Using the MAP D-PHASE database to evaluate the QPF
Improvements of the new SIMM's BOLAM

ISPRA
Istituto Superiore per la Protezione e la Ricerca Ambientale

Stefano Mariani and Marco Casaioli, simm-pre-meteo@isprambiente.it

ISPRA – Institute for Environmental Protection and Research, Rome, Italy

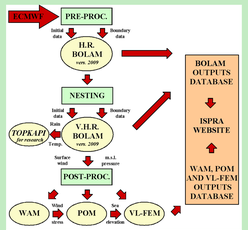
European Geosciences Union General Assembly 2012
Vienna | Austria | 22 – 27 April 2012

BACKGROUND & AIM

- *The improvement of MET forecasts is one of the primary goals of any hydro-meteorological or environmental institution running a NWP model.*
- *MET forecasts can be improved, e.g., by implementing more accurate and advanced physical parameterizations or by providing HI-RES (in time and space) initial and boundary conditions.*

*An **intercomparison** study over a long time period is necessary to statistically evaluate such performance improvements.*

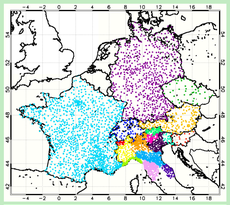
- *A fully updated version of the 0.1° BOLAM MET model is currently implemented into the **ISPRA hydro-meteo-marine forecasting system SIMM** (Speranza et al. 2007).*
- *In addition, experiments on a newer BOLAM version and on different model configurations (HI-RES initial and boundary conditions; different nesting design; increase of the domain size; and decrease of the horizontal grid step) are ongoing.*
- *QPF improvements of such new versions are evaluated with respect to the previous operational version.*



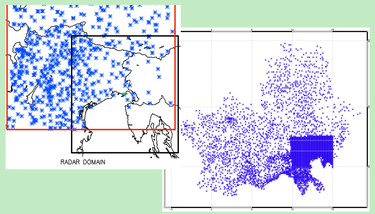
SIMM chain

DATASETS

- *Rain gauge data collected during the Operations Period (DOP: Jun-Nov 2007) of the WMO WWRP project "**MAP D-PHASE**" are considered as observational dataset.*
  - → *Observatinal analysis through a **two-pass Barnes scheme**.*
- *Radar data were also collected after DOP to provide further information on selected case studies (**25–28 Sep. & 22–25 Nov. 2007**).*
  - → *Radar data and rain gauge are combined through a Bayesian-based approach: **RainMusic**.*



DOP db: ca. 3900 rain gauge stations



Fossalon radar (NE Italy – ARPA FVG) & 0.1° verification grid

- *Forecast series: the one originally provided during DOP (QBOLAM); the one obtained (reforecast) with the current operational version (BOLAM11); the ones related to the model experiments.*
  - → *Forecasts remapped over common verifications grids (0.1°; 0.3° and 0.5°).*

METHODOLOGY

- ✓ *Representativeness/structure and scales of the fields compared have been addressed through a spectral analysis (Göber 2008; Lanciani et al. 2008; Weygandt et al. 2004, Chèruy et al. 2004).*
- ✓ *Categorical scores and skill scores (e.g., Wilks, 2006) are calculated over a sum of daily contingency tables (CT) w.r.t. a set of given thresholds.*

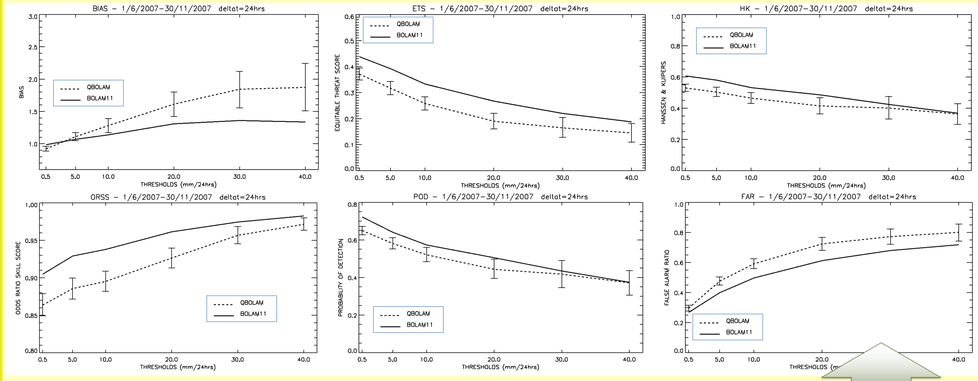
$$\text{BIAS} = \frac{a+b}{a+c}$$

$$\text{ETS} = \frac{a-a_r}{a+b+c-a_r} \quad \text{with} \quad a_r = \frac{(a+b)(a+c)}{a+b+c+d}$$

$$\text{HK} = \frac{(ad-bc)}{(a+c)(b+d)} = \text{POD} - \text{F} = \frac{a}{a+c} - \frac{b}{b+d}$$

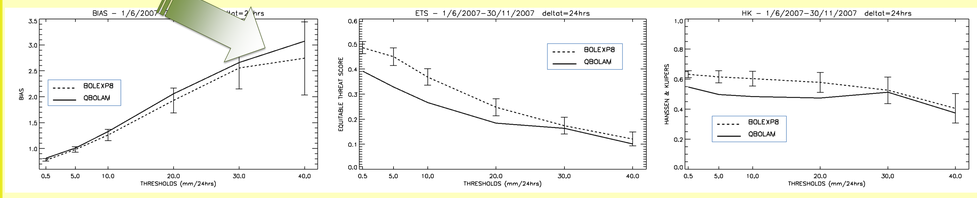
$$\text{ORSS} = \frac{\text{ODDS}-1}{\text{ODDS}+1} = \frac{ad-bc}{ad+bc} \quad \text{where} \quad \text{ODDS} = \frac{ad}{bc}$$

- ✓ *ROC (deterministic) curves (Mason, 1982).*
- ✓ *Bootstrap-based hypotesis test (Hamill, 1999) to provide the score differences between two "competing" models with confidence intervals.*
- ✓ *Geographical mapping (on a 0.5° grid) of CT elements to provide a physical interpretation of the scores.*
- ✓ *Case-study approach: eyeball subjective verification + objected-oriented approach.*

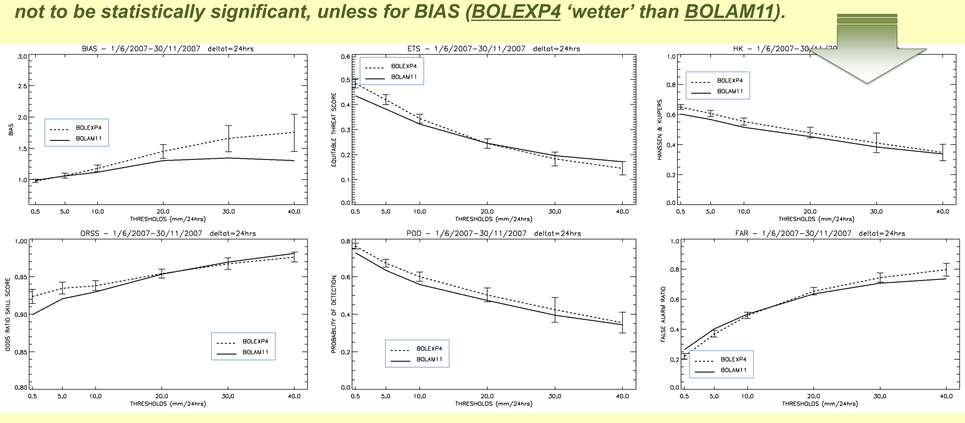


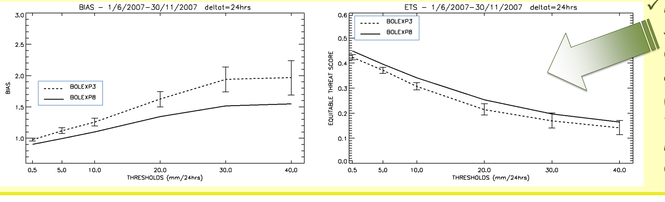
SKILL SCORE COMPARISON

- ✓ *BOLAM11 (solid line) statistically performs better than QBOLAM (dashed line) over a 0.1° verification grid. Since QBOLAM spectra have more small-scale structure than BOLAM11 ones, a fair comparison should be done on a coarser (0.5°) verification grid.*
- ✓ *On a 0.5° grid, BOLAM11 still performs better than QBOLAM in terms of ETS and HK, at least at the low-medium thresholds (not shown). The same result is obtained when comparing QBOLAM against BOLEXP8. However, forecasts remapped over a 0.5° grid result to be wet (BIAS values > 1).*



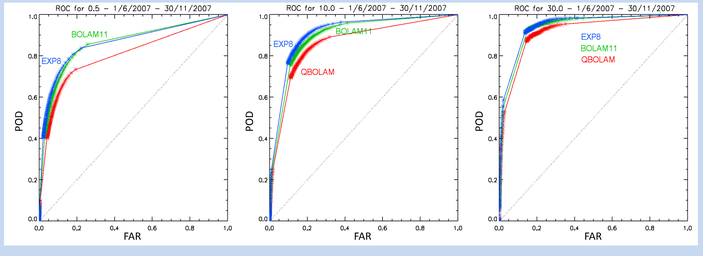
- ✓ *It is also under investigation the extension in LON-LAT of the model domain together with the decrease of the grid size (0.07°) and the use of HI-RES initial and boundary conditions (BOLEXP4). From the preliminary results, score differences between BOLAM11 (solid line) and BOLEXP4 (dashed line) seems not to be statistically significant, unless for BIAS (BOLEXP4 'wetter' than BOLAM11).*



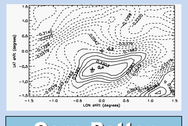


- ✓ *Differences result to be instead statistically significant when comparing BOLEXP8 forecasts against BOLEPX3 forecasts (obtained using the newer version, HI-RES initial and boundary conditions and the original nesting configuration).*

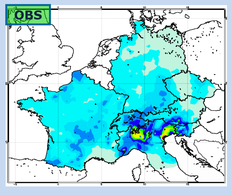
ROCS & CT GEO-LOCALITATION & CRA ANALYSIS

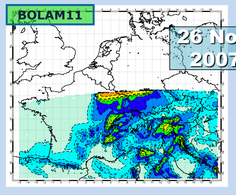


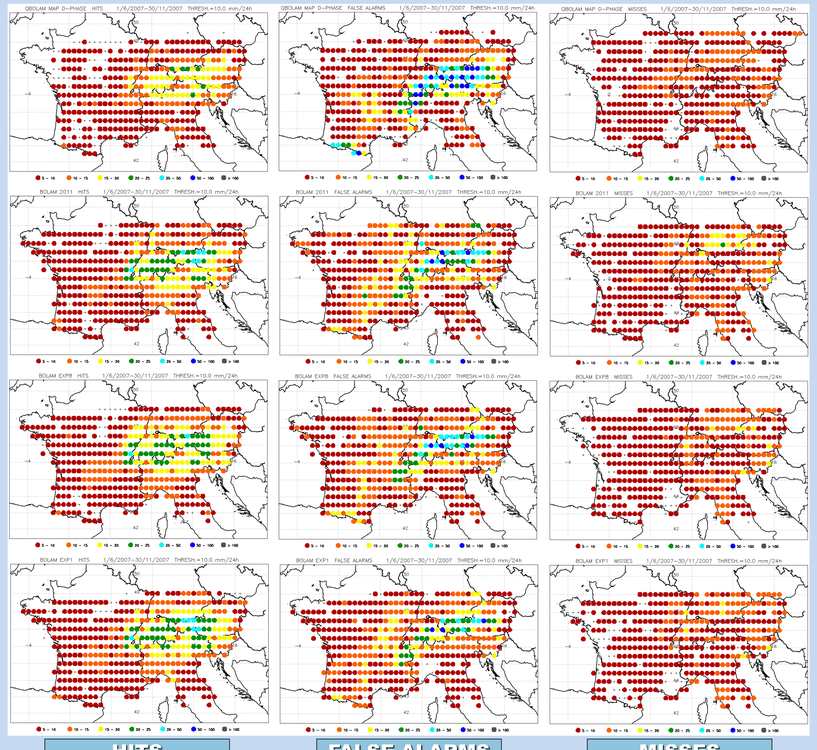
- ✓ *ROC: Increase in model performance moving from QBOLAM to BOLAM11. A slight increase is observed w.r.t. forecasts obtained with the newer BOLAM version, using the same model configuration (BOLEXP8).*
- ✓ *CT geo-location: improvement in BOLAM11 QPF quality especially over the previously-critical areas (high mountains; Accadia et al. 2005) and heavily-flooded areas (NE Italy). When considering the BOLEXP8 forecasts, it is observed an increase in terms of 'HITS' and a decrease in terms of 'MISSES'. A reduction of the 'FALSE ALARMS' is not so evident. BOLEXP1 forecasts – obtained using a newer BOLAM version, HI-RES initial and boundary conditions and a different nesting configuration (no "father" model) – shows a CT geo-location similar to those of BOLEXP8.*



Corr. Patter match 5.0 mm 24–1 thres.







HITS | FALSE ALARMS | MISSES