Strategy on verification tools

WG5 recommendations

Some background info…..

In **2006**, the STC approved the development of a common, unified verification 'library' through the VERSUS project, which was followed by the VERSUS2 project in **2009**.

Objective: develop a common COSMO verification software package that would enable the production of homogeneous and comparable statistical results.

Caveats: Technical limitations of VERSUS, unforeseen delays, and the need to decentralize verification tool development from one member service to more COSMO partners.

WG5's members have considered the possibility of developing separate verification modules that would not necessarily be directly linked to the existing software, something that would assume a **change in the COSMO strategic direction for verification tools** development.

- **SMC meeting (08/2014):** "*The long-term strategy on the COSMO verification software and possible VERSUS system has to be discussed and prepared separately as longer-term action. WG5 should take proactive role in the process*."

- **STC meeting (09/2014):** "*The STC decides to task the WG5 to draft a COSMO strategy for the verification tools to be submitted to the STC (For which purposes shall VERSUS be used? Shall additional COSMO verification tools be foreseen? For which purposes?).*"

STC statements on verification tools

I. *CVS required for „**common needs**“ that need to be clearly defined and précised by WG5*

II. *The possibility to have a **central installation** of CVS needs to be examined (with one partner running it for common tasks, or to be used by partners via web-interface).*

III. *CVS should be supplemented by **additional tools (CVS+)** with the use of local tools or the use of existing external tools. The coordination of this activity will be done by WG5 (additional needs/approval) while rules of this exchange should be given precisely.*

IV. *It is not possible to cover the needs of all partners with a CVS+ therefore, **no replacement of the local tools is targeted**. A strategy on verification tools should include a broader exchange of local tools to cover needs of more than one partner and **defined standards** could facilitate this exchange.*

V. *A CVS+ should target in:*

- *covering most **verification needs addressed by WG5** (diagnostics on operational datasets, conditional verification),*
- *addressing where possible the **verification needs of all groups**, even though it is impossible to cover all research needs*
- ***utilizing additional tools** when necessary (WG5 and others)*
- *compromising between coverage of the needs and the **cost / effort***
- no large own development of verification tools should take place

Roadmap

• After SMC meeting (Jan2015), WG5C requested some boundary conditions to be given from the STC for this proposed strategy

• A questionnaire was given to STC members and based on the answers, the following statements were agreed (March 2015)

• Following this, another more detailed questionnaire was filled by most WG5 participants that included questions in areas as:

- ✓ Evaluation of VERSUS software usage at each service
- ✓ CVS: Common Verification Software
- ✓ Local verification tools

• The final discussion based on all the material gathered, took place during the VUS/WG5 meeting in late May 2015. A report on WG5 recommendations for the strategy on verification tools was prepared in the form of commonly agreed scenarios

Definition of Common Verification Activities (use of a CVS can be advantageous)

Common Plot Seasonal Reports: Verification results of statistical indices for main weather parameters derived using the operational COSMO model implementations in each service. The domain (common or custom), resolution, statistical scores/methods, frequency and graphical representation, are decided on an annual basis from WG5. The main findings of this organized analysis is presented during the GM plenary session together with the long term trend of them, providing a basis to track the performance of COSMO model - ***CVS and possible AVT requirement***

Conditional Verification Tests: Methodical evaluation of model performance in order to reveal the typical shortcomings of a model and to provide information to the model developers as well as to the forecasters with regard to model reliability. Verification software that allows for CVS applications is necessary – ***CVS requirement***

Science Plan Strategic Priorities: Investigation on statistical methods to identify the skill of convection-permitting and near convection-resolving model configurations.,probabilistic and ensemble forecast verification, severe and high impact weather verification. The application of these is closely related to existence of the necessary verification tools – ***CVS and AVT requirement***

The necessity to have Common Verification Software and Tools (CVS+) for Common Verification Activities was underlined by the vast majority (but to be followed by scientific and technical standards).

Centralization of Verification Tools

The installation of the verification tools on a dedicated "central" facility could simplify the accomplishment of COSMO-WG5 common activities. Such tailored unified system could represent a **cost-effective alternative** for COSMO members who don't have specific IT resources or technical staff for maintaining a state-of-art local verification system, unbinding local resources available for research/operational activities outside COSMO framework.

Critical factors:

- A central facility design – although providing a unified reference system – may require an **additional effort** to the members in terms of **data exchange** and networking between the data providers and the central data hub. Moreover, the **monitoring of the data fluxes** should be ensured on daily basis by the members, appointing the responsible focal points.
- In case of a central facility serving several external users responsible for sub-tasks (=transfer of responsibility), the risk of lacking the control on the activities themselves is envisaged.
- The current suite of VERSUS **hosted at ECMWF** in the framework of a dedicated Special Project, even though could represent an example of "central installation" for COSMO verification activities, **is far enough from being marked as an "operational"** (=reliable+prompt+upgradable) system and should not be considered here as prototype for the aforesaid central facility in subject.

After exchange of opinions during WG5 meeting (May 2015), it was decided that such a solution is not favourable thus it is not included in any proposed scenarios for the future verification tools strategy.

C V S +

EXCHANGE RULES

COSMO SW
SCM rules
Maintenance

OR

No technical
support
Documentation
Common
repository

Common Verification Software

(for all point-wise verification)

Deterministic and EPS model systems

VERSUS

DWD SW

Other
(e.g.MOVERO, MET)

Additional (Any Other) Verification Tools (AVT)

Fuzzy logic Spatial Methods

Displacement Object-oriented Spatial Methods

Feedback Files Use

Other

DIST

VAST

SpatialVx

DWD

Other

VERSUS (Main features)

Traditional verification scores for deterministic models (with the use of BUFR and ASCII obs)
Conditional Verification (CV)
Weather Dependant verification
Time series and Daily cycle plots
Upper air verification with TEMPs
Feedback files verification of vertical atmospheric structure
Confidence Intervals
Station-based post processed forecasts
EPS verification module (refinements are expected)

Constraints: Some important VERSUS features and refinements that concern score accuracy (EPS) were included in the last software version (30 July 2015) and not all test results, are not available yet. Spatial methods and verification based on gridded data methods are not included in the software. Few additional minor developments that concern deterministic verification are required for VERSUS. It is decided that no further developments are expected though VERSUS2 project, so these activities will have to be included (when possible) in the Long-Term Maintenance Plan or covered by other tools.

DWD (Main features)

The new verification environment at DWD is meant to replace the existing verification of numerical weather prediction systems against SYNOP observations currently running at DWD. It takes advantage of "**feedback files**" that are produced during the assimilation cycle.
An **R package** of useful functions has been developed aiming to exploit the information contained in feedback files and to quickly and reliably produce verification results. On top, an **interactive (R based) tool for the online visualization** of verification results has been developed. The most striking advantages of the new verification system are the shortfall of data pre-processing, the fast and simple calculation of standard verification scores and the interactive browsability and online production of results.

Constraints: The application of this software is based entirely on Feedback Files (FF) for surface and upper air. The production of FF will be possible with **MEC (Model Equivalent Calculator)** software. MEC just became available by DWD (04/09) but with unknown application complexity.
The timeline for complete software availability is 2016 (U.Pfluger will inform on status) ***Spatial methods are not currently included and the software does not allow for conditional verification applications as it is not linked with a RDB. Also EPS verification is currently not included.***

CVS

Scenario 1: VERSUS+VAST

Scenario 2: VERSUS+VAST with gradual migration to other CVS (DWD-SW) (2 year period)

Scenario 3: VERSUS+VAST+DWD SW (parallel use)

AVT

Scenario 1: Unrestricted exchange of tools to a common repository

Scenario 2: Development or adaptation of tools or verification code as requested from WG5 members. This can be realized as part of COSMO PP/PT(s)

For AVT both scenarios can coexist

Prerequisites/Implementation

- Successful completion of all **VERSUS2** Tasks and improvement of software performance
- 3 month intensive evaluation period after **VERSUS2** Task work completion
- Extension of **VERSUS** Maintenance Plan (I,III)
- **DWD SW** as CVS candidate, conditions and rules of exchange
- **DWD SW** test phase within WG5 (available in 2016)

Prerequisites/Implementation

- New methods once evaluated and requested by users, exchangeable in common repository (definition)
- **Scenario I**, no SCM rules apply or any obligation for technical support, I/O format (only documentation)
- **Scenario II**, AVT follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions).
- Application and adjustment on common I/O interface of already available software packages can be performed as subTasks of COSMO PPs (e.g. SpatialVx in INSPECT)

Prerequisites and Implementation Plan for CVS

✓ Successful completion of all VERSUS Tasks included in phase7 of VERSUS2 project.

✓ Improvement in VERSUS software performance as described in the optimization processes to be included in the preceding software version (final test results available September 2015).

✓Three (3) month intensive evaluation period from WG5 members after the completion of all Tasks of VERSUS2 project (September-November 2015).

✓Extension of VERSUS Maintenance Plan in the case of Scenario I, III. Additional minor required developments, to be included in the Maintenance Plan.

✓ For DWD software as CVS candidate, conditions and rules should be defined for being exchangeable. As it is strongly dependant on MEC software for input Feedback Files preparation, the adaptation as COSMO CVS will be relied on the decision of each service to move towards this approach for verification applications. This can introduce delays for WG5 activities and a parallel CVS usage should be considered for a considerable time period (MEC software just released, DWD software available end of 2016).

✓An evaluation testing phase of DWD-SW within WG5 is recommended.

CVS

Scenario 1: VERSUS+VAST

Scenario 2: VERSUS+VAST with gradual migration to other CVS (DWD-SW) (2 year period)

Scenario 3: VERSUS+VAST+DWD SW (parallel use)

AVT

Scenario 1: Unrestricted exchange of tools to a common repository

Scenario 2: Development or adaptation of tools or verification code as requested from WG5 members. This can be realized as part of COSMO PP/PT(s)

For AVT both scenarios can coexist

Prerequisites/Implementation

- Successful completion of all **VERSUS2** Tasks and improvement of software performance
- 3 month intensive evaluation period after **VERSUS2** Task work completion
- Extension of **VERSUS** Maintenance Plan (I,III)
- **DWD SW** as CVS candidate, conditions and rules of exchange
- **DWD SW** test phase within WG5 (available in 2016)

Prerequisites/Implementation

- New methods once evaluated and requested by users, exchangeable in common repository (definition)
- **Scenario I**, no SCM rules apply or any obligation for technical support, I/O format (only documentation)
- **Scenario II**, AVT follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions).
- Application and adjustment on common I/O interface of already available software packages can be performed as subTasks of COSMO PPs (e.g. SpatialVx in INSPECT)

Prerequisites and Implementation Plan:

✓Scenario I and II can be complementary

✓AVT products following new methods will be presented during WG5 annual meetings and once evaluated and requested by the users, necessary software will be exchanged in a common repository.

✓For scenario I, no SCM rules apply to AVT exchange neither any obligation for technical support. Only commitment is for AVT to be accompanied by adequate documentation.

✓Any software can become part of AVT, developed or adapted in the framework of a PP or PT if it follows certain standards (I/O common interface, scripting for adaptation to existing systems, installation on major linux distributions - Scenario II). This can include any application or adjustment of already available software packages (e.g. SpatialVx applications in INSPECT Task 2).

✓Necessary a common repository definition for all software included in the AVT concept.

I/O data format

Following the **experience of other consortia**, as well as **estimating the effort** that is usually devoted to the **adjustment of verification tools to local DB and I/O formats**, any future verification strategy should be closely related to this aspect.

• **Uniform I/O format standards for exchangeable data** (required from verification tools) should be agreed at least **for the AVT** (CVS can have already a non flexible architecture: e.g. VERSUS).

• Any required (by verification software) **data format adaptation** will have to be performed **externally**, separating in this way the main focus of the verification tools that is the correct application of statistical methods from the proprocessing of data or the graphical representation, that is the focus of other tools.

• Fieldextra is the official COSMO software and its use is suggested for the postprocessing of any gridded forecast or observation field, while LIBSIM and other tools can be utilized for other forms of observations (e.g. BUFR).

• **AVT software** suggested in the scenario II, should be followed by **I/O data format standards**, making in this way easier the adaptation of each tool from most users and permitting the exchange of data for common experiments (e.g. Task2 INSPECT).

According to the feedback obtained from the questionnaire on the verification tools, around 2.5 FTEs can be devoted from partners over the next 1-2 years for the implementation of the strategy (part is attributed to INSPECT project).

Implementation Plan

As VERSUS is proposed in all scenarions for CVS (short or long term) as VERSUS Final Phase of testing after final release is suggested– *STC decision is pending*

Simultaneously, practice on DWD verification Software application is necessary; Feedback from DWD on SW status, MEC installation and production of FF, Test phase Participants

Definition of Software Repository for AVT - Communication with CNMCA

Exchange of software and any verification tools as described in the strategy documentation, when is requested

INSPECT Task 2b "Adaptation of existing free verification packages (in particular, SpatialVx and VAST) to COSMO data and development of local tools with the aim to create scripts for applying the most widely used spatial methods" deliverables available to Common Repository

I/O format for verification tools to be discussed with involved SCA (VERSUS, FIELDEXTRA, VAST)

Current WG5 activities will be continued in the present framework until the final implementation of any new strategy for verification tools