Verification of precipitation forecasts of operational DWD-models against radar data

Ulrich Damrath, Deutscher Wetterdienst
Ulrich.Damrath@dwd.de

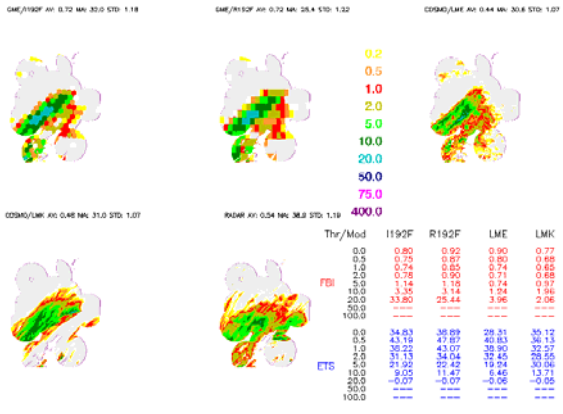
➔ Basic data

➔ Traditional verification

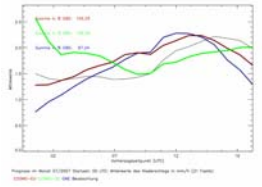
- Diurnal cycles (observation, forecasts of GME, COSMO-EU and COSMO-DE) starting at different forecast times
- Contingency tables

➔ „Fuzzy“-type verification

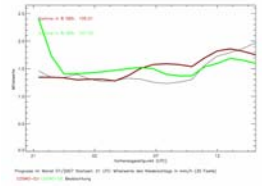
- Application of B. Eberts package
- Coupling intensity scale method and fraction skill score



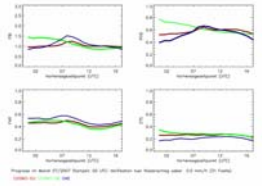
Basic data for comparison with RADARs (example with some missing observations)



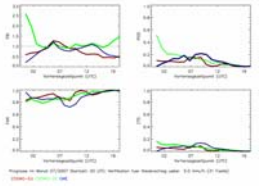
Diurnal cycle of precipitation for forecasts starting at 00 UTC July 2007



Diurnal cycle of precipitation for forecasts starting at 21 UTC July 2007



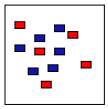
Diurnal cycle of precipitation for forecasts starting at 00 UTC July 2007 - Scores precipitation > 0mm/h



Diurnal cycle of precipitation for forecasts starting at 00 UTC July 2007 - Scores precipitation > 5mm/h

The framework for fuzzy verification

1. Select a set of scales with indeces s=1, 2,..., S and event intensity thresholds with indeces k=1, 2,..., K over which to compute the fuzzy verification results
2. For each scale s:
   a. collect the gridded forecasts within the window of scale s surrounding the observation.
   b. For each intensity threshold k compute scale-dependent quantities $(\ldots)_k$ according to various decision models
   c. For each intensity threshold k compute the desired verification scores over the domain. and event intensity thresholds with indeces k=1, 2,..., K over which to compute the fuzzy verification results.



Observation

Forecast

"Fuzzy" verification methods

(From E. Ebert, Presentation given at a workshop in Zurich November 2006)

- First (?) suggested by H. Brooks at 1998 Mesoscale Verification workshop
  - Brooks et al. (1998)
  - Zepeda-Arce et al. (2000), Weygandt et al. (2004)
  - Atger (2001)
  - Damrath (2004)
  - **Casati et al. (2004)**
  - Germann and Zawadski (2004)
  - Rezacova et al. (2005)
  - Theis et al. (2005)
  - **Roberts and Lean (2005)**

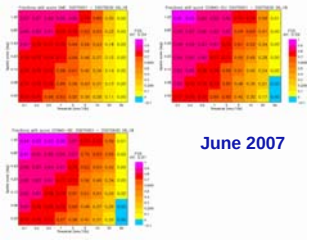
COSMO-DE: FBI=2.5 POD=0.55 FAR~0.8 **FAR=0.78**
COSMO-EU: FBI=0.6 POD=0.1 FAR~0.8 **FAR=0.83**

$$FBI = \frac{A+B}{A+C} \qquad FBI = \frac{POD}{1-FAR}$$

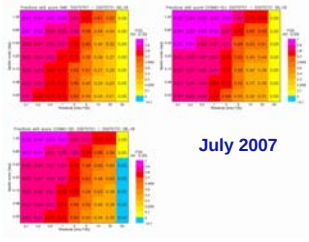
$$POD = \frac{A}{A+C} \qquad POD = FBI\,(1-FAR)$$

$$FAR = \frac{B}{A+B} \qquad FAR = 1-\frac{POD}{FBI}$$

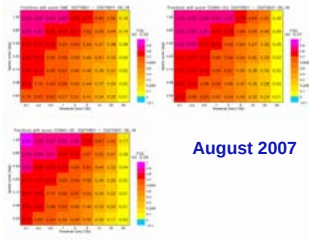
„Fuzzy“-type verification for 12 h forecasts (vv=06 till vv=18) starting at 00 UTC
- fraction skill score



June 2007

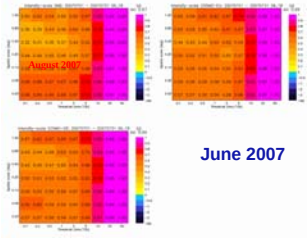


July 2007

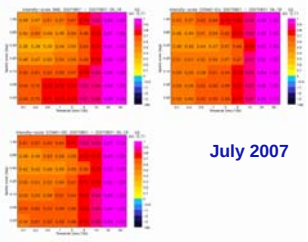


August 2007

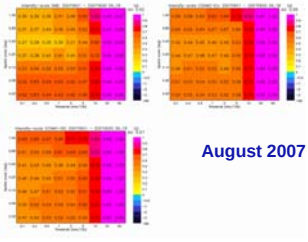
„Fuzzy“-type verification for 12 h forecasts (vv=06 till vv=18) starting at 00 UTC
- intensity scale skill score



June 2007



July 2007



August 2007

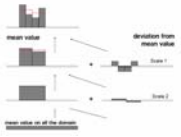
... yet another score

N. Roberts - Fraction skill score (summing up over all scales)

$$FSS = 1 - \frac{FBS}{\frac{1}{N}\left[\sum_{i=1}^{n}(\log)^2 + \sum_{j=1}^{n}(\log)^2\right]}$$

Problem:
Smooth fields give best results concerning rmse-related scores!

B. Casati - Intensity scale skill score (real scale separation)



B. Casati: energy squared Skill Score

1. Energy squared:
$$En^2(X) = \overline{X^2}$$

2. MSE Skill Score:
$$1 - \frac{MSE(Y,X)}{En^2(X) + En^2(Y)}$$

Modification by U. Damrath in order to have the same structure as the FSS

2. MSE Skill Score:
$$1 - \frac{MSE(Y,X)}{En^2(X) + En^2(Y)}$$

$$WSS[scale] = \frac{2*COV(OBS[scale],FCT[scale])}{VAR(OBS[scale])+VAR(FCT[scale])}$$

$$WSS[scale\ x] = 1 - \frac{VAR(ERR[scale\ x])}{VAR(OBS[scale\ x])+VAR(FCT[scale\ x])}$$

$$WSS[scale\ 1] = 1 - \frac{VAR(ERR[scale\ 1])}{VAR(OBS[scale\ 1])+VAR(FCT[scale\ 1])}$$

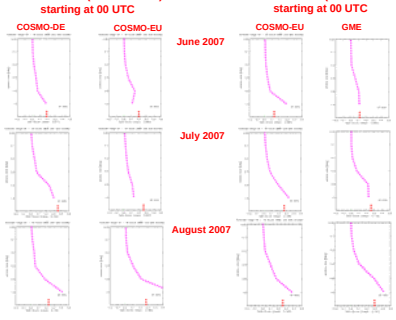
$$WSS[scale\ 2] = 1 - \frac{VAR(ERR[scale\ 2])}{VAR(OBS[scale\ 2])+VAR(FCT[scale\ 2])}$$

$$WSS[mean\ value] = BIAS \quad or \quad ABSE$$

„Fuzzy“-type verification - scale dependent skill score

12 h forecasts (vv=18 -vv=06) starting at 00 UTC

24 h - forecasts (vv=30 -vv=06) starting at 00 UTC



Summary

➔ Traditional verification

- pronounced spin down effect in COSMO-DE which is perhaps connected with some problems during latent heat nudging
- regardless of double penalties - COSMO-DE shows advantages against the other models although there is a frequency bias up to 3 for the first 3 to 5 forecast hours.

➔ „Fuzzy“-type verification

- With fuzzy methods (fraction skill score, intensity scale skill score and a new score based on FSS and ISS) it can be shown that on horizontal scales of around 15 km the limit of predictability is reached.