Statistical properties and validation of Quantitative Precipitation Forecast

Maria Stefania Tesini, Carlo Cacciamani, Tiziana Paccagnella

Hydro meteorological Regional Service ARPASIMC, Bologna, Italy
email: mstesini@arpa.emr.it

1 Introduction

Observed precipitation fields show a high variability both in space and time and the amount of rainfall could vary a lot within a short distance.(Zepeda-Arce et al.,2000). The increasing of horizontal resolution in NWP models seems to enable them to reproduce this variability, even if frequent errors in time and space positioning make difficult a grid-point based employment of models QPF. In order to asses the ability of the models in reproducing the variability of the precipitation fields, we investigated the statistical properties of the observed and forecasted rain values falling within a predefined geographical area and in a specific time period (also called boxes). In particular we studied the distribution function (*pdf*)and evaluated some summarizing quantities, such as the mean, the maximum value and quantiles for each of the selected box. Results for different size of the chosen areas and period of time are used to validate the QPF of the COSMO suites that run operationally at ARPA-SIMC (COSMO-I7 and COSMO-I2) in comparison with the global model IFS-ECMWF.

The aim of this study is to provide an interpretation key for more correct use of models QPF, especially with respect to high precipitation events.

2 Dataset and methodology

Observed precipitation data consisted of about 1500 rain-gauges made available by the Italian National Department of Civil Protection and, in the framework of the COSMO Cooperation, by a number of Italian regions and Meteo-Swiss. The dataset cover almost all the Italian territory, (see figure 1) even if it is not homogeneous both in space and time. In this work we considered only precipitation values accumulated over 24h , starting at 00 UTC or, in same cases , at 6 UTC. The QPF field investigated in this study is provided by COSMO-I7 (7 km horizontal resolution), COSMO-I2 (2.8 km horizontal resolution) and the global model IFS-ECMWF (25 km horizontal resolution, and 50 km horizontal resolution before 2006). In most of the cases the 00 UTC runs have been taken into account, and the grid-point forecasted values refer to the +24h (or +30h) of integration, even if in same tests (not shown in this work) we considered also the precipitation from +24h to +48h. In order to compare the observations with the forecasted rain field we devised a strategy that takes into account all the observed values and grid-point forecasted values falling into a predefined geographical area, several times wider than the horizontal resolution of the model. As a starting point ,the domain is dived in squared areas, with size and position chosen subjectively and depending on the available number of observation stations and geographical features of the territory. We assumed that the precipitation occurrence in each of the stations points and grid-points within the selected area is equi-probable. The size and the dimension of the area play a crucial role in order to consider well-founded this assumption. In this preliminary study we followed a pragmatic approach: we should have enough forecast and observation points to perform significant statistics in each box. The precipitation values of all stations and all

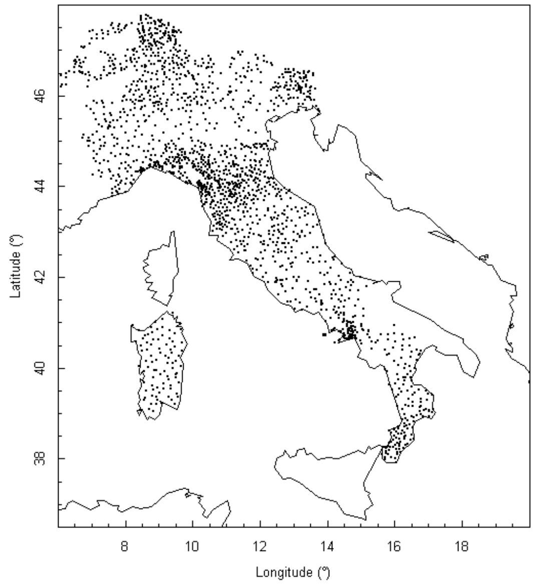

Figure 1: Location of rain-gauges in Italy

model grid-points falling in the same box are aggregated and processed by using two different approaches, schematically described in figure 2. In the first approach, called "*representative value approach*" the evaluation of a summarizing value for the precipitation field in each box, such as the mean or the maximum, has been performed. The second approach consisted in the study of full distribution function of the precipitation values in each area, mainly from a qualitative point of view.

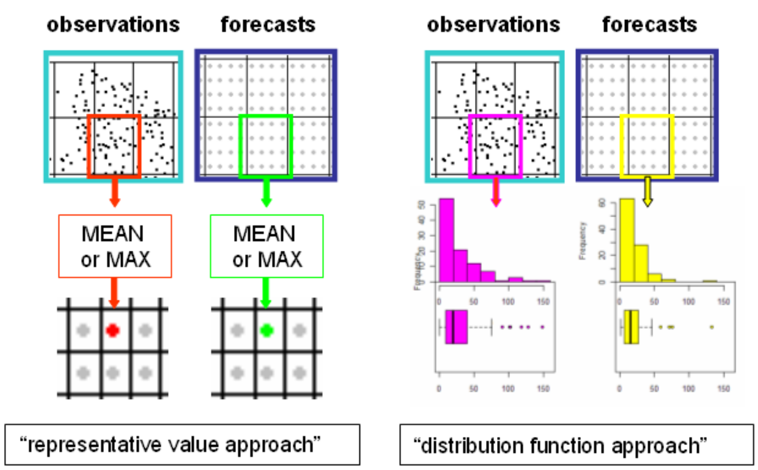

Figure 2: Description of the used approaches to evaluate QPF quality in the selected boxes

3 Results: Validation of QPF using the "representative value approach"

The QPF quality is assessed by making a comparison of a representative value in each box for observations and forecasts, using a methodology developed in the past years at ARPA-SIMC for the verification of COSMO-LEPS high-resolution precipitation forecasts (Marsigli et al.,2008). The verification domain is divided in boxes of 0.5 x 0.5 degree about 50 x 50 Km) in order to contain four IFS-ECMWF grid-points. Each box contains about

45 COSMO-I7 grid-points, 280 COSMO-I2 grid-points, while, as regards the precipitation observation, the number of stations vary between 3 and about 40 (only boxes with at least 3 observations inside have been considered in the computation of the verification scores). Categorical statistics (e.g. Probability of Detection, Frequency Bias, False Alarm Ratio) have been computed from the elements of contingency tables, on the basis of the exceeding of some predefined precipitation thresholds by the representative value (mean or maximum) of forecast and observation for each box over the period of interest. Some results concerning the period Autumn 2008 are presented as example in figure 3 (mean value exceeding the threshold of 1 mm/24h), figure 4 (mean value exceeding the threshold of 20 mm/24h) and figure 5 (maximum value exceeding the threshold of 20 mm/24h), but in outline they are representative of the verification over other periods.

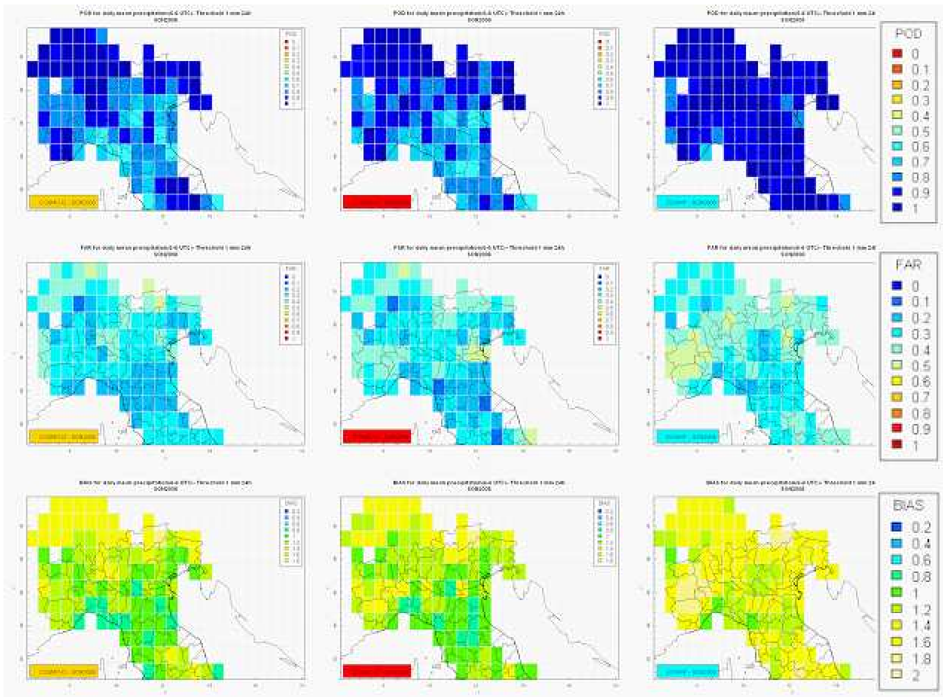

Figure 3: Probability Of Detection (top panels), False Alarme Ratio (middle panels), Frequency Bias (bottom panels), for mean value exceeding the threshold of 1 mm/24h for the models COSMO-I2 (left), COSMO-I7(center),IFS-ECMWF(right)

The results of this type of verification point out a strong dependence on the choice of the thresholds. For low thresholds (e.g. 1 or 5 mm/24h) and in particular for the mean value as indicator, the difference in the scores of the three models are small (except for the BIAS SCORE where IFS-ECMWF has values everywhere greater than one, showing a tendency to over-forecast the precipitation events) and quite independent on geographical location. Increasing the thresholds (e.g. 10 or 20 mm/24h) the scores tend to be very sensitive to the geographical positioning of the considered area and the performance of the models become very different, in particular if the maximum values are considered. In this case the COSMO models seem to have better scores, in particular the POD, even if the number of false alarm cases increases. On the other hand, IFS-ECMWF shows less false alarms but also a noticeably decreases in the BIAS SCORE, moving from a situation of over-forecast at lower thresholds to a situation of under-forecast at higher thresholds, especially when the maximum value is taken into account. The worsening in the scores concerning the maximum values, that means at least one points should exceed the thresholds, suggests that the lower resolution models have difficulties in reproducing high localized precipitation events, such as for example convective events.

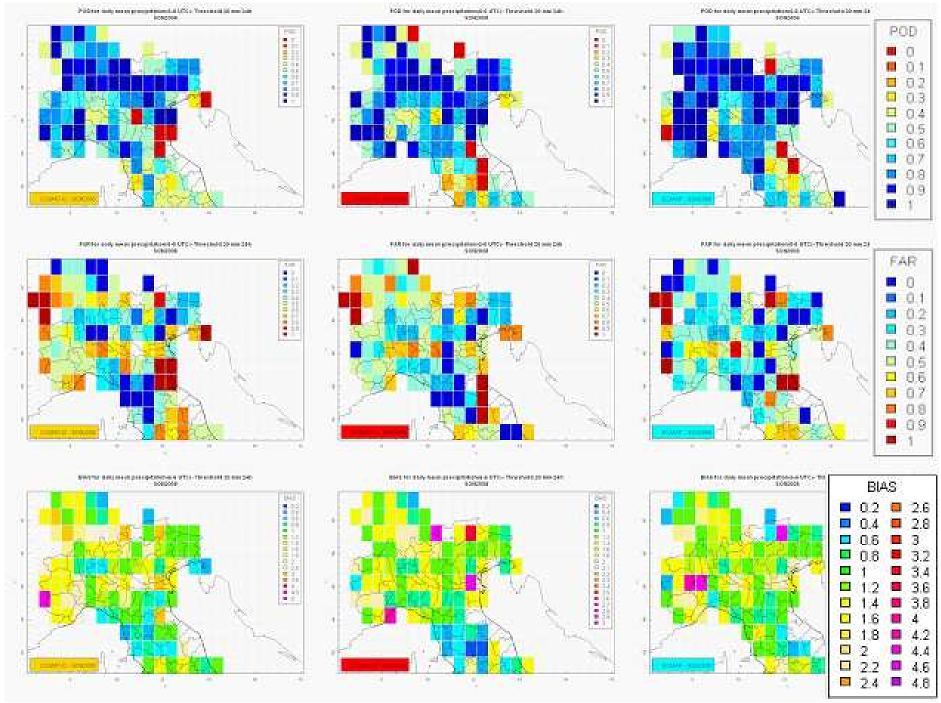

Figure 4: Probability Of Detection (top panels), False Alarme Ratio (middle panels), Frequency Bias (bottom panels), for mean value exceeding the threshold of 20 mm/24h for the models COSMO-I2 (left), COSMO-I7(center),IFS-ECMWF(right)

The "*representative value approach*" has been used also to evaluate the seasonal accumulation of the precipitation, in order to investigate the ability of the models in reproducing the total amount of rain occurred, taking no notice of the daily correspondence of observations and

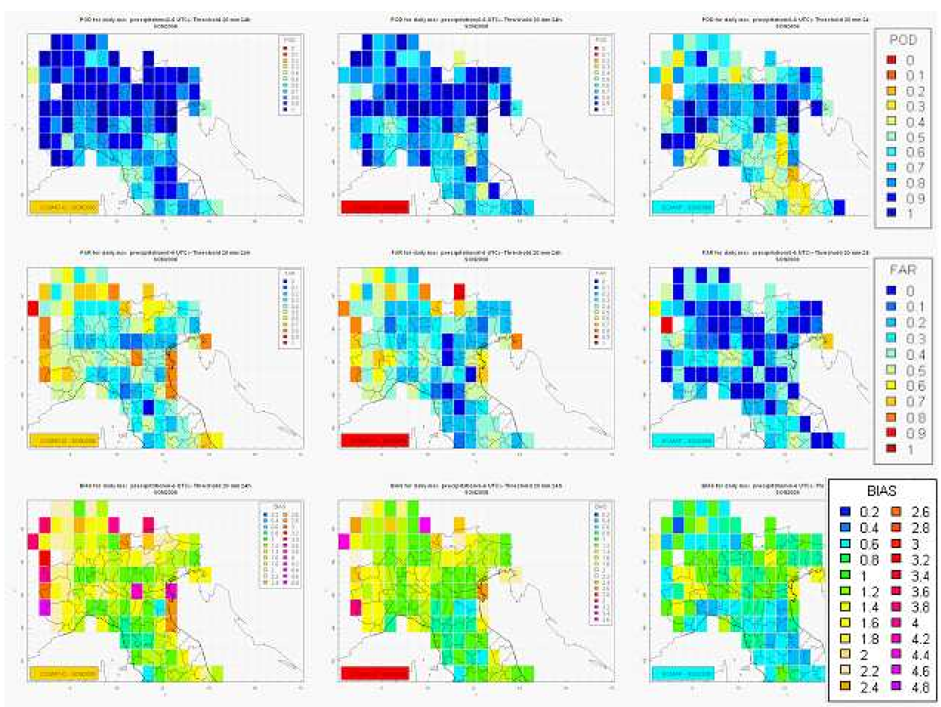

Figure 5: Probability Of Detection (top panels), False Alarme Ratio (middle panels), Frequency Bias (bottom panels), for maximum value exceeding the threshold of 20 mm/24h for the models COSMO-I2 (left), COSMO-I7(center),IFS-ECMWF(right)

forecasts. For each box all the forecasted and observed daily mean (or maximum) values are accumulated over “common days” (depending on the availability of observations). Results concerning the period March-April 2009 with boxes of 50 km x 50 km are presented in figure 6 (sum of the mean values) and figure 7 (sum of the maximum values). The maps of observed total amount, visualized in the bottom-right corners, show a strong dependence on the geographical positions of the boxes, depicting the morphological feature of the Italian territory. The observed precipitation pattern is quite well reproduced by COSMO-I7 and perhaps even better by COSMO-I2, considering both mean and maximum values cases. On the other hand IFS-ECMWF does not show the observed geographical variability in the accumulated precipitation pattern and in the case of the maximum values the amount of rain is definitely underestimated.

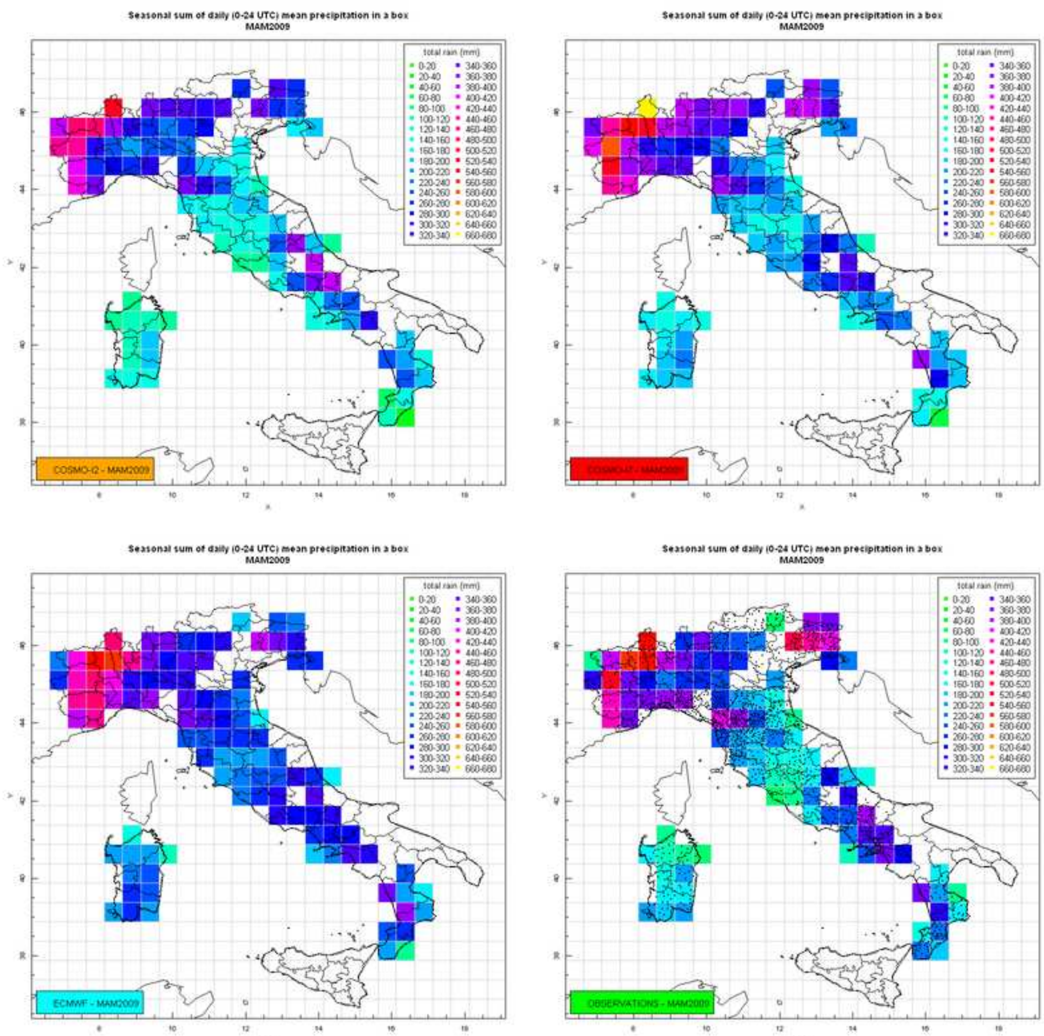

Figure 6: Accumulation of daily mean precipitation over the period March-April 2009: Observations(bottom right),COSMO-I7(top right), COSMO-I2 (top left), IFS-ECFMF (bottom left)

4 Results: Assessment of quality of QPF *pdfs*

In order to asses the ability of models in reproducing the variability of the precipitation fields inside a box over a period of time, we focused on the study of their distribution function (*pdf*), mainly form a qualitative point of view. We used a box-plot representation, that easily provides information about the spread of the distribution and furnishes a roughly

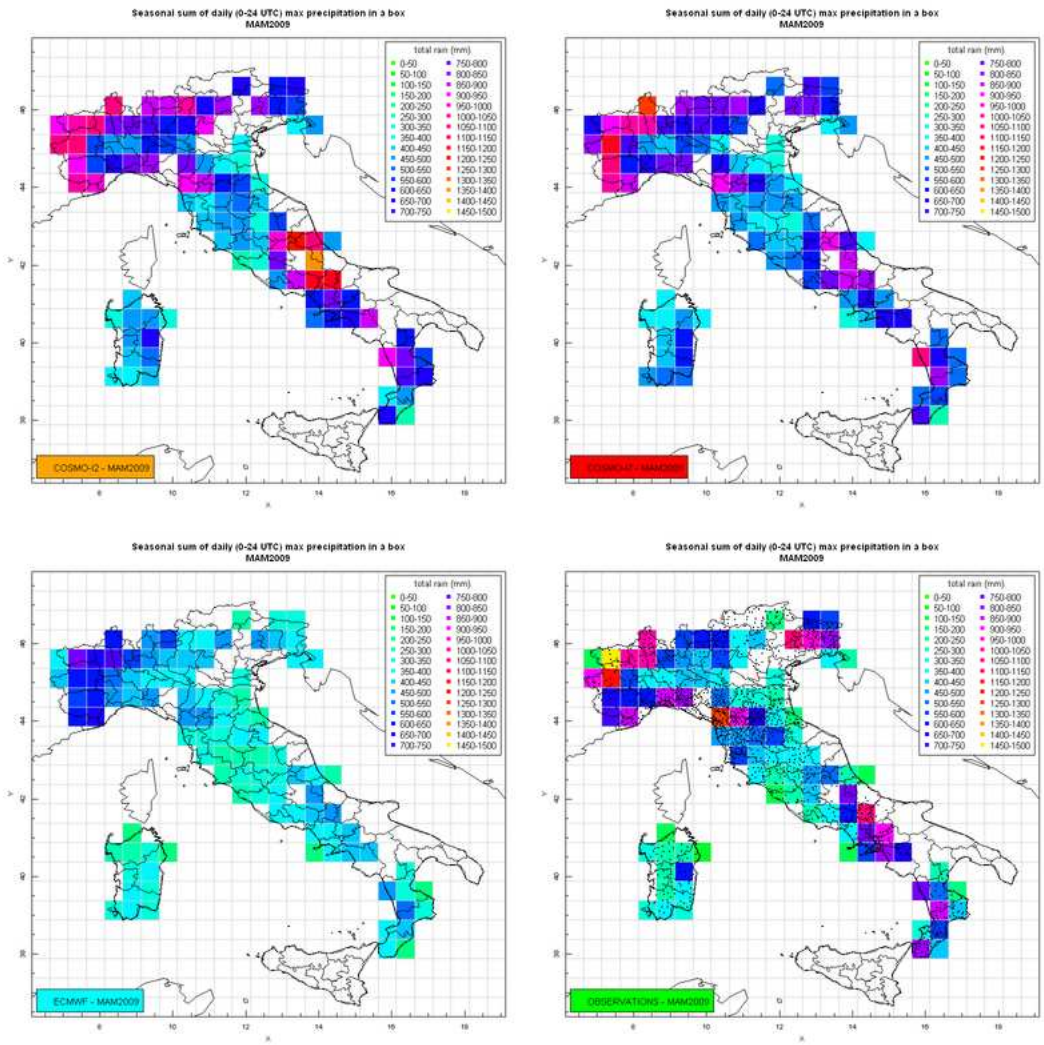

Figure 7: Accumulation of daily maximum precipitation over the period March-April 2009: Observations (bottom right), COSMO-I7 (top right), COSMO-I2(top left), IFS-ECFMF (bottom left)

description of the key measures that define it, such as the median, the quartiles, maximum and minimum values, and a range of values identified as “outliers”. Chronologically our first attempt to investigate the distribution function of the precipitation fields was done comparing the models COSMO-I7 and IFS-ECMWF, considering area of about 100 km x 100 km selected on the basis of the geographical position. The idea was to select, for example, area with homogeneous terrain (almost entirely flat or in the same mountain slope) in order to consider equi-probable, at least in a first approximation, the precipitation falling in each point of the area. The subdivision of the Italian territory in 1x1 degree boxes is shown in 8.

The main aim of that study was the assessment of the “climatology” of the forecasted and observed precipitation in each box over a relative long period of time (e.g. same season for a couple of years), taking into account, for the study of the distribution functions, all the daily precipitation in each points of the area of interest.

In figure 9, as an example, by means of the box-plot representation are displayed the distribution of COSMO-I7 QPF (red), observations (green) and IFS-ECMWF QPF (blue) for each area of the study (along the x-axis), concerning the period Autumn 2005 (named: SON2005) and Autumn 2007 (named: SON2007). Green dotted lines in each panel represent the 99th and 90th percentile of the observed rain distributions, while the red and blue lines refer

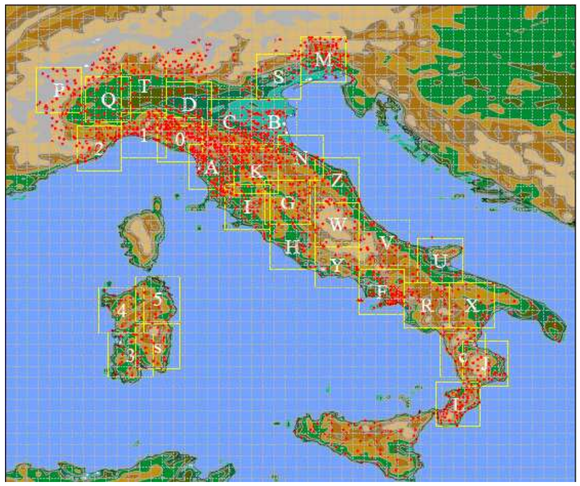

Figure 8: distribution of 1x1 degrees boxes over Italian territory

respectively to the model QPF distributions. The distribution of observed rain shows that the 90 % of rain events throughout SON2005 and SON2007 is generally lower than about 20 mm/24hours, with a large variability due to geographical position of the areas. Events greater than 50 mm/24hours represent only 1% of the total, but it is noteworthy that these are just the events in which forecasters are interested in for the issue of high impact weather warnings.

COSMO-I7 distribution seems to be fairly realistic, at least since the 90th percentile, even if in many areas also the 99th percentile is well described. Greater differences are due to outliers, with large overestimation of the maxima in many areas. However, it should be noted that very often the results could be affected by spatial representation problems, such as in costal areas that present sea points not covered by observation. The features of the observed *pdf*, and in particular the spread of the tail of the distribution, seem to be reasonably well reproduced. On the other hand, the spread of IFS-ECMWF *pdf* does not cover all the range of the observed values, with large underestimation both of 90th and 99th quantiles.

In the "representative approach" the daily comparison among forecasts and observations was performed using only some indicator of the distribution of the precipitation fields, nevertheless, even in a small area (e.g. 2500 Km^2) , the difference in the amount of rain from a station to an other one can be large. In figure 10 is presented ,as an example, the time-series of the daily distribution for the month April 2009 in an area of 50 km x 50 km in the east of Sardinia.

For each day along the x-axis , the box-plot represents the distribution of the amount of rain within the selected area. Considering, for example, the observed rain for the 12^{th} of April (presented in the bottom-right corner): it ranged from a minimum of about 15 mm/24h to more than 60 mm/24h as a maximum, with most of the values around 30 mm/24h. This intra-box variability is quite well reproduced by the COSMO models, in particular COSMO-I2, even if in terms of mean values IFS-ECMWF can be considered the more accurate. Of course the number of grid-points inside a box play an important role (only 4 for IFS-ECMWF), nevertheless this is the major aspect to be taken into account in order to appreciate the additional information that higher resolution models can provide. On the other hand this example pointed out that the precipitation verification based on the comparison of station point against the nearest grid-point is in general very misleading, in particular for non homogeneous terrain where orographic effects can influence the precipitation structures, or in case of convective events.

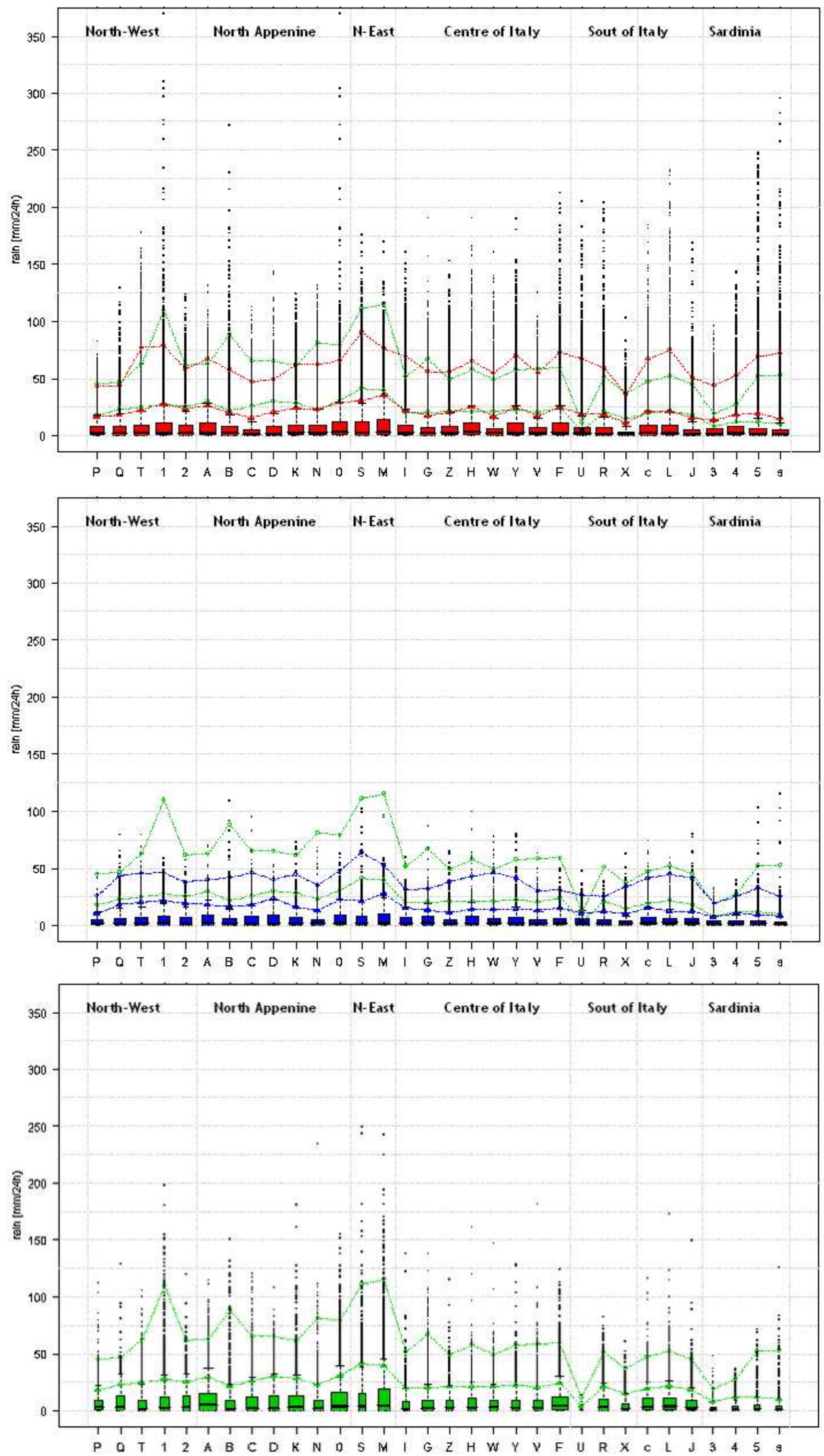

Figure 9: Distribution of COSMO-I7 QPF (red, top panel), IFS-ECMWF QPF (blue, central panel) and observations (green, bottom panel) for the period SON2005 and SON2007 in each area of the study

5 Conclusions

In this study the variability of the observed and forecasted precipitation fields in relatively small areas has been investigated, considering the COSMO model implementations running operationally at ARPA-SIMC (COSMO-I7 and COSMO-I2) in comparison with the global

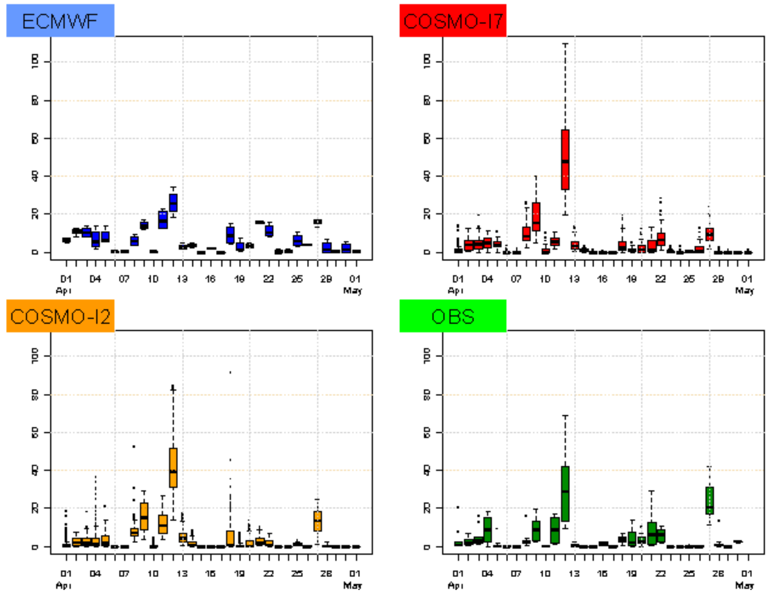

Figure 10: Time-series of the daily distribution for the month April 2009 in an area of 50 km x 50 km in the east of Sardinia

model IFS-ECMWF. The study of the distribution functions over a quite long period (e.g. two seasons) has pointed out, mainly from a qualitative point of view, that the distribution function COSMO-I7(the higher resolution model in the preliminary study) seems to be fairly realistic, at least since the 90th percentile, even if in many areas also the 99th percentile is well described. Greater differences are due to outliers, with large overestimation of the maxima in many areas. The features of the observed *pdf* , and in particular the spread of the tail of the distribution, seem to be reasonably well reproduced. On the other hand, the spread of IFS-ECMWF *pdf* (the lower resolution model in this study) does not cover all the range of the observed values, with large underestimation both of 90th and 99th percentile. The study of the intra-box variability of the daily amount of rain confirms the ability of higher resolution models in reproducing the main features of the distribution function of precipitation fields; in particular COSMO-I2 seems to be the more realistic while COSMO-I7 in many cases presents too high maximum values. These qualitative considerations about the distribution functions of forecasted and observed precipitation fields are confirmed also from a quantitative point of view considering the results of the QPF verification based on the "representative value approach". When the mean value is taken into account for the evaluation of the categorical statistics, the result pointed out the good quality of IFS-ECMWF, especially if lower reference thresholds are considered. Increasing the precipitation thresholds and taking into account the maximum value for the evaluation of the statistical scores, higher resolution models tend to gain respect to lower resolution model, even if the number of false alarm is also relatively high. On the other hand, IFS-ECMWF presents a greater number of missed alarms when we choose high reference thresholds. According to most standard verification measures, high resolution model forecasts would have poor quality, but it might be very valuable to the forecaster since it provides information on the distribution and variability of the rain field over the considered region. It's also interesting to note the strong dependence of the models behaviour on the geographical positioning of the boxes, on the period of the year and on weather type occurrence. In some areas models exhibit different behaviour depending on season, and throughout the same season they perform differently according to a different weather regime. This aspect is even more evident for higher resolution models because of the stronger interaction of the synoptic flow with the orography.

References

[1] Marsigli,C., Montani,A., Paccangnella,T.,2008: A spatial verification method applied to the evaluation of high-resolution ensemble forecasts, *Meteorol. Appls.*, **15**, 125-143

[2] Zepeda-Arce, J, E. Foufoula-Georgiou, and K. Droegemeier, 2000. Space-time rainfall organization and its role in validating quantitative precipitation forecasts, *J. Geophysical Research*, 105(D8), **10**,129-**10**,146.