between the actual latitude dependency and the model area.

It seems like there is no international consensus on how large a data-sample should be. Especially not for a regional model. We therefore do not know how big a data-sample should be in order to be '*big enough*'. The strategy so far has been to collect as much data as we can, calculate the coefficients and then use those as long as the monitoring plots are not drifting.

3.2 Effect Of Bias-Correction

If figure 3, $(\mathbf{y}-\mathbf{Hx}_b)$ or *innovation statistics* are plotted both for raw (uncorrected) radiances and bias-corrected. From there it can be seen that the bias-correction improves the statistical distribution of the first guess departures. That is encouraging, and an indication that the method does what it is supposed to do.

It is, however, also important to see what effect it has on the forecasts. To test that, an impact study has been done where the radiances were assimilated with and without bias-correction, figure4. In that experiment, the RMS errors became larger for temperature, geo-potential and MSLP if the observations were not bias-corrected. It may even be so, that the use of non-corrected radiances gives a negative impact. That has not been tested

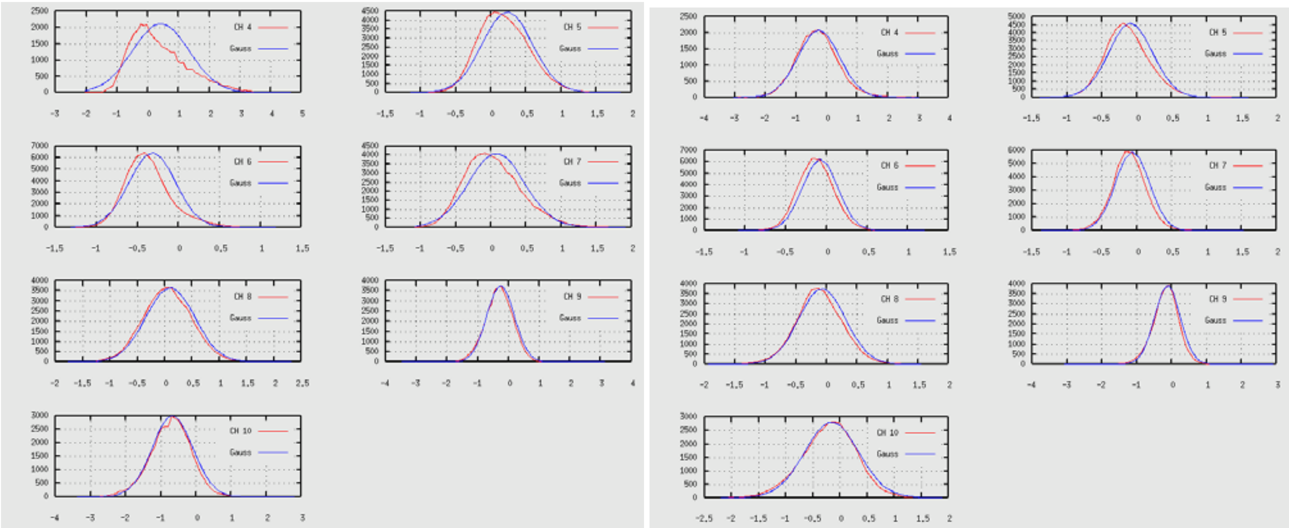

Figure 3: Obs minus first guess, $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$, statistics for AMSU-A channels 4-10. The data is from March 2005, SMHI operational suite. x-axis: $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ 0.1K slots y-axis: Number of observations Left plot: Uncorrected radiances and a Gaussian (blue) curve overlayed. Right plot: Bias-corrected radiances with a Gaussian curve

though. We can at least say that the use of bias-corrected radiances gives positive impact on forecast scores.

4 Monitoring

Monitoring of the $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ departures in the HIRLAM community is mostly done as plotting time-series. At SMHI, daily bias, RMS and variance of the $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ departures are plotted and updated once a day. The aim of the monitoring is to detect instrument errors and/or checking the stability of the bias-correction. The reason for plotting daily averages is that the number of observations available for each HIRLAM run varies a lot. HIRLAM is at SMHI run four times a day and one run can have 1000 active radiances while the next only 50.

The monitoring at SMHI has been very stable the last year, but some things have happened which indicates how difficult it can be to do monitoring in a limited area model. First of all, in late August 2005 the number of NOAA16 data drastically reduced which caused the monitoring plots to behave alarmingly noisy, figure5. That may not be very surprising, but it must be kept in mind that the behavior of the monitoring is dependent on a steady inflow of data.

In January 2005 NOAA16 AMSU-A channels 9-14 were having some kind of problem. This could be seen in the SMHI monitoring for channel 9, but not 10. At the same time, there was some mail-correspondence on the ITWG mail-list that addressed the issue. With the combination of those two factors, channels 9 and 10 could be taken out of operations at SMHI.

Those two incidents show that monitoring can be a bit problematic in a regional model. For example, instrument failures may be concealed in fluctuations due to varying data-samples.

5 Summary

The HIRLAM community have adopted the Harris/Kelly scheme to do bias-correction of AMSU-A. The strategy for calculating the coefficients is to collect as much data as possible and then use the coefficients as long as

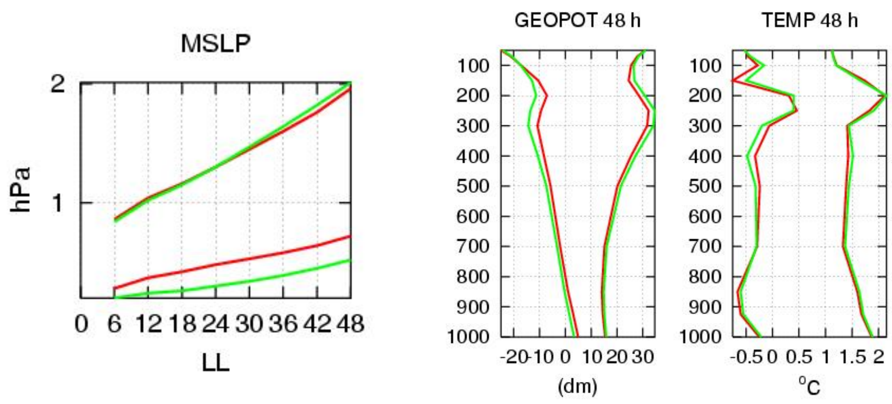

Figure 4: Red: Bias-corrected radiances used. Green: Uncorrected radiances used. Impact study at SMHI covering 1 month: 10:Aug to 11:th Sep 2004. Data comes from NOAA15 and 16.

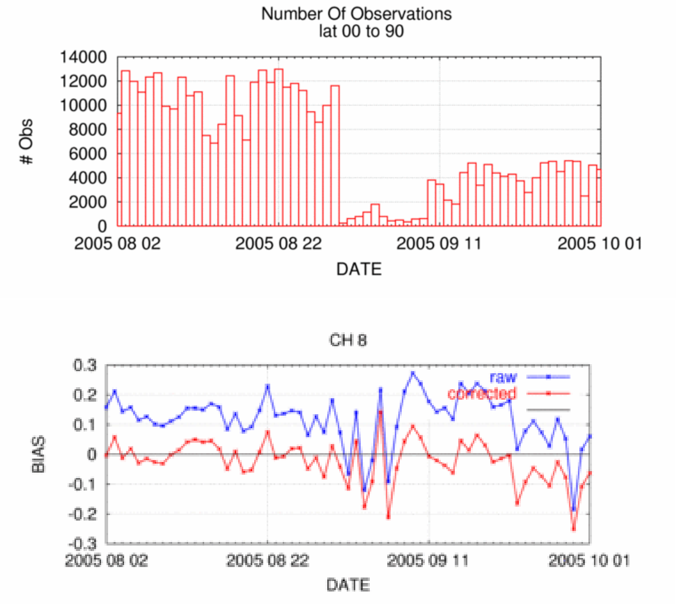

Figure 5: Monitoring plots from SMHI, Sweden. NOAA16, AMSU-A. x-axis on lower plot: the dates. y-axis on lower plot: 24h averages of $(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}_b)$ *statistics.*

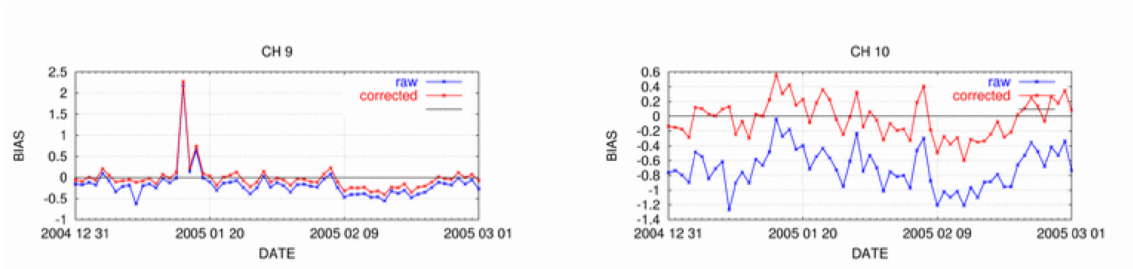

Figure 6: Monitoring plots from SMHI, Sweden. NOAA16, AMSU-A. x-axis: the dates. y-axis: 24h averages of $(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ *statistics.*

the monitoring plots are not drifting. Over all, our bias-correction procedure seems to work quite well:

- The bias-correction improves innovation statistics.
- We get a positive impact on forecast scores if bias-corrected radiances are assimilated.

Monitoring is problematic in a regional model. Since the data-samples vary from run to run and from day to day, the monitoring plots can fluctuate as a consequence of that. This may conceal fluctuations due to instrument failures.

References

[Amstrup, 2004] Amstrup, B. (2004). Impact of noaa15 and noaa16 atovs amsu-a radiance data in the dmi-hirlam 3d-var data assimilation system - november and december 2003. *HIRLAM Newsletter*, 45:235–247.

[Harris and Kelly, 2001] Harris, B. and Kelly, G. (2001). A satellite bias correction scheme for data assimilation. *Quaterly Journal of the Royal Meteorological Society*, 127:1453–1468.

[Schyberg et al., 2003] Schyberg, H., Landelius, T., Thorsteinsson, S., Tveter, F. T., Vignes, O., Amstrup, B., Gustafsson, N., Jarvinen, H., and Lindskog, M. (2003). Assimilation of atovs data in the hirlam 3d-var system. *HIRLAM Technical Report*, 60.