Summary report of the
2020 Joint DAWD and DAsKIT working days,
Online meeting/Vienna 14-16 September

Benedikt Strajnar, Maria Monteiro

In 2020, the Joint DAWD and DAsKIT Working Days were planned similarly to those of last two years: during 2,5 days, the experts on different aspects of DA from the two groups could exchange their expertise in specific topics boosting the progress of local DA systems. Due to the actual pandemic situation the event was held as a visio-meeting which was organized by ZAMG, using the S4B facility.

The meeting started with a welcome message from Benedikt Strajnar (RC LACE Data Assimilation Area Leader) and Maria Monteiro (ALADIN Data Assimilation Core Programme, so called DAsKIT) who expressed their gratification to join together so many experts from the RC LACE DA and DAsKIT groups which deal in the actual moment with crucial topics for LAM DA like radar and humidity data assimilation. The Austrian team (Florian Meier, Phillip Scheffkneckt, Stefan Schneider) was acknowledged for technically hosting the meeting, and also the invited experts from HIRLAM (Roger Randriamampianina) and Météo-France (Claude Fischer) for their participation. Totally, the meeting gathered around 35 participants.

Like in previous years, the meeting consisted in three different parts, split in shared and separated sessions: the national status reports; the dedicated topics reports, which allowed the discussion on issues and further progress; and on practical sessions (DAsKIT). In particular, the practical session this year was dedicated to the Scalable Acquisition and Pre-Processing system (SAPP) data ingestion by BATOR; and to the OBSMON tool recompilation with gmkpack, which has to be adapted to allow ALADIN DA observations monitoring. Special attention was also given to the HIRLAM DA systems status report, to the OPLACE status report and to the exchange of information with Météo-France (the latest under the form of a list of questions and answers). The next sections of this report give more details on the different outcomes of the meeting.

Summary of status talks (Day 1)

National presentations are available from http://www.umr-cnrm.fr/aladin/spip.php?article364 and http://www.rclace.eu/?page=11#workshops. Below is a summary of main presented points per country:

Slovenia
- Developed a prototype of 1.3 km NWCRUC.
- EDA experiment (daily and monthly covariances), mixed impact with tuning needed.
- Evaluated 3 methods for radar dealiasing.
- Validated OSCAT observations (operational) and EMADDC test2 dataset.

Slovakia
- Upgrade to cy43t2 ongoing, validations done for Blending and BlendVar, blacklisting of coastal stations (LSREJ) for CANARI.
- Compared locally received Mode-S MRAR and EHS.
- Validated radar Doppler winds at high Nyquist velocity.

- Developed a QC system for surface data based on A-LAEF system.
- Continued work on implementing/phasing of STD assimilation in higher model cycles.

Romania
- Explained lack of manpower in DA, plans to increase involvement next year.

Hungary
- Upgrade to cy43t2 ongoing, plans to test longer periods.
- AROME-EPS operational from February (no DA).
- Experiment with hourly RUC confirmed that less frequent OI (3h) for surface performed better.
- Ongoing efforts to update Observation Monitor software.
- Performed a recalculation of EDA B-matrix for 90 level setup, tuning needed.
- Validated Czech and Hungarian Mode-S MRAR, positive impact of CZ data over winter period.
- Tested impact of GEOWIND and HRWIND, neutral scores, plans to revise blacklisting.

Czech Republic
- New B-matrix put to operations (EDA) after tuning and tests.
- Inter compared correlation functions for OI (CANARI) for national Synop reports, new correlation function at 40 km length scale outperformed the Mescan setting.
- Operational use of observations extended by wind profilers, HRWIND, ASCAT.
- Extensive work on radar reflectivity (validation of obs. operator and screening) and wind (stays)

Croatia
- Backup of operations at ECMWF (no DA).
- Tested converted from COST to bufr format for ZTD observations.
- Compared different tuning for OI (CANARI), neutral impact.

Austria
- Adaptations in AROME-RUC (increased observational use), several applications of IAU studied. Poor-man ensembles for forecasters are planned.
- A number of wind profiles (no VAD from radar stations) and sodars (wind) added to operations. A fix for height/pressure attribution based on FG proposed for wind profilers.
- A data set of Netatmo surface data to be evaluated soon.
- Worked on resolving a bug in DFS for Doppler winds.
- Studied differences between cy40 and cy43 in reflectivity processing in BATOR, in cy43t2 less observations are used due to applied quality threshold.

Algeria
- CY43T2 is in operations (dynamical adaptation).
- 3D-Var (SYNOP, TEMP and AMDAR) at CY43T2 has been cycled for both ALADIN and AROME: preliminary validation has shown added value on screen level parameters, except for 2-m temperature in opposition to what has been found when running an AROME surface DA. A combined DA solution may be the solution.
- Further progress will include the assimilation of GPS data and with SEVIRI (twin project with MF and FMI).

Belgium
- AROME operational setup with surface DA (1.3km, 3-h cycling, hourly ALARO coupling, local SYNOP data) from June 2020 at CY43T2.
- Development of a combined CANARI + 3D-Var solution for ALARO (SYNOP, TEMP and AMDAR): 2-m temperature RMSE decreases but 2-m relative humidity RMSE

decreases only during the night period (in comparison with dynamical adaptation), besides, vertical mean analysis increments are generally smaller than innovations, except above 300 hPa, for temperature and at surface, for the wind (not shown for other parameters).

- Development of a combined OI_MAIN+3D-Var solution for AROME (SYNOP, TEMP and AMDAR, SHIP, GPS-SOL): GNSS assimilation seems to have a neutral impact, except for 6-h precipitation.
- Local odb is now read by OBSMON.
- Future plans include: further test of both combined solutions, testing of VarBC for GNSS assimilation, improvement of scripts (project IMA): use of ODIM radar data and MODE-S.

Bulgaria

- CY43T2 is in operations (dynamical adaptation) and CY43T2_bf10 has been successfully compiled using Intel Parallel Studio 2020 (milen.tsankov@meteo.bg).
- Local implementation of SAPP has started.
- A new machine will arrive in October 2020.

Morocco

- A new HPC infra-structure has arrived and CY41T1 has been installed.
- 3D-Var has been tested on the new HPC machine.
- Plans include: porting of surfDA, development of a combined OI_MAIN + 3D-Var DA solution, implementation of CY43, moving to 1.3km resolution.

Poland

- CY43T2 is implemented in operations (dynamical adaptation).
- CANARI (ALARO, CY43T2_bf10) is in preparation.

Portugal

- SurfDA is cycled in operations (CY38T2) for diagnostic purposes.
- Local NWP systems are been ported to ECMWF, tested and validated at CY43T2: dynamical adaptation, surface DA, combined (OI_MAIN+3D-Var) DA. Besides local ARPEGE couplings are used; local observations under GTS BUFR format, local OIFS HDF5 radar observations.
- The "createsuites" ecFlow suite generation from ARSO has been successfully adapted at ECMWF.
- Portuguese and Spanish OIFS HDF5 radar data has shown a slight positive impact on short-range forecasts over the Iberian Peninsula during a 20-day winter period in a hybrid 42/43 cycle experiment (beaufix).
- SAPP is being implemented
- OBSMON v3.3.2 (Shiny part) has been locally implemented

Tunisia

- CY43T2 has been installed on the new HPC infra-structure.
- 3D-Var with Jk blending is being implemented for AROME (CY43T2).
- Changing grid point humidity in AROME to spectral.
- Pre-processing tool POP-RMI (BUFR-GTS toolbox) has been implemented.
- New challenges include the assimilation of new observations: AMDAR, local GNSS.

Turkey

- CY43T2_bf10/11 have been installed.
- BATOR is now being in use with locally produced SAPP BUFR data.
- Scripts for pre-operational AROME surface DA are ready (with IFS LBC).
- Future plans include: start pre-operational cycling of AROME surface DA, monitoring of observations; preparation for 3-h upper-air assimilation RUC.

OPLACE system
- Added ATOVS and IASI and ASCAT from Metop-C, performance issues for AMSU-A1 (ch.3) and MHS (chs. 2,3).
- Redesign of the OPLACE scripts which enable parallelization increase robustness and ease monitoring.
- A new IASI selection method adds more IASI data to the dataset.
- Mode-S EHS from EMADDC test stream significantly increase number of observations, quality issues with winds over some regions.

Overview of HIRLAM DA activities
- Most of HIRLAM countries currently run OI_MAIN + 3D-Var systems based on cy40.
- A QC based on similarity (Machine Learning) index for surface stations is being developed
- Norwegian radar department develops/uses methodology to avoid aliasing in wind measurement using several PRF mode.
- 15 min ZTD observations are being tested.
- M. Mile tests supermodding with scatterometer observations, with good progress.
- HIRLAM develops field alignment technique, to be included in their cy43.
- The main focus is on operationalization of 4D-Var.
- To simulate background errors in case of little observations, EDA Brand technique is being developed and should enter the common code, however requires a large number of ensemble members.

Remote teleconference with MF

Claude Fischer joined the discussion session at the end of Day 1 and provided questions. First of all, the issues with external access to MF were discussed and feedback (who can connect to MF via ftp/telnet) was provided. The following questions by LACE and DAsKIT colleagues were then discussed:
- Q: What is the progress of OOPS and how LACE can familiarize with it? A: MF tests 4D-EnVar configurations in research environment, currently not fully aligned with ECMWF. To get familiar with OOPS: 1) through contribution of phasing and using Davai unit testing tool, 2) through dedicated LACE manpower for a specific scientific problem addressed through OOPS-ready codes in cooperation with OOPS development team at MF. One of the options mentioned was also preparation of a sample 3D-EnVar configuration suitable for testing outside MF and/or during DA code training week (currently planned for 2021).
- Q: Where the namelist for cy46 can be obtained (related also to STD implementation by M. Imrišek)? A: answered that the build of 46t1_bf05 is ongoing. P. Brousseau was suggested to be contacted for recent namelists.
- Q: Where to find codes for VarBC correction of raw Mode-S observations? A: Some colleagues are already in contacts with V. Pouret, who will provide related codes and information.
- Q: Is OIFS OPERA data hub really migrating elsewhere? A: Not for the moment, it now seems likely that the service will remain at MF.

Summary of thematic talks (Day 2)

Session on assimilation of humidity information

In the first talk by Fatima Hdidou the Moroccan GNSS network, developed by Maroc Meteo for meteorological purposes, was described. This network is complemented by IGS (International GNSS station Service) stations. A good quality of measurements was demonstrated in comparison with radiosondes and NWP forecast. Impact study was carried out with AROME-Morocco 3-hourly 3D-Var at 2.5 km resolution. Site–specific biased were removed before assimilation using a static method, based on FG departures. It was observed that GNSS ZTD data influence the lower troposphere below 500hPa. A small positive impact on 3h forecast with respect to the radiosondes was detected for specific humidity standard deviation and additionally bias of RH 2m was improved up to 24h into the forecast. ZTD observation influence precipitation scores (positively for high amounts) and it was shown on a case study that it improved distribution and intensity of precipitation.

The second talk by Benedikt Strajnar described the improved one month dataset of GNSS ZTD observations over Slovenia (national provider). In this case, Var-BC was used to detect and correct biases, using 3 predictors (default in cy43t2). It was shown that the convergence of VarBC was not over, and remaining bias contributed to degraded quality of forecast. The scores could be somewhat improved after an additional month of repeated VarBC cycling. It was also shown that ZTD has substantial impact on precipitation on given cases. Testing on the larger dataset is necessary, other options are static bias correction or more frequent (3h) cycling of VarBC (this was recommended by HIRLAM experience).

The session was concluded with a very interesting talk by Zahra Sahlaoui about a novel method to assimilate rain rate observations. For a case of heavy Mediterranean precipitation event over southern France, a 1D-Var+3DVar approach was used to assimilate precipitation analysis Antilope (based on radar and raingauges), and simplified physics packages (large-scale condensation and convection) were used as an external observations operator using the AROME-FR background. Experiment design included a reference (all observations without radar reflectivity) and two additional runs, one with radar reflectivity assimilation and another with aforementioned rain assimilation. From analysis departure statistics it was shown that rain assimilation lead to better agreement of analysis with observations, compared to other two experiments. The analysis of dynamical fields such as vertical velocity showed more favourable conditions for convection in this experiment. Also, the precipitation forecast was better; one of the rain bends, present in the subsequent Antilope analysis, was only correctly represented in this experiment. This was confirmed on average with precipitation categorical verification (improved POD, ETS and HSS). The same methodology could be applied also to the total column water.

Session on radar data assimilation

This session was mostly devoted to summary of the LACE stays performed earlier in 2020. During a stay in Ljubljana, A. Trojáková investigated the use of the radar reflectivity observation operator within ALARO model configuration, studying the effect of graupel in the ALARO microphysics scheme on simulated radar reflectivity. The ALARO version with prognostic graupel provides very similar precipitation structures with slightly smaller intensities compared to the version without prognostic graupel. The prognostic graupel in ALARO are still considered as a very fresh development which needs detailed evaluations in NWP context. The graupel initialization to zero is necessary for the reflectivity observation operator even in case of ALARO without prognostic graupel to avoid

extremely high simulated reflectivity in radar assimilation. The initialization of graupel in both cases should be better understood and eventually fixed. It was found that prognostic graupel in ALARO microphysics has positive effect on simulated reflectivity, namely by a small reduction of random error (STD) of reflectivity innovations, see Fig R1. To make a reference to existing operational radar assimilation, ALARO and AROME-FR reflectivity innovations were compared over Germany. Both setups provide qualitatively similar statistics which give us more confidence for further testing of radar reflectivity data assimilation within ALARO configuration.

Antonín Bučánek presented the outcome of the second stay in Vienna it was screening procedure for radar reflectivity in AROME/ALARO, the construction of pseudo observed relative humidity columns, and thinning. Three minor bugs were spotted during investigation and discussed with MF. A reproducibility problem was identified in thinning if 2 or more input hdf5 files from one station were ingested. A few open questions related to the current data selection setup are still opened. One is the modification in for no-rainy observations as only observations with minimum detectable signal <= 0 are now allowed for assimilation. For French radars, this means observations at distance up to about 100 km, and the setting may be questionable for LACE area. Another yet unresolved issue is assimilation of dry reflectivity observations with zero first guess departure, which can only happen when model is dryer than radar (i.e. below detection threshold) at given observation location and elevation. It would be preferable to discard those observation from assimilation, but by default they keep being assimilated and produce pseudo-humidity columns with resulting positive or negative increments which may be unrealistic.

Katarina Čatlošová summarized her stay in Prague on radar radial wind DA. An overview of the available OPERA data was presented with focus on radar scanning strategies and selection of radars providing radial wind observations with Nyquist velocity larger than 30 m/s. Several bugs in HOOF were found during the investigation. Observation-minus-background (OMG) statistics in passive experiments showed reasonable quality of this wind subset, except for Slovenian and Danish radar network where wind reporting is inappropriate. A first case study showed preliminary positive impact of radar winds.

Session on other upper-air data assimilation developments

The session on upper-air observations continued with a comparison of Mode-S EHS and Mode-S MRAR data originating from the same aircrafts, presented by Katarina Čatlošova. The dataset was retrieved from Buchtuv kopec radar in Czech Republic where both data types were available over the period of Jan-Feb 2018. Significant differences in temperature measurements were observed. MRAR being more accurate. EHS computation formulas were shown to perform similarly, however, the exact algorithm used by the KNMI/EMADDC would be useful for inter-comparison.

In the next presentation by Hungarian colleagues Viktoria Hommonai and Kristof Szanyi, the evaluation of two Mode-S datasets in Hungary was presented: the Czech and Hungarian set of locally received Mode-S MRAR observations. The investigation period was winter, with prevailing anticyclone and the scores were rather neutral overall. It was shown that Czech MRAR data had some positive impact on precipitation for a few involved weather cases. For the Hungarian dataset, a very small number of observations passed the QC, so the data decoding/whitelist procedure has to be revisited. Possible distribution on European level was also discussed from Hungarian and Slovak case.
The last presentation in this session was devoted to optimization of AROME –RUC (Florian Meier). Firstly it was mentioned that unusually high impact of radial winds on DFS score was reduced by fixing

the calculation procedure (obs. error). Next, different setups of IAU were presented. The currently used two-stage IAU seems to be a good compromise between balance and accuracy. Florian also spoke about the proposed correction for height/pressure relation of wind profiler observations, where he proposes to use profile from FG rather than the standard atmosphere profile as currently coded. Two sodar sites (wind only) were recently added to operational assimilation. Then, the radar processing in cy43t2 was compared to currently operational cy40. There are less reflectivity observations due to introduced check of minimal quality and more radial winds, for which the rejection limit was reduced to 8 m/s.

Session on surface data assimilation

In the first talk, Stefan Schneider provided a review of the SEKF algorithm and explained the design of coupled and uncoupled experiments. Next, the status of different remote-sensed observations as input to SEKF was presented. For Land Surface Temperature (LST) assimilation, the best results are obtained by hourly perturbation window and using several soil layers. As input, a high resolution dataset is obtained by CDF matching of MSG and Sentinel observations. The experiments with local observation errors for soil wetness index (SWI) assimilation showed little impact of this local sampling, so it is not advised to be used in operations. It was also found that assimilation of SWI led to too wet surface during summer. The last topic of the talk was leaf area index (LAI) where climatology (PGD fields) can be upgraded with hi-resolution products such as Urban Atlas.

Helga Tóth presented a revitalized work on SEKF in Hungary. Two (summer and winter) cases were studied in detail by conducting impact studies with respect to AROME/HU OI-MAIN operational model. 3-hourly data assimilation cycle was started two weeks before the events. Forcing required to offline SURFEX run were coming from AROME inline forecasts at 9 m (radiation, precipitation, wind, humidity, pressure). For the winter case, the analysis of 2 m temperature and low-level cloudiness using SEKF did not really differ from the result of the operational run: similarly false low-level cloud and overestimated temperature analysis were obtained. For the summer case, however, a notable improvement of daily cycle could be demonstrated. In the near future, experiment are to be repeated on cy43t2 and SURFEX version 8.0.

DAsKIT exercises

Two exercises have been prepared in advance and discussed during this session: the first consisted on the adaptation of BATOR code to allow the lecture of local SAPP BUFR SYNOPs; the second consisted on the compilation of OBSMON library with gmkpack tool so that it could read odb files created with T-version of ALADIN system code. In this way, the two tools SAPP and OBSMON could become part of assimilation chaines of DAsKIT countries to boost their local progress.
 The two exercises have been prepared in beaufix, the HPC platform of Météo-France; the documentation of the exercises is available from beaufix and from (http://www.umr-cnrm.fr/aladin/spip.php?article364). The first exercise was prepared and documented by Y. Cenliz (TSMS); the second was documented and prepared by I. Dehmous (RMI). Due to Météo-France's firewall issues, the week before and during these WD, the exercises had to be limited to a live demonstration on 2/3 users that did not suffered problems. This session was recorded by ZAMG (F. Meier) and is available from him.

Summary of final discussion (Day 3)

The LACE DA AL first briefly presented an overview of discussed topics. The following discussion then first touched verification: the question was if we really capture impact of observations with our current verification procedures. R. Randriamampianina first explained that impact of any particular subset of observations in LAM is expected to be short-lived and small (the impact of actual observational networks is larger on global models). However, on the verification side, a lot of possible improvements are possible, including addition of national LACE stations. This now in most centres involve the "vobs" file format of Harmonie verification tool. A. Deckmyn commented that a reader/decoder from obsoul to SQLite can be constructed within the "harp" R verification package. LACE colleagues would appreciate this solution from a wide variety of solutions. Still related to harp, Alex explained that the work on reading grib2 files (e.g., to read precipitation analysis reference like INCA) is ongoing. Another point in discussion was Var-BC for ZTD, where Roger explained that a tool to merge 3 h and 24 h VARBC.cycle files is available in HIRLAM. Regarding the issue with small humidity background errors provided by EDA (as seen in some countries), Florian Meier explained that it is possible to tune only humidity covariances by REDNMC_Q switch, which is to be verified and tested. It was also mentioned that large-scale constraint (Jk) is worth the investigation especially where obs. coverage is poorer (e.g. over North Africa). The presented 1D+3D-Var rain rate assimilation also seems to be interesting to employ in LACE. It was however raised that we may be missing good quality precipitation products such as Antilope in France.

The meeting concluded with planning in both subgroups. The related RC LACE plan for 2021 can be obtained at the RC LACE website; while DAsKIT plans for 2021 can be found at http://www.umr-cnrm.fr/aladin/spip.php?article364

On the meeting planning and as a conclusion, DAsKIT group found the meeting very fruitful for the topics and the shared experiments; besides it was found important to follow other countries progress. The practical sessions were also acclaim as an efficient learning process, since problems are dealt in real time.

TODOs

A short to-do list can be created out of discussions during the whole working days:

- Provide information on the methodology and tools to compute EDA-based B-matrix (links to forum, Budapest DA training week)
- Alex Deckmyn shared his tool through https://github.com/adeckmyn/bufr_filter_gts
- Alex Deckmyn will prepare a seminar on HARP usage up to the end of 2020
- Alena Trojakova should share some hints on how to tune CANARI, kindly provided by Jean-François Mahfouf; hints on B-matrix computing; and the report on HARP validation at CHMI.
- P. Brousseau (MF) to be contacted for DA namelist in cy46, to help M. Imrišek in phasing of STD developments.
- Roger to provide information on improved radar scanning strategy in Norway.
- Roger provides tool to merge GNSS items in VARBC.cycle for different network times.
- F. Meier tries to coordinate with Claude on his DA contributions (LHN, wind profilers, radar wind rej. limit).

Some of these items were completed during and immediately after the workshop.