Regional Cooperation for Limited Area modeling in Central Europe

Validation of ENS 3DVar within ALADIN-LAEF Phase II

Report on stay at ZAMG 30/07 – 24/8/2018

Ing. Martin Imrišek

Supervisor: Mgr. Martin Belluš
Wien 2018

Contents

Acknowledgement

I would like to thank to my supervisor Martin Belluš for his help and advices during my stay at ZAMG, and also my thanks belongs to Florian Meier, Mile Mate and Alena Trojáková for their support. I can not forget mention help from Geodetic and Cartographic Institute in Bratislava, Slovak University of Technology in Bratislava and Zentralanstalt für Meteorologie und Geodynamik.

1 Foreword

The 3D variational data assimilation is one of the main approaches how to prepare the most reliable initial conditions of upper air for the integration of numerical weather prediction model. The first goal for my stay was the implementation of various data types to 3DVar of ALADIN–Limited Area Ensemble Forecasting (LAEF) system. Following task was impact validation of data assimilation to upper air (separately and combined) and implementation of data perturbation. The main goal was verification of ALADIN–LAEF experiment with 3DVar (*phase II*) against ALADIN–LAEF experiment without 3DVar (*phase I*).

2 Short LAEF description

ALADIN–LAEF system with 16 members runs on cycle 40t1 with 4.8 km horizontal resolution and 60 vertical model levels with ALARO-1 physics. Domain covers wide area $\lambda = (351.80°, 67.06°)$, $\phi = 25.76°, 52.85°)$. The ensemble system consists from this three steps: ESDA (Ensamble of Surface Data Assimilation), upper air spectral Blending (with ECMWF) and Integration (12 or 24 hours and 6 hour coupling). All experiments were performed over the period of two weeks from 16.05.2016 00 UTC to 31.05.2016 12 UTC, with 2 network ranges (00 and 12UTC).

3 Technical verification of 3DVar data assimilation

The first task was to implement and verify 3DVar data assimilation to existing system with all possible kinds of observation. The ordinary SYNOP data (OPLACE), AMDAR (OPLACE), TEMP (OPLACE), GEOWIND (OPLACE) and GNSS zenith total delay (SUT) were used. In common assimilation cycle there were assimilated cca 6200 (37.85%) SYNOP measurements (including GNSS zenith total delays), cca 550 (3.35%) AMDAR measurements, cca 32 (0.20%) GEOWIND measurements and cca 9600 (58.60%) TEMP measurements. There were no difficulties in implementation of SYNOP, AMDAR and TEMP data. Increments are shown in figures 1, 2 and 3 respectively. At assimilation time 00UTC were increments at model level 55 mainly in the area of airports in cities Antalya, Aerhus and Gothenburg. The main impact of AMDAR data is around 15-th model level which has Standard Atmosphere Height around 10332 m.

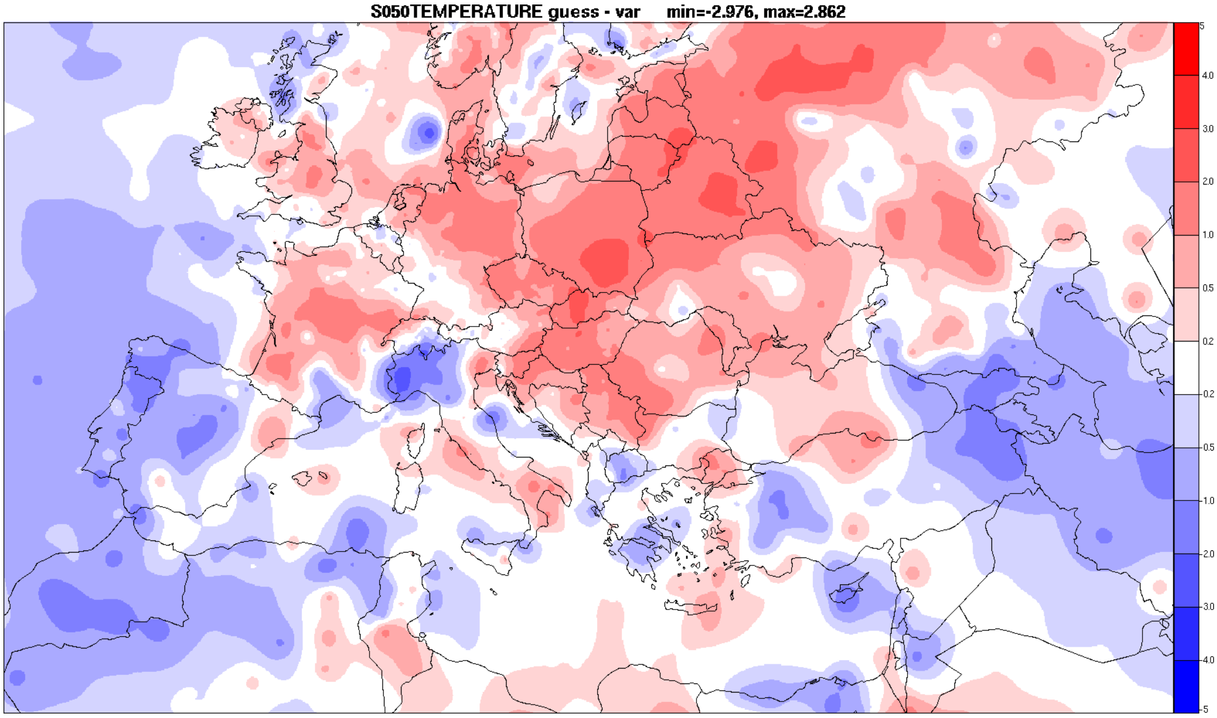

Figure 1: Increments in temperature [°C] at model level 50 only from SYNOP data assimilation.

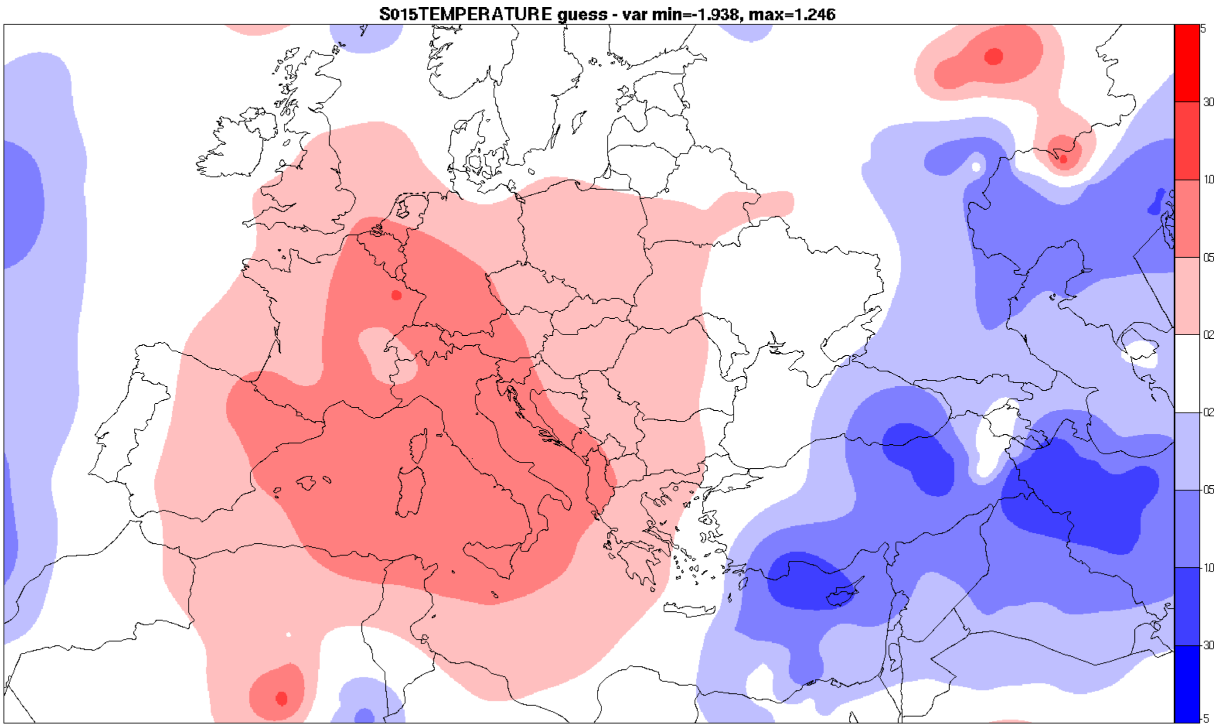

Figure 2: Increments in temperature [°C] at model level 15 only from AMDAR data assimilation.

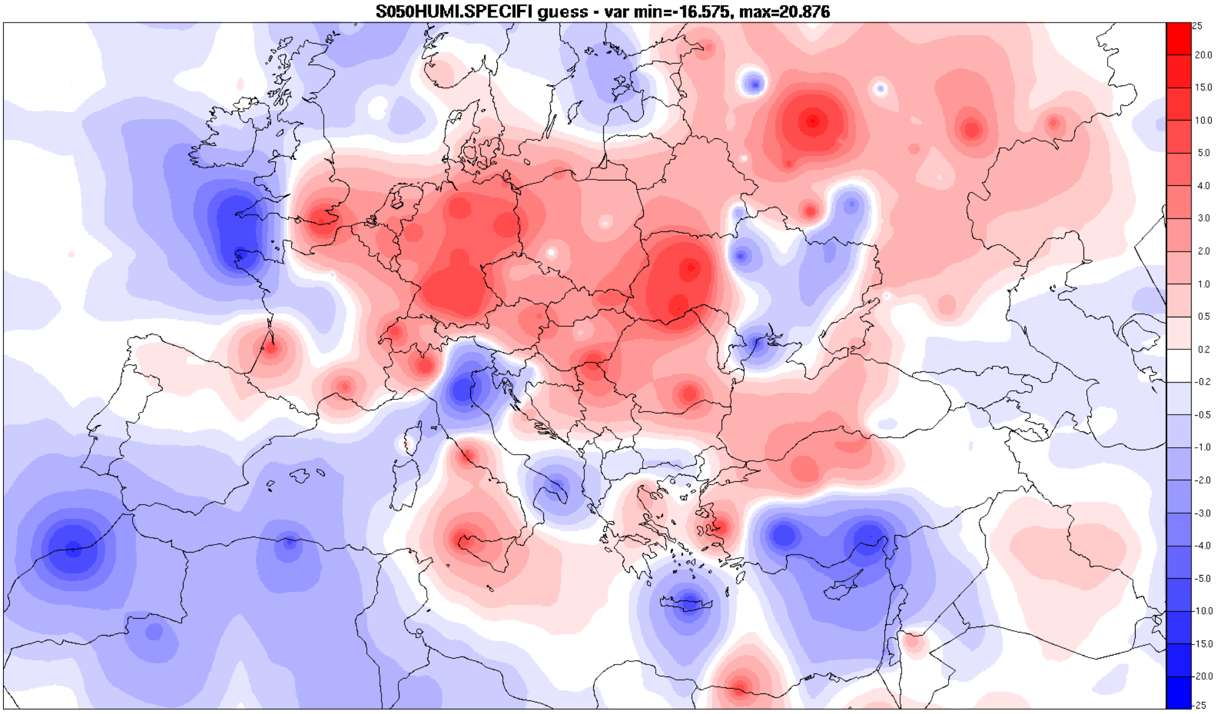

Figure 3: Increments in specific humidity [g kg^{-1}] at model level 50 only from TEMP data assimilation.

As one can see, the values of increments are in reasonable ranges, so this data were used as input in *phase II* assimilation experiment (section 4).

3.1 Assimilation of GEOWIND

More work was done on assimilation of GEOWIND, it was necessary to make some modifications in *mf_blacklist.b* in section *Geographical blacklist by satellites*. The highest increments values are around model level number 45, as one can see on figures 4 and 5. This model level has Standard Atmosphere Height around 862 m. The low count of assimilated data can be increased by tuning of thinning and blacklisting. More reliable analysis can be achieved by assimilation of another atmospheric motion vectors: HRWIND.

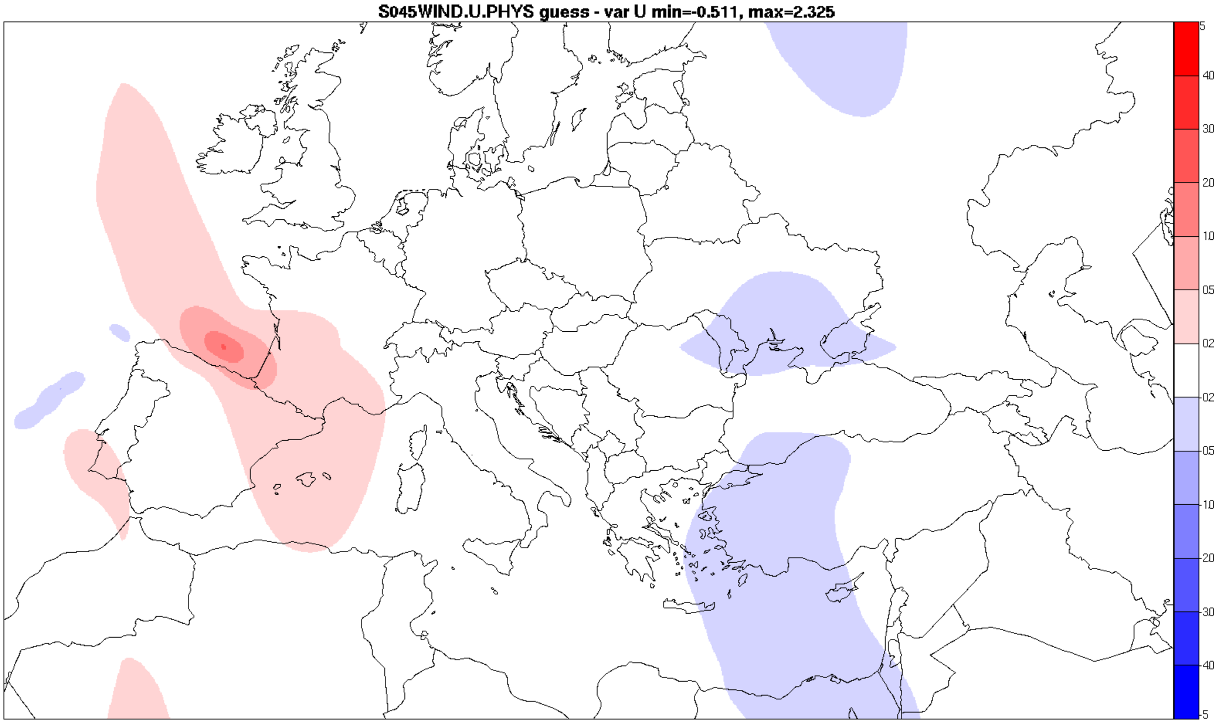

Figure 4: Increments in wind speed [m s^{-1}] in direction U at model level 45 only from GEOWIND data assimilation.

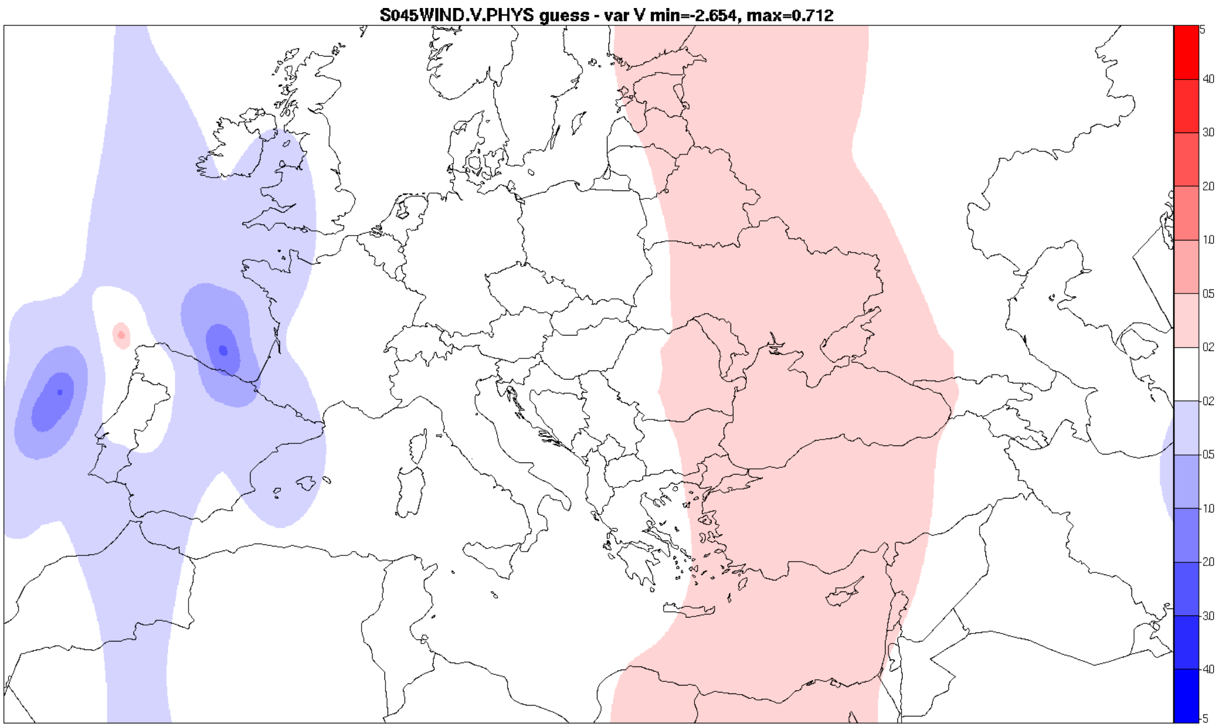

Figure 5: Increments in wind speed [m s^{-1}] in direction V at model level 45 only from GEOWIND data assimilation.

3.2 Assimilation of GNSS zenith total delay

Before GNSS data assimilation the whitelist for GNSS permanent stations has to be assembled. The first step was estimation of first guess (integration without 3DVar) departures for all 16 members and all time steps from whole period. The first idea, was to obtain whitelist for GNSS stations from combination of all first guess departures from all members and all terms. But this approach was not satisfying, because combining all first guess departures for one station will make the histogram of departures too much kurtosis and the Pearson's Chi-squared test (H0: Data set have normal distribution) will reject station for further assimilation. Therefore another approach was investigated, each member was separately tested and the count varies from 5 to 12 of rejected stations per member. Third approach was examined, based on testing all members together just for one day with Jarque–Bera test (if data have skewness and kurtosis matching a normal distribution, H0: Data set have normal distribution). This test also rejects different amount of stations in each day. Based on this results it was decided to use best day (with all members and lowest amount of rejected stations) and best member (with all days and lowest amount of rejected stations) together for whitelist estimation. Two stations were excluded due to big difference in model and orography altitude and two stations were excluded due Pearson's Chi-squared test. The increments in specific humidity after GNSS zenith total delay assimilation are shown in Figure 6.

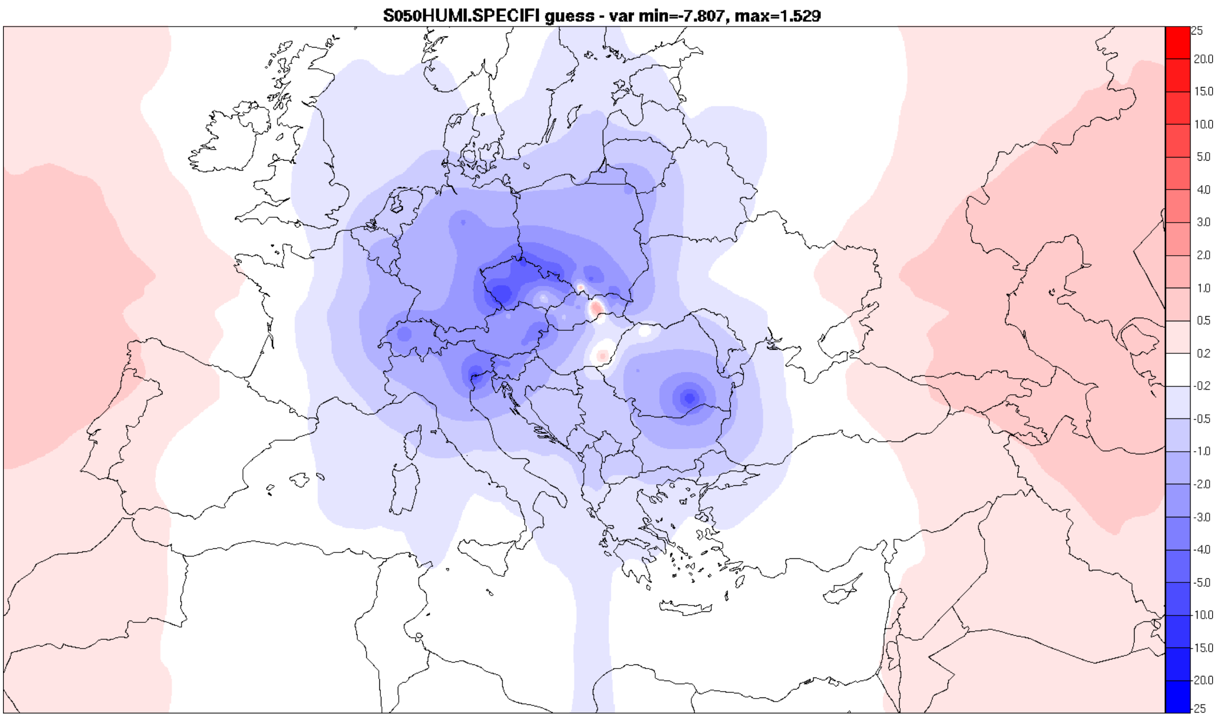

Figure 6: Increments in specific humidity [$\mathrm{g\,kg^{-1}}$] at model level 50 only from GNSS zenith total delay data assimilation.

At first sight the specific humidity increments in central Europe are opposite like in

TEMP assimilation, but GNSS zenith total delay represents whole air mass above station whereas TEMP data are discrete measurements in space while the radiosonde is drifting. The changes in TEMP increments depend on the level of model, as one can see in figures 7 and 8.

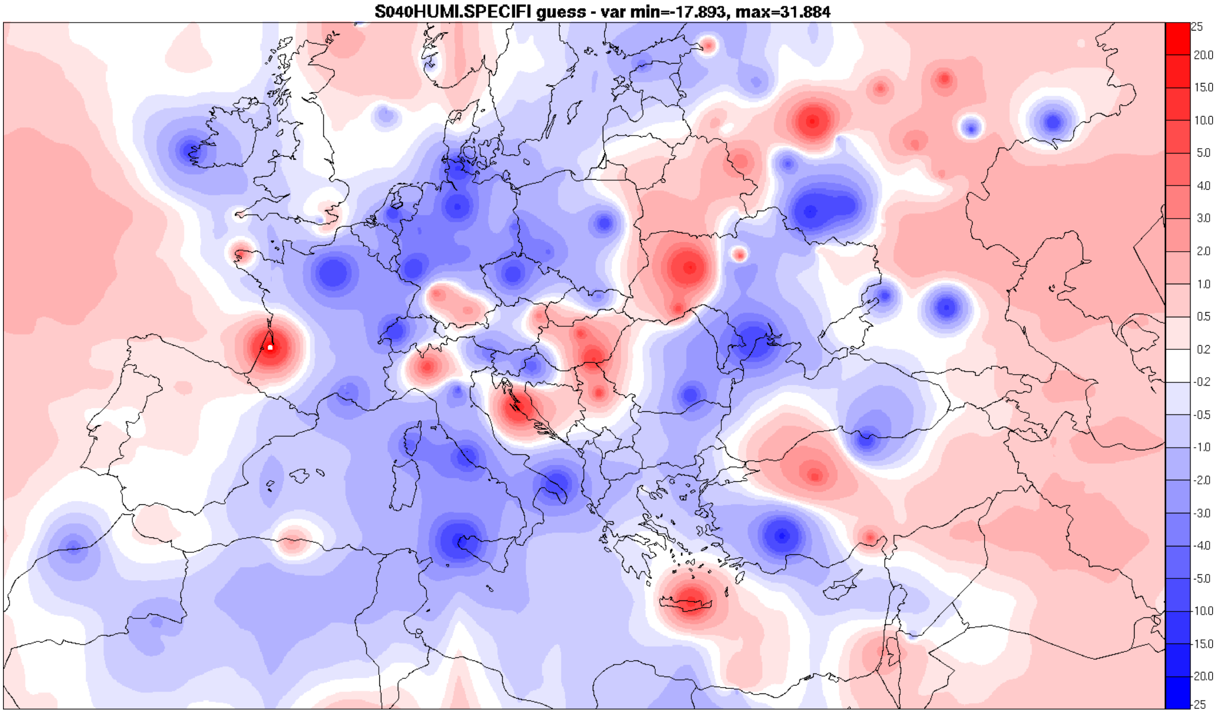

Figure 7: Increments in specific humidity [g kg^{-1}] at model level 40 only from TEMP data assimilation.

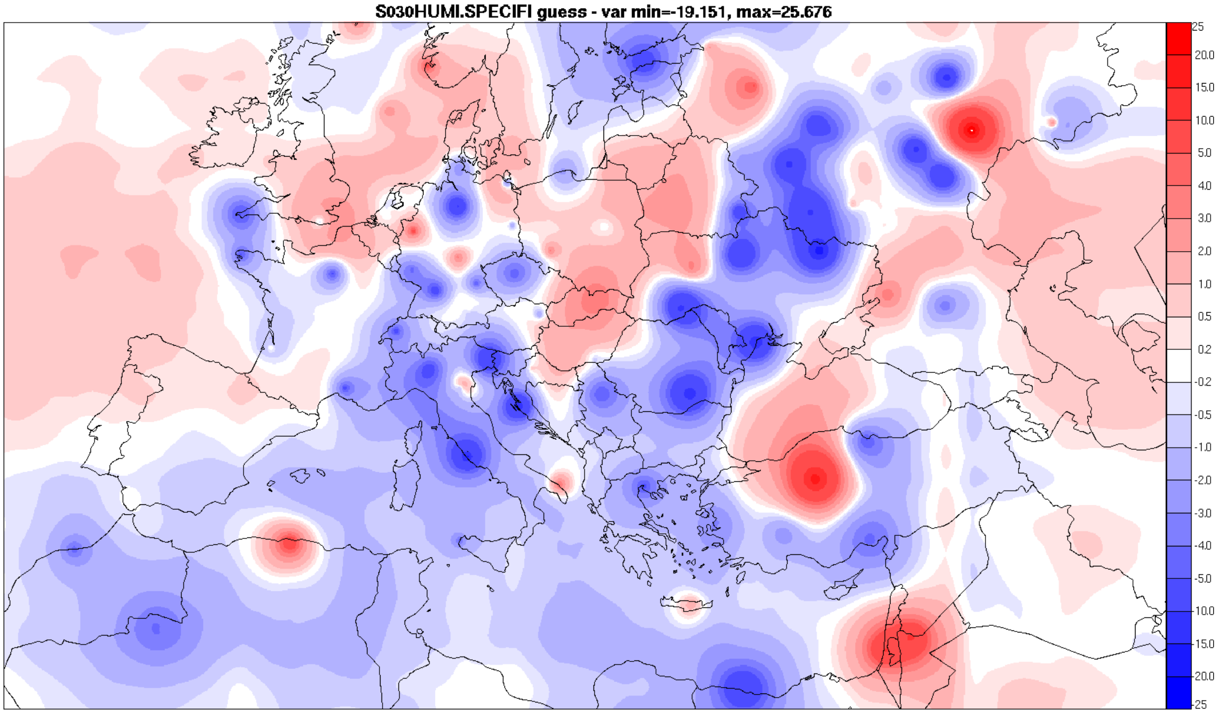

Figure 8: Increments in specific humidity [$\mathrm{g\,kg^{-1}}$] at model level 30 only from TEMP data assimilation.

3.3 Data flow

At this stage it was investigated, whether 3DVar upper air assimilation will be more suitable after or before the Blending step. Two experiments with assimilation of all available data were done, the first was blend–var and second was var–blend. In blend–var is more dominant the effect of assimilation as it is shown in Figure 9, on the other hand in var–blend (Figure 10) are gradients in increments of temperature at model level 50 smaller. The model fields are smoother due to digital filtering in Blending. It is obvious from the comparison of mean values of temperature increments, that the var–blend mean value is smaller (that means hotter) than blend–var (difference is 0.033 °C), but this influence can by neglected.

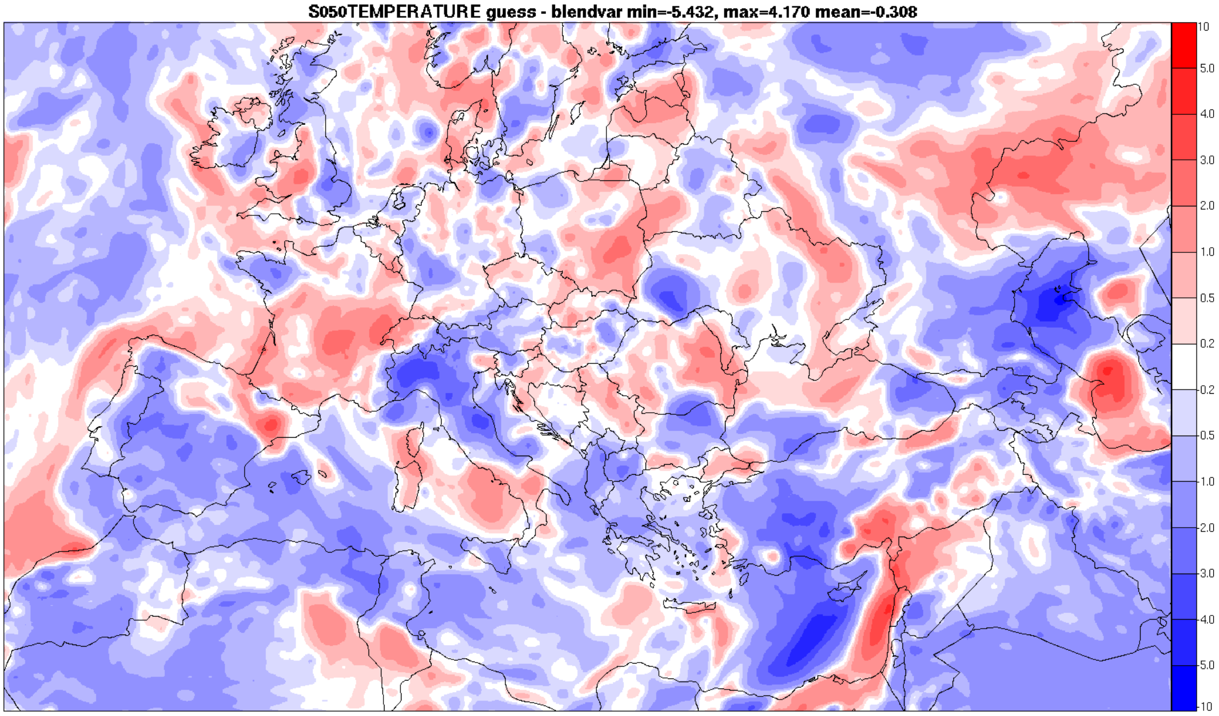

Figure 9: Increments in temperature [°C] at model level 50 from assimilation of all data from blend–var.

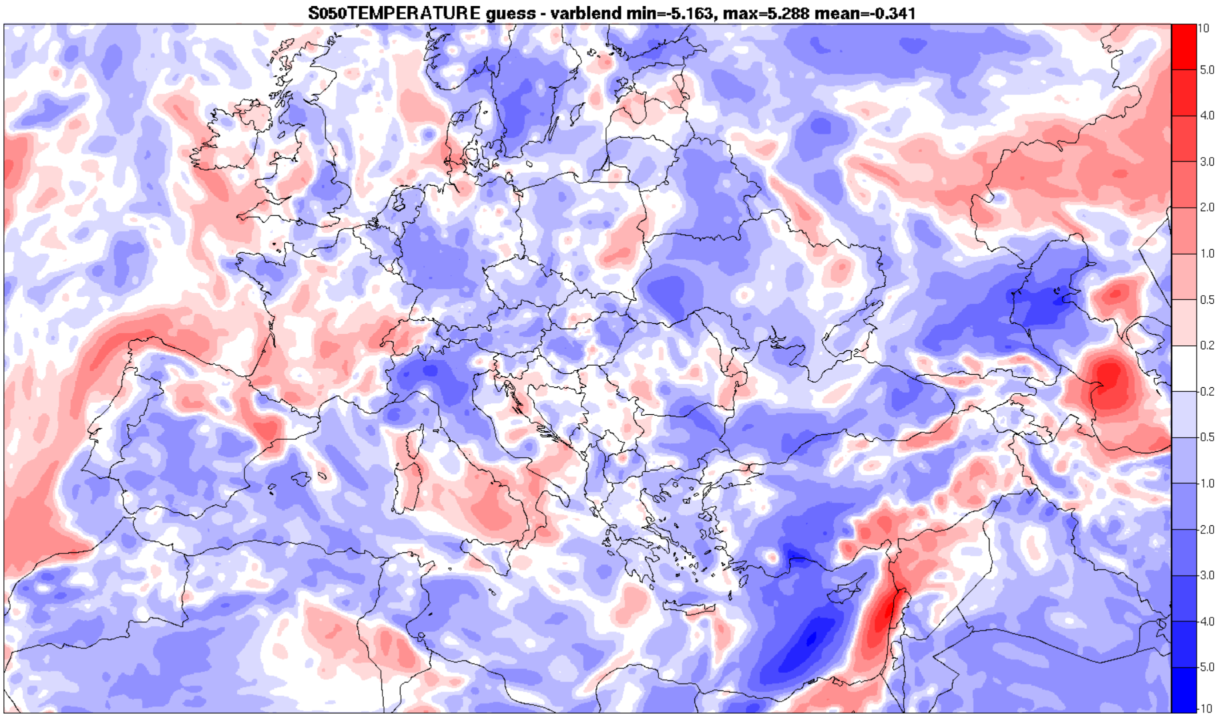

Figure 10: Increments in temperature [°C] at model level 50 from assimilation of all data from var–blend.

It was decided, that the implementation of assimilation will be before Blending i.e. the steps will follow in this sequence: ESDA, ENS 3DVar, Blending and Integration. This chosen data flow is experimental and can be changed in future with simple intervention. The integration length depends on a day time. For 00UTC the integration was set on 12 hours only for cycling needs, but for 12UTC it was set on 24 hours for cycling and validation purpose.

3.4 Perturbation of assimilated data

The next task was implementation of perturbation of assimilated data to 3DVar. The perturbation of assimilated data was done in quality check step of 3DVar (Screening) by adding three entries to the namelist in section *&NAMSCC*:

- *LPERTURB=.T.,* this entry enables perturbation of data,
- *NAENSEMBLE=1,* this entry enables changes of perturbation depending on number of ensemble member,
- *NAEMEMBER={MEMB},* this entry sets the number of ensemble member.

Perturbation is done by adding a small error with normal distribution to all data. The impact of perturbation is satisfying (mean is -0.02 °C) on temperature as is shown in Figure 11.

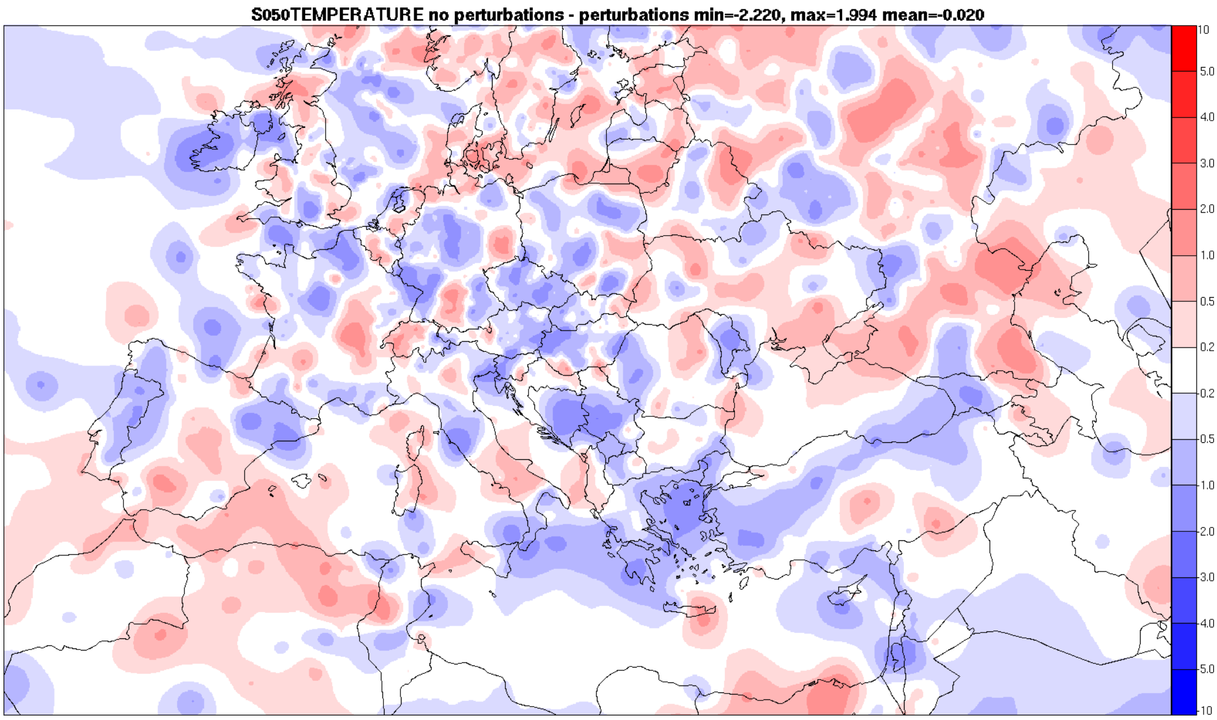

Figure 11: Difference in temperature [°C] at model level 50 between assimilation of all data with and without perturbation.

4 Impact of 3DVar data assimilation

After validation of data perturbation in upper air assimilation the full LAEF system (16 members) 12 (00UTC) or 24 (12UTC) hour forecast was computed for whole time period. This means 16 members * 16 days * 2 network times with these following processing steps:

- ESDA – perturbed SYNOP data,
- ENS 3DVar – perturbed all data,
- upper air Blending – with ECMWF,
- Integration – 12 or 24 hour.

The experiment was verified against the ECMWF reanalysis for the period from 16.05.2016 12UTC to 30.05.2016 12UTC only at 12UTC and 24 hour forecast. These verification scores are denoted with name *phase II*. The *phase I* scores were computed from the same time window but without 3DVar data assimilation. The verification parameters bias, outliers, RMSE and spread were computed for the 925, 850, 500 and 250 hPa pressure levels for temperature, relative humidity, geopotential and wind speed. The interesting and significant results are shown in the following figures.

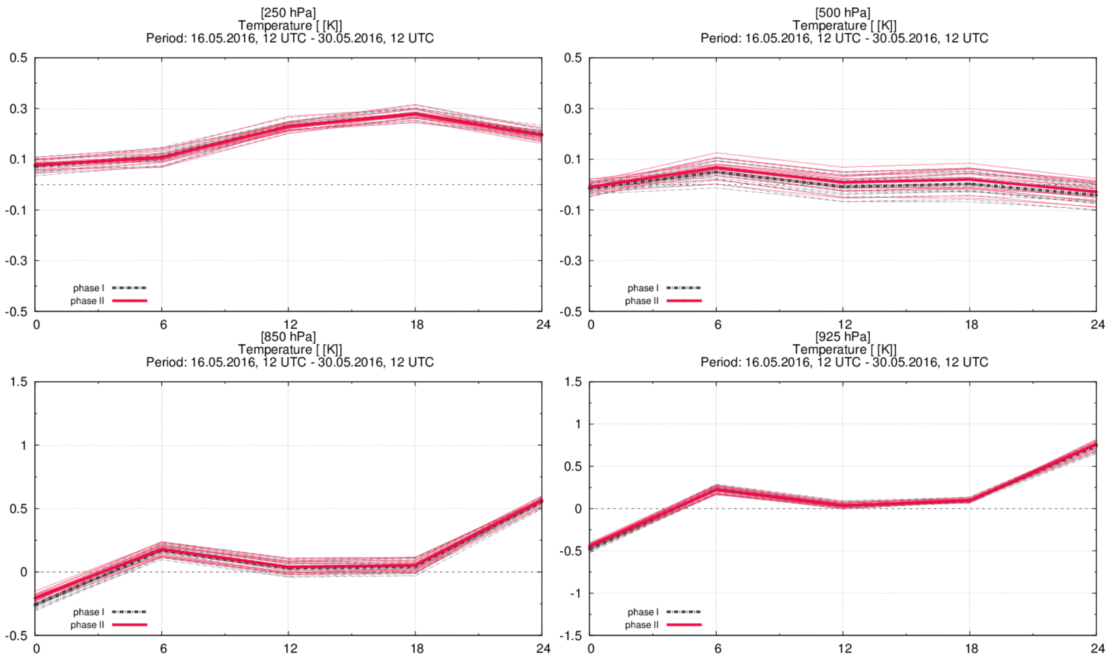

Figure 12: Temperature bias, small improvement was noticed in 850 hPa pressure level at time of analysis (0 hour of forecast).

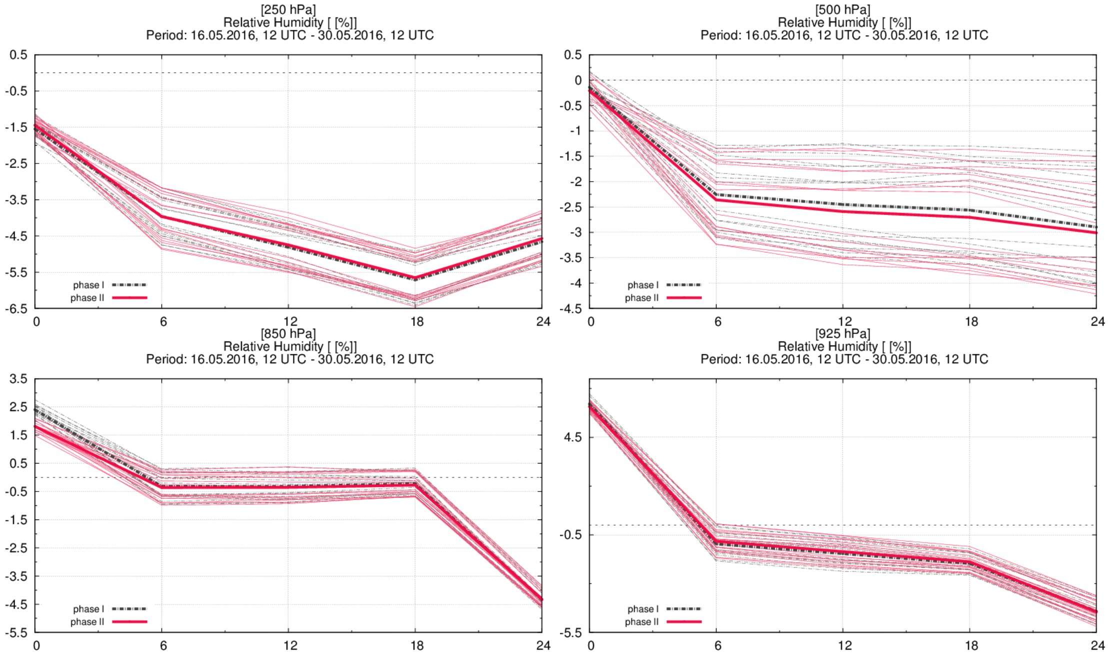

Figure 13: Relative humidity bias, small improvement was noticed in 850 hPa pressure level at time of analysis, but also small degradation at 6,12,18 and 24 hour forecast at 500 hPa level.

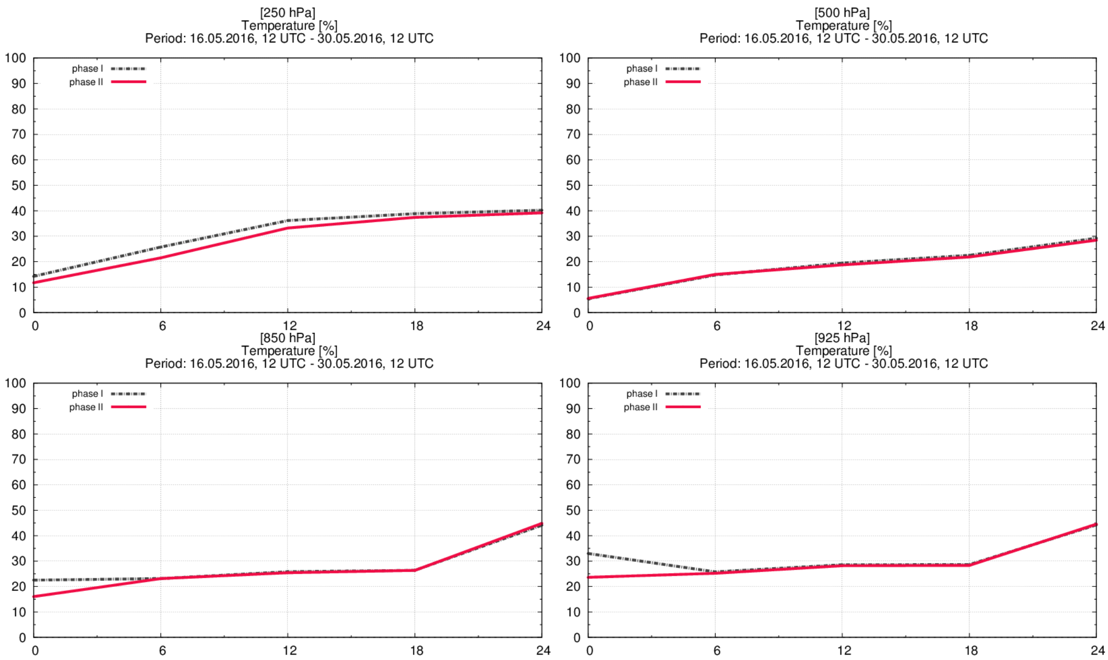

Figure 14: Temperature outliers, all pressure levels show improvement (decrease) in number of outliers at time of analysis, the same results one can see for relative humidity, geopotential and wind speed (not shown)

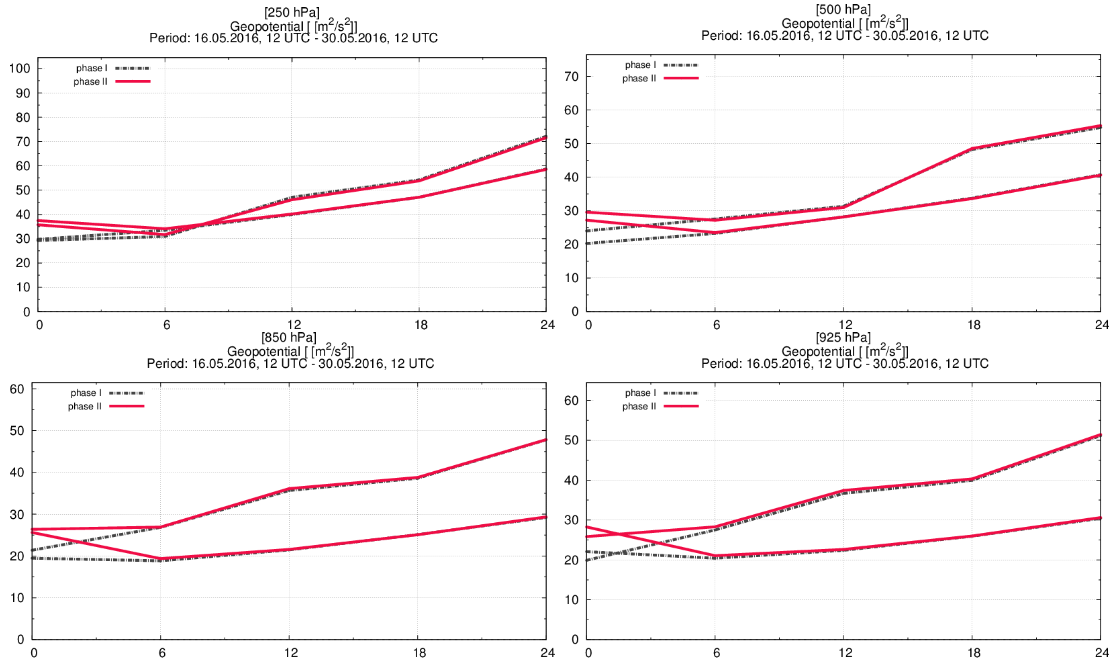

Figure 15: Geopotential RMSE and spread, degradation was found at time of analysis in all pressure levels in RMSE (increase – the upper line), but improvement (increase – the bottom line) of spread is present at the same time. This Geopotential RMSE degradation is the biggest drawback in the verification.

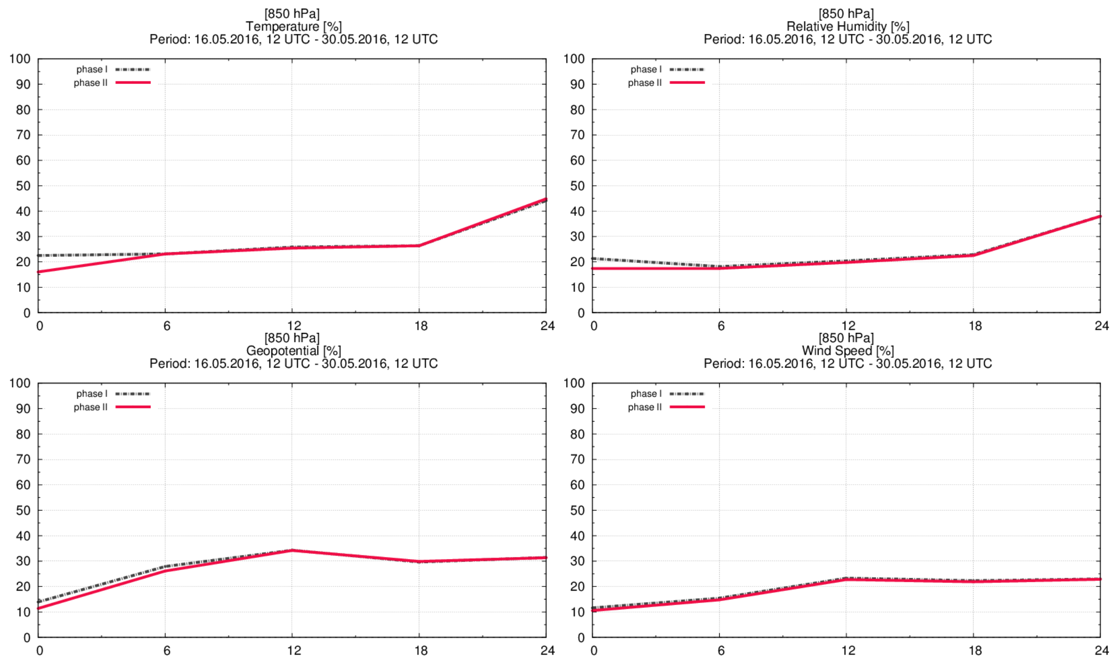

Figure 16: 850 hPa outliers for temperature, relative humidity, geopotential and wind speed, as one can see the improvement in present only at time of analysis.

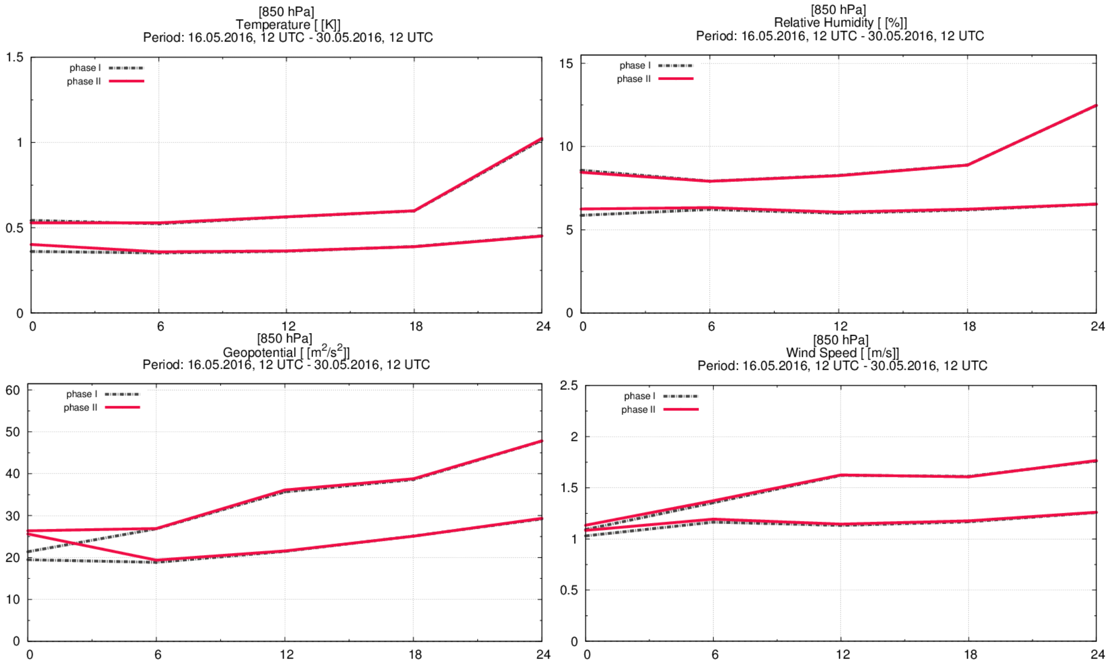

Figure 17: 850 hPa RMSE (upper line) and spread (bottom line), the 3DVar has neutral impact, but spread has been improved (increase) in all cases.

Impact in other cases is rather neutral, or is so small that can be neglected, so other verification cases are not displayed. But overall one can say, that 3DVar have small positive impact in all pressure levels on outliers (decrease) and on spread (increase). On the other side the slight degradation in RMSE of geopotential has been noticed.

5 Conclusions

The 3D variational data assimilation in upper air was successfully implemented in existing LAEF 5km system. The technical correctness was confirmed by examination of the increments in temperature, specific humidity or wind speed in various model levels. At this implementation stage it is possible to assimilate SYNOP data, AMDAR, TEMP, GEOWIND and GNSS zenith total delays. The 3DVar was included into LAEF system after ESDA and before upper air Blending, more in section 3.3. Also the perturbation of assimilated data was successfully implemented in data quality check, section 3.4. After those achievements the forecast verification of temperature, relative humidity, geopotential and wind speed was performed (section 4). The impact was assessed on four pressure levels: 250, 500, 850 and 925 hPa. The bias, outliers, RMSE and spread was computed for ensemble system with (*phase II*) and without (*phase I*) 3DVar. For most cases one can say the impact is neutral, or is so small that can be neglected. Only for the outliers (decreased) and spread (increased) one can say that 3DVar has positive impact, nevertheless the degradation of the RMSE in geopotential was noticed. It is possible, that change of data flow (blend–var scheme) might have more impact on 6 hour forecast. It is recommended to extend this verification by some interesting case studies as well. They may better show the added value of ENS 3DVar method.