Results of precipitation verification over Italy

E. Oberto, M.Turco ARPA Piemonte – Area previsione e monitoraggio ambientale, Torino, Italy

Corresponding authors: elena.oberto@arpa.piemonte.it, marco.turco@arpa.piemonte.it

The verification of forecasts, especially of the QPF (Quantitative Precipitation Forecast), is one of the most important activities of ARPA Piemonte because it allows a better understanding of how the models behave in different meteorological situations, highlights the inner characteristics, and helps in the evaluation of the reliability, whether in average or extreme terms or over the long term and in the current situation. In order to verify the quality of any forecasting model, there must be the highest possible number of observed data so that a comparison is carried out over a long enough period of time and over a wide enough area which can be considered a statistically valid sample.

The ARPA Piemonte decentralised Centro Funzionale makes use of the global IFS model of ECMWF, which is distributed through the national meteorological centre of the Military Air Force, and the Italian version of the COSMO Model, known as LAMI (Local Area Model Italy). This model was developed in the framework of COSMO Consortium among Germany, Switzerland, Italy, Greece and Poland (COnsortium for Small-scale MOdelling, www.cosmo-model.org). In this work its Italian version called LAMI (Limited Area Model Italy), maintained and developed by USAM, ARPA-SIM and ARPA Piemonte, has been taken into account and studied. It has two runs (at 00UTC and 12UTC) and a horizontal resolution of 0.0625° (about 7.5 Km) and is used operationally by the Piedmont Meteorological Service.

Verification aims at ARPA Piemonte

The alert system is based on QPF over meteo-hydrological basins in terms of mean areal and maximum value expected, therefore the model precipitation field plays a fundamental role into the meteo-hydrological risk assessment over the national territory. So, the main verification purposes are oriented to estimate the model improvement in terms of underestimation/overestimation, accuracy into space-timing detection and capability to predict correctly strong and poor rainfall.

In this study we carry out the QPF verification over a common dataset, the high resolution network of rain gauges coming from COSMO dataset and Civil Protection Department, about 1300 stations. We calculate the model's skills and scores considering the 24h averaged cumulated precipitation value (observed and forecasted respectively) over 90 meteo-hydrological basins.

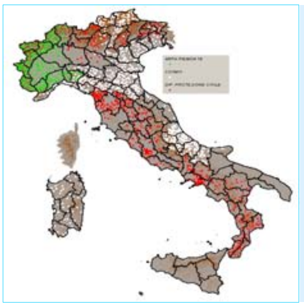


Fig. 1. Common rain gauges dataset with meteo-hydrological basins: red dots are the Civil Protection Department network, white dots are the common COSMO network, green dots are the Piedmont regional network

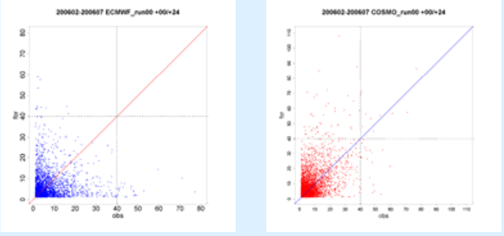


Fig. 2. Scatter plots : 24h cumulated average precipitation over each basin. Period: 6 months from 200602 to 200607. Models: ECMWF run00, COSMO run 00, first 24h (+00 UTC/+24 UTC).

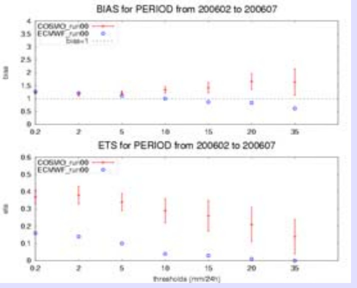
Scatter plot

We choose to show the first 24h (+00 UTC/+24 UTC) scatter plot but similar remarks can be reported also for the second ones. On the one hand, we see for ECMWF (fig. 2 plot on the left) that many points spread close to the axis both for small and large precipitation amounts. On the other hand we see for COSMO model (fig 2 plot on the right) a minor spread of the points, closer to the bisector, but with a general overestimation tendency.

Contingency table and statistical indices

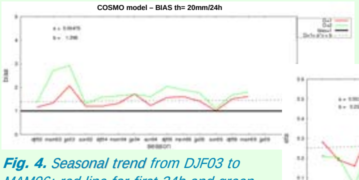
We calculate the statistical indices (fig. 3) considering 24h cumulated average observed/forecast precipitation over the basins versus increasing thresholds. The Hamill Hypothesis test (bootstrap resampling technique) is used to evaluate if the model's performance differences are statistically significant.The error bars (confidence interval) indicates 2.5th and 97.5th percentiles of resampled distribution, applied to the "reference" model. Till 5mm the BIAS differences are not statistically significant and both models tend to overestimate. Above 5mm the gap between the models is more pronounced: ECMWF has a good skill till 15mm then it has a underestimation, COSMO has an increasing overestimation trend. The ETS skill has a statistically significant difference: we obtain a better results for COSMO with respect to ECMWF, but both models worsen with increasing thresholds.

Fig. 3. Statistical indices (BIAS,ETS): 24h cumulated average precipitation over each basin. Period: 6 months from 200602 to 200607. Models: ECMWF run00, COSMO run 00, first 24h (+00/+24), second 24h (+24/+48).



Focus on COSMO model

In this part we consider a long data period to study and evaluate the long term COSMO performance and features. In the following pictures (fig. 4) we show the seasonal trend of statistical indices, starting from winter 2003 to spring 2006, both for the first (red line in the plot) and the second day (green line in the plot), having chosen a high fixed thresholds of 20mm (the most representative). There is no significant trend in time but instead a seasonal periodicity, in which the higher overestimation occurs during the summer and the better performance are obtained in autumn; moreover it appears to have a more or less constant BIAS with respect to the forecast time, and a slightly improving ETS. Additionally, we report in fig. 5 an example of the cumulated average (over each basin) precipitation seasonal maps, produced and analysed in order to deduce some general "climatological" considerations over the alert areas and different domains.



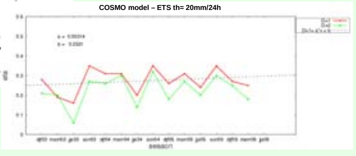
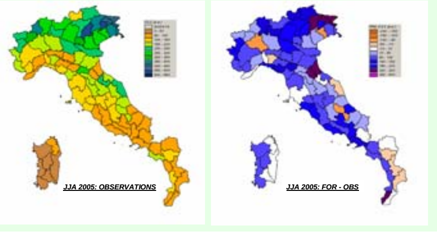


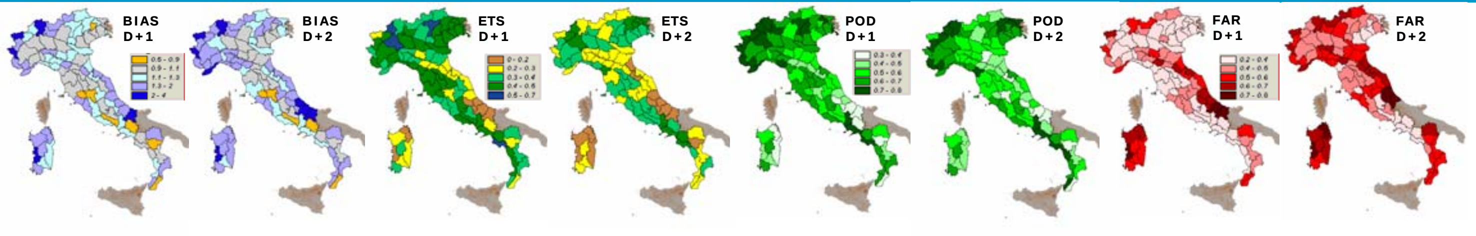
Fig. 4. Seasonal trend from DJF03 to MAM06: red line for first 24h and green line for second one.

Special study of Cosmo model: SKILL=SKILL(forecast time, threshold, space, type and location of the region...)

Fig. 5. Cumulated mean areal seasonal maps (JJA '05): in the left picture is highlighted the pluviometric regime with respect to the season and the territory, the right picture represents the model over/underestimation error in terms of quantitative amount.



Error spatial distribution



Finally, in the last image sequence is shown the spatial error distribution over a long period starting from winter 2003 to spring 2006. We plotted BIAS, ETS, POD and FAR over each of the Italian meteo-hydrological basins with respect to a fixed thresholds of 10mm (the minimum to reach a sufficient statistics) for the first and the second 24h respectively, to evaluate the long term behaviour of the model with respect to territory and orography. For example, there is an overestimation over most of the basins, especially over the mountain areas and Central Italy, but in general a quite good ETS. We obtain an high POD especially in north-western Italy and more false alarms over the alpine basins and over Central Italy. It is noticeable a deterioration for D+2 in all the scores.

Recommendation on verification practices

In order to evaluate the model performance it is important to have a high-representative network, in terms of spatial resolution and high data quality, to better depict the reality. Then, to study the model behaviour, it is necessary to consider the skill as a function of the threshold, forecast time, aggregation time, seasonal, monthly or daily periods, and with relation to the orography and territory. The skill scores and methodologies chosen must be targeted to the specific verification purpose.

References

(1) Accadia C., Casaioli M., Mariani S.,Lavagnini A., Speranza A., De Venere A., Inghilesi R., Ferretti R.,Paolucci T., Cesari D., Patruno P., Boni G., Bovo S., Cremonini R. "Application of a statistical methodology for limited area model intercomparison using a bootstrap technique" Il Nuovo Cimento, vol. 026, Issue 01, p.61, 2003.

(2) Hamill, T.M. "Hypothesis tests for evaluating numerical precipitation forecasts" Wea. Forecasting, 14, 155-167, 1999.

(3) Turco et al. COSMO newsletter n.5 "Comparative high resolution verification over Northern Italy"