Summary report on DAsKIT video-conference, 18 December 2020

Participants: *Alex Deckmyn, Idir Dehmous (Belgium); Andrey Bogatchev, Boryana Tsenova, Milan Tsankov, Rilka Valcheva, Konstantin Mladenov (Bulgaria); Fatima Hdidou, Zahra Sahlaoui (Morocco), Marcin Kolonko (Poland); Maria Monteiro (Portugal); Wafa Khalfaoui (Tunisia); Yelis Cengiz, Meral Sezer (Turkey); Alena Trojakova (OPLACE, Czech Republic)*; *Piet Termonia (ALADIN PM, Belgium); Claude Fischer (ACCORD PM, France). Excused: Mohand Ouali Ait Meziane, Ghiles Chemrouk (Algeria); Malgorzata Szczech-Gajewska (Poland).*

The main topics of this video-conference were:

1. Status on the DAsKIT countries
2. Actual issues
3. Follow up of September's 2020 DA Working Days
4. Perspectives under RWP2021
5. AOB

Short status per country:

ALGERIA

data acquisition:
Conventional: SYNOP (3-hour GTS; 1-hour GTS automatic stations), SYNOP SHIP, non-standard SYNOP (hourly), TEMP (5 stations/day), BUOY, AMDAR (1-hour GTS BUFR);
Non-conventional: ASCAT (MetopA, MetopB/2 times intervals (8-10, 19-22 UTC) GTS BUFR), GNSS (hourly).
data pre-processing:
SYNOP (conversion to BUFR; duplications and amends are tackled); TEMP (filtered over ALADIN domain); AMDAR (filtered over ALADIN domain; selection of template 311010); ASCAT (filtered over ALADIN domain); pre-processing of GNSS data.
monitoring tools:
OBSMON installation is on-going; MANDALAY (CY40T1); odbsql.
verification tools:
HARP not installed.
surface DA:
AROME OI_MAIN (CY40T1_bf07) is being cycled under test mode with GTS SYNOP data (first results available for a 20-day cycling period), but without blendsur.
upper-air DA:
BATOR (CY40T1_bf07): SYNOP, TEMP, AMDAR;
BATOR (CY43T2_bf10): tested for SYNOP, TEMP, AMDAR, ASCAT, GNSS;
CY43T2_bf10: successfully tested for bator, screening, e927, e131 and e001, 3D-Var;
B-matrix computed for ALADIN (6km); AROME B-matrix has been computed from AEARP downscaling with an hybrid cycle (CY40T1_bf07 for BATOR and CT43T2_bf10 for 001).
6-hour pre-operational 3D-Var for ALADIN/AROME at 00,06,12,18UTC (CY43T2, SYNOP, TEMP, AMDAR); on-going validation of ALADIN BATOR and 3D-Var with multiple observations types : ASCAT (Metop A, B and C), GNSS (with updated list GPS)
Plans: bild ALADIN 3D-Var and AROME 3D-Var in belenos.
combined surface+ upper-air DA:
Plans: building a pre-operational version of 3D-Var cycle, combined with OI_MAIN, for AROME at CY43T2.

BELGIUM

data acquisition:
Conventional: SYNOP, TEMP, AMDAR from GTS (BUFR);
Non-conventional: ground-based GNSS (E-GVAP samples from Royal Observatory of Brussels, ROB/GTS, BUFR or COST716), MODE-S EHS (KNMI).
data pre-processing:
Python script that deals with duplications and amends; SAPP (ECMWF) server is installed; GPS data is converted from COST format to GTS BUFR with an UK tool (met Office) from E-GVAP. In a first approach, statistics on the observation error had to be determined, in particular, mean bias and error density, which are used when varBC method is not in use. However, in this case a 'whitelist' has to be used to enable this observation type.
Plans: configuration of SAPP to local needs.
monitoring tools:
OBSMON is technically working (see 2020 LACE DAWD & DAsKIT WD for details); MANDALAY installation (CY43T2_bf10), odbsql.
verification tools:
HARP.
surface DA:
3-hour cycling of an ecflow suite for surface DA (OI_MAIN, AROME 1.3, CY43T2_bf10) was set in operational mode by switching off CANOPY;
CANARI surface analysis in ALARO (4km) is being cycled with a 3-hour frequency.
upper-air DA:
computing B-matrix for ALARO 4.0 and AROME 1.3 by the NMC method (period: 01.02.2019-01.08.2019).
3-hour cycling of 3D-Var setup for ALARO4.0 and AROME1.3.
Plans: Ensemble technique B-matrix
combined surface+ upper-air DA:
BATOR (CY43T2_bf10): tested for SYNOP, TEMP, AMDAR;
3-hour cycling of 3D-Var+OI_MAIN (AROME1.3, CY43T2_bf10);
3-hour cycling of 3D-Var+CANARI (ALARO4.0, CY43T2_bf10).
Plans: to use other data sources like MODE-S and OPERA-ODIM radar; validation of combined DA solution OI_MAIN+ 3DVar.
operational systems:
CY43T2 by surface DA (coupling with ARPEGE).
reported issues:
Météo-France SYNOP stations (RADOME) with template 307096 are rejected in BATOR.

BULGARIA

data acquisition:
Conventional: SYNOP, TEMP from GTS (BUFR), local SYNOP (converted to BUFR).
data pre-processing:
new tools to split compressed BUFR file into single BUFR file have been developed in-doors using eccodes; duplications are then removed using the same tools.
monitoring tools:
OBSMON, MANDALAY ported locally.
verification tools:
HARP ported and some tests just started; local surface verification tool.
surface DA:
BATOR (CY43T2_bf10): SYNOP (locally); AMDAR (tests in beaufix);

OI_MAIN (CY40T1_bf07) for AROME-BG was migrated from beaufix (Météo-France HPC platform) and cycled for 2 weeks with BUFR data and ODB validation; a newcomer has arrived and efforts are being put onto its training;
Plans:migration of surfDA for AROME-BG to CY43T2_bf10 on the old machine using local SYNOP.
operational systems:
CY43T2 by dynamical adaptation is running in operational suite since November 2019:
ALADIN (105L/5km/72h/00,06,12,18UTC runs) and
AROME(60L/2.5km/36h/00,06,12,18UTC runs).
Plans: To restart the work on DA.

MOROCCO

data acquisition:
Conventional: GTS BUFR SYNOP (local and automatic), TEMP and AMDAR from GTS (BUFR);
Non-conventional: local GPS, ATOVS (BUFR).
data pre-processing:
GTS and local BUFR file for conventional observations and GPS data.
monitoring tools:
OBSMON installation: on-going in the local machine; MANDALAY on the local machine.
verification tools:
HARP not yet.
surface DA:
BATOR (CY41T1): SYNOP, TEMP and AMDAR from GTS (BUFR); update of param.cfg (TEMP 309052 from the new radiosondes over Morocco).
surface DA (OI_MAIN) runs for AROME-MOROCCO in beaufix; waiting to port and cycle to the new local machine.
upper-air DA:
3-hour cycling 3D-Var for AROME-MOROCCO has been cycled in the new machine (CY41T1) and should enter in pre-operational suite up to the end of 2020; B-matrix diagnostics have been done, comparing the downscaling with the ensemble method.
combined surface+upper-air DA:
OI_MAIN and 3D-Var are now being cycled in the new machine (CY41T1) and then should be ported at CY43T2 (beginning of 2021).
operational systems:
Plans: CY43T2_bf10 should enter into operations after CY41T1 with the combined algorithm (OI_MAIN+3D-Var).
reported issues:
a segmentation fault at CY41T1 is seen when adding the analysis produced by OI_MAIN.
Plans: to mode to CY43T2

POLAND

data acquisition:
OPLACE data is used; local SYNOP is available.
data pre-processing:
conversion of local SYNOP to BUFR.
monitoring tools:
OBSMON installed and tested with DAsKIT WD data.
verification tools:
HARP-v2 runs for DA cycle but not in use.
surface DA:

6-hour cycling of a surface DA based on CANARI (not SURFEX) for ALARO (CY40T1_bf07 and CY43T2_bf10), new LBC 4 km; 66-hour forecasts, introduction of new precipitation types; AROME CY43T2 under test.
operational systems:
ALARO CY43T2_bf10 (newLBC, 70L, 2.5km; 4x a day, up to 72h);
AROME CY43T2_bf10 (?L, 2.5km; 4x a day; up to 30h).
Plans:
AROME CY43T2_bf10 (?L, 2.5km; 4x a day; up to 36h)
reported issues:
OBSMON: problems when testing local data with graphical/shiny part of OBSMON; the implementation of the conversion tool (from the local experiment's output data to shiny recognised format) is missing.

PORTUGAL

data acquisition:
Conventional: SYNOP, TEMP, AMDAR from GTS (BUFR);
Non-conventional: OIFS radar data.
data pre-processing:
local handling of duplications and amends (FORTRAN): SYNOP and TEMP (missing validation); following local implementation of SAPP (SYNOP ECMWF BUFR data).
monitoring tools:
home-made (metview plotting for SYNOP); local OBSMON_V3.3.2 (shiny part) and MANDALAY (CY40T1_bf07) implementation is on-going.
verification tools:
local (home-made IPRODS-IVERIF) surface verification tool; implementation of pre-MONITOR@ECMWF with the support of Slovenia.
Plans: HARP not implemented.
surface DA:
3-hour cycling of a standalone surface DA scheme (OI_MAIN, CY40T1_bf07, AROME, 60L, 2.5km); on-going validation of 48-hour forecasts of AROME-PT2 (CY40T1_bf07, 60L) initialised by surface DA using as reference the same AROME-PT2 model configuration, initialised by dynamical adaptation for the two periods: WINTER: 10dez2018-10fev2019 (cold and rainy period); SUMMER: 01ago2018-09set2018 (extreme temperatures).
combined surface+ upper-air DA:
BATOR (CY43T2_bf10, ported to ECMWF): SYNOP, TEMP, AMDAR;
B-matrix computed by AEARP downscaling and tested in beaufix for AROME-PT2 OI_MAIN+3D-Var, with AROME DA VarBC (CY40T1); 20-day validation on beaufix platforms of combined AROME_PT2 OI_MAIN+3D-Var (CY42T2) has revealed a slightly improving using conventional + OIFS HDF5 volumetric data, in particular for larger amounts of 24-hour accumulated precipitation (Skill Scores and Probability of Detection), keeping the False Alarm Rates.
Porting beaufix's later experiment to ECMWF machines (at CY43T2_bf10, with an adaptation of Slovenia (ARSO) scripts ported to ecgate): testing cycles have been run for a Summer period; however screening is rejecting all Portuguese radar data; HOOF tool was added for pre-processing; pre-'monitor' verification tool has been plugged-in; other local diagnostic tools (metview based) have been migrated to ECMWF platforms too.
operational systems:
dynamical adaptation of AROME-PT2 (CY40T1_bf07, 60L, 2.5km);
Plans: on-going migration to ECMWF computing platforms (CY43T2_bf10).

reported issues:
installing BATOR CY43T2_bf_09 installation in the local machine (IBM-p7) since the native compiler does not supports FORTRAN2008 features; installation in the local machine (IBM_p7) with gcc since it was not possible to install gcc compiler with older software on the machine which does have maintenance support.

TUNISIA

data acquisition & pre-processing:
OPLACE;
Conventional: local SYNOP; TEMP and WIND PROFILER;
POP-RMI is used to pre-process conventional data.
monitoring tools:
OBSMON and MANDALAY implemented on the local machine.
verification tools:
HARP installed; its test should be started during the first half of 2021.
surface DA:
BATOR (CY43T2_bf10, new HPC): SYNOP;
surface DA (OI_MAIN) has been implemented in beaufix but not yet on the local machine;
Plan: move to CY43T2_bf10 in the new HPC.
upper-air DA:
6-hour DA cycling (ALADIN, AROME).
combined surface+ upper-air DA:
B-matrix has been computed by the ensemble method and tested in beaufix for AROME;
On-going implementation of a combined surface + 3D-Var DA, with a Jk component on the new HPC platform: V-matrix computation; namelist tuning; code modifications (adapting Jk ALADIN existing routines for AROME).
Plans: cycling of AROME 3D-Var on the new HPC; Jk validation blending to overcome the “sparseness” of observations on Tunisian domain; Enhance Local Data Base Observation for DA; use of more observations: AMDAR, loca GNSS.

TURKEY

data acquisition:
Conventional observations: SYNOP, AIREP and TEMP from GTS (BUFR), local SYNOP (conversion to BUFR);
Non-conventional observations: AMSUA (Metop & NOAA), MHS (METOP & NOAA) and AMV (METEOSAT).
data pre-processing:
SAPP BUFR data (SYNOP) was tested with a locally adapted export version of BATOR for local observations (see 2020 LACE DAWD & DAsKIT WD for details).
monitoring tools:
OBSMON has been installed on desktop and tested with provided observations; OBSMON is not installed on supercomputer yet: a python script has been created to visualize MANDALAY output.
verification tools:
Currently HARMONIE verification is used.
Plan: HARP implementation.
surface DA:
BATOR (CY43T2_bf10): tested with local SYNOP observations;
CANARI (CY43T2_bf10): tested successfully with local SYNOP (t2, rh2); diagnostics done for one SYNOP station; CANARI tuning by checking the structure functions MESCAN with default has been performed. This scheme is now in pre-operational mode.

combined surface + upper-air DA:
B-matrix has been calculated from AEARP at CY43T2 by the ensemble method for ALARO-Tr;
Plans: set-up of a joint surface (OI_MAIN) + 3D-Var DA to AROME-Turkey; to compute a new B-matrix for the AROME combined solution with couplings from ECMWF.
operational systems:
CY43T2_bf10 is operational since September 2019 for AROME (72L, 1.7km and 48-h lead time);
CY40T1_bf07 is operational for ALARO (60L, 4.5km and 72-h lead time);
6-hour combined surface + upper-air DA cycling for ALARO CY40T1 was stopped due to lack of computing power (at 00, 06, 12, 18UTC network times), at 4.5km, 60 levels and with LBC from ARPEGE
reported issues:

Acknowledgments: to Alena Trojakova (CHMI, LACE) for always bringing hints and solutions to DAsKIT problems.

Main conclusions:

1. at the last quarter of 2020, the majority of DAsKIT countries have the CY43T2_bf10 implemented in-doors (except Morocco (at CY41T1) and Portugal: the first country is now focused on implementing its local suites in the new machine; and Portugal, has moved its recent developments to ECMWF HPC platforms).

2. concerning the implementation of DA algorithms, efforts for the validation and discussion of the DAsKIT set for surface have slowed down in almost all the countries; however this scheme has entered into operations in Belgium and pre-operational in Turkey, after some tuning of the OI/CANARI structure functions (MESCAN against default); besides,

3. most of the countries are focusing their efforts on the local implementation of a 3D-Var algorithm or a combined algorithm of surface+3D-Var, namely: Algeria, Belgium, Morocco, Portugal, Tunisia and Turkey. This implementation is done by porting a previous set (Algeria, Turkey), by building the scheme from scratch (Belgium), or by migrating it from a reference environment (Morocco, Portugal, Tunisia). Bulgaria and Poland are at this moment blocked due to lack of computer power.

4. The second main concern of DAsKIT countries is the handling of locally available data. Main observation types are: SYNOP, TEMP, AMDAR and WindProfilers (Tunisia) under BUFR format from GTS or from local networks. However, some countries are now concerned with the handling of non-conventional data (for instance, ASCAT, GNSS (Algeria, Morocco), GNSS, MODE-S (Belgium) ODIM volumetric data (Belgium, Portugal)):

5. so far, countries have created pre-processing tools based on eccodes (ECMWF) to handle BUFR data, as it is the example of Belgium and Bulgaria. Pre-processing consists of removing duplicates from corrections and amends and filtering a certain type of BUFR template over a particular geographical area. Pre-processing is applied to SYNOP, TEMP and AMDAR, but removing duplicates on TEMP is a process missing to the most part of the DAsKIT countries.

6. in particular, Tunisia (as a non-ECMWF member country) has successfully implemented and tested the Belgium pre-processing PYTHON tool (POP-RMI), being user friendly and simple to install; while

7. Turkey is successfully using SYNOP BUFR data coming from SAPP on the local (CY43T2-export) BATOR version.

8. MANDALAY (CY40T1 or CY43T2) has been implemented and tested with demo data in almost all countries (still on-going for Portugal) and no issues have been reported so far. Turkey has their own set of scripts to do some data monitoring with MANDALAY;

9. OBSMON has been implemented and tested in almost all the countries with demo data (still on-going for Algeria and Portugal), and Belgium has started to use it on a regular basis;

10. HARP has not been implemented in most of the countries, but some discussion started on the possibility to use it in the short-term (Tunisia); some countries start to invest on MONITOR (HARMONIE system).

11. Issues have been reported (and have been solved or on-going efforts are being developed to solve them) at the level of BATOR digestion of templates: 307096 (SYNOP) and TEMP (309052).

Perspectives and recommendations under RWP2021:

1. Generally speaking, next year, the efforts of DAsKIT group should be conducted under the new package DA8 of the RWP2021 (supported by the actual DAsKIT countries and also Romania);

2. the DA8 should still count with a coordination; the main target will be to support the actual DA capacity building of some of these nine ACCORD members, through the joint effort to set up a common basic kit of combined DA (OI_MAIN+3D-Var) with conventional observations: SYNOP, TEMP and AMDAR.

3. The work of this group should be reported twice a year to the PM; therefore, the number of actual quarter video-meetings on countries' progress/status will be reduced. Instead,

4. major efforts should be put on the organization of dedicated video-meetings on practical sessions (that can be recorded), or opportune topic seminars. One example can be on the use of observations, other than conventional: due to the eventual lack of homogeneity on national observations networks, further observations type (than conventional) will be treated under the corresponding DA working package of the RWP; however, DAsKIT can propose the organization of a seminar to ACCORD DA Area Leader, so that an expert from a more experienced country presents an introductory seminar on this specific observation type.

5. some track should be followed on the possibility to keep an on-line DA training using the latest Common DA training (Budapest) material. Moreover,

6. the annual organization of the DA Working Days inside the new consortium has been found as an excellent opportunity for exchange of information within DA experts.

7. A Wiki Page dedicated to DAsKIT work will be created by Turkey during 2021.

8.Taking into account the RWP2021, the main topics foreseen for 2021 under the new package DA8 can be listed: monitor countries' acquisition and pre-processing efforts; re-visit the possibility to have a common verification tool; tackle the common surface DA settings; tackle the common upper-air DA settings; revisit the common data monitoring tools;

9. what concerns the data acquisition and pre-processing tools, it is expected to follow the SAPP implementation in the DAsKIT countries; besides, some tests for comparison of SAPP with POP-RMI has been proposed;

10. concerning the possibility to have a common verification tool, efforts will be mainly focused on HARP since a dedicated plug-in for FA and BUFR files reading is expected for 2021. An introductory seminar on HARP and its on-line training is then a candidate proposal for a DAsKIT dedicated seminar during 2021;

11. The tuning of CANARI and SURFEX inside OI_MAIN and the validation of the latest algorithm will be a target of surface DA under discussion during 2021.

12. Stabilise a combined DA KIT configuration: CY43T2 namelists, scripts and B-matrix computing (obs perturbations and diagnostics) will be of major concern for the DAsKIT group during 2021. In this way a practical session on B-matrix computation will become a good candidate for a dedicated remote practical session during 2021.

13. Another possibility of a dedicated video-meeting is the implementation of data monitoring procedures using OBSMON.