Feature-based verification of deterministic precipitation forecasts with SAL during COPS

Matthias Zimmer[1] *, Heini Wernli[1], Christoph Frei[2] and Martin Hagen[3]

1) Institute for Atmospheric Physics, University of Mainz
2) Federal Office of Meteorology and Climatology, MeteoSwiss Zürich
3) Institute for Atmospheric Physics, DLR Oberpfaffenhofen

1. Introduction

Verification is a central aspect of quantitative precipitation forecasts (QPFs) to assess their shortcomings and problems. The MAP D-PHASE dataset offers a unique opportunity to investigate the QPF performance of different deterministic regional numerical weather prediction (NWP) models. Due to the complex nature of precipitation fields the verification of QPFs is very challenging. Also for current high-resolution NWP models, including models without parameterization of deep convection, it is difficult to accurately capture observed convective rainfall events in time and space. When using classical grid-point based verification measures, errors in the timing and/or location of the events lead to the so-called "double penalty" problem. In order to measure the skill of QPFs in a more reasonable way, a feature-based or fuzzy verification technique is required. (See *Casati et al. (2008)* for a review of recently proposed verification approaches.) One of the newly developed QPF quality measures, referred to as SAL *(Wernli et al. 2008)*, will be used in the present study.
The goal of this study is to reveal the differences in the QPF performance of various NWP models that, among other factors differ in terms of horizontal resolution, treatment of deep convection and the influence of assimilated data. QPFs are taken from the MAP D-PHASE dataset during Summer 2007, when the COPS (Convective and Orographically-induced Precipitation Study) field experiment took place (*Wulfmeyer et al. 2008*). The area of interest is the German Part of the COPS region (see Fig. 1).

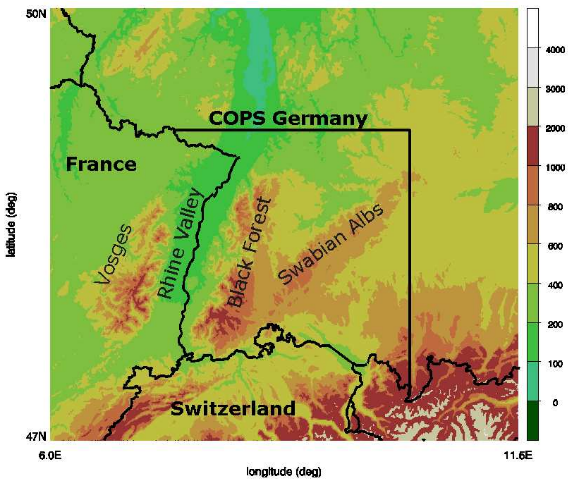

Figure 1: Topography of the German part of the COPS region (referred to as "COPS Germany") with the Black Forest, the Swabian Alb and the upper Rhine valley.

*zimmerm@uni-mainz.de

2. Data and methods

(a) Observations and model data

The observational precipitation field was obtained from about 150 rain gauges in the COPS Germany region, provided by the German Weather Service. These daily accumulated station values were transformed to a grid with a horizontal resolution of 7 km, using the gridding technique of *Frei and Schär (1998)*. More information about the gridding of the German rain gauge data can be found in *Paulat et al. (2008)*.

For this study, 19 deterministic models have been extracted from the D-PHASE "model zoo" (see Table 1), which can be divided into three categories: global models (1), medium-resolution local area models (LAMs) with parameterized deep convection (10), and high-resolution LAMs (7) where deep convection is explicitely calculated.

model	horizontal resolution	parameterization of deep convection	model category
ECMWF	~30 km	yes	global
MM5_60	60 km	yes	medium-resolution
MM5_15	15 km	yes	medium-resolution
QBOLAM11	11 km	yes	medium-resolution
ALADAT	9,6 km	yes	medium-resolution
ALADFR	9,5 km	yes	medium-resolution
MESONH8	8 km	yes	medium-resolution
COSMO-EU	7 km	yes	medium-resolution
COSMOCH7	7 km	yes	medium-resolution
LMEURO	7 km	yes	medium-resolution
LAMI7	7 km	yes	medium-resolution
COSMO-DE	2,8 km	no	high-resolution
LMITA	2,8 km	no	high-resolution
AROME	2,5 km	no	high-resolution
COSMOCH2	2,2 km	no	high-resolution
MESONH2	2 km	no	high-resolution
MM5_CT	2 km	no	high-resolution
MM5_4D	2 km	no	high-resolution
ISACMOL2	2 km	no	high-resolution

Table 1: List of the 19 models with their horizontal resolution (Δx in km), information about the parameterization of deep convection and the corresponding model category.

To calculate daily precipitation sums (from 06 to 06 UTC on the following day), the simulations started at 00 UTC (except for COSMO-DE and COSMOCH2 for which the forecasts starting at 00 and 12 UTC had to be combined) were used. Due to the fact that the 19 NWP models have different horizontal resolutions ranging from 2 to 60 km (see Table 1), the forecasts were transformed onto the same grid as the observations with a horizontal resolution of 7 km.

(b) The SAL technique

For the assessment of the QPFs of the 19 NWP models, the recently developed feature-based verification technique SAL (*Wernli et al. 2008*) has been used. SAL contains three independent components measuring the quality of the (S)tructure, (A)mplitude and (L)ocation of the QPF.

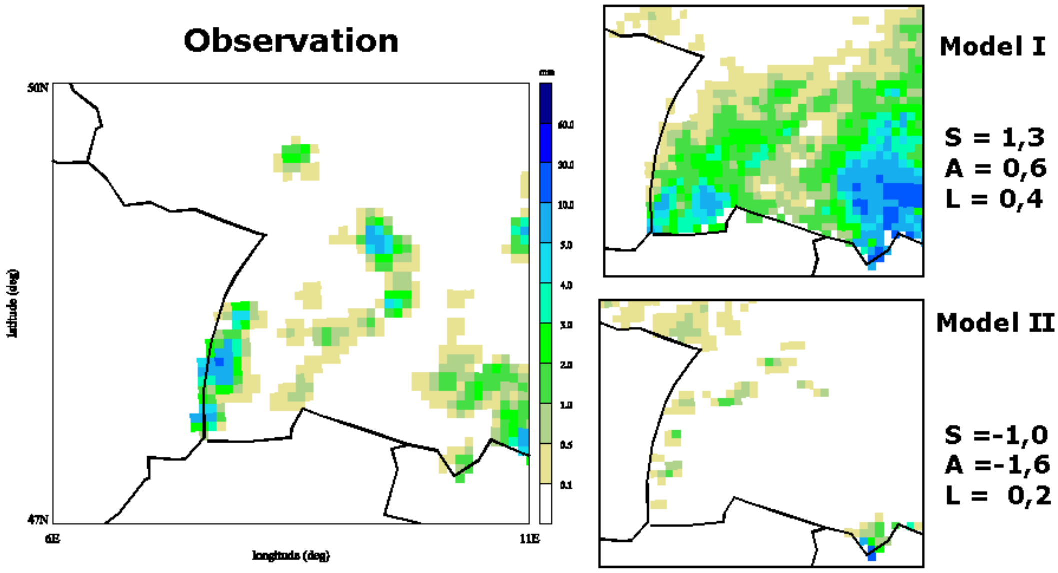

Figure 2: An example for the application of SAL. The left panel shows the observations in mm/day and the right panels show two forecasts, which score different in terms of SAL.

The structure component S investigates the size and shape of precipitation objects. It is defined in the range $[-2 \cdots +2]$, where negative values correspond to too small and/or too peaked objects, while positive values indicate too large and/or too flat simulated precipitation objects. Corresponding to SAL, a value $S = 0$ indicates a perfect structure of the QPF. The amplitude component A evaluates the total amount of precipitation in a predefined region. The values of A are within $[-2 \cdots +2]$, where 0 represents again the perfect value. Negative values of A correspond to too little and positive values to too much predicted precipitation, respectively.
The location component L quantifies the displacement of observed and simulated precipitation objects, relative to their overall centers of mass. The values of L are within $[0 \cdots +2]$ and also here 0 denotes the perfect value.

(c) An example for the application of SAL
To illustrate the working of SAL, an example is shown in Fig. 2 for southwestern Germany. The observations are shown on the left-hand side with a couple of distinct precipitation objects. The upper right panel shows the simulation from model I with intense and widespread precipitation. This leads to high positive values of $S = 1,3$ and $A = 0,6$, i.e. the amount of precipitation is overestimated as well as the typical size of the precipitation objects. The forecast from model II is shown in the lower right and this model simulates almost no precipitation. The negative values of $S = -1,0$ and $A = -1,6$ quantify the significant under-prediction of the amount of precipitation and the size of the objects.

3. Results

For concisely representing the results of all 19 models, the interquartile range (IQR) is plotted against the median in Fig. 3 for the SAL components S and A determined for all 92 QPFs in JJA 2007. In this diagram, a perfect model is represented by a point at the origin (median=0, IQR=0). In contrast, a random forecast, generated by using a “Monte-Carlo” approach to the set of observed precipitation fields (*Wernli et al. 2008*), has a median of zero with a high IQR (> 2).

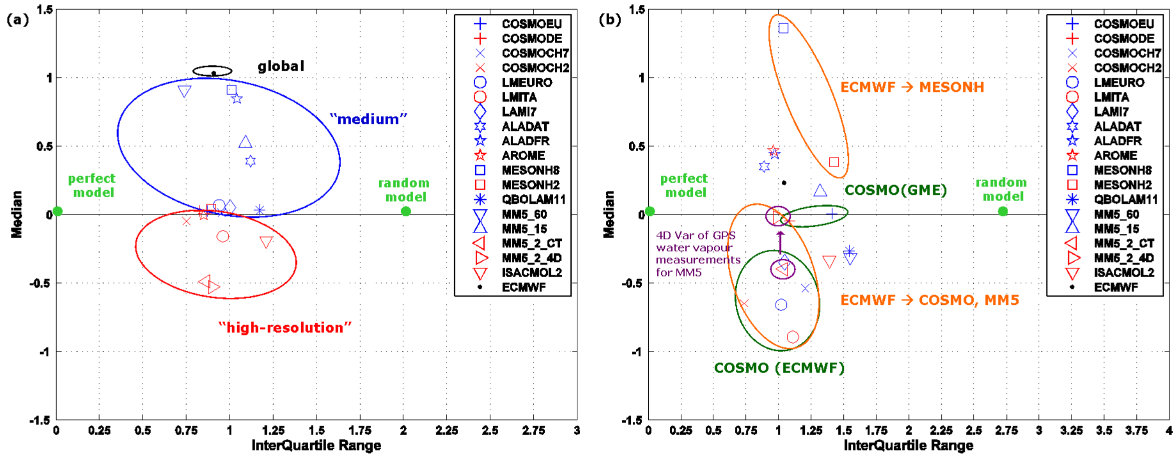

Figure 3: Interquartile range vs. median for the SAL components S (a) and A (b) for JJA 2007 in the COPS Germany domain. Black symbols denote global, blue medium-resolution and red high-resolution models, respectively.

For S, the results for the three model categories are well separated: The medium-resolution models (blue) show a tendency to predict too large and/or too flat rain objects (median of $S \geq 0$), while the high-resolution models (red) simulate too small and/or too peaked rain objects (median of $S \leq 0$). Most of the high-resolution models have a significantly smaller IQR than the medium-resolution models. The global model produces the largest median of S, but a fairly small IQR.
The picture for the A component (Fig. 3b) is less systematic, but three particular aspects are apparent. The first one (dark green) is related to the performance of the COSMO model driven by different global models. The COSMO model driven by the GME (global model from the German Weather Service) simulates the amount of precipitation better than if driven by the ECMWF global model, as revealed by the differing medians. The second aspect corresponds to the fact that different LAMs (orange) driven by the same global model (ECMWF) perform differently. For the MM5 and COSMO models, the simulated amount of precipitation is on average underestimated, while for MESONH the amount is overestimated (MESONH is the only model using a two-way nesting method). The third aspect (purple) is the positive effect of the 4D-Var assimilation of GPS water vapor measurements in the MM5 model for the amplitude component of SAL. In contrast, both with and without assimilated GPS data, precipitation objects in the MM5 forecasts are too small and/or too peaked (as indicated by the negative median of S).

4. Summary & Conclusions

Daily QPFs from a set of 19 models have been verified with the recently developed feature-based technique SAL in the German part of the COPS area during summer 2007. The investigation reveals some important differences between medium-resolution models with parameterized deep convection and high-resolution convection permitting models. On average, the high resolution models simulate the structure of the precipitation field more realistically than the medium-resolution models (as shown by the smaller IQR of S), however some of them show a systematic underestimation of the structure component (median of $S \leq 0$). The quality of the simulated precipitation amount is linked to the following three characteristics. The COSMO model driven by the GME (median of $A \simeq 0$) predicts the amount of precipitation better than if driven by the ECMWF (median of $A < 0$). The IQR of A is smaller for the high-resolution models. For the MM5 model, the use of 4D-Var assimilation of GPS

water vapor measurements leads to a better estimation of the precipitation amount but not of the structure of the precipitation objects.
Probably for the first time such a large dataset of forecasts from different models has been available for a thorough validation of QPF performance. It is interesting that systematic differences occur between models that parameterize deep convection (here referred to as medium-resolution models) and models that don't parameterize deep convection (here referred to as high-resolution models) - if using a feature-based verification measure like SAL. In particular when considering the structure of a precipitation field, the first results from the new generation of convection-permitting models are promising. However, one should keep in mind that this study focused on a relatively small mountaineous region in southwestern Germany during one particular summer season. Further work is required to assess the robustness of these results and to quantify the QPF quality of the different model categories in other regions and during other seasons. Also, the considered summer season in 2007 was characterized by strongly contrasting synoptic weather conditions (*Zimmer and Wernli 2008*) and it will be rewarding to compare the models' QPF performance for different weather regimes.

Acknowledgments: This work has been funded by the priority program "Quantitative Niederschlagsvorhersage" (SPP-1167) of the German Science Foundation (DFG). We are thankful to the German Weather Service (DWD) for providing the rain gauge data.

References

Casati, B., L. J. Wilson, D. B. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. E. Ebert, B. G. Brown and S. Mason (2008). Forecast verification: current status and future directions. *Met. Apps.,* **15,** 3-18.

Frei, C. and C. Schär (1998). A precipitation climatology of the alps from high-resolution rain-gauge observations. *Int. J. Climatol.,* **18,** 873-900.

Paulat, M., C. Frei, M. Hagen and H. Wernli (2008). A gridded dataset of hourly precipitation in Germany: its construction, application and climatology. *Meteorol. Z.,* in press.

Wernli, H., M. Paulat, M. Hagen and C. Frei (2008). SAL – a novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.,* **1**36, 4470-4487.

Wulfmeyer, V., A. Behrendt, H.-S. Bauer, C. Kottmeier, U. Corsmeier, A. Blyth, G. Craig, U. Schumann, M. Hagen, S. Crewell, P. Di Girolamo, C. Flamant, M. Miller, A. Montani, S. Mobbs, E. Richard, M. W. Rotach, M. Arpagaus, H. Russchenberg, P. Schlüssel, M. König, V. Gärtner, R. Steinacker, M. Dorninger, D. D. Turner, T. Weckwerth, A. Hense and C. Simmer (2008). The Convective and Orographically-induced Precipitation Study: a research and development project of the World Weather Research Program for improving quantitative precipitation forecasting in low-mountain regions. *Bull. Amer. Meteor. Soc.,* **89,** 1477-1486.

Zimmer, M. and H. Wernli (2008). COPS Atlas – The meteorological situation from June 1 till August 31, 2007. Internal Report No. 1, Institut für Physik der Atmosphäre, Universität Mainz, 100 pp.