Use of VarBC at Météo-France

Météo-France and CNRS - CNRM/GMAP/OBS

May 19, 2020

Contents

1 Theoretical VarBC formulation

The implementation of VarBC has been described in Dee (2004) and Auligné et al. (2007). The bias is computed following the scheme introduced in Harris and Kelly (2001). It is expressed as a linear combination of bias predictors (such as the scan position, the thickness of given atmospheric layers, ...). For a given satellite, instrument and channel, the estimated bias is written as:

$$b(x, \beta) = \sum_{j=1}^{N} \beta_j p_j(x) \tag{1}$$

where $p_j(x)$ are the N bias predictors and β_j the N bias prediction coefficients.

The reader can refer to Ide et al. (1997) for more information on the notations used.

1.1 Augmented control vector

An augmented control vector that includes bias prediction coefficients can be defined:

$$z^T = [x^T \beta^T] \tag{2}$$

The error covariance matrix of z is defined as $Z =< \tilde{e}^b(\tilde{e}^b)^T >$ where $z^b = z^t + \tilde{e}^b$. It is currently assumed that parameter estimations error and state estimation errors are not correlated which leads to :

$$Z = \begin{bmatrix} B_x & 0 \\ 0 & B_\beta \end{bmatrix} \tag{3}$$

The numerical values of the B_β matrix are empirically tuned in the Météo-France system. The smaller the values are, the more reactive the adaptation of VarBC coefficients will be. Therefore, the tuning of this matrix is a compromise to find between reactivity and stability of the bias correction.

1.2 Extended observation operator

The extended observation operator that accounts for the bias is defined as follows:

$$\tilde{H}(z) = H(x) + b(x, \beta) = H(x) + \sum_{j=1}^{N} \beta_j p_j(x) \tag{4}$$

Assuming that $b(x, \beta) \approx b(x^g, \beta)$ where x^g stands for the atmospheric state guess (for a given outer loop). Its tangent linear is defined as:

$$\tilde{\mathbf{H}}(\delta z) = \mathbf{H}(\delta x) + \mathbf{b}(\delta\beta) = \mathbf{H}(\delta x) + \sum_{i=1}^{N} \delta\beta_i p_i(x^g) \tag{5}$$

1.3 Modified 4D-var cost function

Using the augmented control vector introduced in Eq. 2, one can easily obtain the cost function of the 4D-Var:

$$J(z) = \frac{1}{2}(z^b - z)^T Z^{-1}(z^b - z) + \frac{1}{2}\sum_{i=1}^{n}(y_i^o - \tilde{H}_i(z))^T R^{-1}(y_i^o - \tilde{H}_i(z)) \tag{6}$$

where n is the number of time slots.

From Eq. 3 and 5, it be rewritten as follows:

$$\begin{aligned} J(x, \beta) \; = \; & \tfrac{1}{2}(x^b - x)^T B_x^{-1}(x^b - x) + \\ & \tfrac{1}{2}(\beta^b - \beta)^T B_\beta^{-1}(\beta^b - \beta) + \\ & \tfrac{1}{2}\textstyle\sum_{i=1}^{n}(y_i^o - H_i(x) - b_i(x, \beta))^T R^{-1}(y_i^o - H_i(x) - b_i(x, \beta)) \end{aligned} \tag{7}$$

In the Météo-France data assimilation scheme, an incremental formulation is used in conjunction with two outer loops (see Courtier et al., 1994). The cost function is therefore modified as follows:

$$\begin{aligned} J(\delta x, \delta\beta) \; = \; & \tfrac{1}{2}(\delta x + x^g - x^b)^T B_x^{-1}(\delta x + x^g - x^b) + \\ & \tfrac{1}{2}(\delta\beta + \beta^g - \beta^b)^T B_\beta^{-1}(\delta\beta + \beta^g - \beta^b) + \\ & \tfrac{1}{2}\textstyle\sum_{i=1}^{n}(d_i - \mathbf{H}_i(\delta x) - \mathbf{b}_i(\delta\beta))^T R^{-1}(d_i - \mathbf{H}_i(\delta x) - \mathbf{b}_i(\delta\beta)) \end{aligned} \tag{8}$$

where d_i is the innovation vector for a given timeslot $d_i = y_i^o - H_i(x^g) - b_i(x^g, \beta^g)$.

It should be noted that both parts of the augmented control vector z are pre-conditioned (which leads to further modification in the expression of J). The preconditioning of β is based on bias predictors statistics that are gathered during the non-linear calculations (*i.e.* during the screening and trajectories). The preconditioning of β is fully described in Dee (2004).

2 Practical implementation: a few examples

VarBC is used in both the global and the regional data assimilation systems at Météo-France but in slightly different ways:

- Within the global model ARPEGE, all coefficients are trained for each sensor and each channel within the 4D-Var and are updated during each assimilation cycle.
- Within the limited area model AROME, a different approach is used. For Geostationary satellite data, the data coverage is considered as sufficient over the limited area domain to train VarBC coefficients dedicated to AROME. For Low Earth Orbiting satellites data, the data coverage is considered as too low to train robust coefficients. Therefore a different strategy is applied and it is the coefficients from ARPEGE which are used to bias correct the radiances within AROME data assimilation.

For each sensor (and each individual channel), one have the possibility to select dedicated bias predictors. The current list of predictors available within the IFS/ARPEGE cycle46t1 can be found in Table 1.

1000-300 hPa thickness	1 (constant)	nadir viewing angle	land or sea ice mask
200-50 hPa thickness	T skin	nadir view angle **2	land mask times winds
10-2 hPa thickness	total column water	nadir view angle **3	
50-5 hPa thickness	surface wind speed	nadir view angle **4	
		view angle (land)	
Radiosonde T 100-850	ln(rain rate+1)	view angle **2 (land)	
Radiosonde T 30-200	ln(rain rate+1)**2	view angle **3 (land)	
Radiosonde T 0- 60	ln(rain rate+1)**3		
Radiosonde T solar elevation		day/night	
Radiosonde T solar elevation**2	descent rate (hPa/s)	cos solar zen angle	
Radiosonde log press	ascent rate (hPa/s)	thermal contrast	
		solar elevation	

Table 1: List of VarBC predictors within ARPEGE/IFS cy46t1

The air mass predictors help to adapt the bias correction to various meteorological conditions including differences in biases between the Tropics and the mid-latitudes. The geometrical predictors taking into account the viewing angles can be helpful to correct for angular biases which can arise from the instruments or the radiative transfer simulations. Some predictors are also related to the position of the sun with respect to the satellite and can be useful for some instruments as it will be shown with the SSMIS example. Some predictors are also less used in our system than others. For instance, one do not use the rain rate predictors which are more adapted to the bias correction of rainfall data assimilation.

Three examples of bias corrections are given below, for the MHS and the AMSU-A sounders onboard MetOp-C as well as for SSMIS onboard the DMSP-F18 satellite. They are given for the ARPEGE 4D-Var system and are also applied to AROME as explained above.

2.1 MHS onboard MetOp-C

Within the ARPEGE 4D-Var and the AROME 3D-Var data assimilation systems, channels 3 to 5 of the MHS sounder are assimilated. Figures 1 shows the monitoring of MHS channel 3 over a 2-month period and Figure 2 the variations of its VarBC coefficients for a shorter period of 20 days. For this particular case, one can see that the bias of first guess departures is rather small, below 0.25 K. For this channel, the VarBC predictors used are a constant, two air mass predictors as well as a polynomial of the third degree for the viewing angle. As can be seen, the values of the predictors are very stable in time.

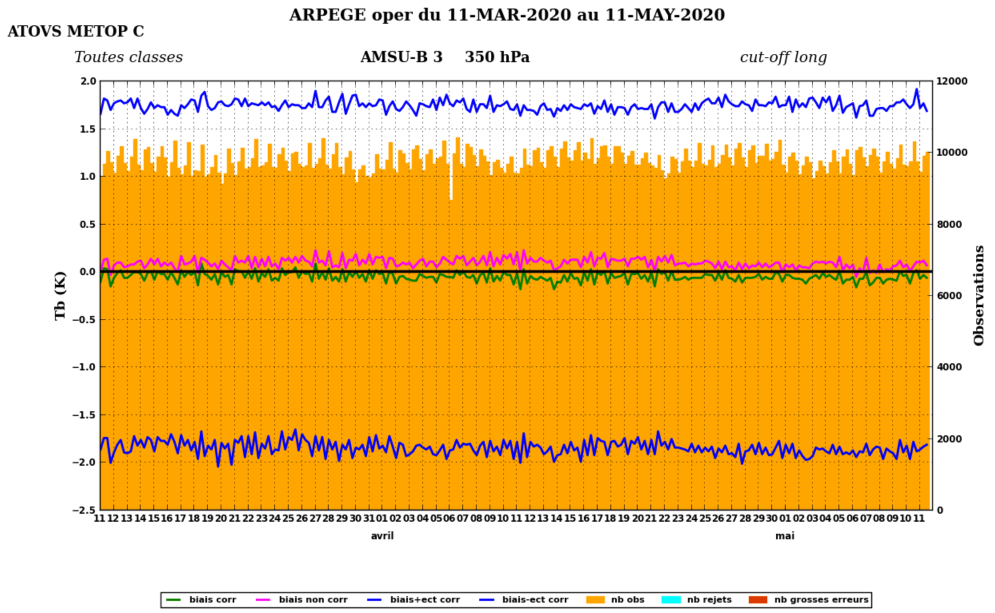

Figure 1: Monitoring of MHS Channel 3 of Metop-C within the ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

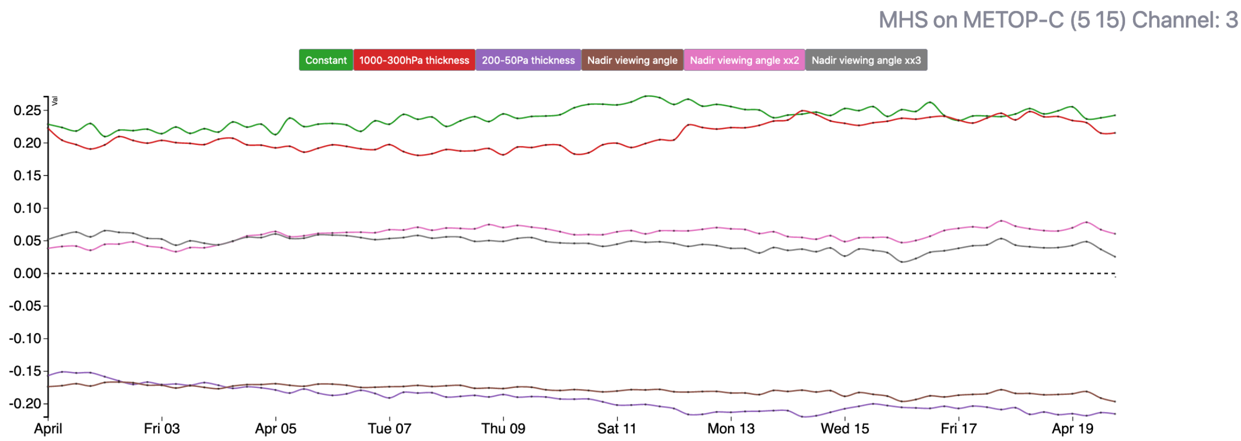

Figure 2: Variations of VarBC coefficients for the predictors selected for MHS Channel 3 from April 1st to April 20th, 2020.

The same plots are shown for MHS channel 4 on Figures 3 and 4. The bias of first guess departures is of roughly -1 K (pink line of Figure 3) and mostly corrected (green line of Figure 3) by the constant value of the predictors (green line on Figure 4). In this case, the difficulty is to not prescribe a too large number of predictors, so that VarBC do not overfit the data. But the risk of setting up a too small number of predictors is that VarBC may not be able to adapt itself in some cases of a drifting bias.

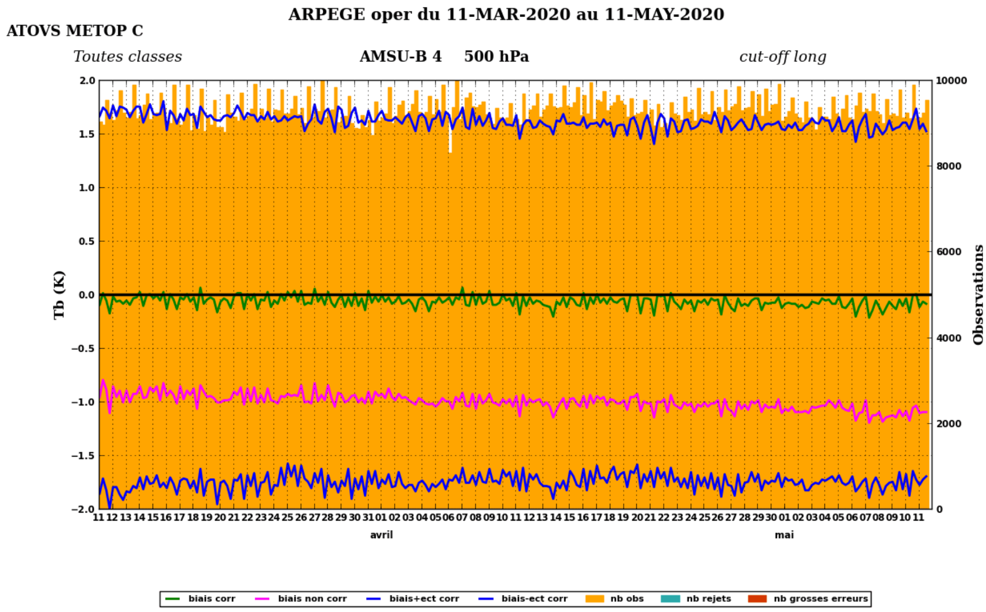

Figure 3: Monitoring of MHS Channel 3 compared to ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

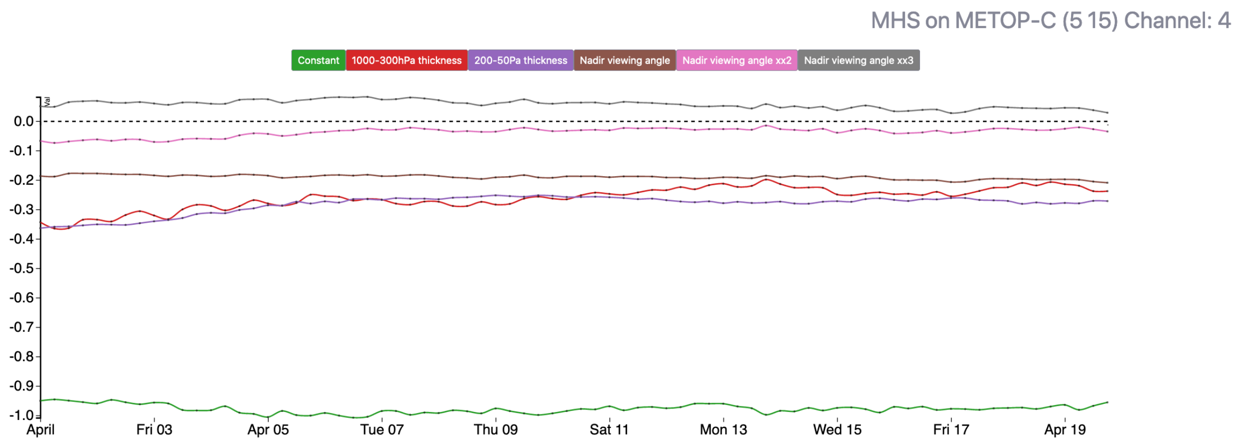

Figure 4: Variations of VarBC coefficients for the predictors selected for MHS Channel 4 from April 1st to April 20th, 2020.

2.2 SSMIS onboard DMSP F18

Figures 5 to 8 display two window channels at 37 GHz for the SSMIS microwave conical imager which are assimilated over oceans. For those two cases, the air mass thickness predictors are not used as for the sounding channels of MHS. Instead, several predictors characterizing the ocean surface are used: the skin temperature and the surface wind speed which are both used as input for computing the surface emissivity over oceans through the FASTEM software included within RTTOV. The cosine of the solar zenith angle is also used to correct for a known bias of the SSMIS instrument related to ascending and descending orbits.

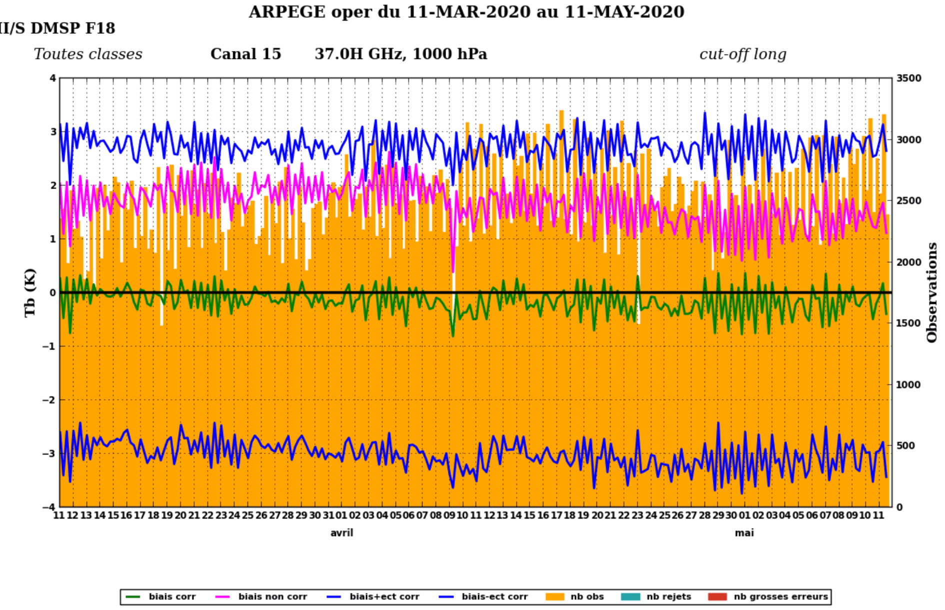

Figure 5: Monitoring of SSMIS Channel 15 compared to ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

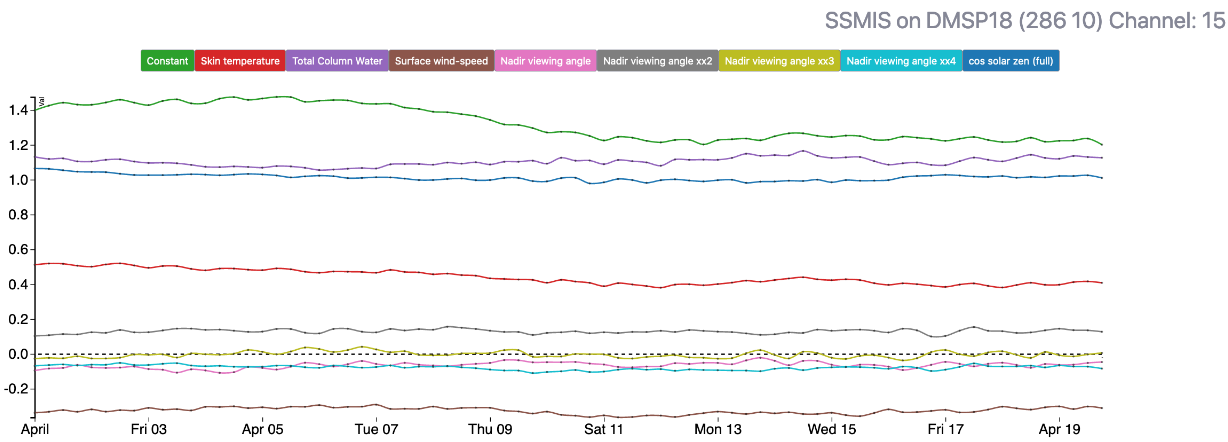

Figure 6: Variations of VarBC coefficients for the predictors selected for SSMIS Channel 15 from April 1st to April 20th, 2020.

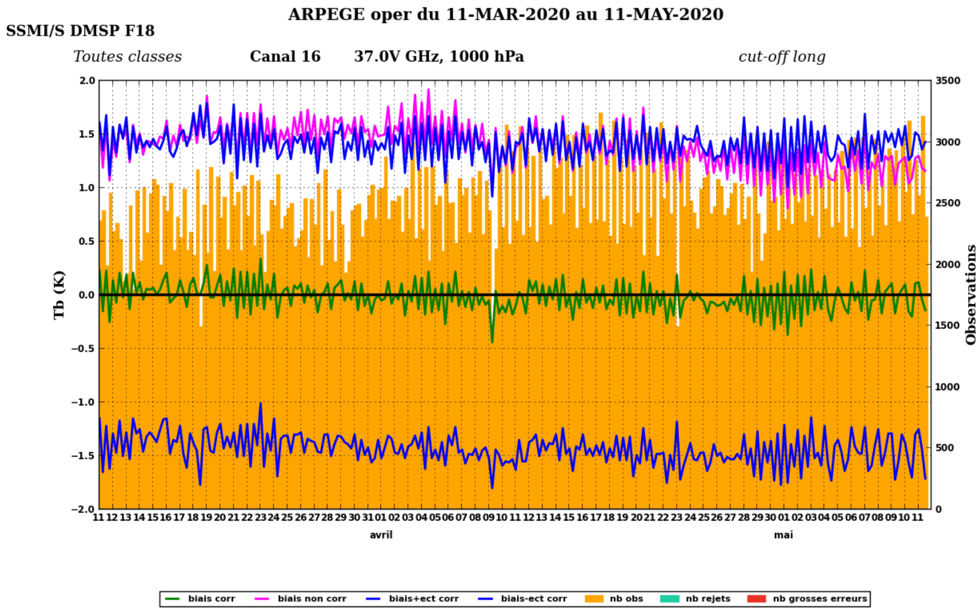

Figure 7: Monitoring of SSMIS Channel 16 compared to ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

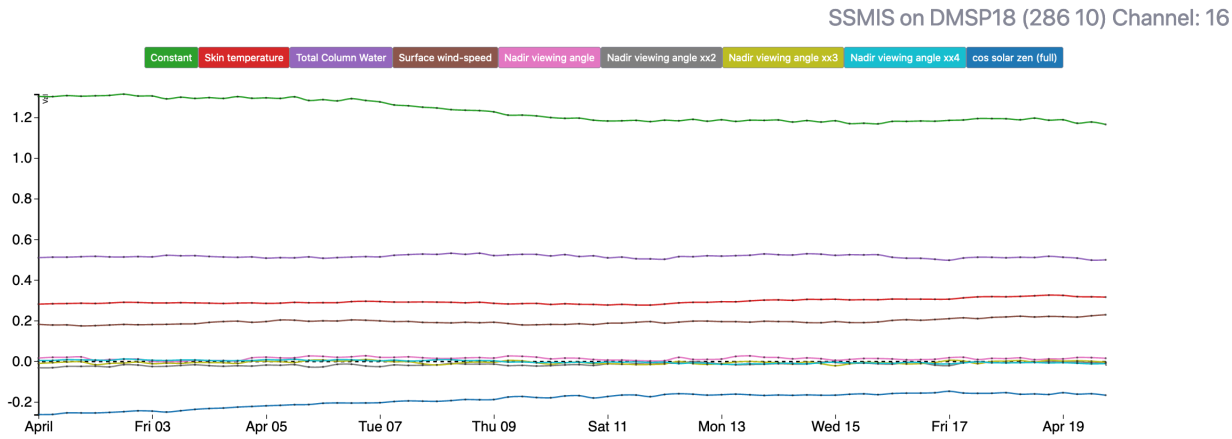

Figure 8: Variations of VarBC coefficients for the predictors selected for SSMIS Channel 16 from April 1st to April 20th, 2020.

2.3 AMSU-A onboard MetOp-C

The example of several AMSU-A channels onboard MetOp-C is interesting because it shows the adaptation of VarBC to a drifting bias, for instance on channels 11 and 12 first guess departures. Over the two-month period, the bias of first guess departures vary by 0.1 K for channel 11 and by 0.2 K for channel 12. One can see in Figures 10 and 12 that this drifting bias is mainly absorbed by the increase of the constant (dark green line). It should be noted in Figures 9 and 11 that the bias of corrected first guess departures is not totally removed by VarBC.

Indeed, VarBC should correct for biases from the observations and the observation operator, but should ideally not correct for a model bias. Anchoring observations within a data assimilation system helps VarBC to distinguish between the biases it should correct from the model ones. Two kinds of data play this role within the Météo-France system: the radiosounding network as well we the GNSS radio occultation constellation. For high peaking temperature channels like channels 11 and 12, it is very likely that the GNSS data prevent VarBC from correcting the full bias of these first guess departures and consider the remaining bias as coming from ARPEGE.

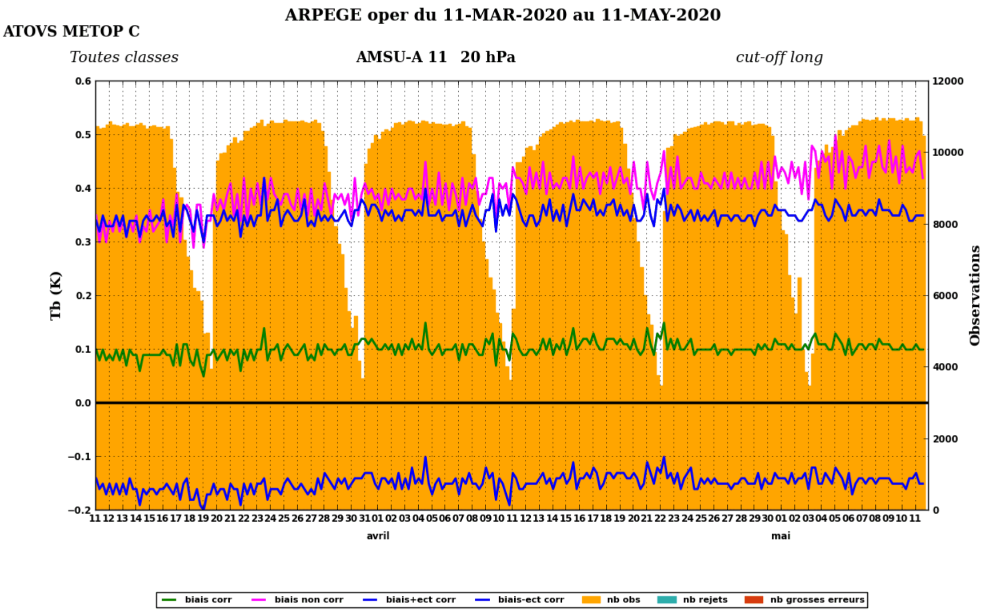

Figure 9: Monitoring of AMSUA Channel 11 compared to ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

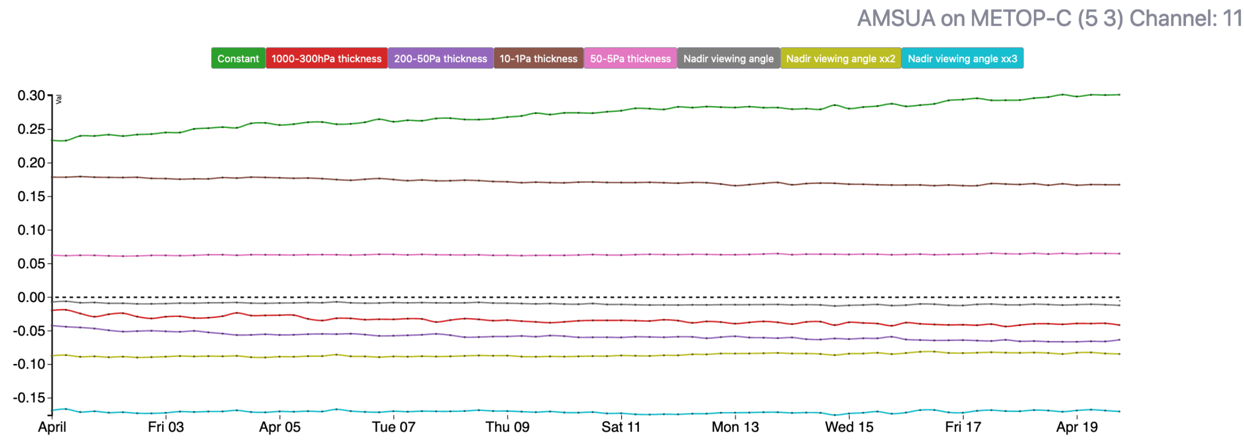

Figure 10: Variations of VarBC coefficients for the predictors selected for AMSUA Channel 11 from April 1st to April 20th, 2020.

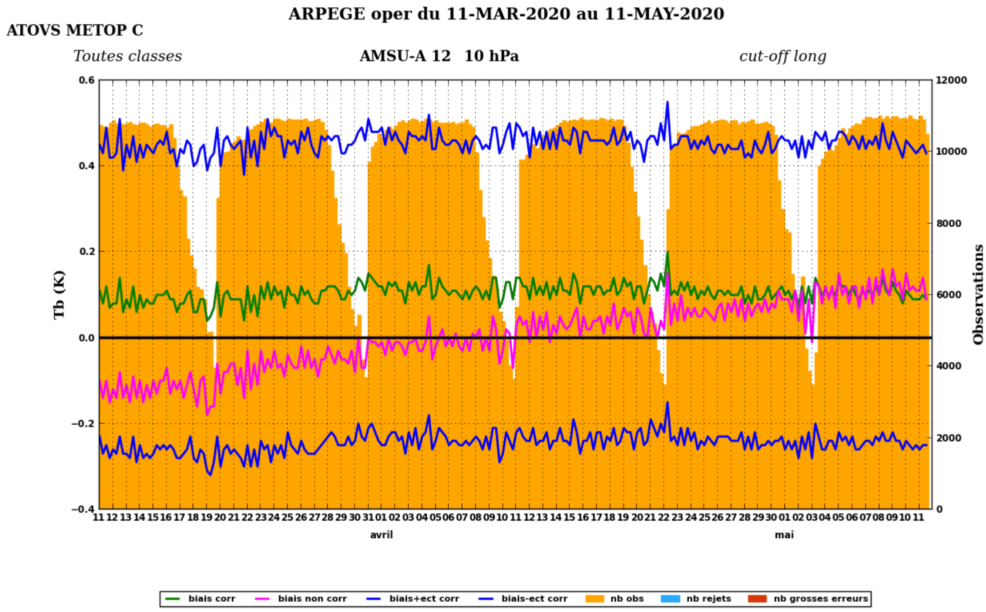

Figure 11: Monitoring of AMSUA Channel 11 compared to ARPEGE 4D-Var system, from March 11th to May 11th, 2020.

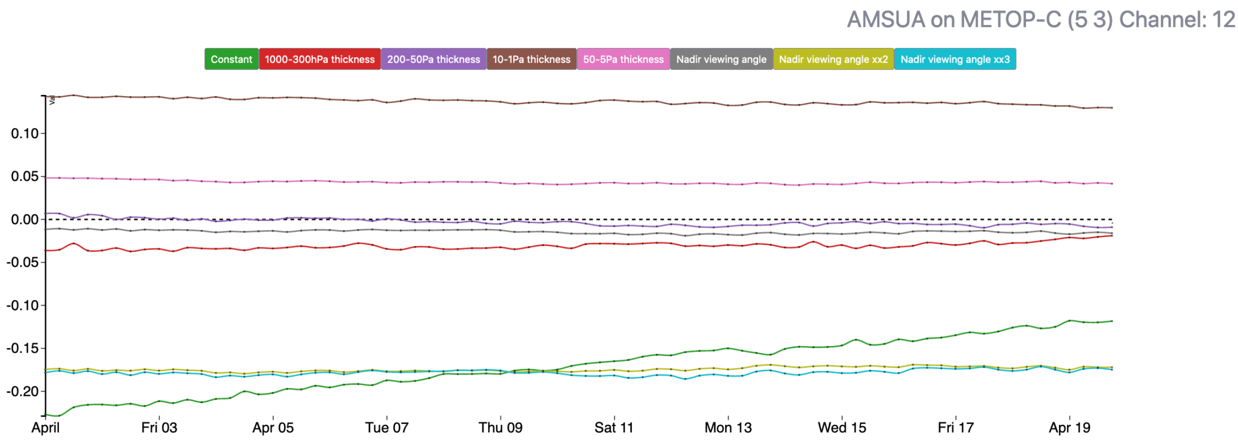

Figure 12: Variations of VarBC coefficients for the predictors selected for AMSUA Channel 11 from April 1st to April 20th, 2020.

References

Aulign√©, T., McNally, A. P., and Dee, D. P. (2007). Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, 133(624):631–642.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4d-var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387.

Dee, D. (2004). Variational bias correction of radiance data in the ECMWF system. In *Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP*, volume 28, pages 97–112.

Harris, B. A. and Kelly, G. (2001). A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Roy. Meteor. Soc.*, 127(574):1453–1468.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C. (1997). Unified Notation for Data Assimilation: Operational, Sequential and Variational. *J. Met. Soc. Japan*, 75(1B):181–189.