Assimilation of ATOVS and GNSS ZTD data in the HARMONIE-AROME model configuration run at AEMET

Joan Campins, Jana Sánchez-Arriola, María Díez, Javier Calvo and Beatriz Navascués

1 Introduction

The present article describes the implementation and evaluation of the assimilation of ATOVS and Ground Based Global Navigation Satellite System (GNSS) zenith total delay (ZTD) data in the HARMONIE-AROME model configuration, which is part of the shared ALADIN-HIRLAM system. This system is run operationally on the AEMET bullx computer over the two domains shown on Figure 1 and these integrations serve as Regular Cycle of Reference (RCR) to monitor the quality of the HARMONIE reference system.

The impact of using ATOVS and GNSS ZTD observations has been assessed separately through two different parallel suites on the domain centered on the Iberian Peninsula. The current operational run, that only assimilates conventional observations, is used as control experiment. Over Canary Islands, only the impact of ATOVS data has been evaluated. All the experiments have been carried out over the same period, from July to November 2016.

The impact studies performed are especially relevant for the atmospheric moisture initial state. In the current operational suite, the only direct humidity measurements assimilated are vertical profiles from radiosondes (TEMP). ATOVS humidity soundings and GNSS ZTD data provide additional sources of information and it is expected that they can contribute to better simulation of the mesoscale atmospheric distribution of moisture.

This paper shortly describes the main features from the collection of this data at AEMET, to the observations selection and its usage within the data assimilation system. Both types of observations show biases with respect to model background. As observation bias can systematically damage the data assimilation process, and finally the quality of the forecasting system, the bias correction is an essential step before the assimilation. To correct the observation biases for both types of observations, the adaptive variational technique VarBC (Dee 2005, Auligné et al. 2007) has been used. An objective verification of all the experiments has allowed to assess the impact on forecasts, with emphasis on the shorter lead times to better understand the influence of the assimilation of these data types.

The control and experimental suites are presented in Section 2. The observation handling and the variational bias correction procedure, for both ATOVS and GNSS ZTD, are described in Sections 3 and 4, respectively. In Section 5, the forecast impact is assessed. Finally, some concluding remarks and future work are presented in Section 6.

2 Control and Experimental Suites

Control suite

The control experiment used for this study is the cycle 40h1.1 of the HARMONIE-AROME configuration that is the operational AEMET NWP suite run on the local computer. The model is run at 2.5 horizontal resolution and 65 vertical model levels extending up to 10 hPa, over two domains: one centered on the Iberian Peninsula that includes the Balearic Islands (called AIB), and other centered on the Canary Islands (called AIC). Figure 1 shows the geographical areas covered by both suites. Boundary conditions from ECMWF are applied at 1-hr intervals.

The main differences compared to the AROME-France cycle 40t1 set-up are described in Bengtsson et al. (2017), and references there in. These include radiation settings and updates in cloud and condensation schemes known as OCND2 concerning specially the separation between cloud water and cloud ice treatment. Also it includes a major update of the turbulence scheme known as HARATU which produces a significant reduction of the 10m wind speed bias, a reduction of the cloud cover and an increase in clouds base height compared to the CBR scheme. The turbulence scheme works in combination with dual mass flux scheme EDMFm for the shallow convection.

The upper-air assimilation is run with a 3-hr cycle using 70 minutes cut-off time for the observations. For the upper levels a 3DVar scheme is applied including for the base configuration the following observations: TEMP, AIREP, AMDAR, SYNOP, BUOY and SHIP. For the surface analysis, CANARI OI is used followed by SURFEX Offline DA (SODA) using the optimal interpolation option with conventional observations, to update SURFEX variables. The large scale from the host model is included in the analysis through a scale selection method (LSMIX).

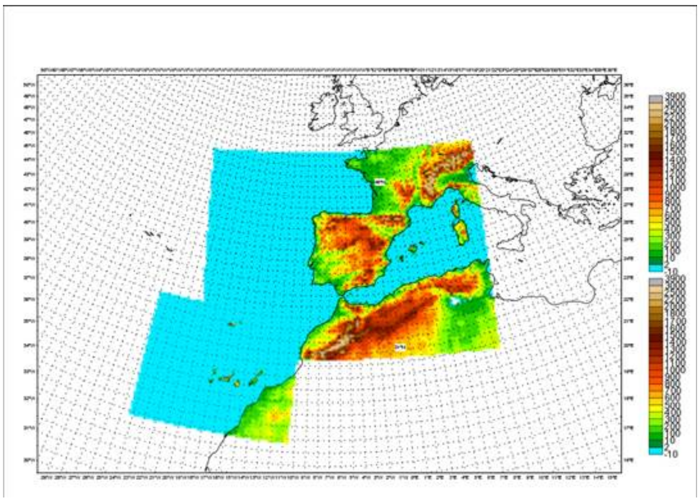

Figure 1*: Domain of Iberian Peninsula and Canary Islands Suites.*

ATOVS experiment description

The HARMONIE configuration is able to assimilate, among many other observation types, microwave radiances from the so-called Advanced TIROS Operational Vertical Sounders (ATOVS) on board of several polar-orbiting satellites. ATOVS comprises AMSU-A, AMSU-B (Advanced Microwave Sounding Unit A and B) and MHS (Microwave Humidity Sounder; which has replaced the old AMSU-B) instruments. In AEMET, ATOVS radiances are available via EUMETCast (local) for NOAA-18, NOAA-19, METOP-A and METOP-B. All these satellites are equipped with AMSU-A and MHS instruments.

In order to explore the impact of the ATOVS assimilation in the current operational suite, two parallel experiments, one for the Iberian Peninsula and the other for the Canary Islands domains, have been prepared. These experiments, named AIB_ATOVS and AIC_ATOVS respectively, are identical to their control counterparts, but they also assimilate ATOVS radiances from different satellites. In Section 3, a detailed description of the data selection, and particularly to the variational bias correction, is presented.

GNSS ZTD experiment description

This experiment is based in previous efforts on the HARMONIE-AROME configuration to assimilate ground based GNSS ZTD observations (Sánchez Arriola et al., 2016). A suite parallel to AIB run has been prepared for the operational assimilation of GNSS ZTD observations in addition to the conventional data. This run has been called AIB_GNSS**.** GNSS ZTD observations are fetched from the EUMETNET E-GVAP European program server via ftp in ASCII format.

3 ATOVS observation handling

The AMSU-A instrument measures the radiance that reaches the top of the atmosphere in 15 channels, that are sensitive mainly to atmospheric temperature in different vertical layers. As the uppermost model level is 10 hPa, the assimilation of those channels with weighting functions having a significant contribution from vertical layer above that value is not recommended. For this reason, the HARMONIE reference system exclude AMSU-A channels 11-15, and it is questioned if AMSU-A channel 10 should be used. AMSU-A channels 1-5 are not used as their weighting functions are close to the surface, so they are affected by surface emissivity and model biases at surface.

MHS instrument measures the radiance at the top of the atmosphere in 5 channels, 3 of them (channels 3 to 5) are atmospheric sounding channels, sensitive to atmospheric humidity in different vertical layers. Weighting functions for channels 3-5 are peaking at middle and high levels (above the model top), being channel 5 the lowest peaking and 3 the highest peaking of the atmospheric sounding channels.

As default, we utilize channels 6-10 from the AMSU-A and channels 3-5 from the MHS instruments. However, due to the poor quality of some channels (this information is confirmed by different satellite monitoring websites as http://www.star.nesdis.noaa.gov/icvs) we exclude channels 7-8 AMSU-A from NOAA-19 and METOP-A, and MHS NOAA-19 channel 3. The selection of satellites-instruments-channels is done within *mf_blacklist.b* routine.

Satellite data are subjected to a horizontal thinning for two main reasons: to reduce the data volume and to avoid the effects of observation error spatial correlation. In HARMONIE two thinning distances are introduced: RMIND_RADC1 and RFIND_RAD1C. The first one is the minimum horizontal distance allowed between two observations, and the second one is the resulting average horizontal thinning distance between two observations after thinning. Due to the different characteristics of the AMSU-A and MHS instruments, RMIND_RAD1C are different, lower for MHS (40 km) than for AMSU-A (60 km). On the contrary, RFIND_RAD1C is the same (80 km) for both instruments.

ATOVS Variational Bias Correction

Background departures for ATOVS observations present biases that can be due to systematic errors in the satellite instrument itself, deficiencies in the radiative transfer model, or bias in the first guess. In HARMONIE, bias correction for ATOVS radiances is carried out using the Variational Bias Correction scheme (VarBC), which is a particular adaptive scheme that is embedded inside the assimilation system and then the bias correction coefficients are continuously updated as part of the assimilation. The evaluation and tuning of VarBC for HARMONIE-AROME is extensively explained in Lindskog et al. (2012), and next we only will underline the outstanding issues.

Bias is estimated by means of a linear combination of predictors (p_i). In the reference system, a set of 5 predictors (i= 0, 1, 8, 9 and 10) are used for AMSU-A channels 6-10 and for MHS channels 3-5, where p_0 = 1 to allow a constant component for the bias, p_1 depends on the atmospheric state at the observed location and p_8-p_{10} depend on viewing angle relative to nadir. The convergence of the predictors to a certain timescale is set by means of the NBG_AMSUA and NBG_MHS parameters. In the HARMONIE AROME 2.5 reference system NBG value is set to 2000 for AMSU-A and MHS.

Before the assimilation of the satellite data, the spin-up of the VarBC coefficients for each assimilation cycle HH is needed. This is achieved by running the data assimilation with ATOVS data in passive mode. As it was mentioned before, with a NBG = 2000, a month is enough to get steady VarBC coefficients for the Iberian domain (this result is similar to that obtained in Lindskog et al., 2012 for a different region).

The calibration of the VarBC coefficients for the Canary Islands domain during the spin-up period has taken about 10 days longer than for the Iberian Peninsula and Balearic Islands domain, as can been seen in the Figures 2 and 3. This could be related with the southernmost latitude and the smaller area of the AIC domain and with the lower observations density in this area.

During this spin-up period, the passively assimilated observations were monitored, that is, the time series of number of observations, bias correction and the background and analysis departures were examined. This monitoring allows to prevent the utilization of satellite data with a small sample of data (for instance those from paths at the edge of the domain), which may create unstable bias correction coefficients. In this case, the blacklisting of those data can be performed by means of the LISTE_LOC_HH files in the observations pre-processing step.

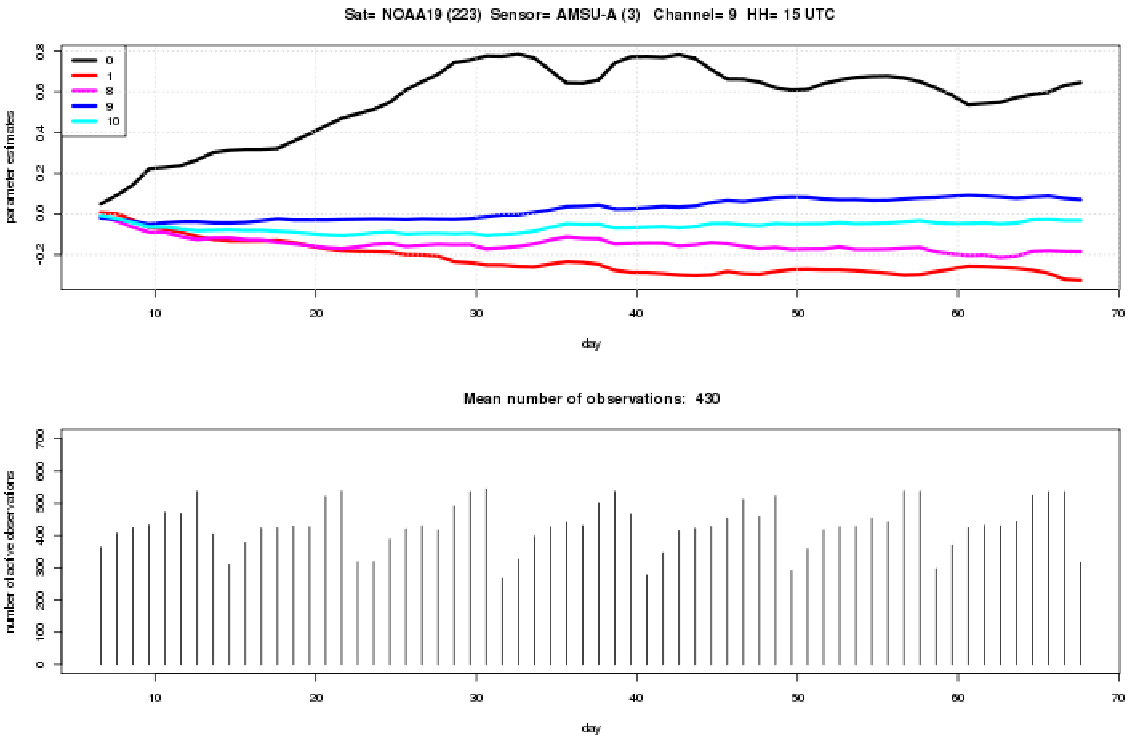

***Figure 2**: Evolution of the VarBC coefficients for the AIB_ATOVS suite (Iberian Peninsula and Balearic Islands) during the spin-up period*

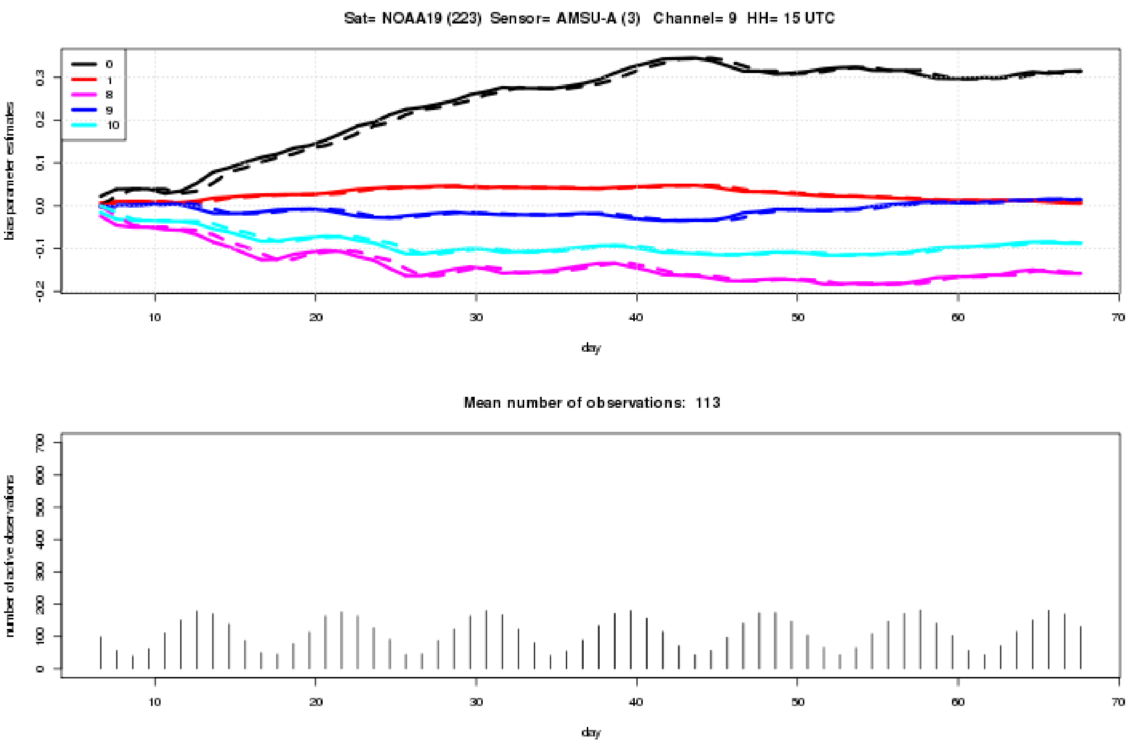

***Figure 3**: Evolution of the VarBC coefficients for the AIC_ATOVS suite (Canary Islands) during the spin-up period showing that it needs a longer period to stabilize the VarBC coefficients.*

Also, the monitoring can reveal large departures that can worsen the analysis, a channel with poor quality or not properly. In this case it is better to keep this channel as passive.

After this spin-up period, those satellites-instruments-channels, that the bias correction seems to properly reduce the first-guess systematic departures, can be assimilated as active. For AIB_ATOVS and AIC_ATOVS, the active assimilation channels are all the aforementioned, except AMSU-A channel 10, that we have decided to keep as passive. Figure 4 displays the first-guess (corrected and non-corrected) and analysis departures (top panel) and number of assimilated observations (bottom panel) for channel 6 on NOAA-18 AMSU-A, from 1st to 30th September 2016 at 18 UTC for AIB_ATOVS experiment (observations over land are displayed in green and over sea in blue). It shows that bias correction is able to significantly reduce departures (non-corrected departures in dashed lines with triangles and corrected in solid lines with circles), close to analysis departures (solid line with squares). Data availability oscillates daily. Overall the number of observations over land is lower than over sea, among other reasons because observations over high orography are rejected.

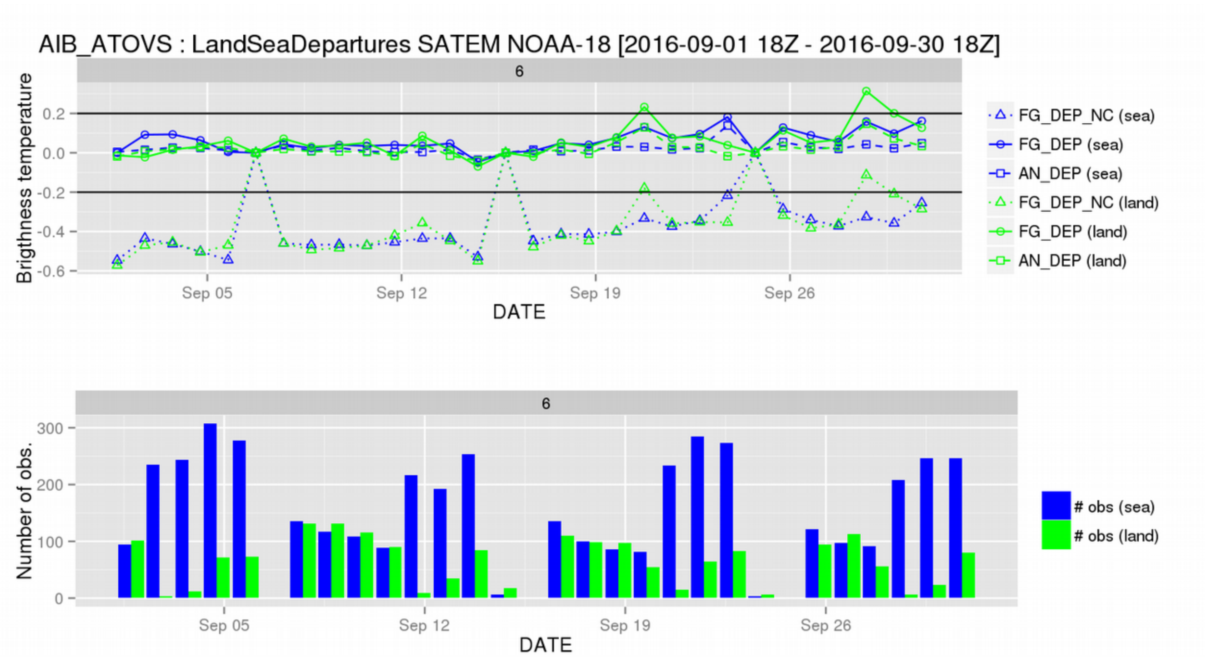

Figure 4: *AIB_ATOVS departures statistics (top) and number of observations (bottom) for NOAA-18 AMSU-A Channel 6 from 1st to 30th September 2016 at 18 UTC for data over sea (in blue) and over land (in green).*

4 GNSS ZTD observation handling

The GNSS ZTD observation processing includes data selection, quality control, and also a Variational Bias Correction to handle the systematic discrepancies between model equivalents and observed values. The available files within the cut-off time of 70 min, have been fetched by ftp, filtered, and prepared to be assimilated by the model at every assimilation cycle.

There are many ground based GNSS sites which raw data are processed by more than one GNSS ZTD analysis center. So, the first step consists on creating a “White List” containing the best quality of station-analysis center pairs. The White List is based on the statistics of the

ZTD departures, the observation-background counterpart in observation space. The criteria taken to select the best pair station-analysis center into the White List was the smallest standard deviation, provided that the skewness did not exceed a predefined threshold. Time series of background departures for this White List construction were obtained through a one month long run before the starting date of the experiment, where the GNSS ZTD data entered in passive mode to the data-assimilation system.This White List used contains 783 sites all around Iberian Peninsula and the sites chosen are from ASI_, ROBH, SGN_, IGE2 and METO GNSS ZTD analysis centers; its distribution can be seen at Figure 5.

In the pre-processing step, the available data from a selection of the best quality station-analysis center pairs, chosen through the White List, is further reduced by applying a temporal thinning in order to retain from each station the ZTD observation closest to the analysis time (ZTD time frequency is about 15 min). No spatial thinning has been applied to these data.

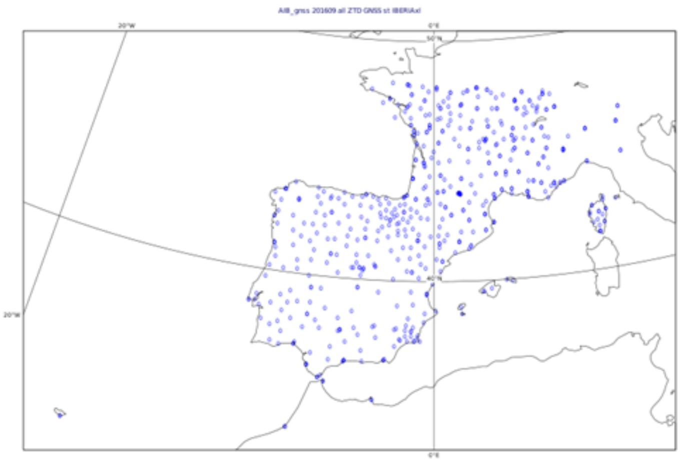

Figure 5*: Distribution of the stations contained in the White List used for this study.*

GNSS ZTD observations Variational Bias Correction

The sources of bias of background departures for ZTD observations come from GNSS data processing algorithms (that may even differ from one Analysis Centre to another), systematic errors in both the ZTD observations operator and the model field, the relatively low model top (10 hPa) used by the HARMONIE AROME 2.5 model configuration that leads to an underestimation of the ZTD model equivalent value, and also the interpolation and extrapolation from the model orography to the real one. The bias derived from all of these sources is estimated by applying a VarBC method to these observations. It is based on a single constant predictor. So, the systematic differences between the observations and the model equivalent are parameterized as site dependent offset parameter, one value per site that is updated every 3 hours. These biases have been first tuned in a 5-6 weeks spin-up period previous to the assimilation of GNSS ZTD observations, during which these data entered to the analysis in passive mode only and so, did not influenced the model state. Variational bias-correction coefficients are in general very sensitive to the bias present in closely located sites during the spin-up period, and that is the reason why the spin-up period used here has been long enough, a bit more than a month, to be able to remove the bias of all stations, taking into account that the adaptivity corresponds to a stiffness coefficient in the variational bias correction scheme (nbg_sfcobs_ndays110) equal to 15.

The functionality of this bias correction is represented by the time evolution (here a month is shown) of the GNSS ZTD observation (obs raw), the bias corrected observation (ob), the first guess (fg) and the analysis (an), shown in Figure 6, where it is possible to see how the variational data assimilation scheme has managed to correct the observation from the systematic bias along the previous period of spin-up, for this particular site of MALLIGE2 that is shown as an example.

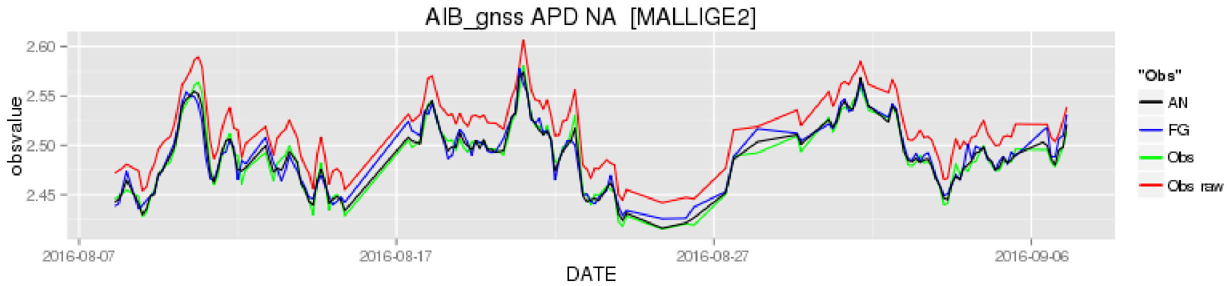

***Figure 6**: One moth time series of bias correction for MALLIGE2 site. Red line is raw observation with no bias correction, green line is bias corrected observation, blue line is first guess and black line is the analysis.*

5 Impact on forecast

The impact of the assimilation of ATOVS and GNSS ZTD data has been assessed through the forecast objective verification of all the experiments against SYNOP and TEMP observations during the test period (July-November 2016). Only the shortest lead times are shown here to draw more clear conclusions about the influence of the extra observations assimilation on the forecast skill.

Overall, the impact found is rather small in most of the variables. It has to be taken into account that, among other reasons, the large scale mixing with the host model prevents a larger influence of these observations in the assimilation cycle. Nevertheless, some rather consistent features seem to indicate a benefit of the assimilation of both types of data.

Figure 7 displays the vertical distribution of bias and rmse of temperature and relative humidity for the different experiments in the Iberian Peninsula and Balearic Islands domain. Both types of observations moisten the model (GNSS ZTD the most), and a positive impact is observed at the vertical profiles of humidity rmse due to the assimilation of either ATOVS (from 700 hPa upwards), or GNSS ZTD (up to 700 hPa) data. This last data type also produces a slight reduction of the cold bias in the lowest atmospheric layers.

The smaller cold bias due to the assimilation of GNSS ZTD data is also visible in T2m scores, accompanied by a reduction of the positive/negative biases of mslp and rh2m, respectively (see Figure 8). Precipitation skill scores (see KSS in Figure 8) are consistent with those obtained in humidity and show the benefit of the assimilation of both data types.

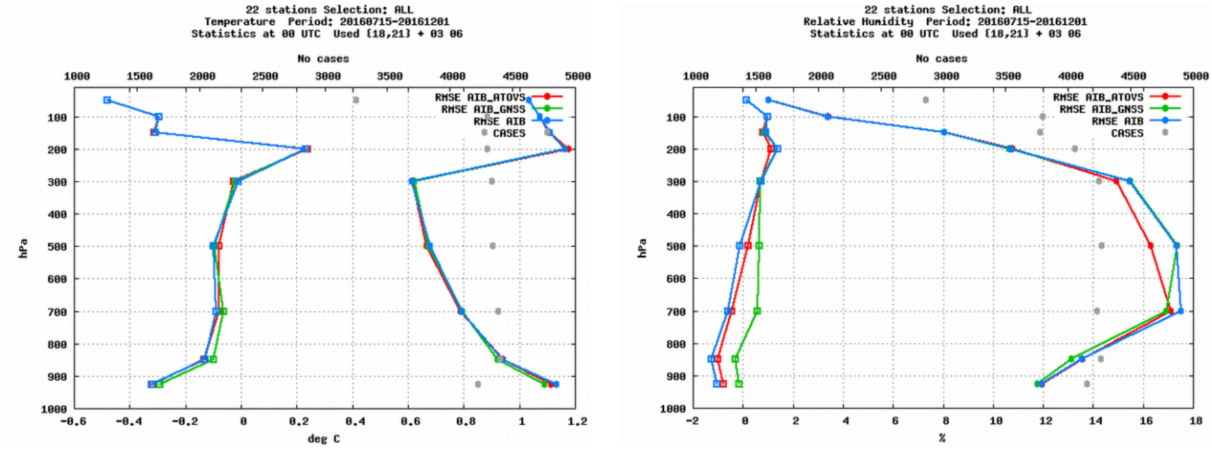

***Figure 7**: RMSE and BIAS against soundings at 00 UTC for the AIB domain comparing ATOVS (red), GNSS (green) and control (blue). RMSE and BIAS for temperature (left) and relative humidity (right)*

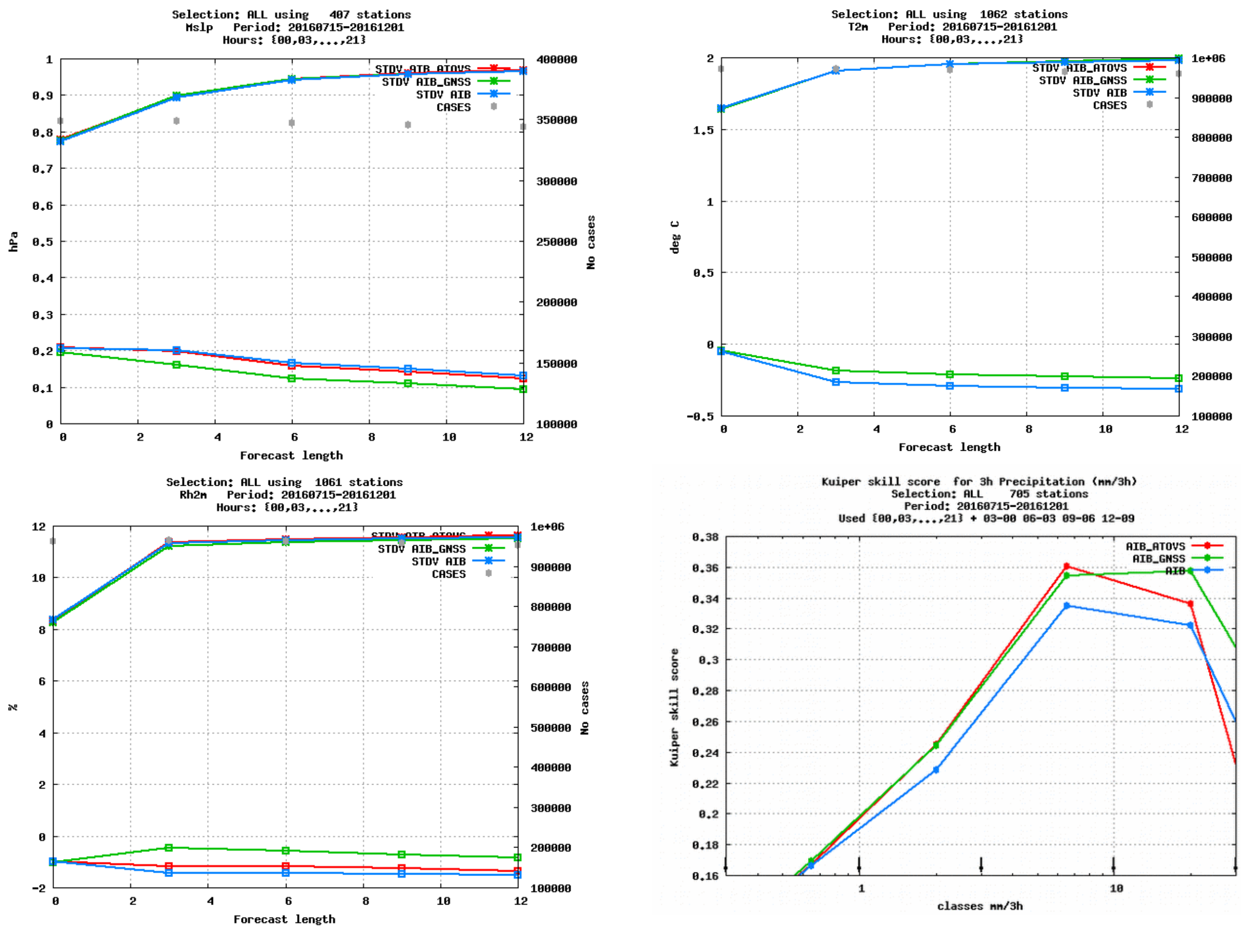

***Figure 8**: STDV and BIAS against synop function of the forecast length for the AIB domain comparing ATOVS (red), GNSS (green) and control (blue). MSLP (top left), 2m temperature (top right) and 2m relative humidity (bottom left). The KSS for different precipitation categories is shown on bottom right plot.*

The influence of ATOVS assimilation in the Canary Islands region is qualitatively very similar to that found in the Iberian Peninsula domain. Only the humidity vertical profile seems to be affected, being AIC_ATOVS slightly moister than AIC, and having a smaller rmse, as it is displayed in Figure 9. This Figure also shows that AIC model also presents in this area a cold bias in the lowest atmospheric levels, that ATOVS assimilation is not able to correct (similarly to the northernmost geographic domain). The impact found in other surface parameters (mslp, T2m, rh2m) is also neutral. Precipitation events have been rare in this area during the test period, and the number of observations and the complex orography of these islands makes difficult to draw conclusions of the objective verification performed for this variable.

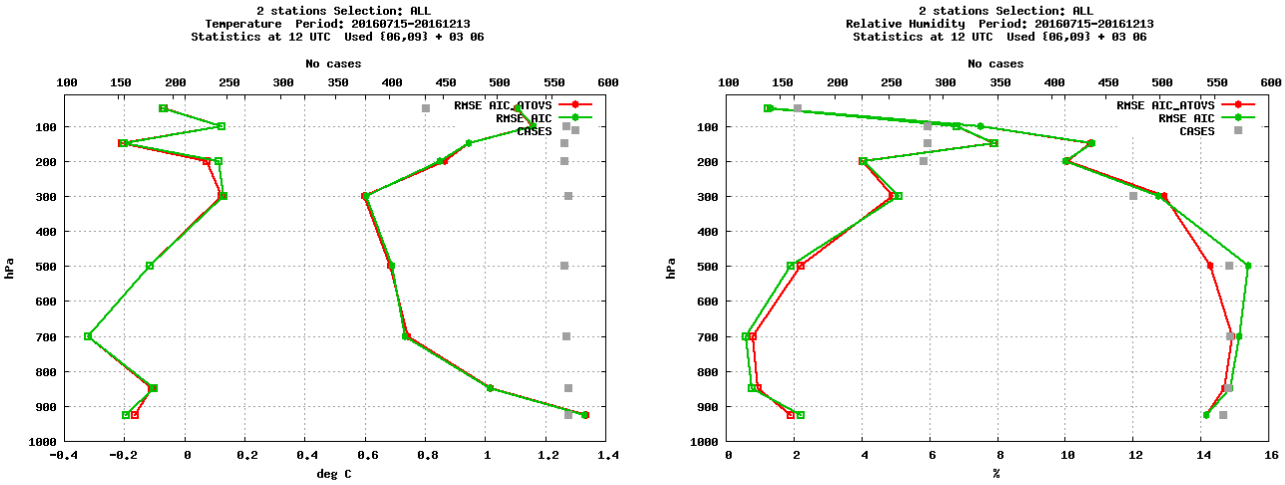

***Figure 9**: RMSE and BIAS against soundings at 12 UTC for the Canary Islands domain (AIC) domain comparing ATOVS (red) and control (green). RMSE and BIAS for temperature (left) and relative humidity (right). Note that only 2 soundings are available for this domain.*

6 Conclusions and future work

A series of parallel experiments have been carried out at AEMET to prepare the assimilation of ATOVS and ground based GNSS data in the operational NWP suite. This system comprises two independent HARMONIE runs covering the Iberian Peninsula and Balearic Islands, and Canary Islands. Apart from the effort needed for a local reception of the different data types, considerable work has been devoted to the setup of the different assimilation experiments. It included, among others, the creation of a White List for GNSS ZTD data, further pre-processing of these data, and three HARMONIE-AROME suites running spin-up periods during which data was introduced in passive mode. These spin-up periods have allowed to calibrate the corresponding VarBC coefficients, to monitor the different observations, and to select the satellites-instruments-channels to be assimilated at each assimilation cycle. Once this training step has been completed, the three experiments assimilating ATOVS in the two domains, and GNSS data in the Iberian Peninsula region, have been run and monitored in parallel to the operational ones over the period July to November 2016.

The impact on forecast produced by the assimilation of these extra observations has been assessed by means of an objective verification against surface and vertical profile observations. The results obtained indicate an overall neutral impact in most of variables. However, and in both geographical regions, the assimilation of ATOVS clearly improves the atmospheric humidity in middle-upper levels. The assimilation of GNSS data produces a positive impact on the humidity, up to 700hPa, and it is also able to decrease the bias observed in temperature at or close to the surface, and in rh2m and mslp. Both data types improve the forecasted precipitation in the Iberian Peninsula and Balearic Islands domain.

The results obtained show a statistically significant sensitivity of the humidity linked variables in the HARMONIE-AROME model, to the initial state of the atmospheric moisture.

Work is on-going for the joint assimilation of both data types over the Iberian Peninsula and Balearic Islands domain. In the Canary Islands region, the next step is to start the assimilation of ground based GNSS observations.

7 Acknowledgements

Gema Morales is responsible for the installation and maintenance in AEMET of the obsmon tool which results have been widely used in this work. Daniel Martín support for the local installation of HARMONIE-AROME model is very much appreciated.

8 References

Auligné T., McNally A.P., Dee D.P., 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.,* **133**: 631-642.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Sass, B. H.,Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu,L., Samuelsson, P., Santos Muñoz, D., Subias, A., Tijm, S., Toll, V., Yang, X., Køltzow, M. Ø., 2017:The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system, Monthly Weather Weather Review, accepted for publication, January 2017

Dee D. 2005: Bias and Data assimilation. *Q. J. R. Meteorol. Soc*. **131**: 3323-3343.

Lindskog M., Dahlbom M., Thorsteinsson S., Dahlgren P., Randriamampianina R., Bojarova J., 2012: ATOVS processing and usage in the HARMONIE reference system. *HIRLAM Newsletter*, **59**: 33-43

Sánchez Arriola J., Lindskog M., Thorsteinsson S., Bojarova J., 2016: “Variational bias correction of GNSS ZTD in the HARMONIE modeling system”. *Journal of Applied Meteorology and Climatolog*y 55(5), 1259-1276; DOI: 10.1175/JAMC-D-15-0137.1.