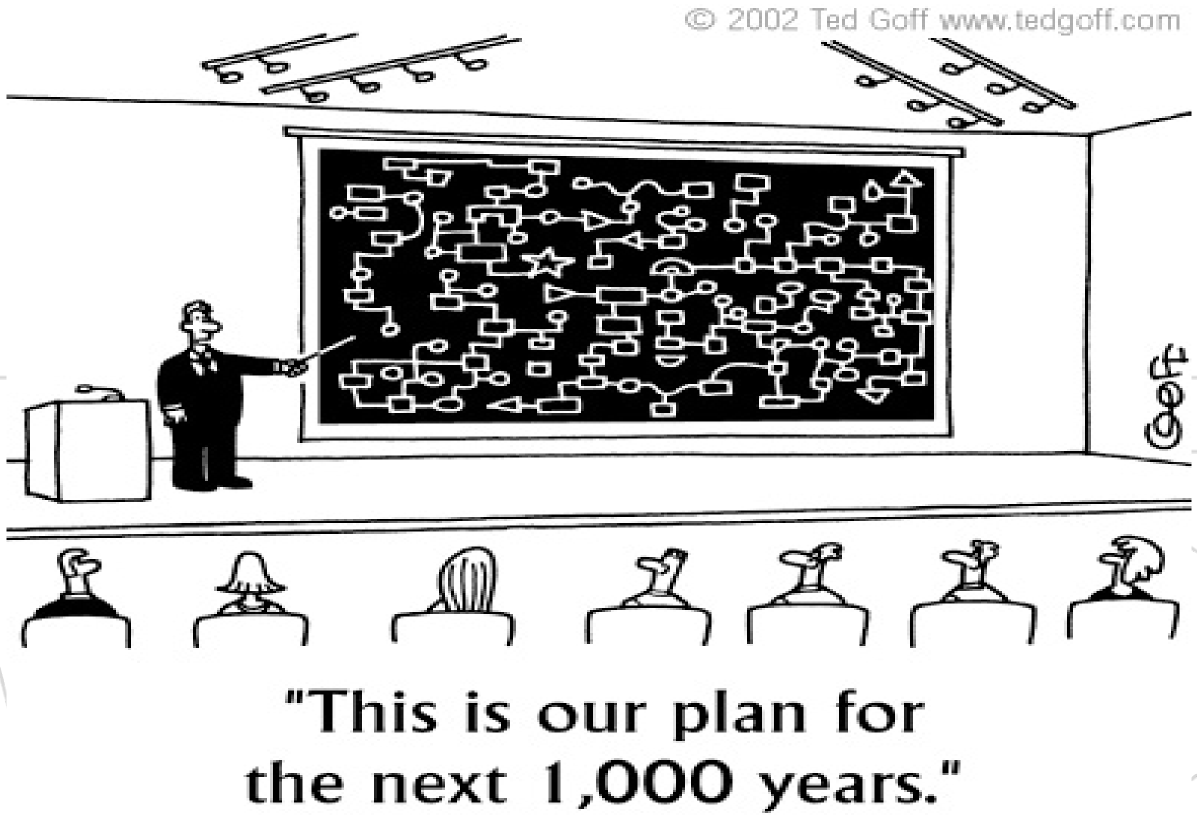
WG5: Planning a Science Plan...



WG5: Science Plan - Next Phase

"STC considers VERSUS as the tool for both operational and scientific verification and recommends that the verification strategy focuses on high-resolution and ensemble applications"

- Further developments based on the current 'state of the art of verification' of mesoscale models need to be addressed and decide on their suitability with COSMO strategy

- VERSUS project long term extension in the form of a priority task supervised by the SCA, will guide and coordinate the developments on a technical level that will be in line with the COSMO Science Plan

- Incorporation of tools developed outside the framework of the project (inside or outside COSMO community) are also needed due to limited human resources.

- Enhanced cooperation with other WGs for cross-cutting issues that mainly concern model development, needs to be addressed in the next phase of the COSMO Science Plan

- Alignment with national plans on verification developments

Work organization

At the end of 2013, the SP text should be ready to be given to external reviewers. For each WG, the SP chapter should contain (similar to the current version):

· ‘*state of the art, current developments*’: up to 2 pages

· ‘*status and expertise in COSMO*’: up to 1 page

· ‘*strategy of COSMO and actions proposed*’: up to 2 – 3 pages; this should include discussion of COSMO resources, involvement of external resources (if planned), expected outcome, and risk assessment

For each of the 3 remaining cross-cutting interdisciplinary issues, the SP should contain 1 -2 pages focusing on ‘strategy of COSMO and actions proposed’:

Verification for physics: first paper exists; WG3 coordinator will prepare draft.

SP Outline: Chapter on Diagnostics and Validation

Main areas were decided during VUS2013 and

the dedicated SMC in May

....minutes from the last SMC

WG5: FG reports about WG5 strategy; selected main points are:

- conditional verification and cooperation with WG3a for model improvement
- implementation of verification methods for convection-permitting models and EPS
- implementation of object-oriented and neighborhood methods
- use of new and more numerous observation data
- verification of the model for the important severe (and therefore rare) events
- besides verification of high impact weather, user-oriented verification and statistical inferences of verification scores are other future tasks

Discussion: AW notes that particularly for precipitation, statements on the quality of models may depend critically on the observations used to verify. The COSMO-CLM community has a quality controlled homogenized data set covering only Germany.
FG answers that a high-quality data set over Emilia Romagna is used in WG5, and for other applications, Uli Damrath has selected 90 stations of good quality for comparing results. It would be needed to correct the radar data by rain gauges.
PE notes additionally that user-oriented verification implies that COSMO has to produce corresponding user-oriented products, and this is related to the discussion on Tuesday how much post-processing should be within the scope of COSMO.

Decision: It is beyond the scope of COSMO to implement algorithms to correct the radar data by rain gauges. However, WG5 members should stimulate in their home services the production of such datasets and should exchange and use them.

Science Plan Overview of activities

Tackling model performance improvement issues through the use of conditional verification (CV)

Once model errors are estimated for a certain variable through "standard" verification, errors should be related to specific inaccurately simulated processes. This can only be achieved by verification under specific conditions (CV) that have to be chosen in order to extract as few as possible model uncertainties that can be due to simulation errors. This procedure aims to systematically isolate single processes or uncertainties that are primarily responsible for measured simulation errors. This procedure is consequently based on the selection of forecast products and associated "mask variables" (model variables, observations or external variables) and the possibility to formulate arbitrary thresholds (conditions) for the product verification. Further to this approach, some interaction with modelers is necessary in order to identify the most selective conditions and also to correctly interpret the results.

SHORT TERM FOCUS – Priotity for COSMO (ConSAT)

Statistical methods to identify the skill (added benefit) of convection permitting and near-convection-resolving model configurations

While it has been relatively easy to show an increase in skill when comparing forecasts from models with greater than 30 km to less than 20 km grid size it has been more difficult to do so for horizontal resolution greater than 10 km to so-called convection-permitting resolutions. This has been especially true for precipitation where the forecast detail is realistic but inaccurate, and for this reason neighborhood methods were employed to compare forecasts in appropriate selected size neighborhoods with the gridded radar data for precipitation. Experience has shown that even for continuous surface parameters, it is hard to consistently prove that higher resolution is more skillful. For continuous surface parameters a verification framework needs to be defined (even probabilistic, BS, RPS). As the effect of both double penalty and representativeness should be estimated, one approach could be to compare surface observations against forecast neighborhoods of varying size centered on the observation site.
Statistical methods proposed for decision-making should lead to the estimation of the relative skill gained when higher resolution is employed, to the assistance in the decision-making process for model upgrades for similar horizontal resolution and to the successful comparison between the determinist forecasts with ensemble ones.

Development of tools for probabilistic and ensemble forecast verification

The challenges in verifying "convection-permitting" ensembles are principally the same as in mesoscale "convection-parameterisation" ensembles, with some minor differences and some added complexities. Due to their nature, convection-permitting ensembles focus on the shortest range (0-24h) and large error growth in such systems which are correlated strongly to the highly non-linear physical processes of convection, thus verification measures must focus on the relevant gain of the use of such systems toward better representation of convection-based parameters. As for deterministic forecasts, neighborhood methods are proposed to be employed to account for the spatial mismatches between forecasts and observations, especially for precipitation, even though ensemble forecasts can address uncertainties of small-scale processes more adequately.

LONG TERM FOCUS

Exploitation of available observational dataset for operational and scientific purposes

For model-oriented verification, processing of the observation data needs to be done to match the spatial and temporal scales resolvable by the model. This requires the availability of high spatial resolution observations such as satellite or radar post-processed data that can be used to produce vertical profiles or gridded surface analysis. Finally, particularly important is the exploitation of any kind of existing, controlled and possibly homogenous set of observations, mainly in the PBL, concerning fluxes, radiation and soil characteristics, such as those available from the SRWNP Data Pool Exchange. Modellers would benefit greatly from these results in their efforts to improve and tune the model's physical aspects as with the employment of conditional verification.

SHORT TERM FOCUS

Severe and High Impact Weather

As there is an increased demand that meteorological services provide accurate forecasts of extreme weather, it is therefore important to be able to objectively evaluate the model performance in these cases. Scores that demonstrate the investment in forecast quality on extreme events prediction. Severe events are rare (top/bottom 5-10% of climatological distribution) and this is the reason that standard skill scores are not useful as they depend on base rate. Dependency scores like SEDS and SEDI have been extensively used by the NWP community for some time, but the use of other scores and methods will also be evaluated.

SHORT TERM FOCUS

User-oriented Verification products

With increasing model resolution, the number of products the users will ask, as well as their objective performance in terms of their expected quality is only going to rise. Different users might have needs for different verification information (e.g. administrative decisions may depend on model performance), so different verification strategies have to be chosen. It will be necessary to diversify verification methodologies to match the different needs and to this end, the scientific community will have to work more closely with the user community in the design of such verification strategies. These main activities could be reviewed and updated in the light of future developments in the main fields of model improvements concerning physics and data assimilation, in order to respond to the actual needs of developers and users alike.

Customer focused products?

SHORT TERM FOCUS

Verification of extreme, high-impact weather

- **EDS – EDI – SEDS - SEDI** ⇔ **Novelty categorical measures!**

Standard scores tend to zero for rare events

Event forecast	Event observed		
	Yes	No	Marginal total
Yes	a	b	a + b
No	c	d	c + d
Marginal total	a + c	b + d	a + b + c + d = n

H = a / (a+c), hit rate

F = b / (b+d), false alarm rate

p = (a+c) / n, base rate

q = (a+b) / n, relative frequency of forecasted events

$$\text{EDS} = \frac{\log p - \log H}{\log p + \log H} \qquad \text{SEDS} = \frac{\log q - \log H}{\log p + \log H}$$

[redacted]: Improved verification measures for deterministic forecasts of rare, binary events. *Wea. and Forecasting*

Base rate independence ⇔ Functions of H and F

$$\text{EDI} = \frac{\log F - \log H}{\log F + \log H}$$

Extremal Dependency Index - EDI

Symmetric Extremal Dependency Index - SEDI

$$\text{SEDI} = \frac{\log F - \log H - \log(1-F) + \log(1-H)}{\log F + \log H + \log(1-F) + \log(1-H)}$$

Availability of resources e.g.: need for development of new resources, resources outside COSMO (academia)

The experience gained in the past years indicates skilled but limited resources in the COSMO community regarding operational verification activities and implementation of new approaches and methodologies. For this reason, in order to optimize the available resources, a recommended strategy is to monitor the efforts of the various European Consortia (e.g. through SRNWP collaboration) in the field of verification, namely to use or adapt what has already been developed and exchange amongst the scientific and operational communities, new methodologies, research results and approaches to verification issues. The long term continuation of VERSUS project in the framework of PT-Support will contribute to the realization of the actions planned and will allow the monitoring of the development. Nevertheless in order to finalize efficiently the planned actions in the foreseen timeframe, it is necessary to increase in the next future the available human resources dedicated to verification activities.

Amended Work Group 5 Task List

1. Common Verification Framework

1.1 Operational Verification *Responsible: ALL*
1.2 Responsibility for Common Plots Reports *Responsible: J.Linkowska,IMGW*
1.3 Verification of vertical profiles using TEMP observations, aircraft data (AMDAR) and wind-profiler data *Responsible: ALL*
1.4 Dissemination of daily Grib model output Files *Responsible: De Morsier, MCH*

2. Exploitation of observational dataset for operational and scientific purposes

2.1 High density verification of precipitation over Italy *Responsible: E.Oberto, ARPA-PT*
2.2 Exchange of a common data set of non-GTS data DWD *Responsible:U.Damrath*
2.3 Evaluation of COSMO models in the lower PBL Responsible: Raspanti, Gofa, Kaufmann

3. Evaluation of convection permitting models performance

3.1 Long Term Trend Verification *Responsible: ALL*
3.2 Conditional Verification *Responsible: ALL*
3.3 Weather Dependant Verification (WDV) *Responsible: ALL*
3.4 Severe and High Impact Weather *Responsible:*

Amended Work Group 5 Task List

4. Neighborhood method techniques

4.1 Verification of COSMO-7 precipitation forecast using Radar composite network

Responsible: D. Leuenberger, MCH

4.2 Precipitation verification using radar composite network with neighborhood methods *Responsible: N. Vela, ARPA-PT*

5. Verification of EPS products (Cooperation with WG7)

6. Other

6.1 Annual Workshop/Tutorial on VERSUS2 & WG5