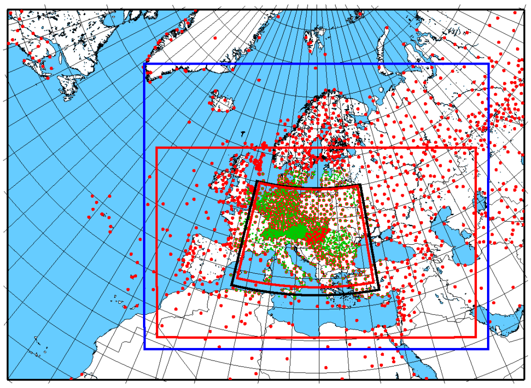
Regional Cooperation for
Limited Area Modeling in Central Europe

IC perturbations by 3DVAR in ALADIN-LAEF



report from RC LACE stay in Vienna, 25/04/ - 20/05/2016

Alena Trojáková, CHMI, alena.trojakova@chmi.cz
in collaboration with Martin Belluš, SHMU, martin.bellus@shmu.sk

Contents

1 Introduction

The regional ensemble prediction system ALADIN-LAEF has been developed within RC LACE to provide a reliable short-range probabilistic forecast. The ALADIN-LAEF uses the spectral Breeding-Blending technique for the upper-air initial condition (IC) perturbations. We aim to further extend the methodology by Ensemble of upper-air Data assimilation (EDA) using the 3DVAR technique to capture IC uncertainty.

This report summarizes the work done during the one month LACE stay dedicated to an implementation of technical means for the ALADIN-LAEF EDA. Following sections give a summary of tackled items. Section 2 describes an estimation of background errors. Section 3 sums up the technical implementation of the 3DVAR. A summary and future plans are provided in the last section.

2 Background error statistics

The background error statistics (further referred to as **B** matrix) are fundamental component of the 3DVAR. Firstly the programs and scripts for **B** matrix computation were validated at ECMWF for cy40t1. Technical details can be found in Appendix A.

The **B** matrix was computed by ensemble approach following Berre (2000) for one month period. The period was constrained by available ALADIN-LAEF LBCs to 15 May - 15 June 2011 12 UTC runs only. Altogether 256 ALARO downscaled 12H forecast differences from 16 members ($m_i - m_{i+1}, i = 1, 3, ..., 15$) were used for **B** matrix sampling. The 12H differences were considered as ALADIN-LAEF uses 12H production forecast as the first guess for production analysis.

Typical diagnostics were briefly cross-checked with respect to the operational background errors used at ALADIN/CHMI, derived from 6H downscaling of 4 members of Assimilation Ensemble of global model ARPEGE (AEARP) to resolution 4.7km for period of February - May 2011. Considering the different setting of both matrices (e.g. driving global ensemble, forecast length of the differences, vertical resolution, etc), the aim is not to compare the two **B** matrices in details, but to qualitatively check the new **B** matrix for the ALADIN-LAEF system.

The ALADIN-LAEF standard deviations, which correspond to the expected amplitude of background errors, are similar to CHMI ones, see Figure 1. A large decrease of temperature standard deviations close to the model top is a bit unexpected, but haven't been investigated further.

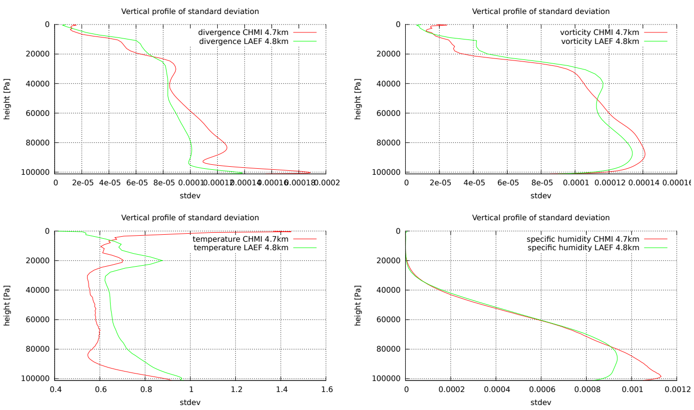


Figure 1: Vertical profile of the standard deviation for divergence, vorticity, temperature and specific humidity for ALADIN/CHMI in red and ALADIN-LAEF 5km in green color.

Spatial propagation of increments is determined by length-scales, which are comparable for both matrices, see Figure 2. Also physical coupling between variables (cross-covariances between different variables) showed similar structure, see Figure 3. Only the scaling is 6 orders smaller for ALADIN-LAEF **B** matrix derived on cy40t1, which is a feature of cy40t1.

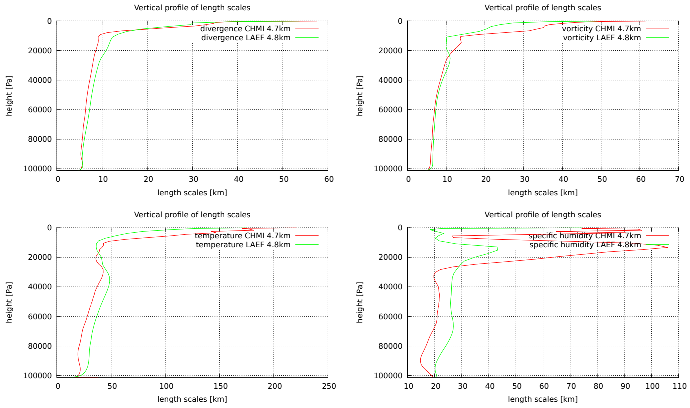


Figure 2: Vertical profile of the length-scales for divergence, vorticity, temperature and specific humidity for ALADIN/CHMI in red and ALADIN-LAEF 5km in green color.

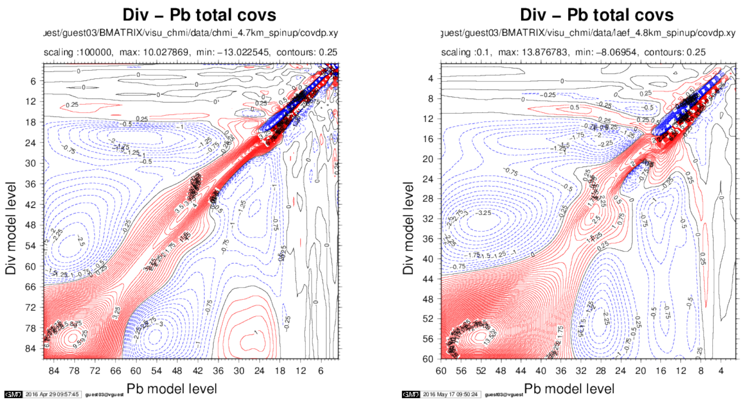


Figure 3: Mean vertical cross-covariance between divergence and vorticity-balanced for ALADIN/CHMI on the left and ALADIN-LAEF 5km on the right.

3 3DVAR implementation

Next step of the 3DVAR implementation was a technical validation of its main components on cy40t1 and a drafting of its scripts. Given the constrains (e.g. length of the stay, the new domain on 5km not yet tested, etc) the 3DVAR implementation was limited to conventional observations (SYNOP, TEMP and AMDAR) and testing started on 11km domain.

The conversion of observation to the ODB (BATOR and SHUFFLE) was tested at first, then surface analysis was technically tried on single and more CPUs. A performance issues were encountered, such as surprisingly long execution of create_ioassign (2–3 minutes), which take only a few seconds on other platforms.

New components of the 3DVAR were drafted in ALADIN-LAEF scripts, which are written in PERL. Finally, Martin Belluš cleaned the draft and extended fetching of observation for AMDAR and TEMP data. Moreover, simple 3DVAR scripts were installed and tested, which can be useful for a running of a simple 3DVAR test on other platforms, e.g. in Toulouse or Prague. Technical details can be found in Appendix B.

4 Summary and future plans

Previous sections provided more details on items tackled during the stay and here follows a brief summary and an outlook for the future work.

The 3DVAR was technically tested and added to the research ALADIN-LAEF environment. More detailed scientific validations and tunings of the 3DVAR is essential for a further progress. Following areas of the future work are suggested:

- B matrix modeling, e.g.

  - recompute B matrix using Blending and/or BlendVAR 12H forecasts to better represent analysis and background errors,

  - considering that ALADIN-LAEF provides downscaling of several EPS members implicitly, a flow-dependent B matrix can be recomputed regularly (e.g. every couple of weeks) with very little costs. Another interesting possibility is to "mix/weight" the background differences sample from a set of meteorologically interesting situations from all seasons and the recent differences of the day. Although the flow-dependent aspect is very important for data assimilation, a practical impact on ALADIN-LAEF performance has to be still investigated.

- tuning and scientific evaluation of the 3DVAR performance, e.g.

  - a revision of namelists and blacklists,

  - an evaluation of the 3DVAR impact to preclude errors in the first implementation,

  - a tuning of observation and background errors following Desroziers *et al.* (2005).

- enhancement of quality of analysis (initial conditions), e.g.

  - by data assimilation of non-conventional observations, e.g. satellite and GPS data. An elimination of systematic biases is key aspect and it has to be implemented,

  - via implementation of an assimilation cycle using the non-lagged coupling, which entails an extension of the number of assimilated observations and a better quality of LBCs for the first guess computation. Eventually, an increase of the assimilation cycle frequency from 12H to 6H or even 3H can be considered. Enhanced quality of analysis and the space-consistent coupling could, at least partially, compensate the use of the lagged coupling in the production.

The list of suggestions is far from complete and it was seen mainly from my deterministic data assimilation point of view. The stay involved mainly technical issues, such as the porting of cy40t1 and the 3DVAR implementation, and scientific challenges and benefits of a closer link of the data assimilation and EPS are still to be explored.

Acknowledgment
The author wishes to thank Martin Belluš for his kind cooperation and all NWP group for their warm welcome and hospitality. This stay was made possible by the financial support of the LACE (Limited Area model for Central Europe) project.

References

Berre L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly Weather Review* **128**(3): 644–667.

Derkova M. 2016. ACNA CY40T1_bf06. Technical report, ALADIN *available online https://docs.google.com/document/d/1Y3rjAlo8d-u_Njl4dHNky8iAScBny9ywp6WjDYvkpM8/edit?pref=2&pli=1*.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society* **131**(613): 3385–3396.

A B matrix computation and diagnostics

This appendix contains technical details of **B** matrix computation and diagnostics visualization.

The **B matrix computations** were performed at ECMWF server **cca**. It consists of following steps:

- computation of forecast differences,
- **B** matrix calculation and diagnostics derivation.

The forecast differences were computed using an adapted forecast model configuration (e001). The model forecasts to be substracted are provided as input under following names: ICMSH$EXPFGIN, ICMSH$EXPANIN. The namelist switch LFEMARSD=.T. ensure writing the differences of prognostic variables into an output file (by default in GRIB format). Furthermore, it was necessary to suppress in-line fullpos and computation of fluxes. It is also essential to set LSPRT=.F. to compute the differences of temperature. The **B** matrix and its diagnostics are computed using a dedicated programs **FESTAT** and **FEDIACOV**, which are included in an export model package using the gmkpack **cca:/home/ms/at/kah/bellus/HOMEPACK/40t1_bf05_assim**.

The **B** matrix was produced for the operational ALADIN-LAEF_11km and a new ALADIN-LAEF_5km domain. Forecasts from the reference experiment for cy40t1 with only dynamical adaptation for all ALADIN-LAEF members over the period of 32 days (20110515 - 20110615, 12UTC only) were used. Experiments were based on already existing ALADIN-LAEF scripts and new components were drafted. The drafts were further cleaned and elaborated by Martin Belluš for a pre-operational implementation. All experiments are stored on **cca** and main results were archived on **ECFS** or **$PERM** file system. Here follows a summary of performed **B** matrix tests:

- CY40_BMATRIX **cca:/home/ms/at/kah/trojakova/exp/**

  - draft of **B** computation on 11km

  - firstly the **B** matrix was computed from 6H differences by accident, later 12H differences were considered - no scientific intercomparison of both matrices was done

  - results archived manually to **ec:/kah/trojakova/LAEF/CY40_BMATRIX/**
- CY40_BMATRIX5 **cca:/home/ms/at/kah/trojakova/exp/**

  - draft of **B** computation on 5km

  - a problem was encountered on 5km to compute the differences (*Error : too many words, enlarge array IGRIB.*) and the MASTERODB had to be recompiled with increased JPDIM in arpifs/var/grbspa_mf.F90. Furthermore, FESTAT seems to be very memory consuming. We had to significantly reduce number of forecast differences (to 40) or to use 8 upgraded nodes (128GB per node) on **ccb** to compute the **B** matrix for the full set of 256 differences.

  - results archived manually to **ec:/kah/trojakova/LAEF/CY40_BMATRIX5/**

domain	FC length	name in $PERM/mbell/const/B	experiment
ALADIN-LAEF_5km	12H	stabfiltn32_2011_5km	exp/CY40_BMATRIX5
ALADIN-LAEF_5km	12H	stabfiltn5_2011_5km	exp/CY40_BMATRIX5
ALADIN-LAEF_11km	12H	stabfiltn32_2011_11km	exp/CY40_BMATRIX
ALADIN-LAEF_11km	06H	stabfiltn32_2011_11km_6H	exp/CY40_BMATRIX

Table 1: Summary of derived B matrices.

Simple tools for visualization of **B matrix diagnostics** were installed on ZAMG server **vguest**, see the section 2 for an illustration, here follows a brief summary.

```
[guest03@vguest:~/BMATRIX]$
# standard deviation and length scales for CHMI and LAEF5km
std/README
std/data                # input data
std/plot_Jb_stats.pl    # plotting script
std/CHMI_LAEF5          # output

# further diagnostics
visu_chmu/README
visu_chmu/data             # input & output data
visu_chmu/rungmt4tex_chmu  # plotting scripts
visu_chmu/rungmt4tex_laef
visu_chmu/rungmt4tex_laef_4.8
visu_chmu/rgnuplot
visu_chmu/gmt.ksh.cross_cov
visu_chmu/gnuplot2ps
visu_chmu/chmu_laef2.tex
visu_chmu/chmu_laef2.pdf   # a selection of diagnostics
visu_chmu/log*             # logs
```

B Technical experiments

This appendix contains technical details of the 3DVAR implementation.

The implementation was performed at ECMWF server **cca**. The export version of cy40t1_bf05 compiled with Intel compiler was used. The user pack with a final set of modifications contains

- known bug fixies gathered by ALADIN ACNA (Derkova, 2016) ,
- modification for B matrix differences calculation on 5km (*arpifs/var/grbspa_mf.F90*)

The user pack can be found on **cca:/home/ms/at/kah/bellus/HOMEPACK/40t1_bf05_assim**.

Various tests were performed and all experiments are stored on **cca:/home/ms/at/kah/trojakova**. Considering that tests were mostly technical **no outputs were stored permanently**! Here follows a summary of performed tests:

- CANASPPT

  - reference experiment for CANARI on cy36 with suppressed SPPT and obs perturbation;

- CY40_TEST1 - **CANARI works TECHNICALLY on cy40t1**

  - technical test for CANARI on cy40t1_bf05 (suppressed SPPT and obs perturbation);

  - PERFORMANCE: the execution of create_ioassign takes 2-3minutes, while at Toulouse (BULL) and Prague (NEC SX) it takes just a couple of seconds!

- CY40_TEST2

  - technical test for CANARI on cy40t1_bf05 for more CPUs

  - PERFORMANCE: preliminary results does not show a clear benefit of using more CPUs, detailed performance analysis might be needed;

- CY40_TEST3 - **3DVAR works TECHNICALLY on cy40t1 on single CPU**

  - technical implementation of 3DVAR on cy40t1_bf05 (suppressed obs perturbation);

- CY40_TEST4

  - 3DVAR testing on more CPUs and trials to fetch/use observations

- CY40LAEF5BVAR

  - the first merge of 3DVAR and Blending (suppressed obs perturbation);

- CY40_L5BVAR - **BlendVAR works TECHNICALLY on cy40t1**

  - the first ALADIN-LAEF BlendVAR test including cycling (suppressed obs perturbation);

  - BUGGED CYCLING: verif scores for 4 days showed identical results! The blending analysis was used insted of BlendVAR one.

- CY40_L5BVAR_01 - BlendVAR evaluation proposal - NOT RUN YET

  - based on CY40_L5BVAR

  - namelist adapted to follow CHMI operational setting

  - cycling: BlendVAR analysis used by the forecast

  - TODO: extend assimilation window for AMDAR to +/- 90min, the adaptation of obsoul_merge.pl needed and please check carefully the cycling !

- sample_3dvar

  - set of simple KSH scripts suitable for testing on other platforms;

  - uses input from CHMI - BEWARE: B matrix, guess and BATOR namelist are domain specific !

  - big data and constants are archived on ec:/kah/trojakova/sample_3dvar/

  - not ready for satellite data;

C ODBVIEWER

A stand-alone ODB package was installed and tested on ZAMG server **vguest**. It allows to run a viewer to quickly check an ODB content. Compilation of tools using BUFR libraries was suppressed due to difficulties with their compilation. The core tool (odbsql) was compiled and tested successfully. Further details and basic usage information can be found in the corresponding README files.

```
[guest03@vguest:]$
#
ODB/odb_CY33R1.007.tar.gz    # original package
ODB/odb_CY33R1.007           # compilation directory
ODB/README.installation      #
ODB/README.usage
ODB/example                  # test data

ODBVIEWER                    # installation directory
```

Following setting have to be exported (or set to a user .bash_profile).

```
# ODBVIEWER
export ARCH=linuxgfortran
export CPU_TYPE=Intel
export OS_VERSION=GNU_Linux
export OBJECT_MODE=x86_64
export ODB_DIR=/home/guest/guest03/ODBVIEWER
export ODB_VERSION=CY33R1.007
alias use_odb=". $ODB_DIR/$ARCH/$CPU_TYPE/$OS_VERSION/$OBJECT_MODE/$ODB_VERSION/bin/use_odb.sh"
```