Computation of the background-error covariances

report from RC LACE stay in Prague, 22/8/ - 9/9/2022

Alina Dumitru, NMA, alina.dumitru@meteoromania.ro
in collaboration with Antonín Bučánek and Alena Trojáková, CHMI

Contents

1 Introduction

In numerical weather prediction models the evolution of the state of the atmosphere is related to accuracy of initial conditions. In the last years many different methods and approaches for improving the initial conditions were developed, and are known as a part of atmospheric science as data assimilation (DA) *(Stanesic et al., 2019)*. Data assimilation merge model state (forecast or backgrounds) and observations in order to obtain the best analysis (known as estimate of atmospheric state). The precision of the analysis is related to the accuracy of the background and observation error statistics *(Monteiro et al., 2010)*.

Observation-error covariances contain information on errors in the observation process and background-error covariances represent errors in the background state (the difference between the background state and the true atmospheric state). The error covariances are used to scale, spatially filter and propagate away information from observations. The background-error statistics (B matrix) are also used for corrections of different model variables according to balance properties of the atmosphere *(Daley, 1991)*. They are dependent on the model and its resolution, the geographical area, the weather regime and the density of the observation network *(Brousseau et al., 2011)*.

A representation with high accuracy of background-error statistics remains a major problem in data assimilation. The focus in the present work is on creation of background-error statistics for three-dimensional variational method (3D-Var) in limited area model. For the moment, multiple methods which can be used to compute the statistics exist. One method known as NMC (National Meteorological Center) consists in computing differences between forecasts started from successive analyses and valid at the same time *(Berre et al., 2006)*. Another technique uses the members of an ensemble of assimilation, in our case we will focus on so-called spin-up method, see *Brousseau et al., 2011*. We need to create limited area ensemble of assimilation for that we can take advantage of existing global ensemble of assimilation. Global ensemble members are downscaled to resolution of limited area model (LAM) and then 6h model forecast are run for each member to let the ensemble adapt to model resolution and its errors. The background errors are then simulated by differences between ensemble members as in *Berre et al., 2006*. Using this method, the analysis error is equal with the differences between different perturbed analyses of the ensemble *(Fisher, 2003)*.

The purpose of this stay was to prepare the necessary steps required for the setup of DA in Romanian operational suite. For this goal the following were needed:

- creation of lateral boundary conditions (LBC) from global assimilation ensemble;
- computation of the background-error covariances by spin-up ensemble method;
- fetching the observation data from OPLACE database;
- technical adaptation for experiments with CANARI and 3DVar;
- tuning of B matrix.

2 Operational settings

Starting from this year in Romania we have a new operational configuration based on cy43t2_bf10, where the horizontal resolution was increased at 4 km compared with the previous 6.5 km. The vertical resolution remains 60 vertical levels, the integration domain covers the Black Sea as you can observe in Figure 1. The lateral boundary conditions are from ARPEGE with a 3 hour frequency. The new configuration uses ALARO-1 vB physical package and fullpos inline with the following settings in NAMFPD: NLAT=373, NLON=501, RLONC(1)=29, RLATC(1)=44.5, RDELX(1)=0.05, RDELY(1)=0.035. It runs 4 times per day, with 78 hours forecasts provided at 00 and 12 UTC runs and 54 hours forecasts provided at 06 and 18 UTC with TSTEP=180. For the moment we run without data assimilation system. To implement a 3DVar assimilation system, one of the necessary steps is related to computation of the background-error statistics.

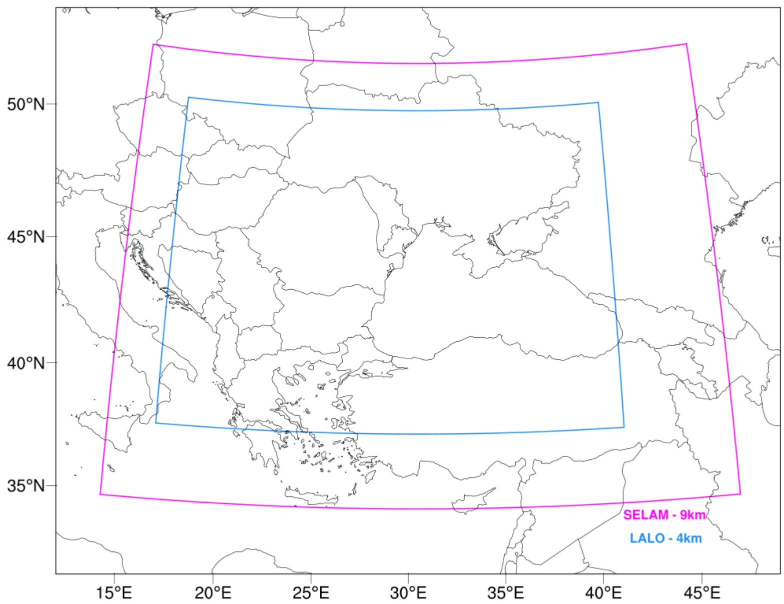

Figure 1: Representation of integration domain at 4km horizontal resolution

3 Computation procedures

3.1 Creation of the background-error statistics

The background-error statistics are essential in a 3DVar assimilation process. The method chosen taking into account our needs and our infrastructure was the spin-up ensemble method. For this purpose were chosen 30 days in two different seasons: summer (20210611 - 20210625) and winter (20220108 - 20220122).

The first step was to create the LBC from global assimilation ensemble and we use AEARP model configuration. AEARP (Assimilation d'Ensemble ARPège) is an ensemble formed by 50 members, without control member and used a 4D-Var and run 4 times per day. LBC were created in MF by running E927 configuration for the Romanian domain. First 6 AEARP members are used. The scripts used were provided by Antonín Bučánek and the settings were adapted for the Romanian SELAM (telecom) domain. After completing these steps on Meteo-France HPC, the files were transferred on our local HPC for the next processes.

The LAM operational configurations of EE927, DFI Initialisation and E001 was used for creation of 6 hours forecast of each member of AEARP. For the next procedures, the package for computation of B matrix for CY43t2_bf10 was downloaded from RC-LACE forum: www.rclace.eu/forum/Bmatrix and the steps described in the README file were followed.

Next step is to compute differences between members:

$$mem_2 - mem_1 \tag{1}$$

$$mem_4 - mem_3 \tag{2}$$

$$mem_6 - mem_5 \tag{3}$$

valid at 00, 06, 12 and 18 UTC for all days of selected period. For this procedure the configuration E001 have to be run on only 1 CPU hence the following parameters were modified in the namelist: LFEMARSD=.T. and LSPRT=.F., also taking care to suppress in-line fullpos (NFPOS=0) and computation of fluxes (NAMXFU - LXFU=.F.). For the chosen period a total of 360 differences were computed and the outputs were stored in grib files.

Next step is computation of B matrix by program festat using already created differences. Afterwards program fediacov was used to compute diagnostics. Being the first time of computing on this HPC, these binaries were compiled using gmkpack.6.6.6. The input variables mandatory for namelist were adapted in the script runall (in stat directory), and the lines which run the festat and fediacov binaries were adapted to our machine:

mpirun –mca btl openib,self,vader –report-bindings –display-map -np 1 FESTAT

mpirun –mca btl openib,self,vader –report-bindings –display-map -np 1 FEDIACOV

and the files stabfiltn360_2021061100.bal, stabfiltn360_2021061100.cv and stabfiltn360_2021061100.cvt were created.

3.2 Visualisation of B matrix diagnostic files

In the same package used for computing the Background Error Statistic is a dedicated directory (visu) which contains a set of KSH and Perl scripts for plotting the diagnostic files. They require some packages of perl, gfortran and gnuplot. Pictures with the B matrix diagnostic files are presented below.

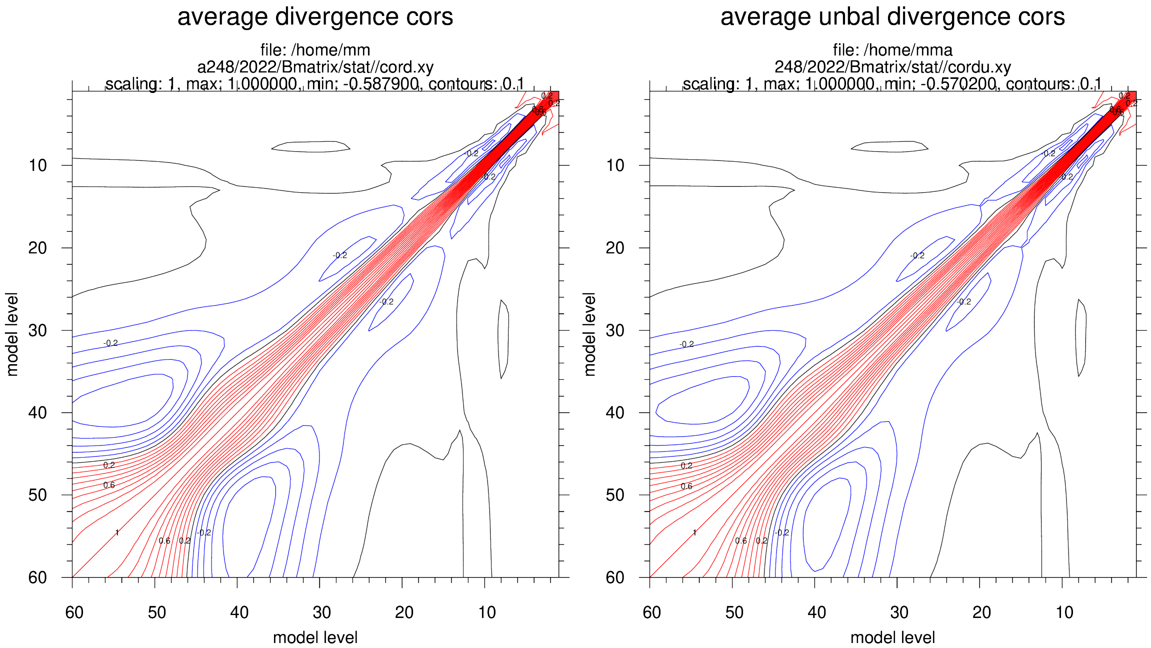

Figure 2: Mean vertical correlation of divergence

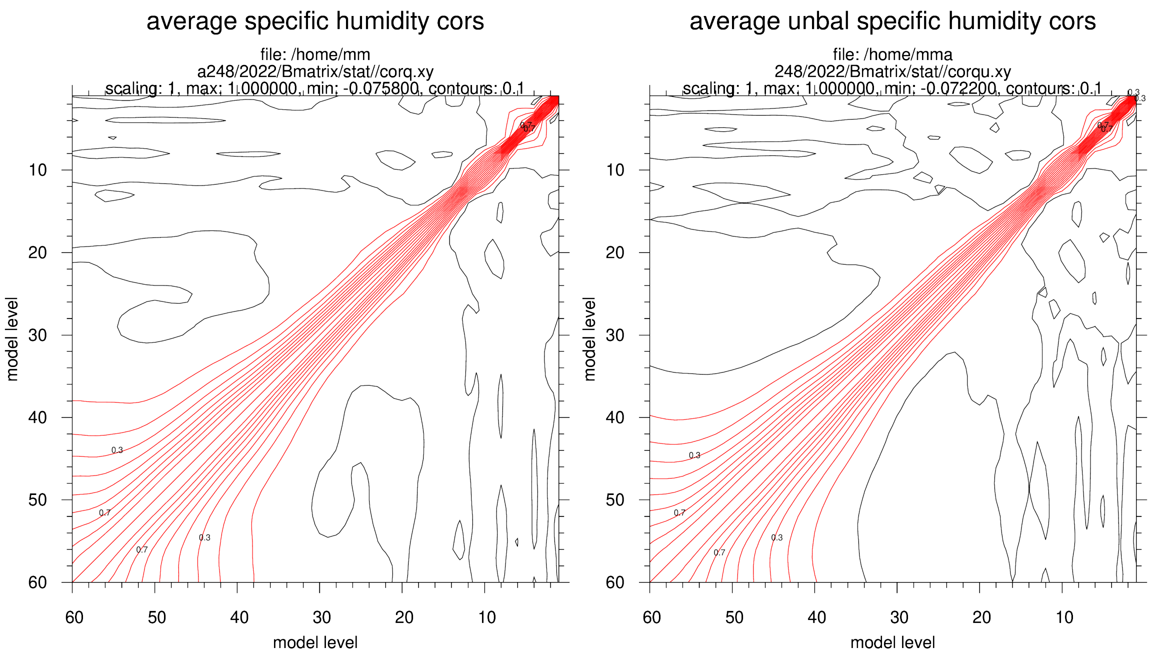

Figure 3: Mean vertical correlation of specific humidity

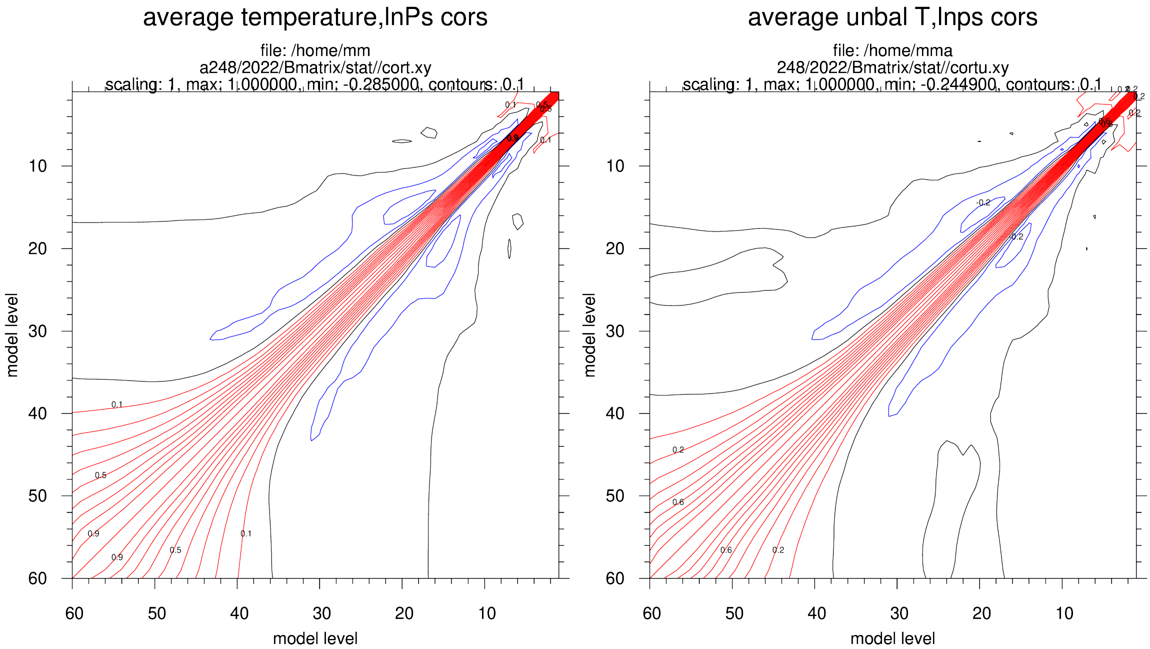

Figure 4: Mean vertical correlation of temperature

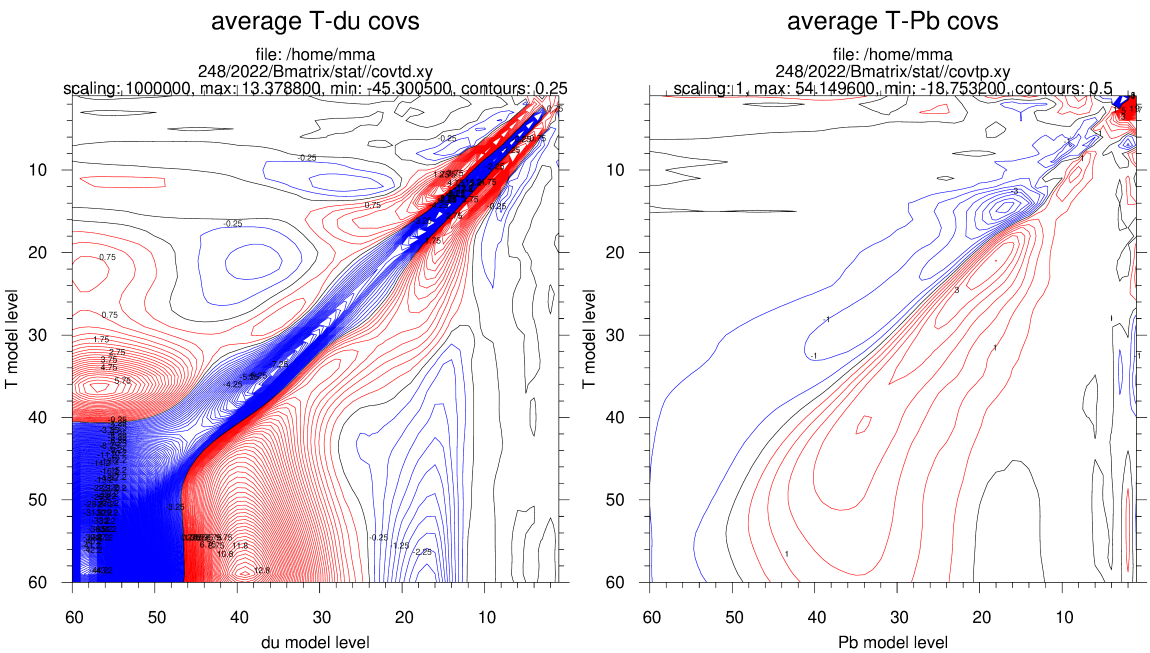

Figure 5: Mean vertical cross-covariance between temperature and unbalanced divergence (left) and vorticity-balanced

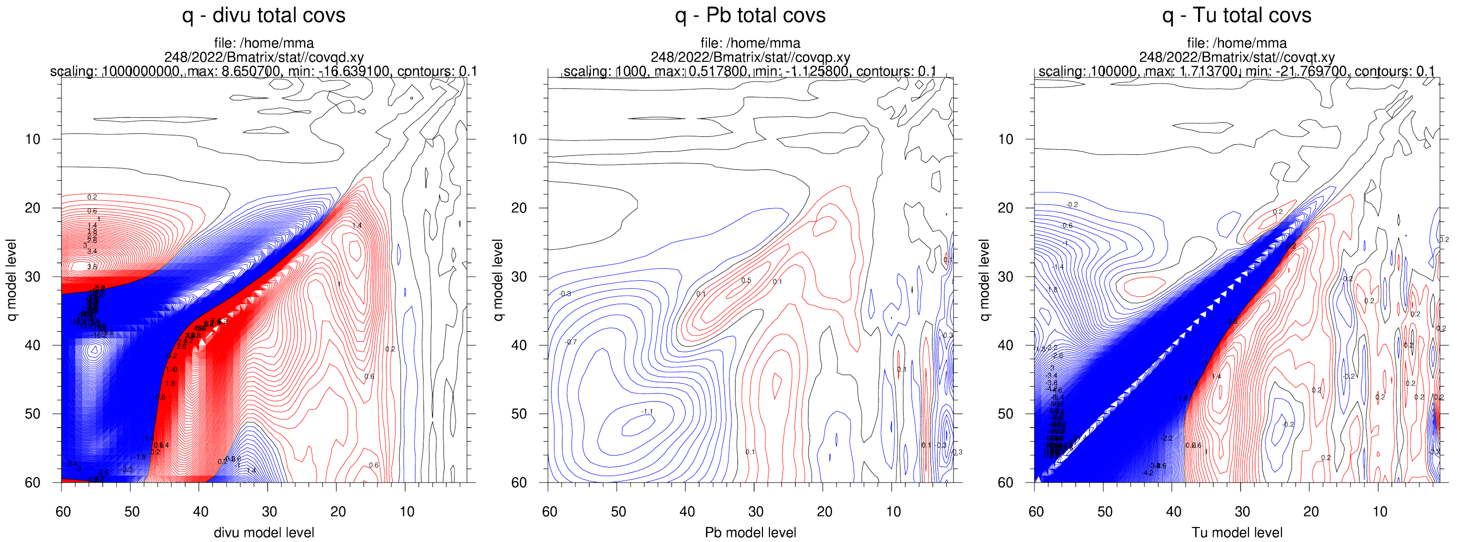

Figure 6: Mean vertical cross-covariance between specific humidity and unbalanced divergence (left), vorticity-balanced (middle), and unbalanced temperature (right)

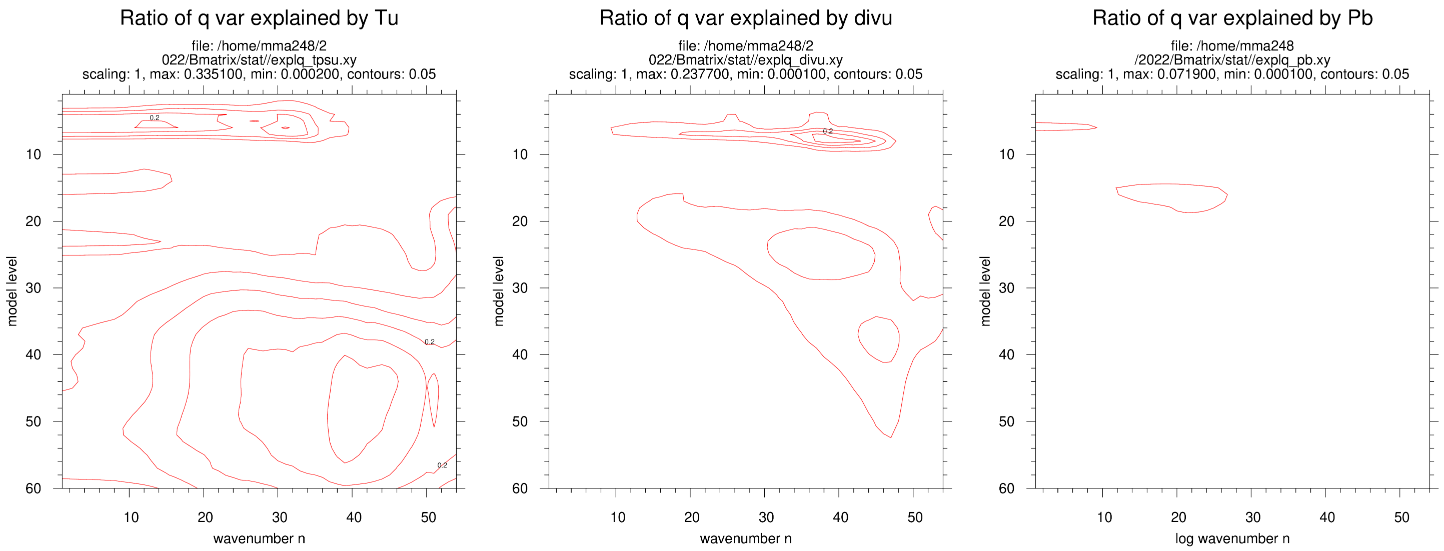

Figure 7: Ratios of specific humidity explained by unbalanced temperature (left), unbalanced divergence (middle) and surface pressure (right)

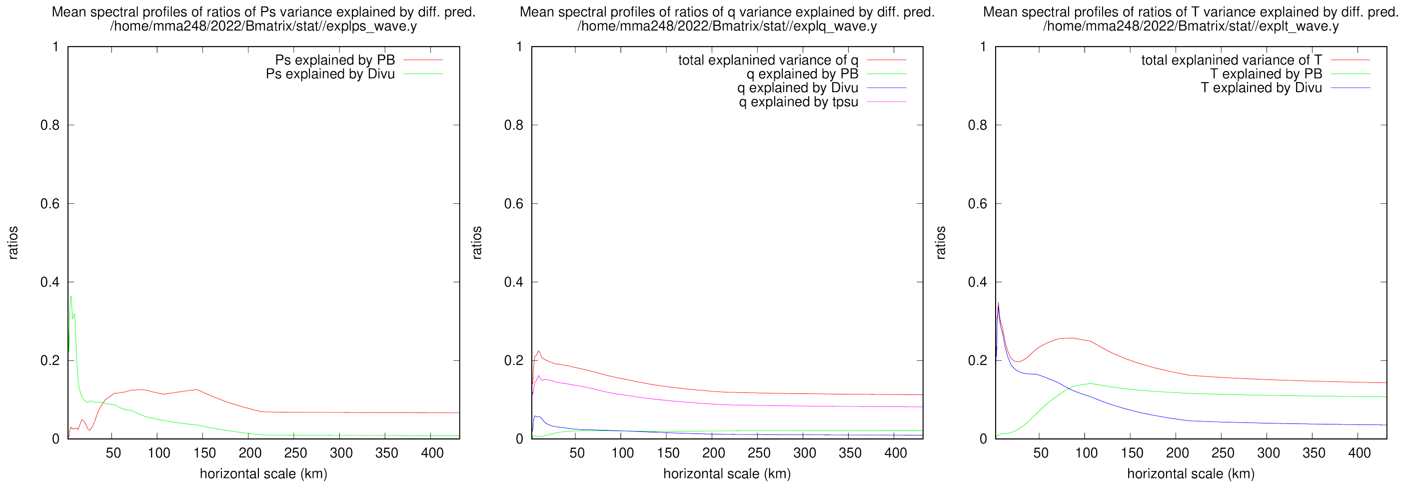

Figure 8: Ratios of explained variances of surface pressure/specific humidity/temperature as a function of horizontal scale and their predictors

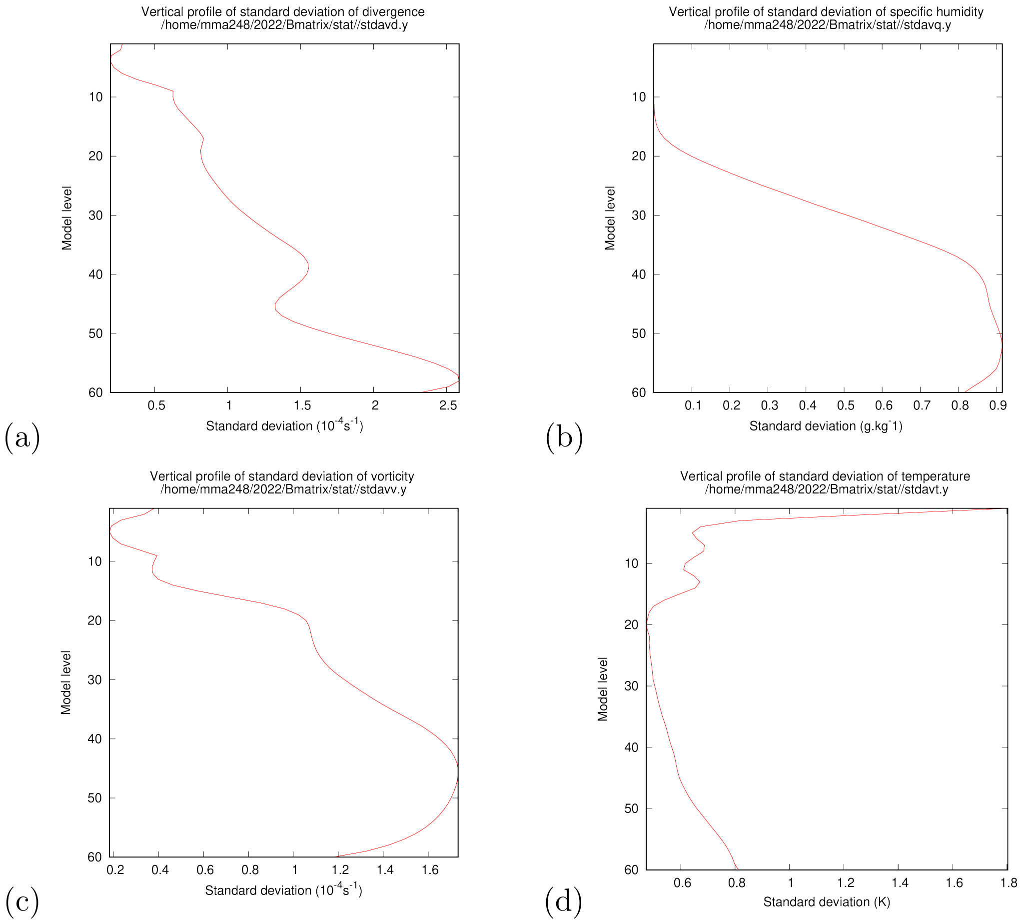

Figure 9: Vertical profile of standard deviation of divergence (a), specific humidity (b), vorticity (c) and temperature (d)

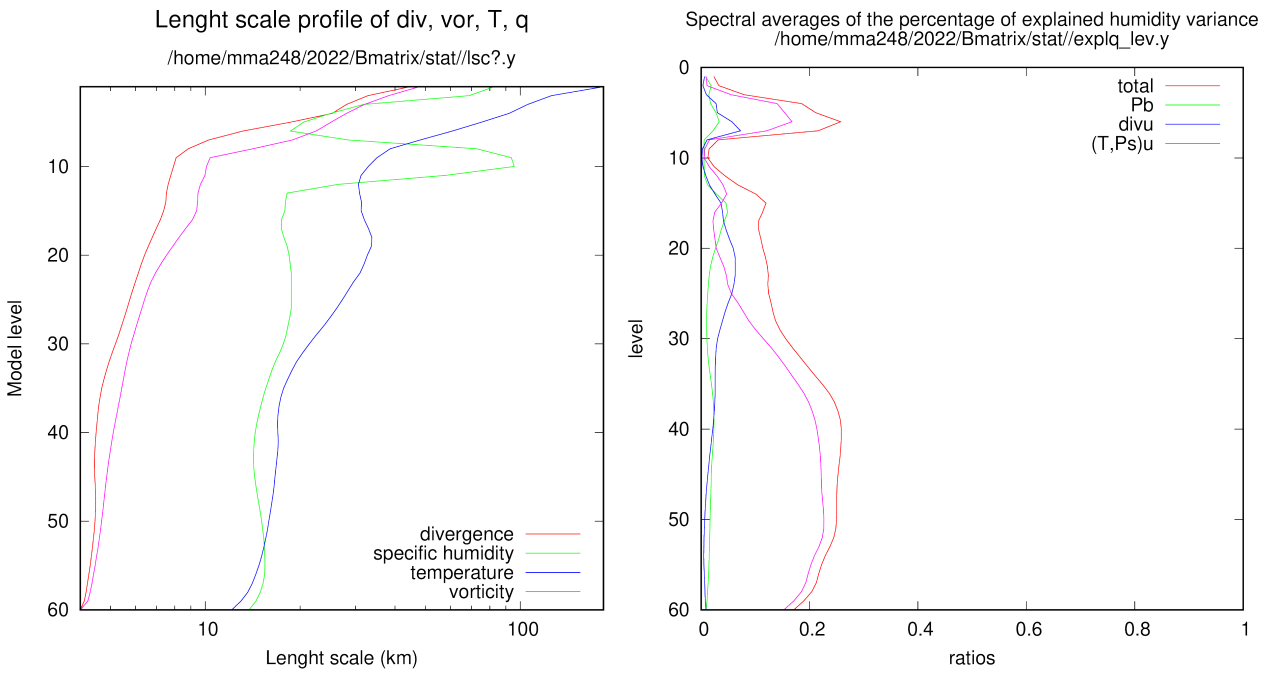

Figure 10: Lenght scale profile of divergence, vorticity, temperature and specific humidity (left) and spectral averages of the percentage of explained humidity variance (right)

3.3 Error diagnostic and tuning

The outputs obtained were transferred on CHMI HPC (kazi) for tests. Using the set of scripts the Czech team provided a set of tests was made to see the impact of assimilation of observations on weather forecast. For these tests the following observation type were used: synop, temp and amdar. A 10 days period between 20181101 - 20181110 was chosen. For the first day at 00 UTC the analysis from the previous day at 18 UTC was used, and taking into account it was a cold start, the reintegration with a set of parameters in namelist (NAMGFL) modified was necessary, in order to contain hydrometeors (see Appendix A). For the beginning the default value for REDNMC was used (REDNMC=0.7). Experiments which contain only CANARI or CANARI and 3DVAR (gradually introducing each type of observation) were performed. The results obtained for 10 days were used for *a posteriori* diagnostics proposed by *Desroziers et al., 2005* which suppose to represent the real standard deviations of observations and background errors in the data assimilation system chosen. For tuning of analysis the ratio between diagnosed values and the predefined ones is used:

$$r = \frac{\sigma_{diagnosed}}{\sigma_{predefined}} \tag{4}$$

The computation was done on kazi HPC and the package was provided by Antonín and can be found on RC-LACE forum: TuneBR. The results are showed in Figure 11 and it can be observed that using default value for REDNMC and SIGMAO_COEF are optimal according to TuneBR package.

Obstype:		all		1		2		5	
Var	Ratio_b	Ratio_o	Cases	Ratio_o	Cases	Ratio_o	Cases	Ratio_o	Cases
bt	0.000	0.000	0	0.000	0	0.000	0	0.000	0
g	0.000	0.482	13150	0.482	13150	0.000	0	0.000	0
ke	0.786	0.762	53593	0.000	0	0.738	36712	0.820	16881
q	1.662	0.725	9837	0.000	0	0.000	0	0.725	9837
t	1.575	0.875	53649	0.000	0	0.793	36816	1.030	16833
rav_pow	1.285	0.786	130229	0.482	13150	0.766	73528	0.888	43551
rav	1.222	0.777	130229	0.482	13150	0.765	73528	0.880	43551
rav_uv	1.085	0.773	130229	0.482	13150	0.756	73528	0.863	43551
rav_sigm	0.959	0.484	130229	0.482	13150	0.745	73528	0.855	43551

Figure 11: Results obtained with TuneBR tool

4 Conclusions

During the stay many results were obtained which facilitates the future implementation regarding data assimilation system in Romania. Having this amount of information the development of the system it is less demanding.

Next step which has to be done it's related to implementation of the surface data assimilation and testing for at least one month. The results obtained will be used for a objective verification and will be analysed what will need to be improved.

After the surface assimilation it's done and the results are improving the analysis the next step will be to start with 3DVar assimilation. The observations will be gradually introduced by their types in the experiments for a better understanding of the impact they have on the analysis and forecast.

Acknowledgement

I would like to thank to Antonín Bučánek and Alena Trojáková for their help and constant support before, during and after the stay. Also, I would like to thank to the NWP colleagues from Prague for their hospitality and kindness.

References

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

[2] Daley, R.; *Atmospheric Data Analysis. Cambridge Atmospheric and Space Science Series*; Cambridge University Press. ISBN 0-521-38215-7, 457 pages, 1991.

[3] Bouttier F., Courtier P.; *Data Assimilation Concepts and Methods March 1999.*

[4] Trojáková A.; *Progress report on 3DVAR testing at CHMI;* 2013., 19 pp available online https://www.rclace.eu/media/files/Data_Assimilation/2013/progress_2013.pdf

[5] Fisher M.; *Background error covariance modelling* Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, United Kingdom, ECMWF, 45–63; 2003.

[6] Stanesic A., Horvath K., Keresturi E.; *Comparison of NMC and Ensemble-Based Climatological Background-Error Covariances in an Operational Limited-Area Data Assimilation System* Atmosphere, 10, 570. https://doi.org/10.3390/atmos10100570, 570, 2019.

[7] Brousseau P., Berre L., Bouttier F., Desroziers G.; *Background-Error Covariances for a Convective-Scale Data-Assimilation System: AROME–France 3D-Var* Quarterly Journal of the Royal Meteorological Society. 137. 409 - 422. 10.1002/qj.750., 2011.

[8] Monteiro M., Berre L.; *A diagnostic study of time variations of regionally averaged background error covariances* Journal of Geophysical Research. 115. 10.1029/2010JD014095, 2010.

Appendix A: Experiments and data on kazi

Bmatrix plots
/home/mma248/2022/Bmatrix/plots

Bmatrix covariances
/home/mma248/2022/Bmatrix/stat

Directories with scripts and namelists for every experiment
/home/mma248/2022/Bmatrix/ro_canari
/home/mma248/2022/Bmatrix/ro_dyna
/home/mma248/2022/Bmatrix/rotest
/home/mma248/2022/Bmatrix/rotest_rednmc_sigmaocoef
/home/mma248/2022/Bmatrix/rotest_synop_temp_amdar

Rundir directories
/work/mma248/2022/exp/ro_canari
/work/mma248/2022/exp/ro_dyna
/work/mma248/2022/exp/rotest
/work/mma248/2022/exp/rotest_rednmc_sigmaocoef
/work/mma248/2022/exp/rotest_synop_temp_amdar

Results from a posteriori diagnostic
/home/mma248/2022/Bmatrix/TuneBR.06_zi94_rotest_rednmc_sigmaocoef

Changes made in namelist in order to contain hydrometeors (NAMGFL)

Normal	Cold Start
YDAL_NL%LREQOUT=.FALSE.	YDAL_NL%LREQOUT=.TRUE.
YDOM_NL%LREQOUT=.FALSE.	YDOM_NL%LREQOUT=.TRUE.
YI_NL%LREQOUT=.FALSE.	YI_NL%LREQOUT=.TRUE.
YL_NL%LREQOUT=.FALSE.	YL_NL%LREQOUT=.TRUE.
YR_NL%LREQOUT=.FALSE.	YR_NL%LREQOUT=.TRUE.
YS_NL%LREQOUT=.FALSE.	YS_NL%LREQOUT=.TRUE.
YTKE_NL%LREQOUT=.FALSE.	YTKE_NL%LREQOUT=.TRUE.
YTKE_NL%NREQIN=-1	YTKE_NL%NREQIN=0
YTTE_NL%LREQOUT=.FALSE.	YTTE_NL%LREQOUT=.TRUE.
YUAL_NL%LREQOUT=.FALSE.	YUAL_NL%LREQOUT=.TRUE.
YUEN_NL%LREQOUT=.FALSE.	YUEN_NL%LREQOUT=.TRUE.
YUNEBH_NL%LREQOUT=.FALSE.	YUNEBH_NL%LREQOUT=.TRUE.
YUOM_NL%LREQOUT=.FALSE.	YUOM_NL%LREQOUT=.TRUE.