RUC Validation

Mohamed Zied SASSI
INM, Tunisia

ARSO, Slovenia

October 2014

OUTLINE

Introduction

The development of data assimilation systems had improved the numerical weather prediction models, increasing their accuracy. Thus, the regular validation of these systems allows to verify their contribution and avoids the deterioration of the numerical model forecasts.

It's in the frame of this principle that stay was dedicated to the validation of the Slovenian data assimilation system.

This report summarizes the obtained results during this stay. In fact, several verification tools were applied namely Variation Bias Correction (VarBC) Monitoring, Degree of Freedom of a Signal (DFS), Background and Observation error Tuning (TuneBR) and Moist Total Energy Norm (MTEN).

A brief description of the applied tools is included in this report, with reference to more detailed papers. The results of verification experiments, which were very beneficial, will be detailed through this report.

I. VarBC Monitoring

Satellite observations are an important component in data assimilation systems. It can provide a huge amount of information which covers a wide range of high levels and channels, however, it's not obvious that it's always improving the analysis. Thus, the user is invited to proceed to the necessary verifications in order to extract the most useful information and avoid any perturbing signals.

The current paragraph deals with the verification and monitoring of available satellite observations in Slovenia, namely NOAA-16, NOAA-18, NOAA-19, METEOSAT_10, METOP-A and METOP-B. During this verification, we could investigate the performance of AMSU-A, AMSU-B, MSG-HR and MHS sensors for the available channels (from channel 2 to channel 13).

1) Inventory of operational suite results:

A first experiment including the operational settings has been run along a period of seven months (From 20th February to 7th of October). The main results related to available signals are described in the following paragraphs:

a) Increase in oma variability in May 2014:

Since the 6th of May 2014, the oma values have known a higher amplitude variability as shown in the figures I.1 and I.2. This discontinuity has been noticed for AMSU-A (Satellites 209, 223, 004, 003) and AMSU-B (Satellites 209, 223, 003).
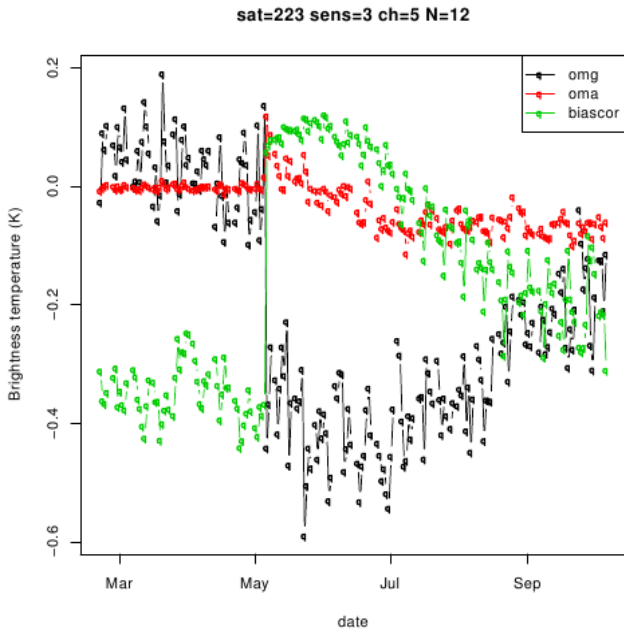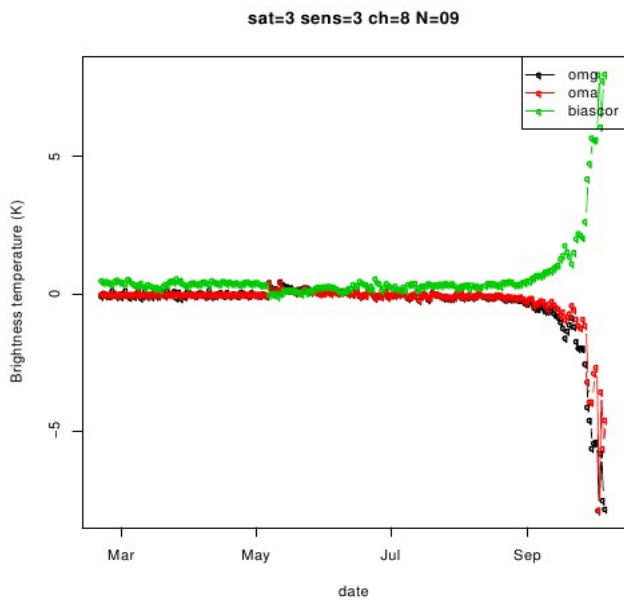
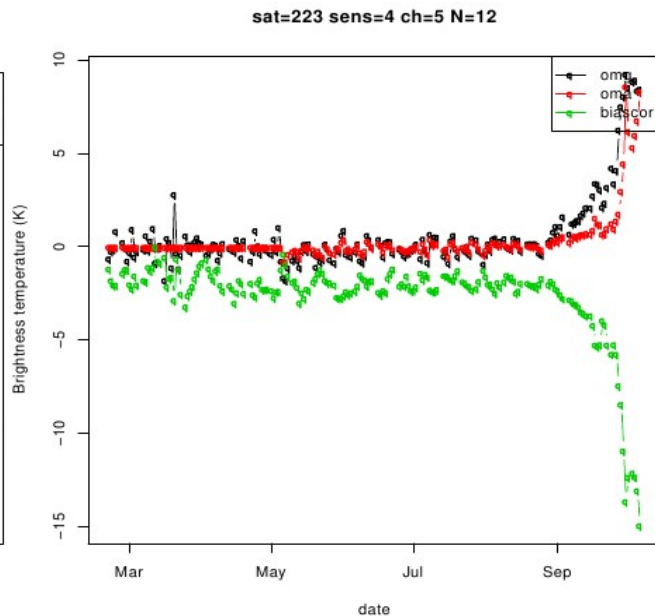**Figure I.1:** Variability of omg, oma and biascor for operational experiment
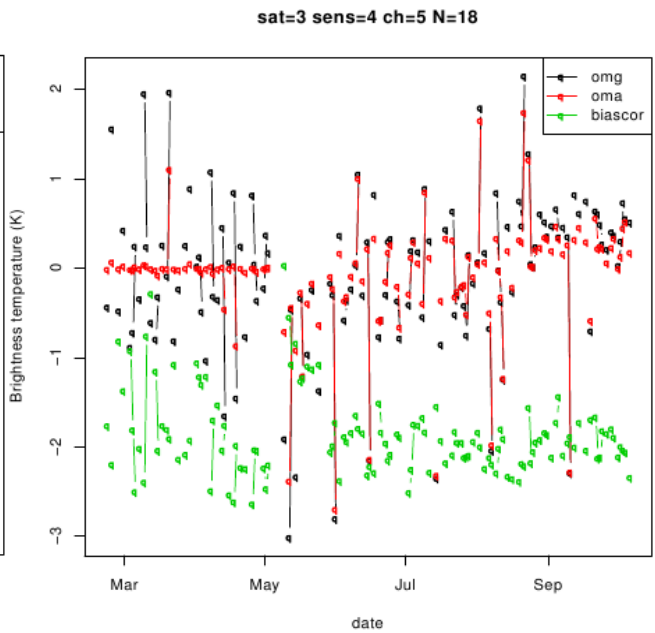
Figure I.2: Variability of omg, oma and biascor for operational experiment

b) Increase in oma amplitude in September 2014:

The amplitude of oma has increased notably since September 2014 as described by figures I.3 and I.4. This issue has been noticed with the sensors AMSU-A (Satellite 003) and AMSU-B (Satellites 223, 004, 003).
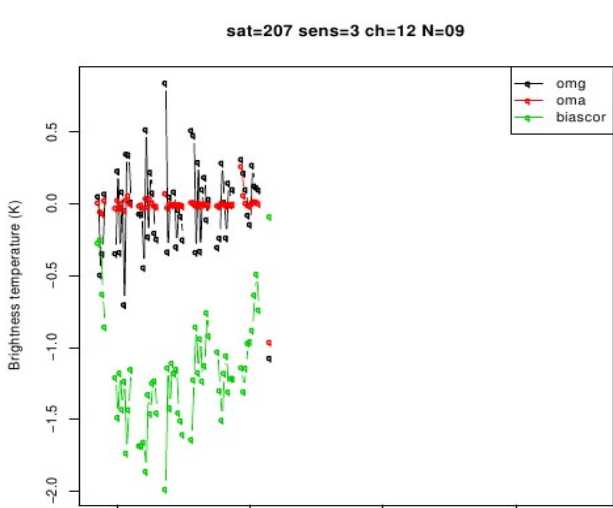

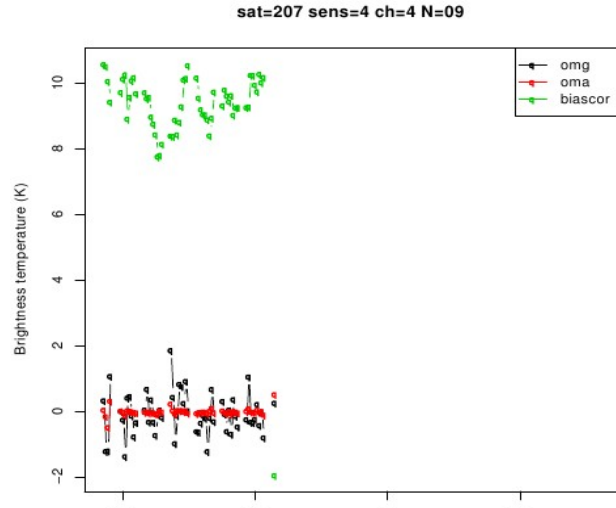**Figure I.3:** Variability of omg, oma and biascor for operational experiment

Figure I.4: Variability of omg, oma and biascor for operational experiment

c) NOAA-16 observation cessation:

According to figures I.5 and I.6, the NOAA-16 observations were no more considered in the data assimilation system since May 8th. This stop was due to some technical problems related to the satellite sensor.

Figure I.5: Variability of omg, oma and biascor for operational experiment

Figure I.6: Variability of omg, oma and biascor for operational experiment

d) New channels enabled:

According to figures I.7 and I.8, the channel 13 of satellites 209, 223, 004 and 003 has been enabled for sensor AMSU-A since September 2014.
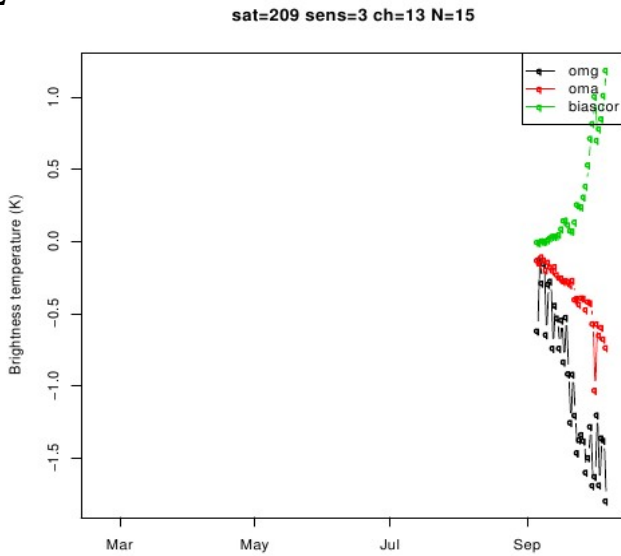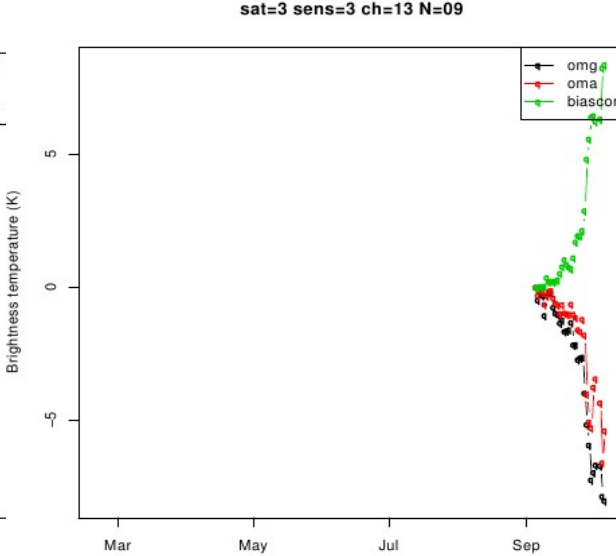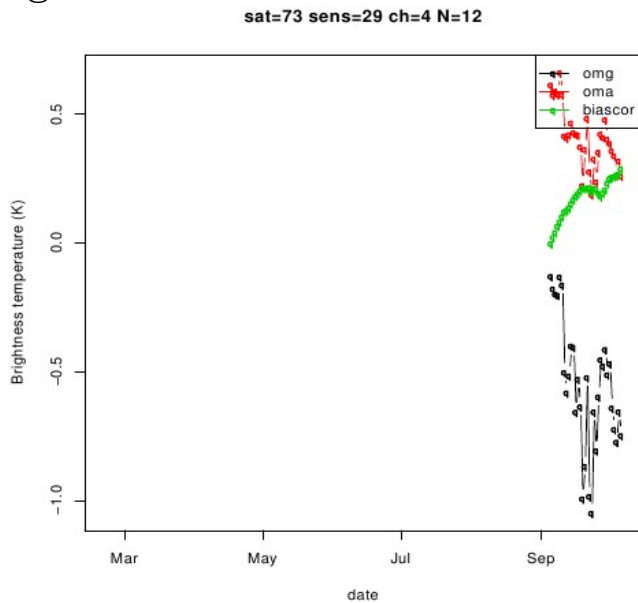

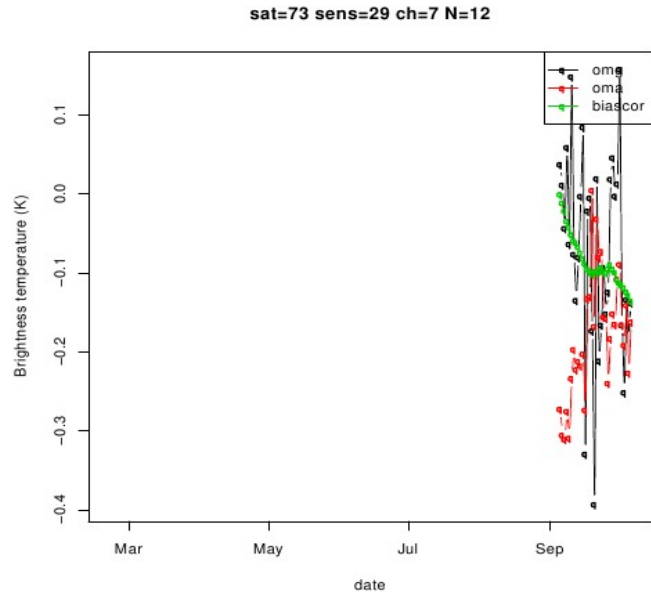**Figure I.7:** Variability of omg, oma and biascor for operational experiment

Figure I.8: Variability of omg, oma and biascor for operational experiment

Concerning the MSG HR sensor of satellite 073, the channels 4, 5 and 6 have been enabled also since September 2013 as described in figures I.9 and I.10:

Figure I.9: Variability of omg, oma and biascor for operational experiment

Figure I.10: Variability of omg, oma and biascor for operational experiment

2) **Verification and tuning:**

Through the following paragraph, we will describe the proceeded investigation of the notable increase of oma amplitude since September 2014.

As a first step, an update of the blacklisted satellite observations has been done, through it all the recently allowed channels had been disabled. A new experiment of five weeks (from September 1st to October 7th) has been run including the new compiled binary.

As a first result, the new experiment oma amplitudes decreased in comparison to the operational behavior, but a considerable oscillation remained, as shown in the following figures:
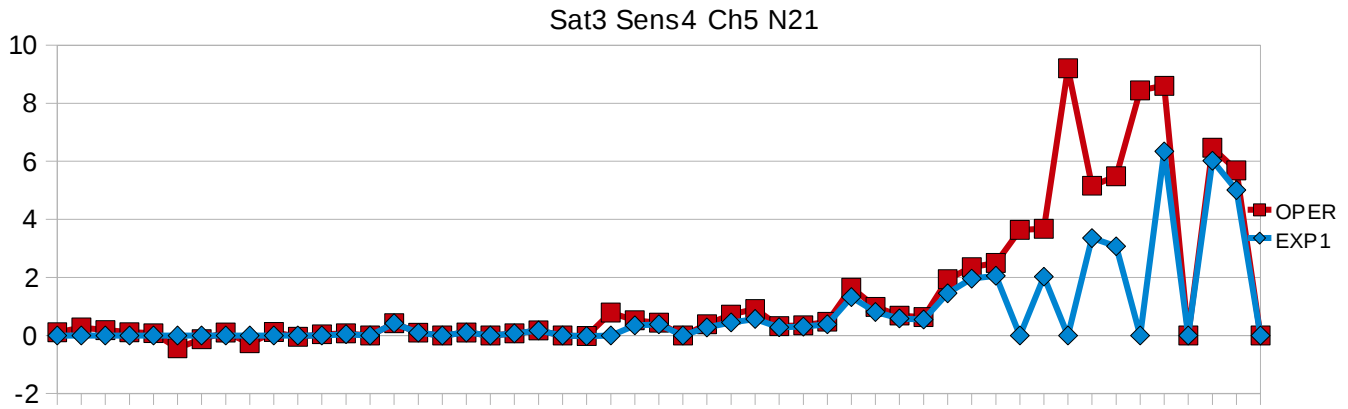
Figure I.11: Variability of oma for operational (Red) and new binary (Blue) experiments
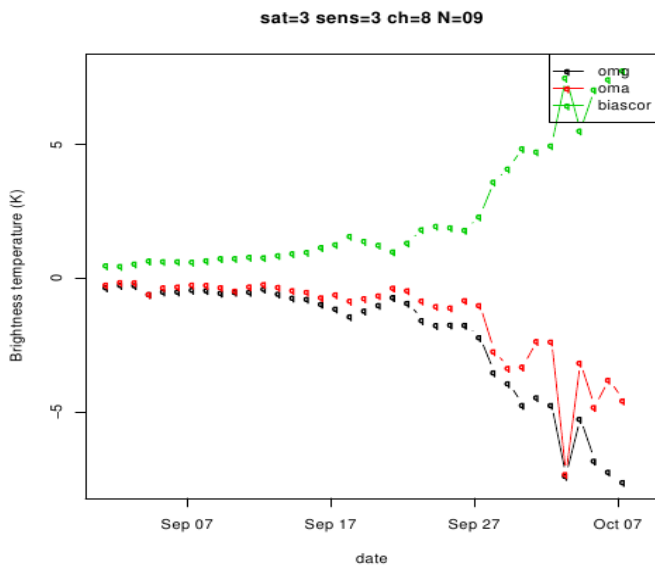
Figure I.12: Variability of omg, oma and biascor for new binary experiment
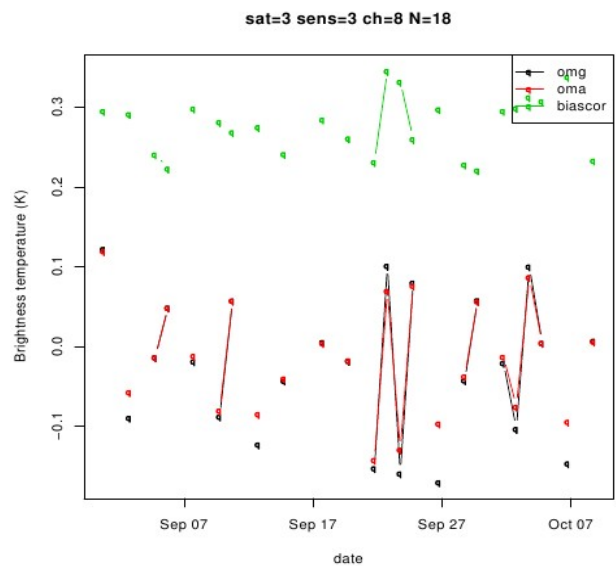
Figure I.13: Variability of omg, oma and biascor for new binary experiment

The next correction step was to tune the VarBC values. Thus, a new experiment has been run including the new VarBC values for the 09h and 21h networks.

In the following examples, we will focus on the behavior of the AMSU-A sensor of the METOP-B satellite. During this step, the VarBC values related to channel 5 to 12 were tuned.

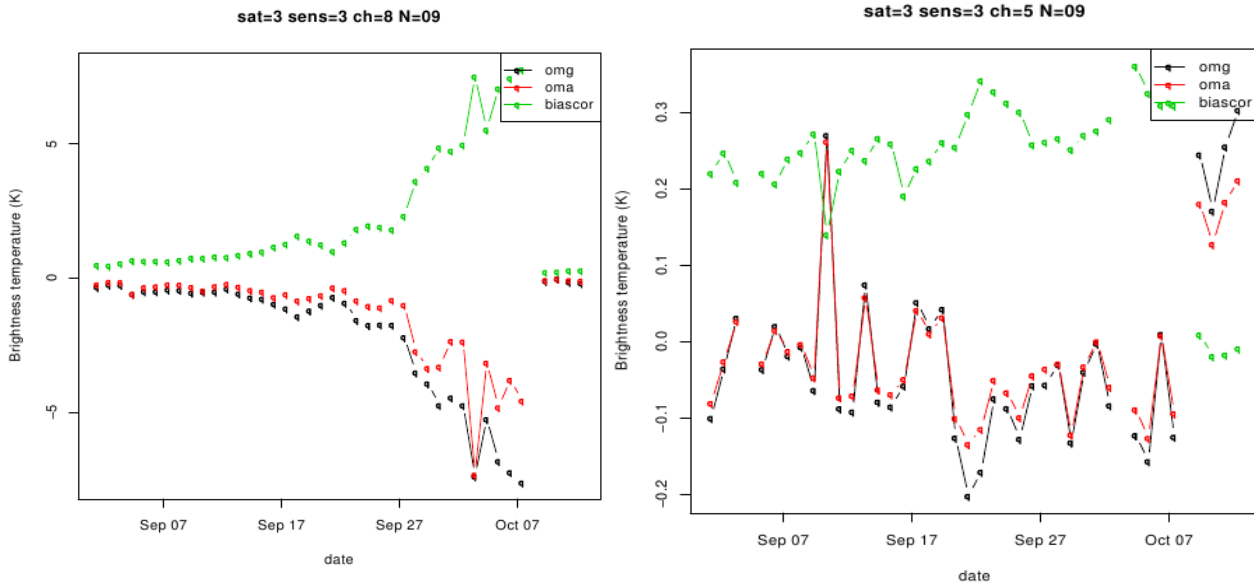The new values correspond to the 18h network were the oma values followed a small amplitude oscillation around 0, which was considered satisfying.

The following figures describe the new oma behavior after this tuning:

Figure I.14: Variability of omg, oma and biascor for new binary experiment

Figure I.15: Variability of omg, oma and biascor for new binary experiment

According to the figure I.14, the new oma behavior is satisfying, although, the figure I.15 shows higher oscillation amplitudes.

The investigation of the tuned channels showed an improvement in channels 7 to 12, but a high increase of amplitudes for the channels 5 and 6. Thus, the VarBCs related to channels 5 and 6 were tuned back to their operational values, however, the new VarBC values for channels 7 to 12 were maintained.

A new experiment including the updated VarBC values was run as described in the following figures:

Figure I.15: Variability of omg, oma and biascor for new binary experiment

Figure I.16: Variability of omg, oma and biascor for new binary experiment

One more run over longer period has been done to verify the behavior of oma after the last tunings:
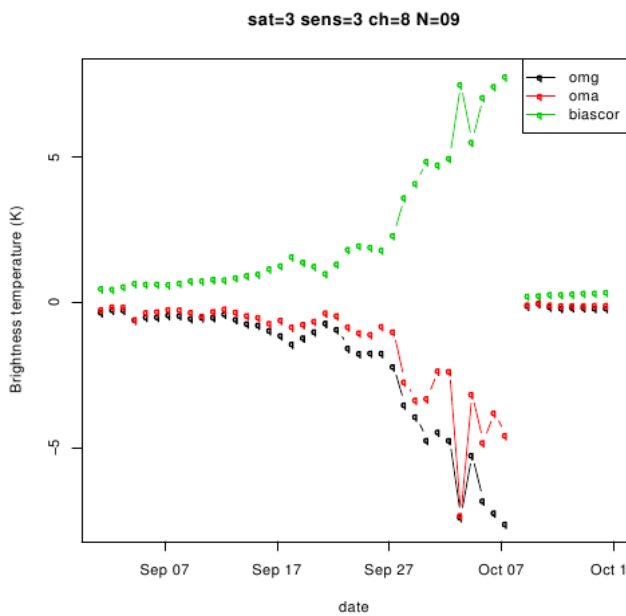

Figure I.17: Variability of omg, oma and biascor for new binary experiment

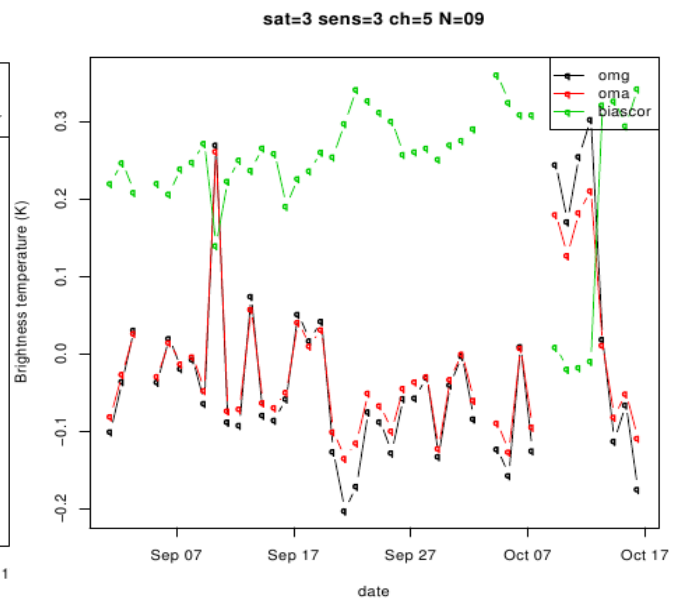**Figure I.18:** Variability of omg, oma and biascor for new binary experiment

After the successful tuning of the AMSU-A sensor VarBCs, we proceeded to the tuning of the AMSU-B sensor VarBCs by initializing them to zero. This tuning concerned the METOP-A, METOP-B and NOAA-19 satellites for the 00h, 03h and 12h networks. The new oma values present a low oscillation amplitude around zero which was considered satisfying.
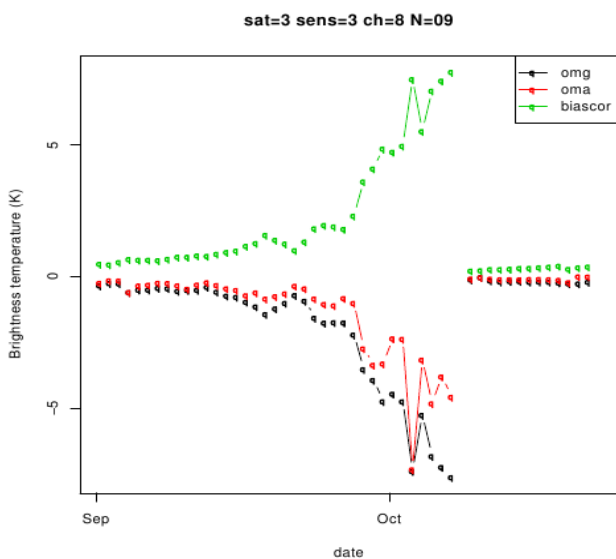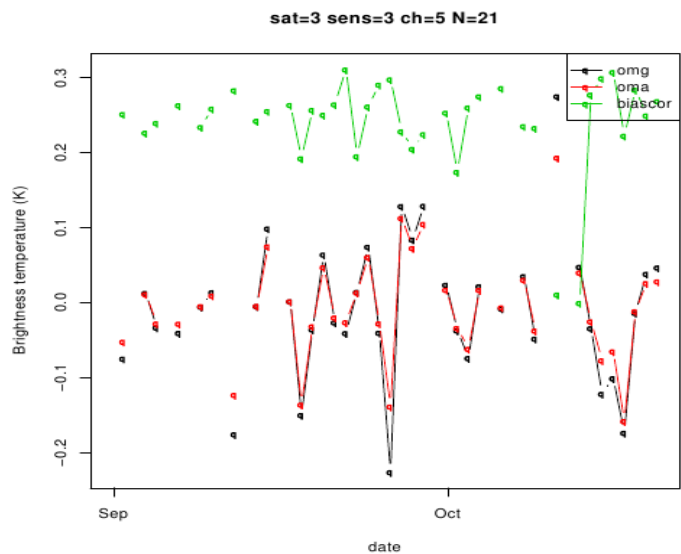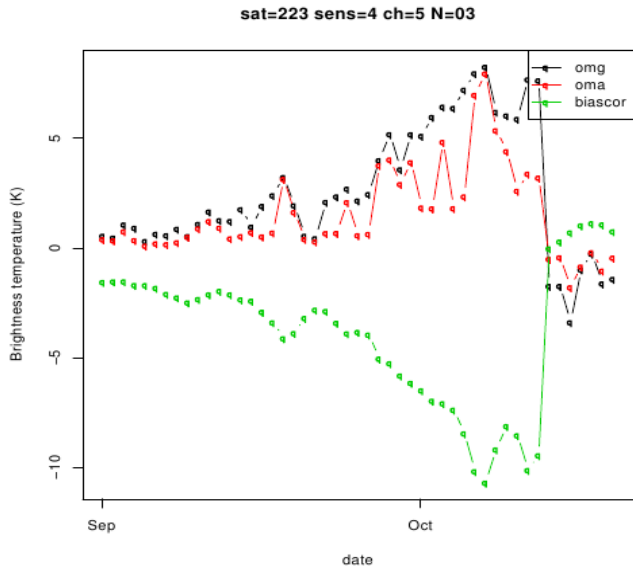

Figure I.17: Variability of omg, oma and biascor for new binary experiment
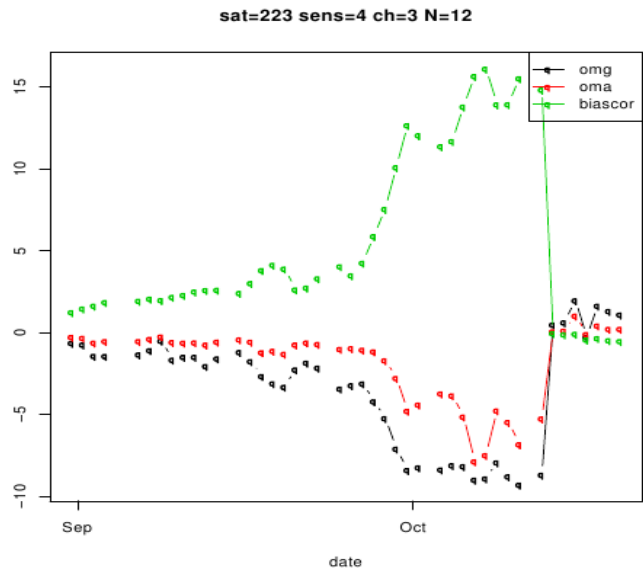
Figure I.17: Variability of omg, oma and biascor for new binary experiment

The new behavior allowed to test turning the AMSU-B observation into active mode for satellites 003, 004 and 223. This action were verified through a new run over 4 days as described in the following figure:
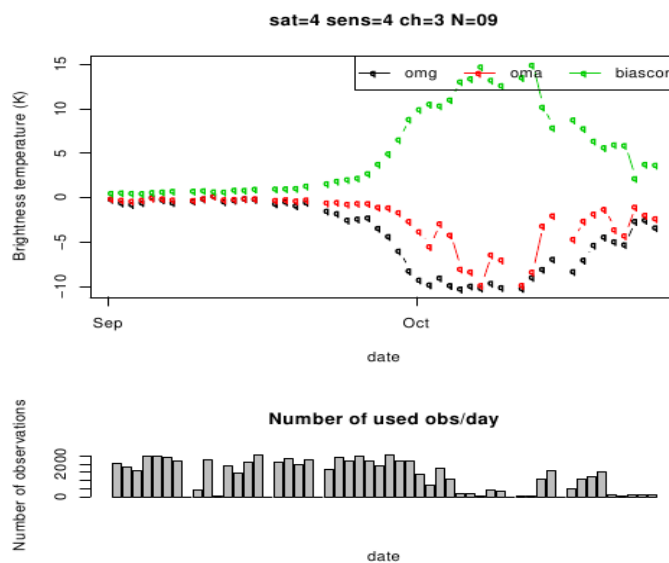
Figure I.11: Variability of omg, oma and biascor before and after the AMSU-B VarBC tuning And number of used observations
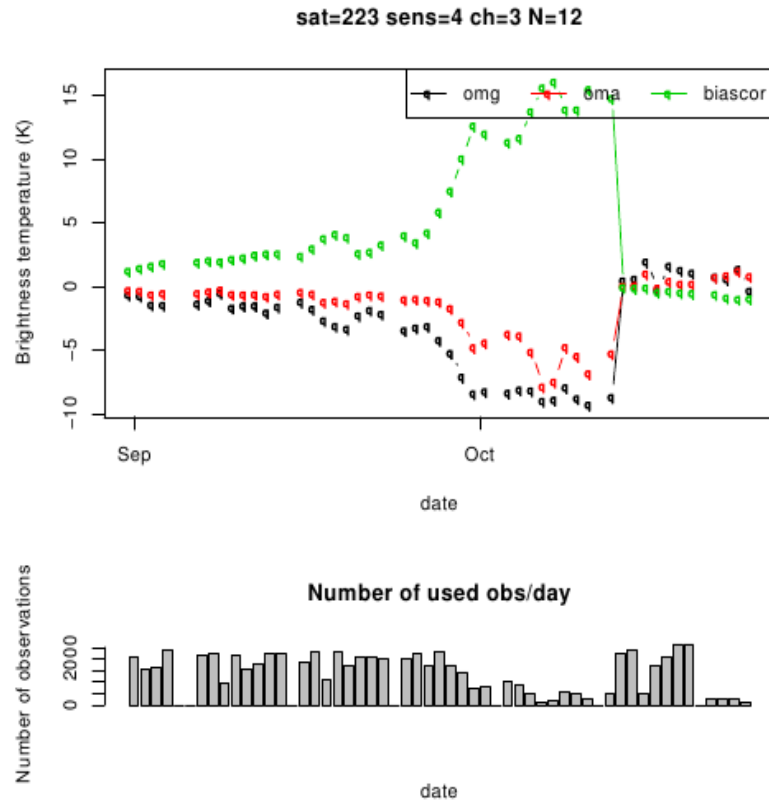
Figure I.11: Variability of omg, oma and biascor before and after the AMSU-B VarBC tuning
And number of used observations

II. Degree of Freedom of Signal (DFS)

The DFS (Degree of Freedom of Signal) is a statistical method which quantifies the weight attributed to an observation type by the assimilation system. It measures the impact of an observation on analysis.

The DFS can be quantified as the trace of the gradient of the analysis in observation space with respect to the observations :

$$DFS = \mathrm{Tr}\left\{ \frac{\partial (\mathbf{H}\mathbf{x}^a)}{\partial \mathbf{y}^o} \right\} = \sum_{y_i^o \in \text{observations}} \frac{\partial (\mathbf{H}_i \mathbf{x}^a)}{\partial y_i^o}.$$

In the linear case, the DFS can be quantified as:

$$DFS = Tr(\mathbf{KT}\ \mathbf{HT}\ ) = Tr(\mathbf{HK})$$

Where K is the Kalman gain matrix and H is the observation operator.

More theoretical aspects and properties of DFS can be found in Rodgers (1996), Fisher (2003) or Cardinali et al. (2004).

The method consists on running a perturbed and an unperturbed experiments. In order to apply the DFS method on the Slovenian data assimilation system, we considered an experiment of 10 days (from September 01st to 10th 2014). The new experiment was modified in a way that it takes into account the Mode-S (Strajnar 2012, JGR) observations.

Than, two domains were considered, first the LACE domain and second a zoomed domain over Slovenia. This methods allows the comparison of the observations weights according to the available types over both domains.

The absolute DFS of available observation types are described in the following figures:
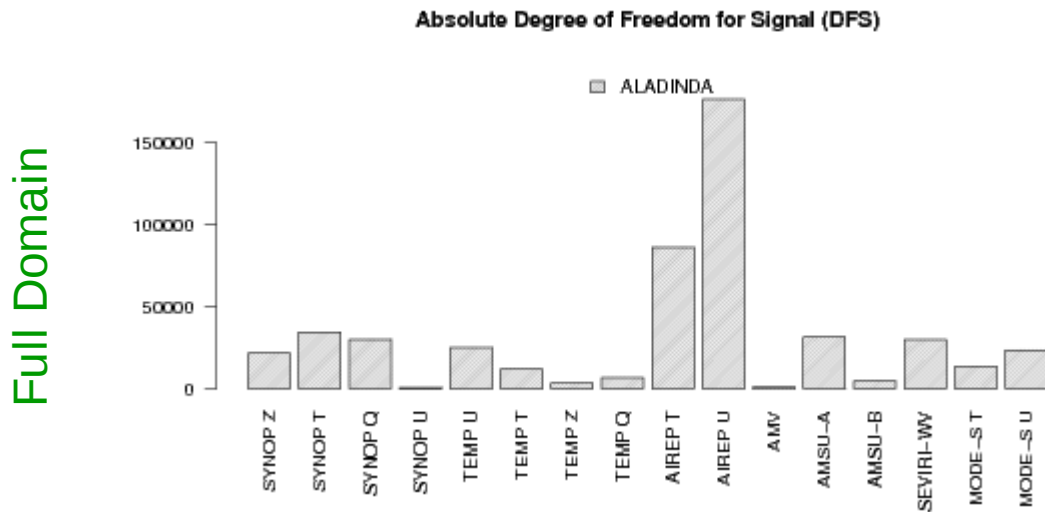
Figure II.1: Absolute DFS for available observation types/parameters on full domain
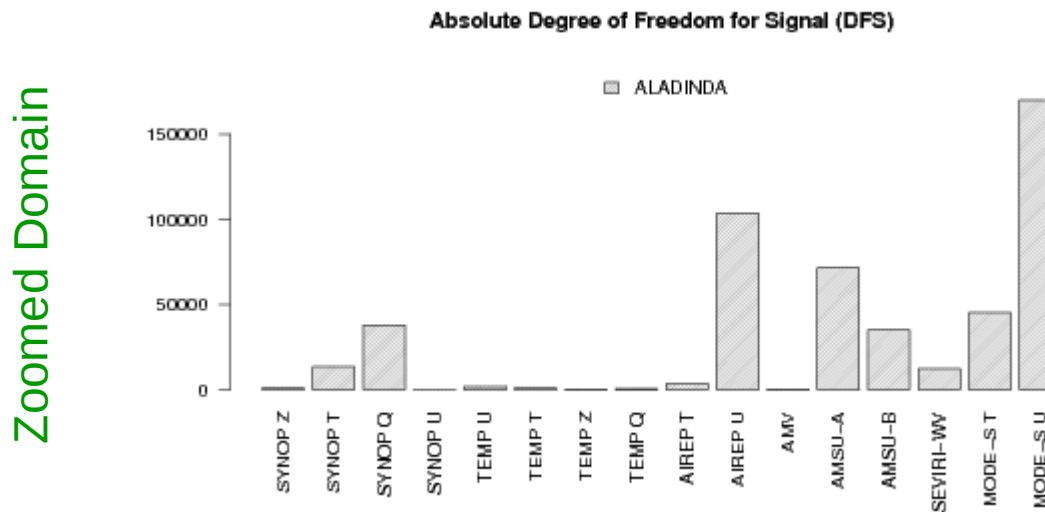
Figure II.2: Absolute DFS for available observation types/parameters on zoomed domain

The figures II.1 and II.2 show a dominant weight of aircraft data over both domains, especially for the wind component.

A second result may be concluded from the Mode-S type, which is available over Slovenia, thus it shows a significant impact on analysis in the zoomed domain, and less impact for the LACE domain, where it can

be dominated by the other observation types which are available over all the LACE domain.

This conclusion can be also shown by the following figures which describes the weight by observation type:
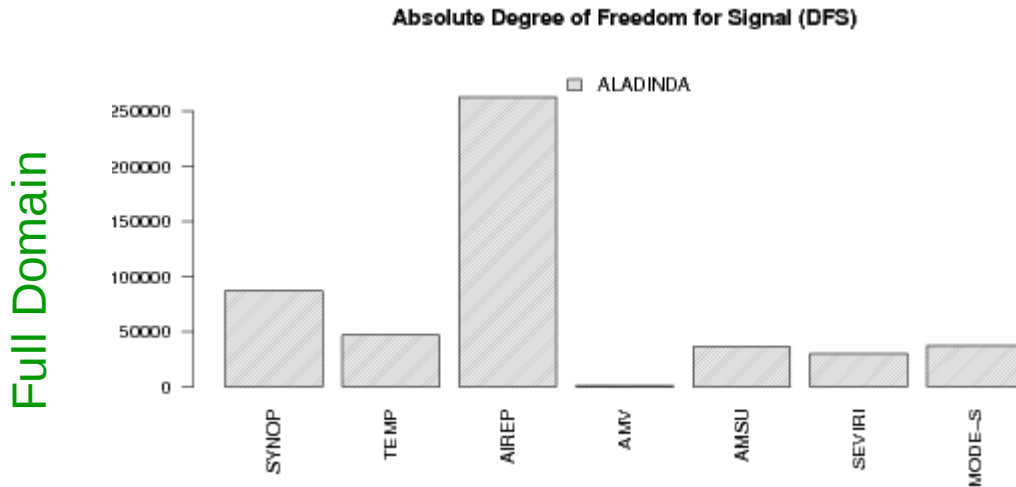
Figure II.3: Absolute DFS for available observation types on full domain
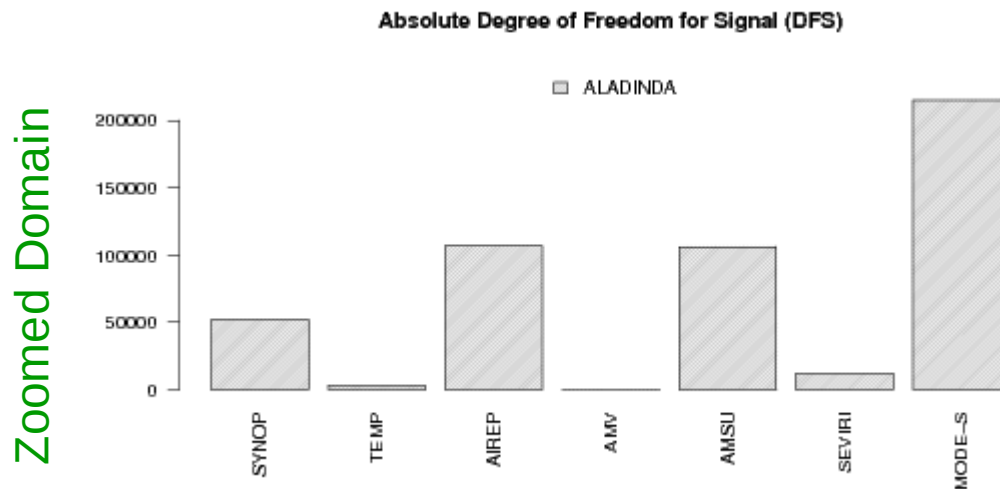
Figure II.4: Absolute DFS for available observation types on zoomed domain

The weight of different observation parameters can be described by the following figures where the wind parameter show a dominant impact over both domains:

Figure II.5: Absolute DFS for available observation parameters on full domain
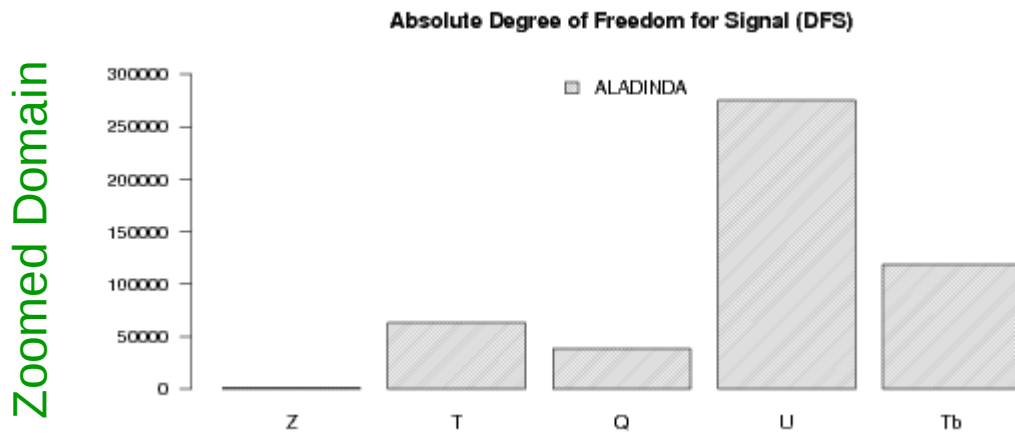
Figure II.6: Absolute DFS for available observation parameters on zoomed domain

The following figures describe the time evolution of DFS according to every observation type. It shows a dependency to the network time which can be notable namely with the aircraft data (including Mode-S):
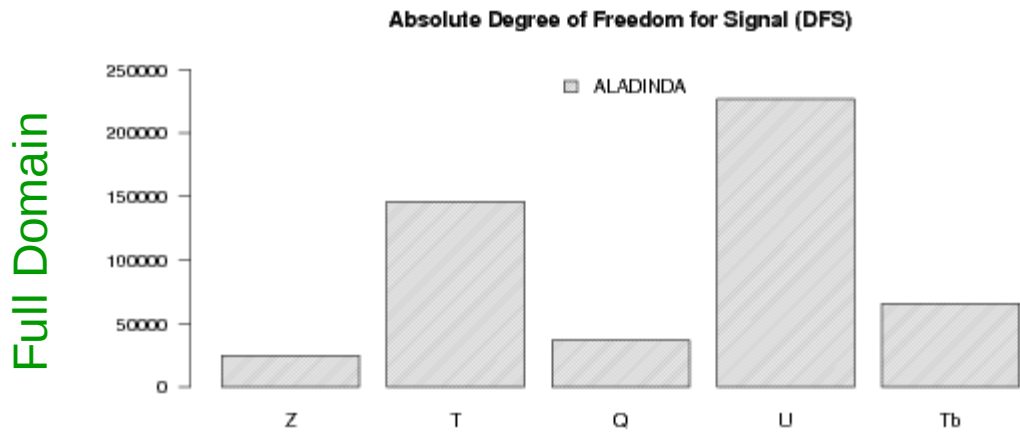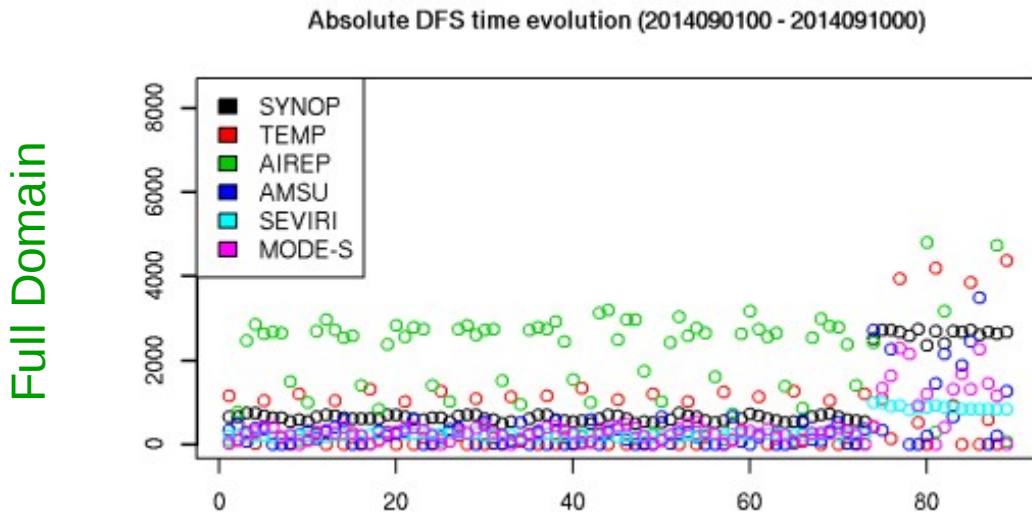

Figure II.7: Absolute DFS time evolution for available observation parameters on full domain
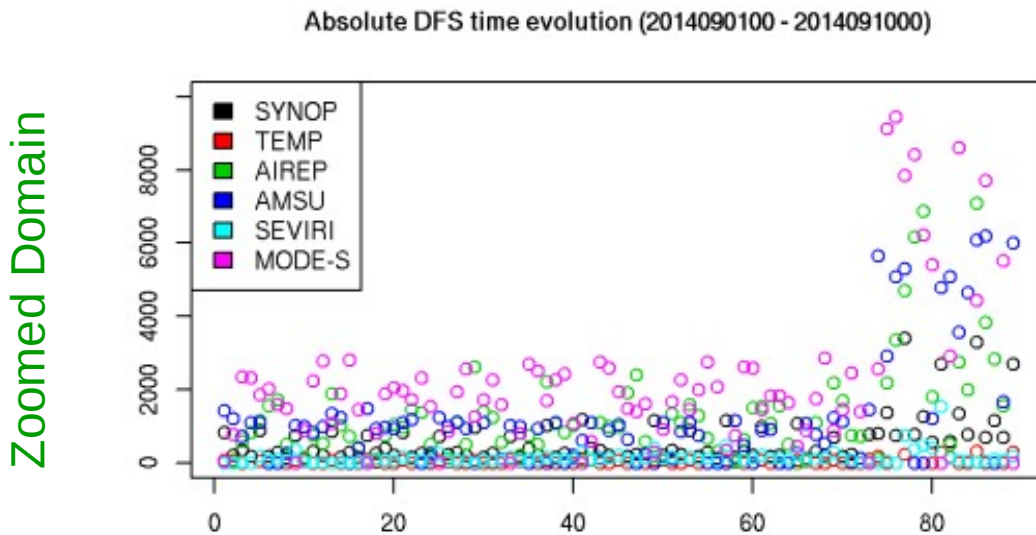
Figure II.8: Absolute DFS time evolution for available observation parameters on zoomed domain

In term of number of observations, the following figures show a high frequency of aircraft and TEMP observations. However, in the zoomed domain, the Mode-S observations are the most available although a high network time dependency.

Nr of OBS evolution (2014090100 - 2014091000)
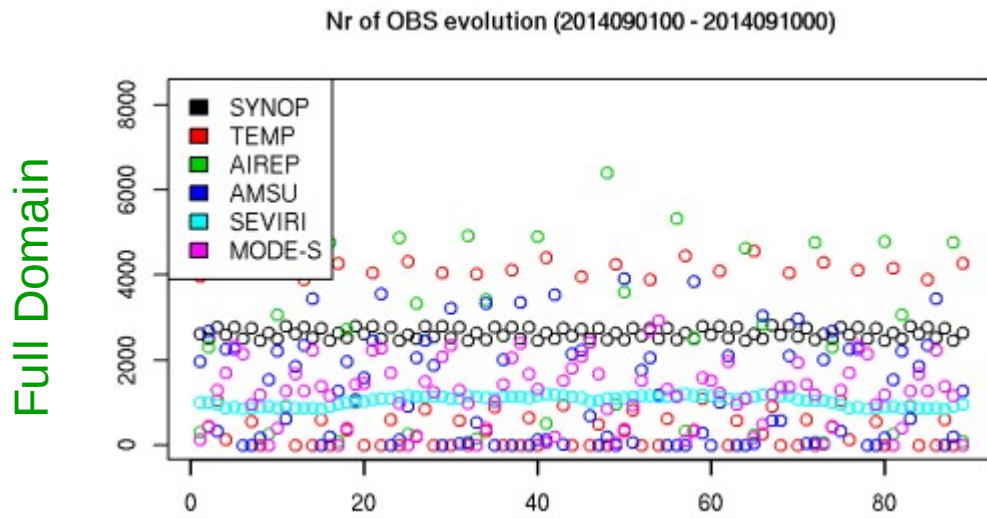
Figure II.9: Number of observations time evolution
for observation type on full domain

Nr of OBS evolution (2014090100 - 2014091000)
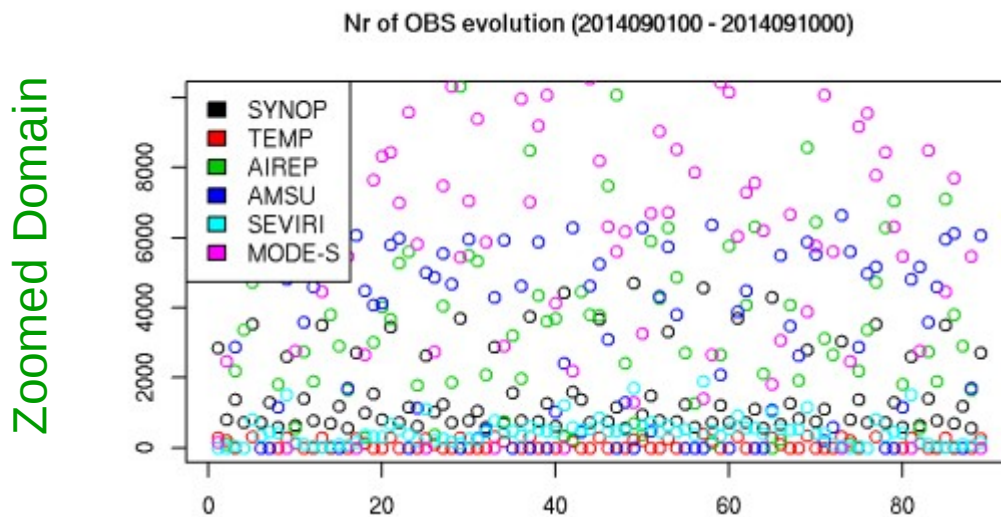
Figure II.10: Number of observations time evolution
for observation type on zoomed domain

The following figures show the accumulated weight and accumulated number of observations for every considered type:
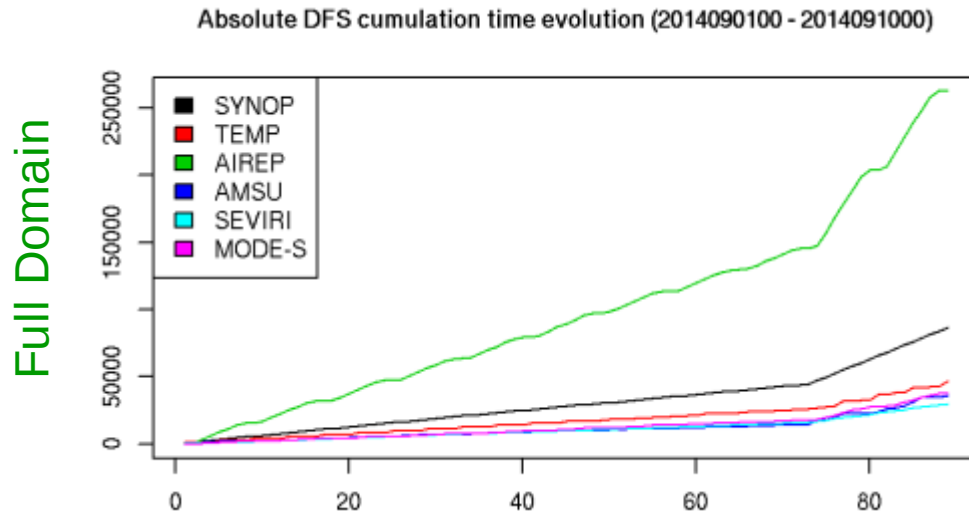
Full Domain

Figure II.11: cumulated DFS of observation types on full domain
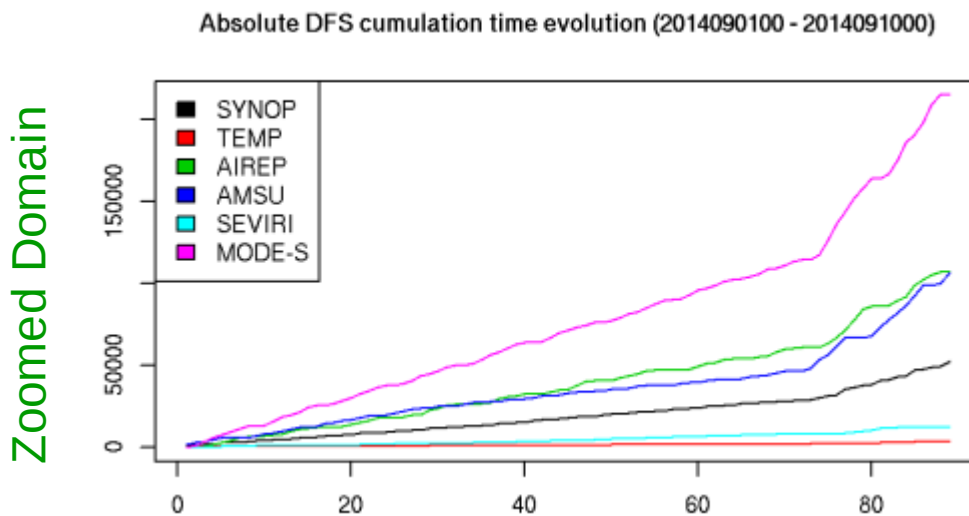
Zoomed Domain

Figure II.12: cumulated DFS of observation types on zoomed domain
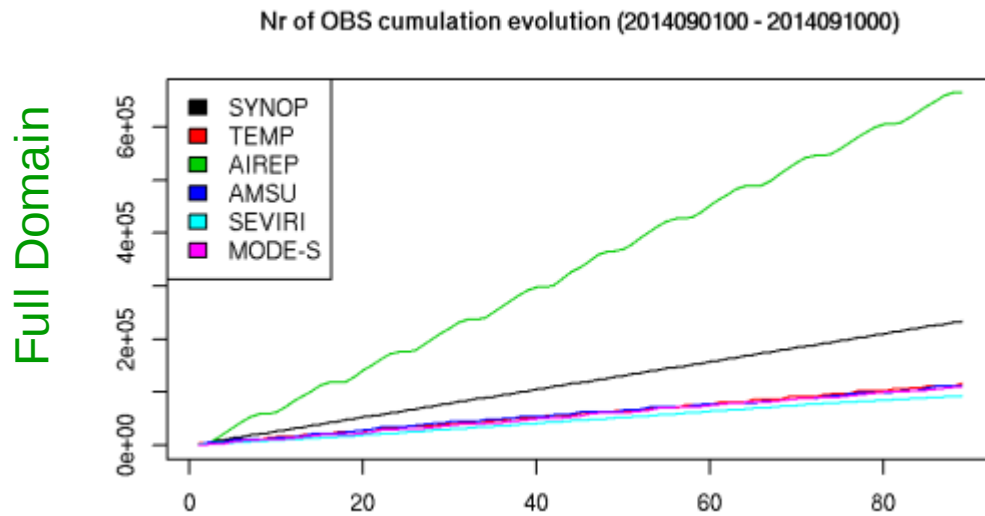
Figure II.13: cumulated number of observations of available types on full domain
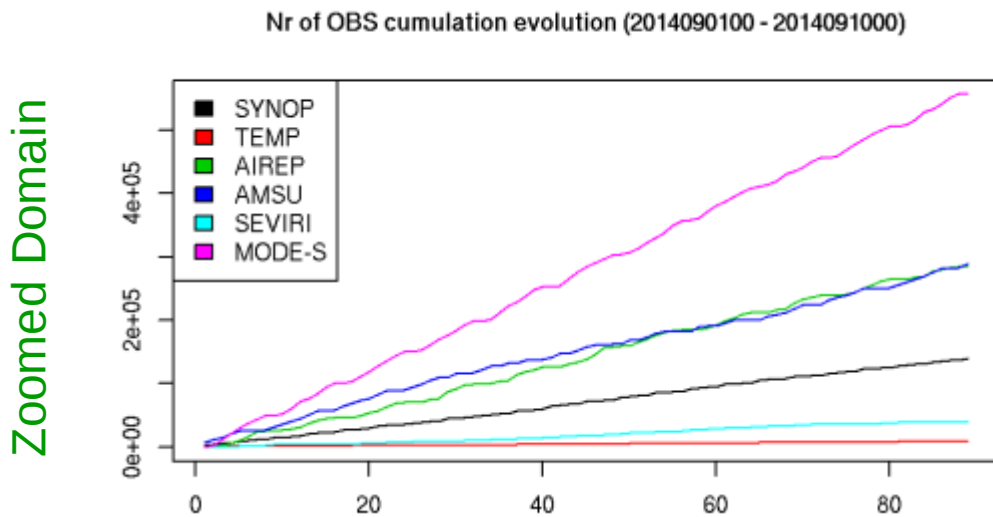
Figure II.14: cumulated number of observations of available types on zoomed domain

III. Moist Total Energy Norms (MTEN)

The Moist Total Energy Norms (MTEN) method is based on the comparison of the moist total energy norm-based cost functions. It evaluates the forecasts sensitivity to observations.

The cost function of a subset of observations i can be described as:

$$J^i = <x_t^i - x_t^{ctr}, x_t^i - x_t^{ctr}>$$

where x_t^i is the forecast initialized without the subset i and x_t^{ctr} is the full observations assimilation forecast.

According to Ehrendorfer et al. (1999), this norm can be described as:

$$<x_t^i - x_t^{ctr}, x_t^i - x_t^{ctr}> = \int_{\eta_0}^{\eta_1} \int \left(u^2 + v^2 + \frac{c_p}{T_r}T_2 + \frac{RT_r}{p_r^2}p^2 + \frac{L^2}{c_pT_r}q^2 \right) \frac{\delta p_r}{\delta \eta} d\eta dD$$

where:

- u, v, T, p and q: The difference between control forecast and forecast without i[th] subset of observation respectively for u and v components of wind, temperature, surface pressure and specific humidity;
- c_p: Specific heat at constant pressure;
- R: Gas constant of dry air ;
- L: Latent heat condensation;
- T_r: Reference temperature;
- p_r: Reference pressure;
- η: Vertical coordinate.

Further detailed description of this method can be found in Storto A, Randriamampianina R. (2010).

The MTEN method was applied over several observation types namely TEMP, SYNOP, AMDAR, AMSU-A and SEVIRI observations.

The experiment covered the period from September 05th to 10th. It's to note that the following results concern the level 80 to 87.

The results showed a dependency to the network time, thus 4 networks were taken into account; 00h, 06h, 12h and 18h as described in the following figures:
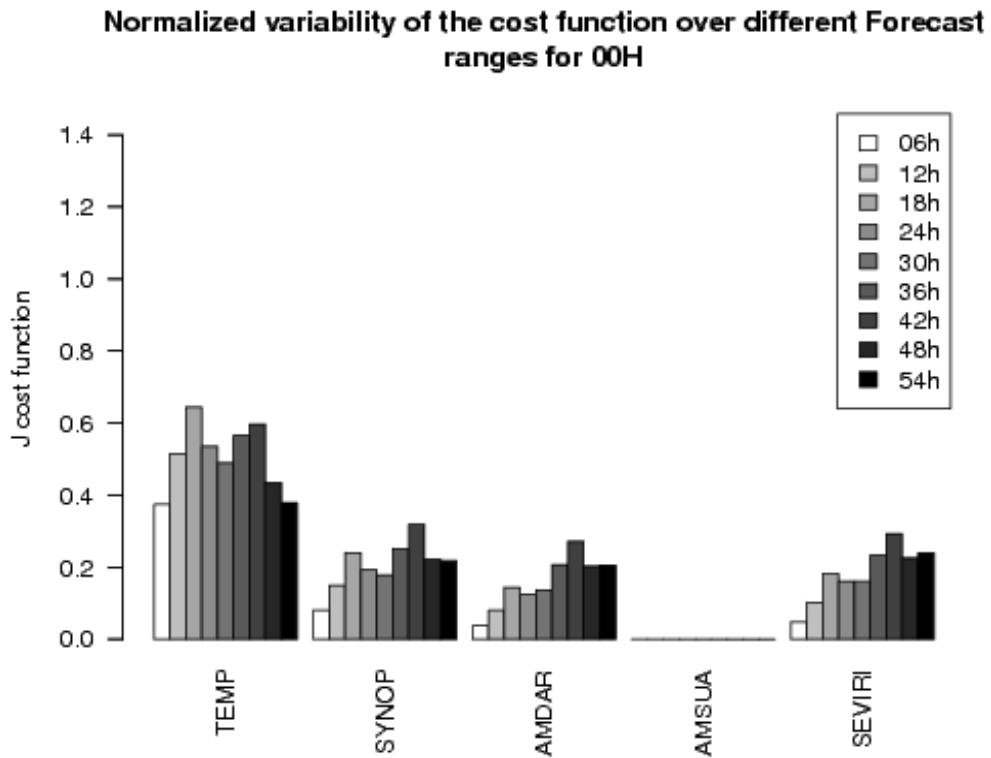
Figure III.1: Normalized variability of the cost function over different forecast ranges for 00H
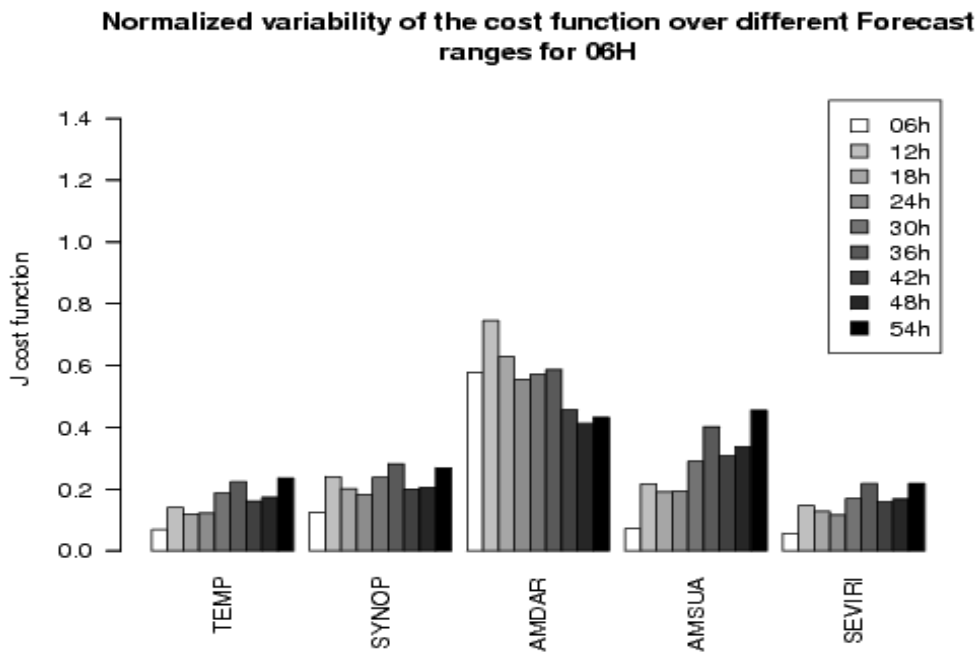
Figure III.2: Normalized variability of the cost function over different forecast ranges for 06H
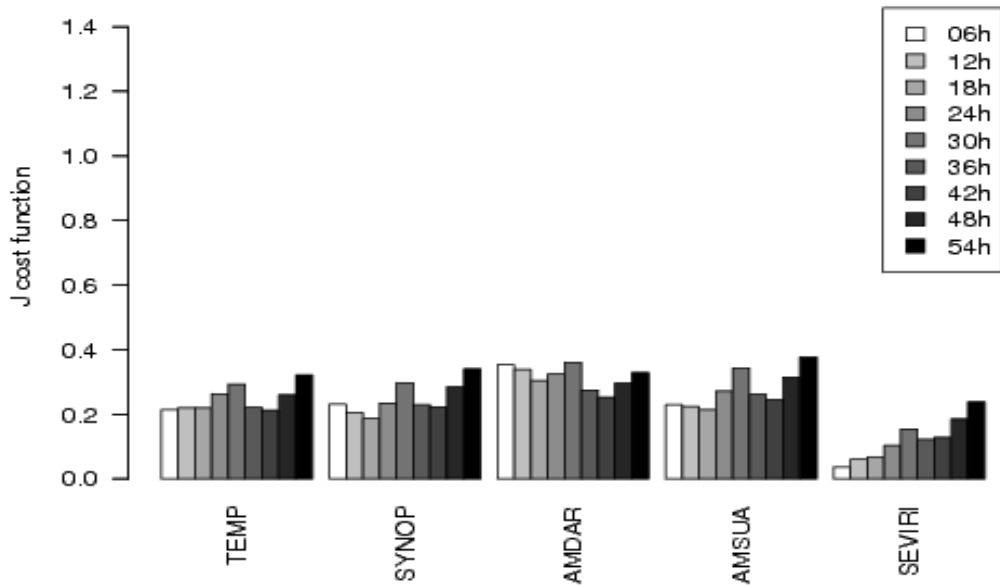
Figure III.3: Normalized variability of the cost function over different forecast ranges for 00H
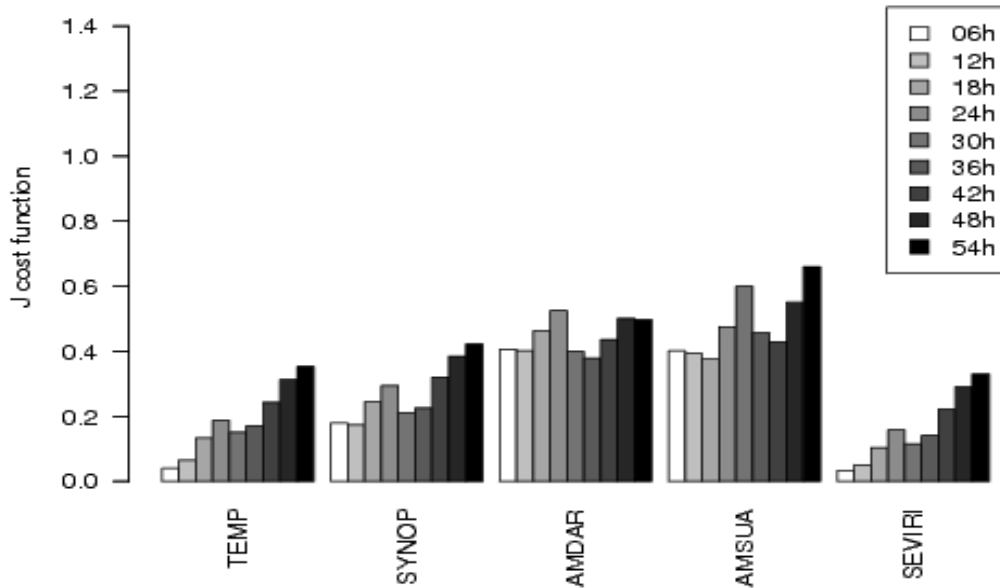
Figure III.4: Normalized variability of the cost function over different forecast ranges for 00H

The results show a high impact of the TEMP observation for 00H network, while an important impact of AMDAR observations was noticed at 06H. This may be related to observation availability which vary notably with time.

The results show also a limited influence of the Seviri observation in comparison to conventional observations and mainly in comparison to aircraft and AMSU-A observations.

IV. TuneBR

This method involves the background and observation error statistics which can be tuned a posteriori in order to adjust the predefined errors.

In fact, the diagnosed errors indicate the real standard deviations of the observation and background errors, and they can be compared with the prescribed ones. The difference between the diagnosed and predefined standard deviations leads to consider the following ratios:

$$r_o = so_d/so_p \text{ and } r_b = sb_d/sb_p$$

where:

so_d: Diagnosed standard deviation of observation;
so_p: Predefined standard deviation of observation;
sb_d: Diagnosed standard deviation of background;
sb_p: Predefined standard deviation of background.

The so_p and so_b predefined standard deviations should than be multiplied by r_o and r_b ratios respectively. These ratios concern the following parameters:

q: specific humidity;
t: temperature;
bt: brightness temperature (just for observation ratio);
ke: kinetic energy (the squared average of u and v wind components).

The tuning of the predefined observation values is possible through the namelist aldnml_screen where it's possible to set the SIGMAO_COEF.

The predefined background values can also be tuned through the namelist aldnml_minim by setting REDNMC. It's also possible since the cy36t1 to set the specific humidity standard deviation values separately through the key REDNMC_Q. In such a case, two logical variables should be set to true namely LREDNMCQ and LREDQERR.

More detailed information about the calculation of r_o and r_b can be found in Desroziers et al. (2005).

The TuneBR method was applied to the operational suite in Slovenia. As a first result, the system had to be tuned within the following mean observation and background ratios:

Table IV.1: Diagnosed observation and background ratios for operational suite
(February 20th to October 7th)

Variable	Cases	Ratio_o	Ratio_b
q	284644	0.60149	1.09751
t	4363447	0.78626	1.32258
bt	9848072	1.02208	
ke	4270991	0.79862	0.61264
Mean	8919082	0.91790	1.03630

These ratios were calculated over the period from February 20th to October 7th. For a cost reduction, we considered the period from September 1st to September 25th to proceed to tunings. Thus, the new calculated ratios are the following:

Table IV.2: Diagnosed observation and background ratios for operational suite
(September 1st to September 25th)

Variable	Cases	Ratio_o	Ratio_b
q	38363	0.69265	1.38649
t	619896	0.73094	1.34506
bt	1340447	1.16326	
ke	600949	0.74928	0.63674
Mean	1259208	0.98081	1.06899

The ratios ro and rb were then used to tune the observation and background standard deviations as explained above, and the new values were set as following:

– SIGMAO_COEF: 0.9 ==> 0.8
– REDNMC_Q: 1.6 ==> 1.45

A new experiment including the new SIGMAO_COEF and REDNMC_Q values was run over the same period, the new calculated ratios are summarized in the following table:

Table IV.3: Diagnosed observation and background ratios for experimental suite
(September 1st to September 25th)

Variable	Cases	Ratio_o	Ratio_b
q	38405	0.67865	1.35098
t	538567	0.70961	1.25525
bt	1247940	1.11331	
ke	533498	0.72514	0.60269
Mean	1110470	0.94726	1.00090

The new diagnosed values show a background ratio which is much closer to 1. This indicates a satisfying improvement in the background standard deviation values. However, more tunings need to be proceeded for the observation standard deviations.

Conclusions

This stay was very enriching experience, during which I could learn and use new verification methods. The different applied tools helped improving different components of the data assimilation system in Slovenia.

The VarBC Monitoring tool results showed the necessity to revise the enabled satellite channels. A first step was to blacklist some satellite signals which were shown none beneficial. As a second step, we could tune some VarBC values for the AMSU-A and AMSU-B sensors. This tuning improved the analysis scores as was explained in details in the section I, but further investigation can be done in order to understand the increase of oma variability amplitudes since May 8th 2014.

The second applied tool, namely the Degree of Freedom of Signal, helped understanding the weight of the available observations on analysis over the LACE and Slovenian domains. The reading of both domains results showed a noticeable impact of aircraft observation, and especially the Mode-S observations.

Then, we tried to identify the different impact of observation types on forecasts by applying the Moist Total Energy Norms method over different network times. The aircraft observations showed a major weight seconded by the AMSU-A observations. A less impact was demonstrated for SEVIRI observations. A further work can be proceeded by investigating the weight of Mode-S observations especially over Slovenia domain.

The last verification tool, TuneBR, was applied to check the balance between the background and observation standard deviations values, and improve the predefined values by suggesting multiplication ratio. A first iteration could help improving the quality of the background values, but further work may be done in order to improve the observation values.

References

Ehrendorfer M, Errico RM, Raeder KD.   1999 Singular-vector perturbation growth in a primitive equation model with moist physics. Journal of Atmospheric Sciences 56: 1627–1678.

Desroziers G., Berre L., Chapnik B., Poli P. 2005 Diagnosis of observation, background and analysis error statistics in observation space, QJRM, 131, 3385-3396

Storto A., Randriamampianina R.   2010 The relative impact of meteorological observations in the Norwegian regional model as determined using an energy norm-based approach

C. D. Rodgers.   1996 Information content and optimisation of high spectral resolution measurements.

M. Fisher.   2003 Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. Technical report, ECMWF

C.Cardinali, S.Pezzulli, and E.Andersson. 2004 Influence-matrix diagnostics of a data assimilation system.

Strajnar B.   2012 Validation of Mode-S Meteorological Routine Air Report aircraft observations