Note on the use of bias predictors for bias correction of AMSU-A data in DMI-HIRLAM

Bjarne Amstrup
Danish Meteorological Institute
Lyngbyvej 100, DK-2100 Copenhagen Ø
e-mail: bja@dmi.dk

1. Introduction

Several impact studies with the DMI-HIRLAM analysis and forecasting system have shown a positive (initially a neutral) impact when using ATOVS[1] AMSU-A[2] radiances in terms of brightness temperatures with data from NOAA16[3], alone or a combination of data from NOAA16 and NOAA17 or NOAA15 ((Amstrup, 2001; Schyberg *et al.*, 2003; Amstrup, 2002; Amstrup, 2003; Amstrup, 2004)). Accordingly, this type of observational data have been used operationally at DMI since December 2002, after a long period of testing the impact of the data in the pre-operational suite.

An important part of assimilating some types of satellite radiances, such as AMSU-A, is the bias correction since the differences between observed radiances and the model derived radiances are neither bias-free nor have a Gaussian distribution. By applying a bias correction it turns out that the differences between observed and model derived radiances behaves much better by becoming approximately bias-free and have a distribution much closer to a Gaussian shape (see, e.g., plots in (Schyberg *et al.*, 2003)). In connection to the implementation of the bias correction in the HIRLAM 3D-VAR system, HIRVDA, it was decided to use 7 predictors in a Harris-Kelly scheme ((Harris and Kelly, 2001), see next section) for use over open sea ((Schyberg *et al.*, 2003)). However, some of these predictors become questionable when assimilating radiances over sea-ice and land, and based on the discussions at the recent "ECMWF/EUMETSAT NWP-SAF Workshop on Bias estimation and correction in data assimilation" (see `http://www.ecmwf.int/newsevents/meetings/workshops/2005/NWP_SAF/index.html`) 8-11 November 2005, it was decided to investigate the bias correction scheme, and evaluate the importance of the different predictors. In addition I wanted to test if a change in one of the air mass predictors might be a good idea. The latter is also important if the same predictors should be used for the channels to be used over land surfaces. The present study was inspired by the presentation made by Nigel Atkinson (see (Atkinson *et al.*, 2006)) at the above mentioned workshop.

2. Bias correction at DMI

In order to use satellite radiances in the assimilation system, a forward model to calculate radiances from model data is needed. The radiative forward model presently used at DMI for calculating

[1] Advanced TIROS (Television Infra-Red Observation Satellite) Operational Vertical Sounder
[2] Advanced Microwave Sounding Unit-A
[3] National Oceanic and Atmospheric Administration

brightness temperatures is based on RTTOV-8[4], available from from the Numerical Weather Prediction SAF[5] (see `http://www.metoffice.com/research/interproj/nwpsaf/index.html`, Saunders *et al.*, 1999 and Matricardi *et al.*, 2004).

The bias correction function used in the operational DMI-HIRLAM 3D-VAR is given by

$$b(\mathbf{x}_b, y^i_{\rm raw}) = c_0 + \sum_{j=1}^{N} c_j P_j(\mathbf{x}_b, y^i_{\rm raw}) \tag{1}$$

where c_j are constant bias coefficients and $P_j(\mathbf{x}_b, y^i_{\rm raw})$ are a set of predictor variables. The implementation of bias correction uses 7 predictor variables:

Pred. 1: a constant displacement,

Pred. 2: a measure (see Eq. 2) of the mean forecast temperature between the 1000 hPa and 300 hPa pressure levels,

Pred. 3: a measure (see Eq. 3) of the mean temperature between 200 hPa and 50 hPa,

Pred. 4: the surface temperature,

Pred. 5: the integrated water vapor content per area from the surface up to the top of the atmosphere,

Pred. 6: the square of the observation zenith angle and

Pred. 7: the observation zenith angle.

Predictor 2 is the measure for the mean forecast temperature between the 1000 hPa and 300 hPa pressure levels and is given by

$$P_2 = \sum_{j=b_2}^{e_2} \frac{T_j - T_{j-1}}{\log(p_j/p_{j-1})} \tag{2}$$

and similar for predictor 3:

$$P_3 = \sum_{j=b_3}^{e_3} \frac{T_j - T_{j-1}}{\log(p_j/p_{j-1})} \tag{3}$$

In eqs. 2 and 3 T_j is the temperature at the target pressure p_j (taken from the RTTOV specified levels) and the summation over j is determined by the number of target pressures between the given limits. Presently, $b_2 = 25$, $e_2 = 43$, $b_3 = 14$ and $e_3 = 21$. A test version has $b_2 = 25$ and $e_2 = 37$ corresponding to pressures between 840 hPa and 321 hPa. For data over sea-ice used operationally and for data over land used in tests, the coefficients for predictors 4 (surface temperature) and 5 (integrated water vapor) is set to 0, and therefore not used as predictors. Given the present state of the accuracy of the model surface temperature over sea-ice and land, and the from time to time rather strong diurnal cycle of surface temperature over land, it has been natural not to use predictor 4 over sea-ice and land at DMI. Furthermore, it often happens that the water vapor is highly variable in the atmosphere. This

[4]Radiative Transfer model for TOVS, release 8

[5]Satellite Application Facility

is not always encountered for in the model and the integrated water vapor predictor was therefore also excluded over sea-ice and over land. The coefficients presently used in operations and in the preliminary tests for use over land have been derived from a 5 month period from January 1 2005 to May 31 2005, except for the period January 15 to January 18. The data from this period have also been used in the "revised" sets of coefficients that have been made in this study of the importance of the different predictors. In the following we will denote the collection of predictors 1, 6 and 7 as the scan angle predictors.

The bias corrections are based on a longer time series of innovations, which are differences between the modeled (by the radiative transfer model) data and the observed data. Here, the innovations are based on archived operational DMI-HIRLAM-T15 3 h forecasts (see (Sass *et al.*, 2005) for a short description on this model) made from the 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC and 21 UTC analyses in the 5 month period mentioned above for NOAA15 and NOAA16 AMSU-A data. HIRVDA version 6.3.6 modified to use RTTOV8 has been used for calculating the innovations, and the thinning of data has been reduced compared to the thinning in the operational 3D-VAR analyses.

Figure 1 shows the bias correction as function of scan angle when only the scan angle predictors are used for NOAA15 and NOAA16 data. Similarly, Figure 2 shows the bias correction due to the predictors 6 and 7 as function of scan angle (note that the y-axis scale is different than in Figure 1) when all predictors are used for NOAA15 and NOAA16 data. The figures show that the scan angle corrections have similarities between NOAA15 and NOAA16 and also that the shapes of the curves are similar in the two figures.

Table 1 shows the statistics for the innovations for the period from January through May 2005 for channels 4 to 10 for NOAA15 and NOAA16 AMSU-A data and the statistics for the innovations for a 108 day period starting early August 2005 for NOAA18. The two left most columns are for uncorrected data and show the overall bias and standard deviation for all the data. The other columns show the standard deviation given the set of bias predictors listed in the top row (as an example 'σ_{1467}' is short for the test in which the bias predictors used are 1, 4, 6 and 7). The biases are not given for the bias corrected data since almost all biases as expected are ± 0.000 K.

It is quite clear from this table that the scan angle correction alone to a very large extend make the reduction in standard deviation compared to the reduction when all predictors are used. Furthermore, by adding the two air-mass predictors 2 and 3 the full reduction of standard deviation is accomplished, or almost accomplished, for all the channels used. It is also seen that even though the overall bias for the uncorrected data are close to zero for some of the channels (such as for channel 4 on NOAA15 and channels 4 and 7 on NOAA16), the standard deviations are still reduced substantially by making a bias correction – by more than 40 % for channel 7 on NOAA16 as the largest. Furthermore, it is seen that predictor 3 seems to be the most important predictor when adding just one extra predictor to the scan angle predictors as the columns with titles 'σ_{1267}', 'σ_{1367}', 'σ_{1467}' and 'σ_{1567}' show. For channels 4 and 5 the effect of adding an extra predictor to the scan angle predictors is rather small. Table 2 shows similar results for NOAA15 and NOAA16 data when the revised pressure limit of 840 hPa instead of 1000 hPa is used for predictor 2. Note that the number of used data is slightly different in the two sets of data causing the very small deviations in the bias of the raw data given in the tables.

Figures 3-5 show the distribution of the number of innovations within 0.05 K intervals for channels 4 through 10. The best fit Gaussian based on the lowest squared differences between a Gaussian

Table 1: Standard deviation (in kelvin) depending upon number of predictors. The first two columns denoted by '$\mu_{\rm raw}$' and '$\sigma_{\rm raw}$' are bias and standard deviation (also in kelvin), respectively, for the raw data. The column denoted σ_{12367} (as an example) gives the standard deviations for the bias corrected data using predictors 1, 2, 3, 6 and 7.

NOAA15 (1561501 data points)										
chnl	$\mu_{\rm raw}$	$\sigma_{\rm raw}$	σ_{167}	σ_{1467}	σ_{1567}	σ_{1267}	σ_{1367}	σ_{12367}	σ_{123467}	$\sigma_{\rm all}$
4	−0.009	0.577	0.547	0.546	0.546	0.544	0.546	0.543	0.543	0.540
5	0.252	0.520	0.492	0.491	0.491	0.491	0.490	0.489	0.489	0.489
6	−0.631	0.590	0.472	0.471	0.472	0.471	0.452	0.451	0.451	0.451
7	0.172	0.500	0.358	0.357	0.357	0.356	0.352	0.351	0.351	0.350
8	0.090	0.359	0.338	0.337	0.337	0.336	0.325	0.324	0.324	0.324
9	−0.234	0.456	0.376	0.375	0.374	0.372	0.365	0.362	0.361	0.361
10	−0.794	0.716	0.568	0.558	0.558	0.551	0.505	0.493	0.490	0.490
NOAA16 (1794608 data points)										
4	−0.025	0.511	0.484	0.484	0.483	0.483	0.484	0.483	0.483	0.477
5	0.060	0.261	0.254	0.254	0.254	0.252	0.253	0.251	0.251	0.250
6	−0.349	0.283	0.192	0.191	0.191	0.190	0.188	0.187	0.187	0.187
7	0.010	0.408	0.249	0.245	0.245	0.238	0.216	0.212	0.211	0.211
8	0.051	0.439	0.417	0.414	0.414	0.411	0.395	0.394	0.394	0.394
9	−0.178	0.457	0.348	0.347	0.346	0.346	0.341	0.339	0.338	0.338
10	−0.654	0.659	0.525	0.520	0.520	0.518	0.483	0.478	0.475	0.475
NOAA18 (1239835 data points)										
4	0.132	0.538	0.493	0.492	0.493	0.491	0.490	0.489	0.489	0.485
5	0.278	0.294	0.272	0.272	0.272	0.270	0.270	0.269	0.269	0.269
6	−0.529	0.220	0.188	0.188	0.188	0.188	0.185	0.184	0.184	0.184
7	−0.309	0.371	0.222	0.222	0.222	0.219	0.206	0.205	0.204	0.204
8	−0.106	0.303	0.276	0.275	0.275	0.273	0.268	0.267	0.267	0.267
9	−0.697	0.354	0.290	0.289	0.289	0.289	0.281	0.280	0.279	0.279
10	−1.151	0.544	0.427	0.414	0.417	0.413	0.398	0.381	0.379	0.378

Table 2: Standard deviation (in kelvin) depending upon number of predictors. Similar to Table 1 except for predictor 2 is the revised one for which the pressure limit is 840 hPa instead of 1000 hPa.

NOAA15 (1558873 data points)							
chnl	μ_{raw}	σ_{raw}	σ_{167}	σ_{1267}	σ_{1367}	σ_{12367}	σ_{all}
4	−0.009	0.577	0.547	0.544	0.546	0.543	0.541
5	0.252	0.520	0.491	0.490	0.490	0.489	0.489
6	−0.632	0.590	0.472	0.470	0.452	0.451	0.450
7	0.172	0.500	0.358	0.356	0.352	0.351	0.350
8	0.089	0.359	0.339	0.336	0.325	0.324	0.324
9	−0.235	0.456	0.377	0.373	0.366	0.362	0.361
10	−0.796	0.717	0.569	0.552	0.506	0.494	0.490
NOAA16 (1796950 data points)							
4	−0.024	0.511	0.485	0.484	0.484	0.483	0.478
5	0.060	0.261	0.255	0.252	0.253	0.251	0.251
6	−0.349	0.283	0.192	0.190	0.188	0.187	0.187
7	0.010	0.408	0.249	0.237	0.216	0.212	0.211
8	0.051	0.439	0.417	0.411	0.395	0.394	0.394
9	−0.179	0.457	0.348	0.346	0.341	0.339	0.338
10	−0.656	0.660	0.525	0.518	0.483	0.478	0.475

function with the same bias and the actual numbers are also overlayed. Figures 3 and 4 are by using all predictors in the bias corrected innovations for NOAA15 and NOAA16, respectively. Figure 5 shows the distribution for the bias corrected data using the scan angle predictors 1, 6 and 7. The introduction of the bias correction do make a good job in producing nearly Gaussian statistics. The "shoulders" seen in some of the channels in the plots with uncorrected data are basically removed by the bias correction and the width of the best fit Gaussian are somewhat smaller for most of the channels as indicated in Table 1 as well. The scan angle predictors alone seem to be sufficient in this respect.

Figure 6 shows similar distributions of bias corrected innovations for data over sea-ice. The figure shows that coefficients based on data over open sea alone cannot produce bias-free distributions.

3. Conclusion

It has been shown that by reducing the number of bias predictors from 7 to 5, excluding the surface temperature and integrated water vapor predictors, lead to as good statistics of innovations for the relative long period studied here. Provided that impact tests with use of air mass bias predictors 2 and 3 in addition to the scan angle bias predictors 1, 6 and 7 gives as good an impact as when using all

the predictors, it should be made operational when updates of bias correction coefficients are made. Furthermore, if this is the case, then it should be tested if a common set of bias predictors for different surface types can be used even though Figure 6 indicates that it may not be possible.

Acknowledgment

The author wants to thank Nigel Atkinson (UKMO) for supplying updated material from his lecture in the "ECMWF/EUMETSAT NWP-SAF Workshop on Bias estimation and correction in data assimilation".

References

Amstrup, Bjarne. 2001. *Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system*. DMI Scientific Report 01-06. Danish Meteorological Institute.

Amstrup, Bjarne. 2002 (March). Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system - February 2002. *Pages 68–76 of: HIRLAM Workshop Report of HIRLAM Workshop on Variational Data Assimilation and Remote Sensing*.

Amstrup, Bjarne. 2003. *Impact of NOAA16 and NOAA17 ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-VAR analysis and forecasting system — January and February 2003*. DMI Scientific Report 03-06. Danish Meteorological Institute.

Amstrup, Bjarne. 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.

Atkinson, Nigel, Cameron, James, Candy, Brett and English, Stephen. 2006 (February). Bias correction of satellite data at the Met Office. *Pages 169–176 of: ECMWF/EUMETSAT NWP-SAF Workshop on Bias estimation and correction in data assimilation. 8-11 November 2005*.

Harris, B. A. and Kelly, G. 2001. A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Roy. Meteorol. Soc.*, **127**, 1453–1468.

Matricardi, Marco, Chevallier, Frederic, Kelly, Graeme and Thépaut, Jean-Noël. 2004. An improved general fast radiative transfer model for the assimilation of radiance observations. *Quart. J. Roy. Meteorol. Soc.*, **130**, 153–173.

Sass, Bent Hansen, Kmit, Maryanne and Yang, Xiaohua. 2005. Operational DMI-HIRLAM in 2004/2005. *Hirlam Newsletter*, **48**, 10–13.

Saunders, R., Matricardi, M. and Brunel, P. 1999. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteorol. Soc.*, **125**, 1407–1425.

Schyberg, Harald, Landelius, Tomas, Thorsteinsson, Sigurdur, Tveter, Frank Thomas, Vignes, Ole, Amstrup, Bjarne, Gustafsson, Nils, Järvinen, Heikki and Lindskog, Magnus. 2003. Assimilation of ATOVS data in the HIRLAM 3D-VAR System. *HIRLAM Technical Report*, **60**.

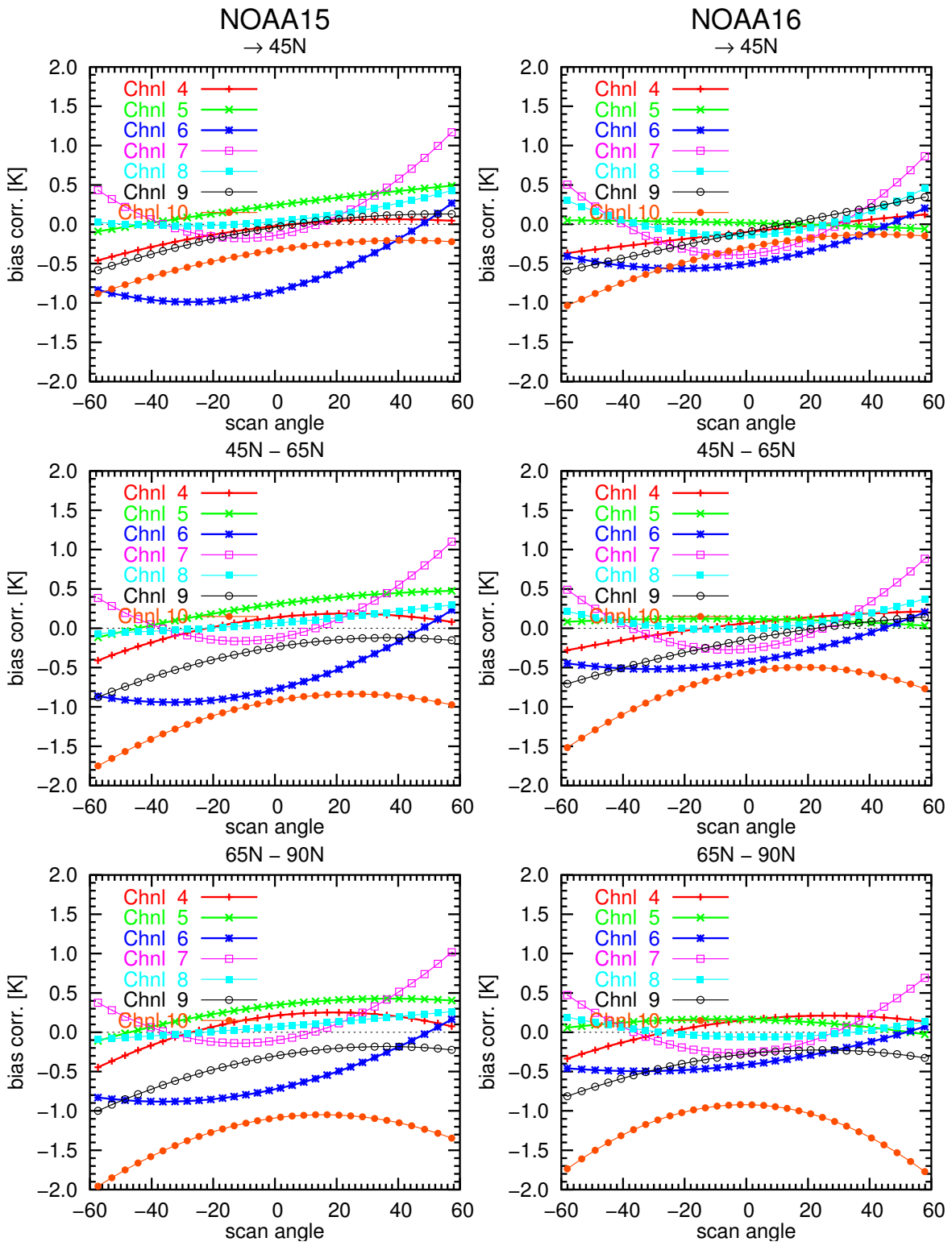

Figure 1: The bias correction as function of scan angle when only predictors 1 (constant), 7 (scan angle) and 6 (squared scan angle) are used. The corrections are for three latitude bands: south of 45°N (upper), between 45°N and 65°N (middle) and north of 65°N (bottom).

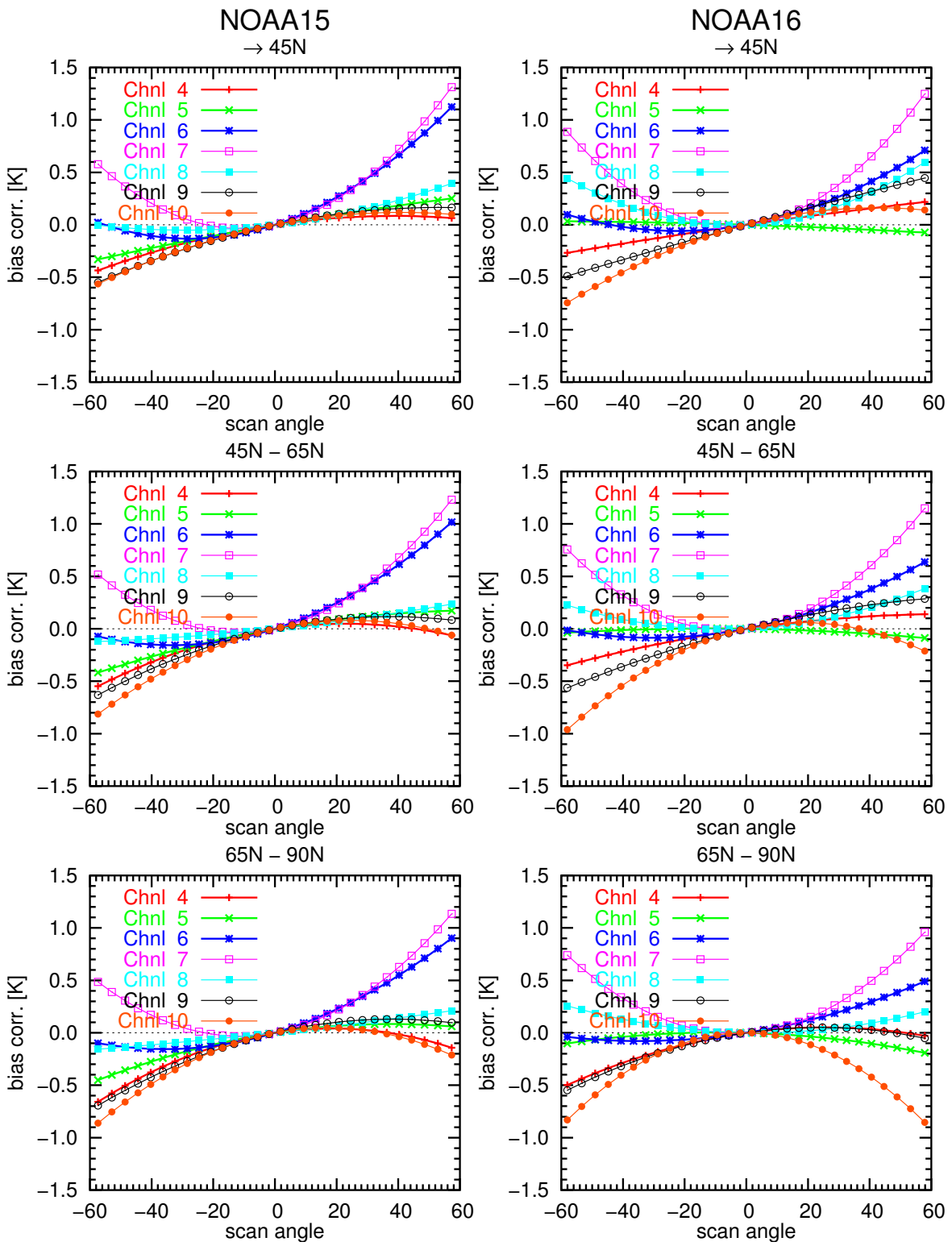

Figure 2: The scan angle part (excluding the constant) of the bias correction as function of scan angle when all predictors are used. The corrections are for three latitude bands: south of 45°N (upper), between 45°N and 65°N (middle) and north of 65°N (bottom).

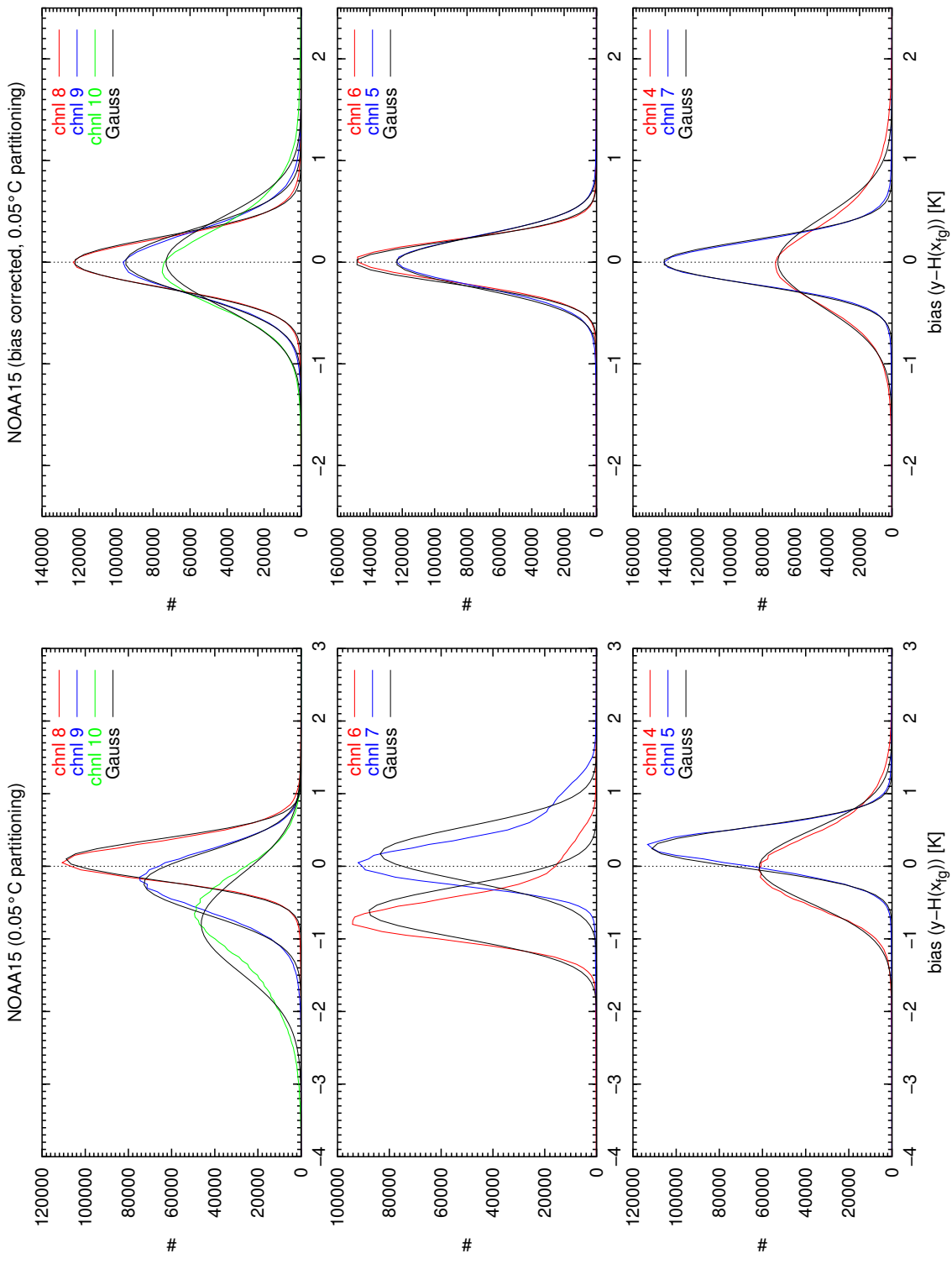

Figure 3: Partitioning of NOAA15 data from the January through May 2005 period according to the number of data with given differences (innovations) between observed brightness temperature (y_{raw}) and the modelled one ($H(\mathbf{x}_{\mathrm{mod}})$). Left hand side is for uncorrected data and the right hand side is for bias-corrected data using all predictors. Overlayed is the "best fit" Gaussian distribution.

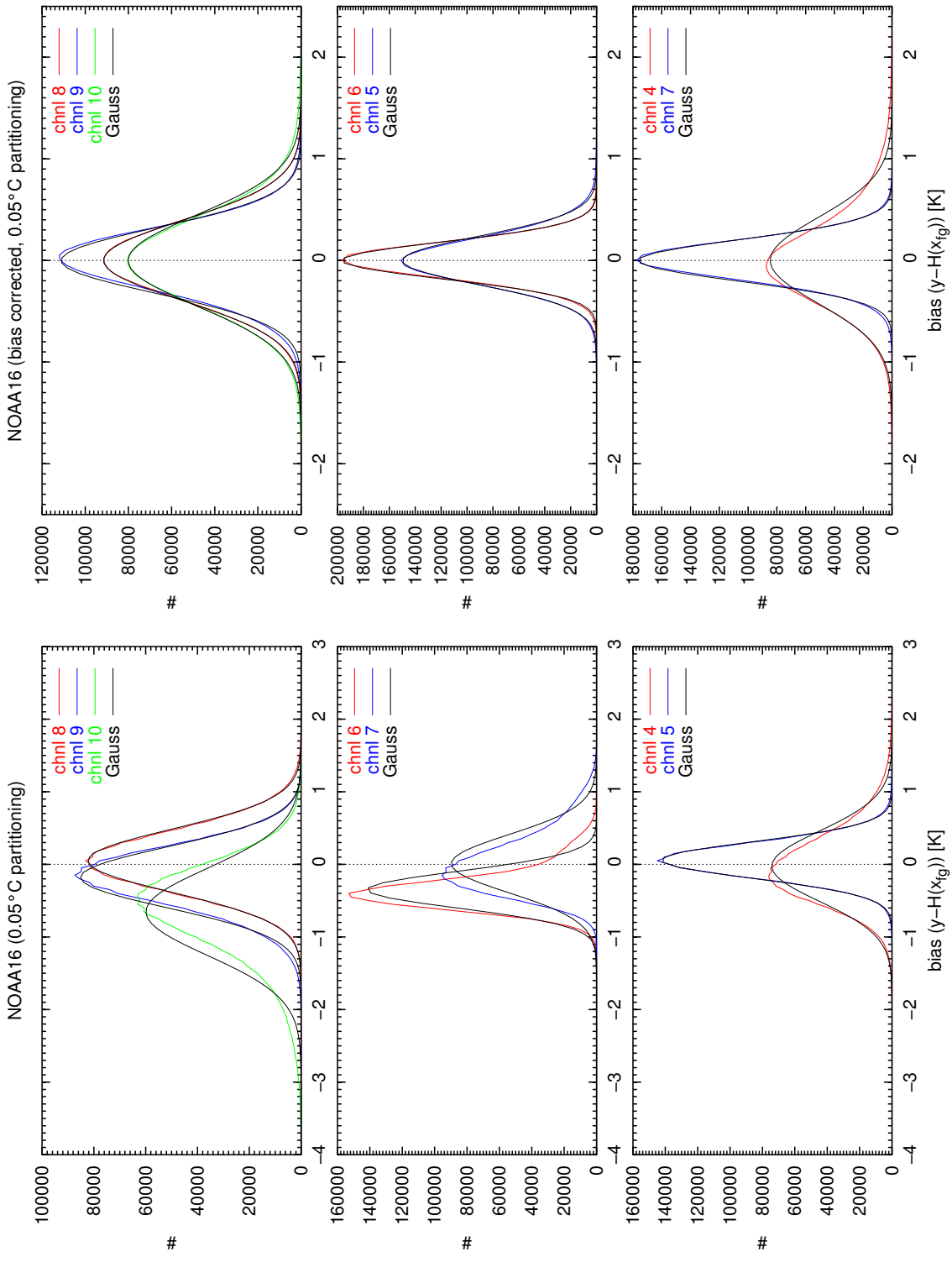

Figure 4: Partitioning of NOAA16 data from the January through May 2005 period according to the number of data with given differences (innovations) between observed brightness temperature (y_{raw}) and the modelled one ($H(\mathbf{x}_{\mathrm{mod}})$). Left hand side is for uncorrected data and the right hand side is for bias-corrected data using all predictors. Overlayed is the "best fit" Gaussian distribution.

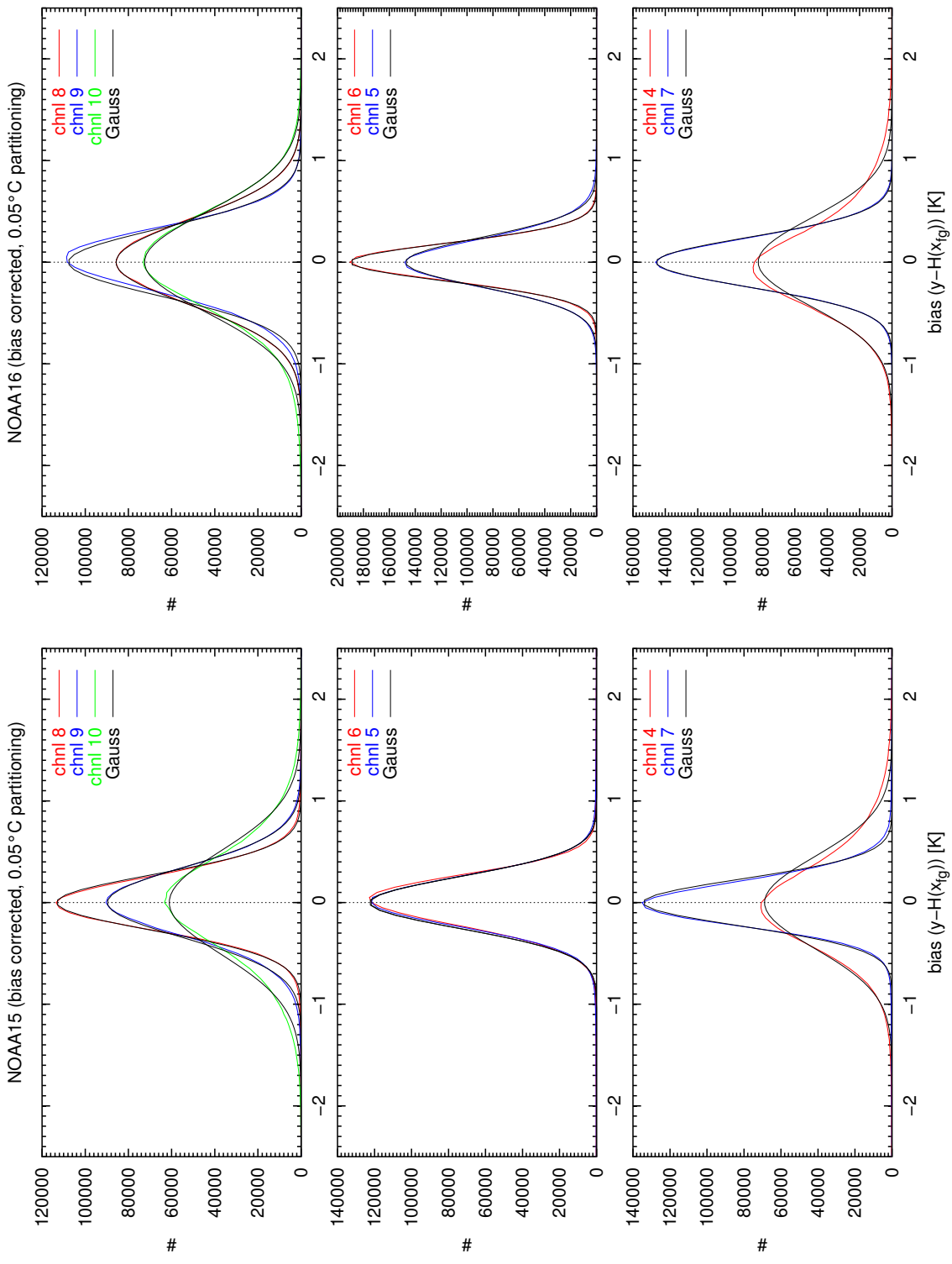

Figure 5: Partitioning of NOAA15 (left) and NOAA16 (right) data from the January through May 2005 period according to the number of data with given differences (innovations) between observed brightness temperature (y_{raw}) and the modelled one ($H(\mathbf{x}_{\mathrm{mod}})$). The plots are with bias-corrected data using predictors 1, 6 and 7 (scan dependent including constant bias predictors). Overlayed is the "best fit" Gaussian distribution.

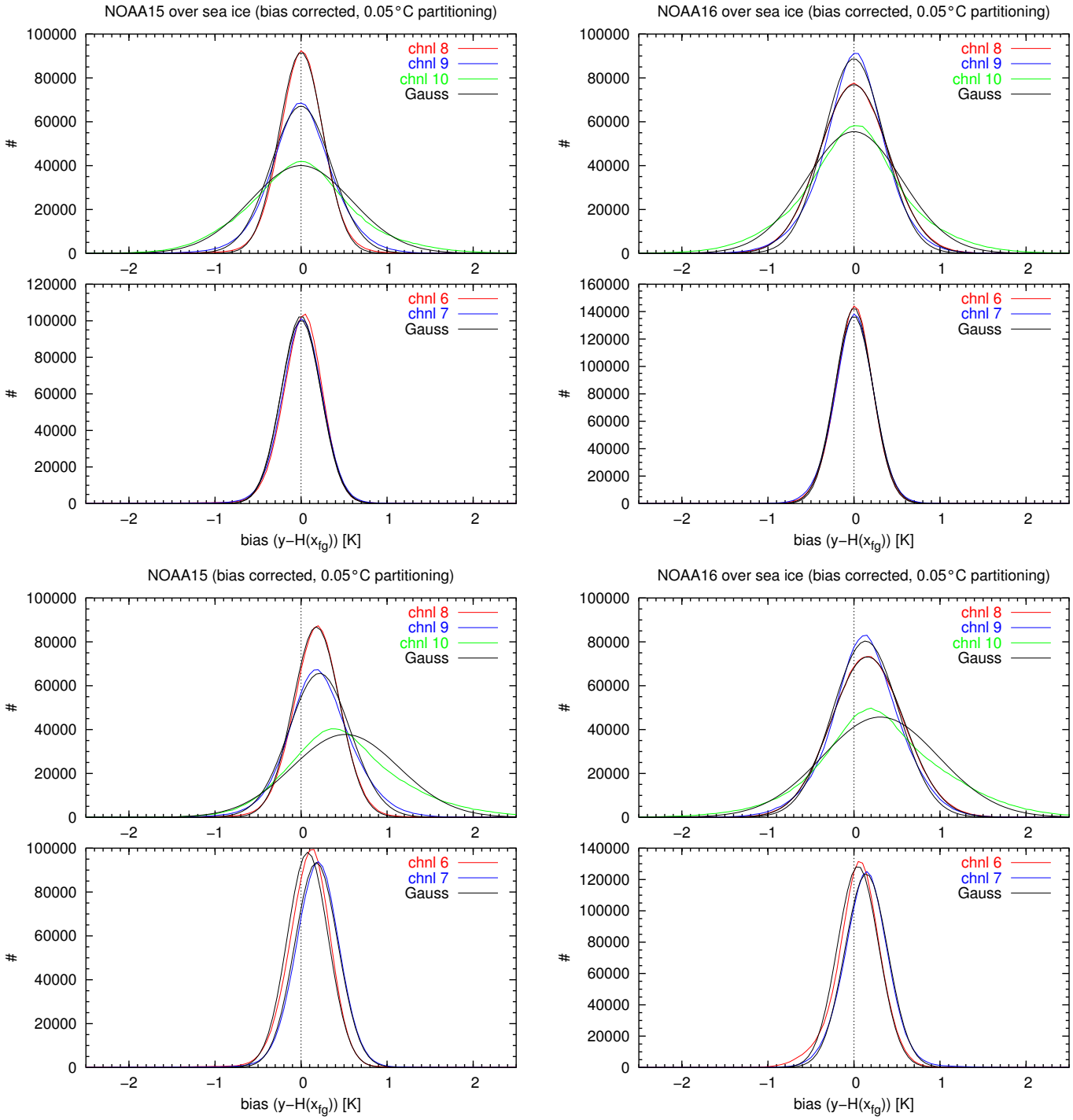

Figure 6: Partitioning of NOAA15 (left) and NOAA16 (right) data from the January through May 2005 period according to the number of data with given differences (innovations) between observed brightness temperature (y_{raw}) and the modelled one ($H(\mathbf{x}_{\mathrm{mod}})$) for data over ice. The plots are with bias-corrected data using predictors 1, 2, 3, 6 and 7 (scan dependent bias predictors and two air mass predictors). Upper plots are with coefficients based on data over sea-ice and lower plots are with coefficients based on data over open sea. Overlayed is the "best fit" Gaussian distribution.