VERSUS Technical Users Seminar- WG 5 Meeting

Date: 27-29.05.2015

Place: Torino, ARPA-PIEMONTE, Italy

Participants:

Angela Celozzi-AC (USAM), Antonio Vocino-AV (USAM), Flora Gofa-FG (HNMS), Dimitra Boukouvala-DB (HNMS), Andrea Montani-AM (ARPA-SIMC), Maria Stefania Tesini-MST (ARPA-SIMC), Chiara Marsigli-CM (ARPA-SIMC), Joanna Linkowska-JL (IMGW), Anastasia Bundel-AB (RHM), Mihaela Bogdan_MB (NMA), Ulrich Pflüger-UP (DWD), Xavier Lapillonne_XL (MCH),   Naima Vela-NV (ARPA-PT), Massimo Milelli-MM (ARPA-PT), Elena Oberto-EO (ARPA-PT), Renata Pelosini-RP (ARPA-PT)

Minutes (per session): Flora Gofa, Angela Celozzi, Elena Oberto and Maria Stefania Tesini

Item 1: VERSUS PP phase 7 Overview

AC presents an overview of PPVERSUS2 Phase 7 (presentation available in the forum Documentation area. Tasks 0.a,b,c (Help Desk, Documentation updates, new VERSUS releases) are ongoing while 0.e (Test of VERSUS3.2extra), O.f (Concurrent use of VERSUS) and 0.g (Performance Optimization) are realized in version 4.1 (under testing). For Task1 (Confidence Intervals (Integration), RHM provided the code and PLteam is integrating it into VERSUS, Tasks 1.b,c (EPS refinements, Automatic programs) are ongoing (version 4.1) while Task 1.d (VERSUS portability) is finalized. Task 2 (implementation of Feedback Files) is under testing phase. Task 3 (GRIB2 Integration) is not performed yet and PL team cannot guarantee that it will be completed until the end of current COSMO year. Task4 (XML integration) is performed partially by IMS, but the development has to be inserted in VERSUS code. AC asked MCH to participate to the test phase but XL informed that MCH will no longer participate in VERSUS tests. Task 5 (Additional statistical techniques) is almost completed. Feedback from tester (NMA) was given and it is being incorporated in VAST.

[ACTION]: Tester needs to be found for Task 4.

Item 2: Follow-up from GM meeting and WG7/VERSUS meeting. Performance Optimization. Verification in background

During the last GM, some malfunctions in the EPS verification have been highlighted. In order to include the feedback emerged, a meeting was organized with the WG7 coordinator. It was held on October 29th in Pratica di Mare and the document VERSUS_CNMCA-ARPAER_29102014.docx describes the agreed activities. AC presented the status of actions divided in Test activities, Functional details, Optimization, EPS refinements. AC informs that most of them are realized or ongoing apart from Action 12: Verify 12h acc. prec for day and night (deterministic verification) from SYNOP and Action 15: Possibility to create a stratification using a txt file.

[DECISION]: For Action 11d: Verify 12h acc. precipitation for day and night, it was requested to be included for deterministic verification.

MS asked if it is possible to run verification with different starting hour (e.g. 03, 09, 15UTC). AC replied that it is implemented but never checked in VERSUS.

[ACTION]: Feature will be tested by MS and information will be disseminated.

Following AC informed on the Performance optimization actions that were performed by the developers. For Optimization on DB Structure it can be done in two ways and according to the user needs:
A. If user needs to keep fcs and obs (only data not scores) in table, the patches VERSUS_3.2.extra VERSUS_3.2.extra_OBS shall be performed. B. If the user has no need to keep fcs and obs in table, the script create_table.sql shall be performed. For Optimization of Loading, "Block inserting" is implemented for scalar fields and "3D Optimize method" (1 block= number of stations in the stratification). This enhancement provides 60% time saving. Parallelism in EPS score execution was also performed (details in presentation). Instructions on TB Tuning are also available in forum. AC also reported for the necessary Meteorological station update, in order to account for stations with wrong data (lat, lon, height) or different occurrence of stations with same WMO block and station number. Two scripts are implemented (delete and update station) and should be run by users. Results from some tests were presented using MCH data after the optimization, with significant reduction of loading time (see graphs). For scores execution, improvements are present but less sound.

[ACTION]: It is suggested that each user it should install patch 4.1, execute table partitioning, execute update of the station table and analyze the best use practices for VERSUS.

Following, AC presented the work performed in VERSUS for adding verification in background activity. This feature improves the multiuser and multitasking that was introduced in VERSUS 4.1 version. Major implementations include: Capability of running more configured verifications simultaneously in background; Web GUI active during the verification execution; Activation of controls for shared resource from different access (e.g. case of running the same process from different users), Restructuring of the Queue Management; Possibility to kill a running process from Queue Manager page; Different management of verification log files: to every verification is associated a unique log.

Item 3: Follow-up from GM meeting and WG7/VERSUS meeting: Suggested best practices with VERSUS

User experiences from tests

AC presented instructions to obtain the best VERSUS configuration. For loading suggestions are: Do not use more than a method for the same FE, it is best to have two FE with separated methods; When the system seems slow, it is strongly suggested to split the grib file into single grib files and get a test on this configuration; "Radius" method is heavier, the greater is the radius more time Versus uses for loading GRIB files. The number of stations in the stratification is another issue to arrange and it is strongly connected to the method. It follows some suggestions: a. If the stratification has more than 1000 connected stations it is recommended to have only one method per FE, b. Be careful to have a stratification adequate of grib area. it is useless to have stations that fall outside the grid, c. If the stratification has more than 1000 connected stations it is preferable to load the observations as first, delete the stations in the stratification that have not data and then load the forecasts, in order to avoid the management of useless forecast data.

Another critical issue is the running of EPS scores for stratification configured on the entire COSMO Domain. It is strongly suggested to: 1.Create the verification on stratification where the updated station procedure has been performed, 2.Evaluate most significant threshold values in order to optimize the intervals number, 3.If the VERSUS machine is modest it is preferable to load a forecast parameter at a time, run the verification on this parameter, remove it from DB and switch toward another parameter.

[ACTION]: Necessary for all users to implement these steps in order to obtain best performance with VERSUS.

The last part of the day was spent with a discussion with ARPA-SIMC participants and their experience with the ongoing tests on EPS (performance comparison with in-house software, and comparison of statistical indexes calculations from both SW). AM, CM and MT reported that loading times are decreased with VERSUS compared with their software but more tests need to be performed. Discrepancies in scores still exist (mainly RPS, RPSS).

[ACTION]: Tests will be continued with ARPA-SIMC but also with RHM. For the rank histogram, instructions will be given for the user to be able to implement various methods to specify ties. Tests needs to be performed to parameters other than precipitation. Results of tests phases (ARPA and RHM) will be disseminated before the end of the project (September).

Item 4: Confidence Interval Integration

AB presents her work for CI and the possible solutions for integrating of CI in VERSUS (details in presentation available in forum). The code was built in R and it could be integrated directly in VERSUS by creating new indexes. XL asks why we used two different methods to calculate the CI. This choice can be confusing. AB states that the basic method was the bootstraping as it was proven to show the best results in the literature. For the dichotomic scores, the bootstrapping was found difficult to apply due to VERSUS structure (only the contingency table outcomes being available only, and not the raw values), so the normal approximation was applied in most cases (but it is possible, as most of the proportion scores, such as the hit rate, follow the binomial distyribution, which tends asymptotically to the normal distribution. Thus, the CIs based on the assumption of normal distribution are calculated for most of the dichotomic scores. For some dichotomic scores, the booootstrap CIs are also available (the scores that are available in R function table.stats.boot: FBI,POD, FAR, and ETS). Note: Following this comment, a method for calculating ME and RMSE CIs based on a normal apprimation was added. A list of CI for each index has been provided through the forum but it is unclear as to supply WG. There are two solutions: VERSUS users integrate them independently on system or else PL team will add in the next patch the entire list and each user will delete unnecessary indexes. The second solution was chosen. AB provided solutions for dichotomic parameters but some graphical issues are raised from the users due to VERSUS graphical limitations. For the dichotomous parameters, it is not possible to represent different indices in the same graph. This is a different philosophy of the current representation of indices in VERSUS so it cannot be changed. FG noted that the graphical representation for continuous parameters is not the usual one (additional lines for lower and upper limits and not with bars) and therefore not convenient at all for use. As PL informed that this cannot change, AB showed graphs to overcome the problem of using CI that are calculated in VERSUS and plotted externally. AB presents through example graphs the need to calculate and present CIs for the scores of the difference between two models (to construct the confidence intervals for the differences themselves). AC explained that such plot is not possible in VERSUS, as the software does not perform calculations of indices for model differences.

[ACTION]: PL team will feed the whole DB with the complete list of CI in the next patch.

[ACTION]: MS will try to create a script to plot the CI outside VERSUS using R. In this way it could be possible to obtain the graphs for the common plots CI for continuous and dichotomic parameters as a dependence on a leadtime or step. Note that for the dichotomic scores, performance diagrams already contain similar information.

[ACTION]: The implementation of CI for EPS is still pending and it should be completed before the GM.

[DECISION]: Graphical representation will not be improved in VERSUS for CI. No priority for WG5 to implement difference in scores between two models with respected CI.

Item 5: VAST Overview

For details presentation ITEM5_VAST.pdf is available in the forum Documentation area. NV gave an overall description of the software (input data, necessary components, configuration files, specifications for LIBSIM and VAST users). She also gave an overview of the installation procedure. Comparison of the scores was performed with VAST and with the DWD software (Beth Ebert Toolbox in IDL) is in progress (final phase). Some slight differences were observed in the index of Multi Event Contingency Table, but the bug has been found and will be corrected soon. A test phase for the whole VERSUS Task was performed by the Romanian colleagues and the problems encountered were caused by the pre-processing of data with LIBSIM (detailed review document available). FG mentions that to declare close the test phase completed it is necessary to have a point to point reply to the tester comments. Also she mentions that a reply is pending about the use of Fieldextra originated CVS input data to VAST and JMB had provided such example dataset. Then, XL mentioned that the Swiss colleagues are interested in carrying out a test of VAST with their radar data.

[ACTION]: In mid-June the final version of VAST will be released and copied in ftp.meteoam.it with associated documentation and data set for first applications. Answer to the tester review will be given by NV.

[ACTION]: A test will be performed by NV with data provided by the MCH as soon as possible. Also a test with Fieldextra CVS will be performed.

Item 6: Portability of VERSUS (Fedora)

For details, presentation ITEM6_Portability.pdf is available in the forum Documentation area.

MB presented the Task work. VERSUS 3.0 has been installed on FEDORA 15 and FEDORA 20 (32 bit). The 3.0 version of VERSUS was chosen because it is the operational version in NMA and it is easier to compare the implementations of the same VERSUS system on different operating systems. The following activities to be performed in the machine configuration have been described: User and group, Additional software (MySQL installation, php-mysql, Apache Installation, Installation and further Top PHP additional packages), php.ini configuration, Apache optimization. No activity should be carried out within the software code. The activities to be carried out are only in the machine configuration. AC mentioned that at CNMCA a further test is not possible as the Fedora operating system is not available.

28 May 2015

Item 7- FEEDBACK FILES implementation in VERSUS (AC):

FF implementation has been introduced in VERSUS versions >4.0. The test phase: dataset available 1 month (August 2014), performed by DWD. Two documents were prepared as reference: "Feedback File Definition Document" and the "Design decisions" that describe the type of data, the scores implemented etc. An extension of DB has been created, in order to manage FF, a new FE, new php GUI for a new area for FF, new R script for the scores.

It has been asked by the participants that some explanation about the version will be given by VERSUS PL. AC explained: at the moment DWD is testing the 4.0 version (containing also FF developments). RHM and ARPA-SIMC are testing 4.1 versions (containing FF developments and EPS optimization). In detail: to install 4.1→ 3.3.1 has to be installed, skipping 4.0.

[ACTION]: The official realise will be 4.2 (after bug fixes). In September a new installation package 5.0 with confidence interval and XML integration will be disseminated. A script will be provided to check the correct installation.

Stratification for AIREP: A new kind of stratification has been introduced, defining an area in which all the points considered (for AIREP and TEMP) will fall: insert lat/lon for P1 and lat/lon for P2 to define "Area FF". Then, the system automatically loads into the database the model described in FF. In FF parameter are defined a list of available parameter. New FE for FF: for TEMP, PILOT, AIREP. UP asks for the possibility to choose only TEMP in the area defined. This is possible trough classic stratification; instead for PILOT/AIREP, a dedicated area without this possibility has been created.

[ACTION]: UP reported problems with the output of data availability report. It needs to be investigated what is displayed in the downloaded text.

Item 7.7 – VERSUS AND FEEDBACK FILES: CURRENT STATE AT DWD (UP):

UP gives overview of FF and for Model Equivalent Calculator (MEC) software, that generates uniform netcdf FFs for every observation system (SYNOP, TEMP, AIREP, etc.) and different verification reference times (normally 00,06,12,18 UTC). He gave an overview of the test he performed with VERSUS and DWD verification system on the same dataset (COSMO and ICON). UP asks also the possibility to verify the analysis and AC reports that this was not required in any specifications. The test showed good agreement between the two software.

Suggestions/Problems:
Stratification by VRT should be implemented (add stratification). [ACTION]: AC declared that it will be included.
A model entry should disappear, if all model data are removed from VERSUS DB –already solved
A verification of a model must be possible over a period, in which the DAS Version has changed
Cross model for FF scores [ACTION]: AC reported that this possibility will have to be evaluated if feasible
Modify information inside FF source model – already solved
Time series plots [DECISION]: only externally from VERSUS such plots can be created
Area definition also for TEMP data (discussed before)

The session has been concluded with a short discussion about FF use inside WG5 partners and the real importance to implement the FF verification in VERSUS. UP reported that MEC software will be available in late June.

Item 8 – TECHNICAL ISSUES AND MINOR REFINEMENTS (AC):

AC presented some technical issues. it has been created a script to clear the DB: inside the GUI you fix data and all the data before will be removed. MS has tested the script but her DB remained the same size. AC mentioned that it is also necessary to optimize the DB after the cleanup.

[ACTION]: In order to limit the data for deleting, it was requested the script to include not only an upper but a lower date limit.

Item 10 – OVERVIEW OF VERSUS FUNCTIONALITY (AC):

AC gives an overview of all implementations in VERSUS. AB asks about the possibility to have the zoom in the map and AC answers that maps are not dynamic but someone can insert a more detailed map). AB also asked if there will be the option for suspect obs for EPS but at the moment the answer is negative. MS mentions that in BUFR files there is the possibility to know how reliable the data is and this could be readable in VERSUS but the software is not able to ingest this information.

[ACTION]: AC will investigate and report on the possibility to include suspect obs for EPS or the ability to delete the suspect obs in the current view for deterministic verification and thus excluding them from EPS verification calculations too.

For the possibility to verify the same month/season over several years, AC informed that this can be achieved through WDV with the definition of a class called "month X" and the verification will be considering all the specific months/seasons selected.

Item 9 – LONG TERM MAINTENANCE (AC):

AC presented the CNMCA plans for the long term maintenance of VERSUS system. It is estimated that two releases per year will be created based on the current plans (only if it is necessary). For small bugs the code will be immediately corrected (release of subversion) while for major bugs, various requests/problems will be collected and release in the reference versions. CNMCA ceases any further development but other partners can contribute to future WG5 requests/needs. It is necessary to organize better the release way but in any case the source code admin must be informed.

AC defines what is included as part of the "maintenance":
- Bug fixes and new patches
- Adapt the SW to new library versions and operative systems
- Code optimization and minor refinements (under SCA)
- User support requested through forum
- Any other maintenance activities according to SCA's term of reference

AC proposes to insert the GRIB2 implementation (Task 3 PP phase 7) into maintenance. FG informs that GRIB2 implementation cannot be considered a maintenance activity and as it was a priority given by the STC in the current project phase7, it should be realised before the project ends. FG also explains that a short term extension of the project can be requested if there are time constraints. A discussion also took place for the proper way to introduce GRIB2 in VERSUS and FG informed that such decisions concern TAC and SCA can consult with them. A discussion was also held regarding VAST and LIBSIM (VAST part) maintenance as there are part of VERSUS developments.

[ACTION]: VERSUS PL will reevaluate the implementation plans for GRIB2 and will inform SPM if any amendement to the project plan is required.

[DECISION]: VAST and LIBSIM (VAST part related) maintenance will be coordinated by VERSUS SCA with communication with developers (VAST→ ARPA-PT, LIBSIM → ARPA-SIMC).

VERSUS WISH LIST

The outcome of VUS meeting was finalized with the preparation of a list of VERSUS to do-wish list in agreement with all the participants and PL team. (blue: ongoing activity, green: completed activity, red: not realized yet (or cancelled)

1. Well working EPS verification – Scores accuracy - verification based on obs period (VERSUS 4.1)
2. GRIB2 format as forecast input format (see DECISION above)
3. Completion of FeedbackFile tests/improvements, possibly with use of more observation types beside TEMPs (VERSUS 4.0 – Airep and Pilot)
4. Improvements in loading speed for operational duties (VERSUS 4.1)
5. Technical improvements in installations problems, upgrade problems, portability on different linux distributions – Virtual box installation
[DECISION]: Virtual box installation will be added to maintenance plan
6. Stabilization of multiple user access/background processing (VERSUS 4.1)
7. Background processing EPS verification [DECISION]: to be performed by next GM
8. Stop/control status of the verification while running (VERSUS 4.1)
9. Improved quality and flexibility of graphs
[DECISION]: As graphical improvements are not possible, will have to be performed externally of VERSUS
10. Automatic saving of scores through scripts executables [DECISION]: to be concluded by next GM
11. Improvement of technical manual and description of implemented methods/plots
[DECISION]: Include in Maintenance Plan
12. Error messages should be more informative (case of no data!)
[DECISION]: Some representative cases will be added to last version of VERSUS by next GM
13. Clear queue warnings [DECISION]: next VERSUS version
14. Support for SQL data base maintenance [DECISION]: partially through Maintenance Plan
15. Loading and verification of forecasts of 03, 09, 15, 21 UTC. [ACTION]: To be checked for PRECI by PLteam
16. BUFR with only 1h-precip accumulating to 6h preci verification
[DECISION]: It was decided that such processing activities can be performed externally (e.g. LIBSIM)
17. Delete suspect obs from DB "automatically" [DECISION]: Possibly by next GM
18. More observation types (remote sensing) [DECISION]: Not possible as no further major developments will be performed in VERSUS
19. CI for models difference statistics [DECISION]: Request deleted (not priority). List of scores (CI) to be included by default in next version (4.2)
20. Extreme weather scores (EDI, SEDI, etc.) [DECISION]: Possibility to be included in Maintenance Plan as minor developments
21. Suspect obs for EPS and TS [DECISION]: To be investigated by next GM -feasibility and effort definition.
22. New complete installation on latest version [DECISION]: By next GM, a final VERSUS version will be submitted along with a complete new installation (including station cleanup, verification-model cleanup).

WG5 session – 28 (afternoon)/29 (morning) May 2015

Status of common plots

DB presented some problems with this activity. Specifically:
- Necessary a comparison between GME/IFS
- Necessary more effort to be given by all members for comments on the results
- Good to add time series or daily cycle plots to compare different years (from each user)

- Useful to compare with the climatology of the area in order to investigate the link with the model behaviour

[ACTION]: During the next GM, CP activity will be discussed again for their necessity/usability/content. Importance of national plans for change in the operation model resolution (from 7km to 2-3km). Evaluate the conditional verification results and plan for a meeting with WG3 during the coming COSMO year.

Next GM Preparation

WG5 will hold parallel meeting at upcoming GM meeting, for PPINSPECT. AB gave a short overview for the submitted revised project plan. Project participants will present their work that is ongoing in Wroclaw. A poster is submitted for INSPECT work and will be presented during the upcoming EMS meeting in Sofia (Sept 2015). During this meeting a special session will be devoted to MESOVICT international project.

PPVERSUS2 parallel session will also take place during GM, with the presentation of the remaining implementations and the outcome of the test phases.

[ACTION]: An intensive use of VERSUS software is requested following the official release of VERSUS patch (last official version 3.3 in October 2014), for possible minor bugs or problems in implemented features.

WG5 session during GM will give special focus on Conditional Verification and Common Plots activity. A session for NWP meteorological test suite was agreed to take place with all involved people.

Preparation of recommendation for future strategy on verification tools

FG presented the feedback from the questionnaire that was distributed in spring to all WG5 members and it concerns the current usage of verification tools but also questions related to the future needs from a Common Verification Software (CVS). A long discussion led to majority decisions for the recommendations for the future strategy on verification tools that need to be submitted to the STC. A report will be prepared with contribution of assigned persons from each service and will conclude the thoughts of WG5 members on this issue.