Towards Operational Probabilistic Precipitation Forecast

Marco Turco, Massimo Milelli

ARPA Piemonte, Via Pio VII 9, I-10135 Torino, Italy

1 Aim of the work

Precipitation forecast from meso-scale numerical weather prediction (NWP) models often contains features that are not deterministically predictable and require a probabilistic approach. Therefore, a post-processing method has been developed in order to derive probabilistic precipitation forecasts from deterministic NWP model output. This method derives a Postprocessed Probabilistic Precipitation Forecast (PPPF) from a deterministic Direct Model Outputs (DMO) by using a spatio-temporal neighborhood method and it is based on the work of Theis et al., 2005 (see also Theis et al., 2002, Theis et al., 2003, Kaufmann, 2007). The procedure is applied to the output of the meso-scale model COSMO-I2, the regional very-high resolution version of the operational modelling system run in the framework of the COSMO consortium.

2 Implementation set-up

The procedure has been implemented on the intranet web page of ARPA Piemonte and at the moment is a tool for the ARPA forecasters only, but in its fully operational implementation it will be available for the Italian Department of Civil Protection. The procedure gives the probability of exceeding a certain threshold (1 mm/6h, 5 mm/6h, 15 mm/6h and 30 mm/6h) and two kind of maps are produced every 6h: over Piemonte region and over Italy. The probability in each grid point and in each forecast time is calculated considering the precipitation forecasted in that point and in the space-time neighborhood, using a certain radius in space and the previous and next forecast in time.

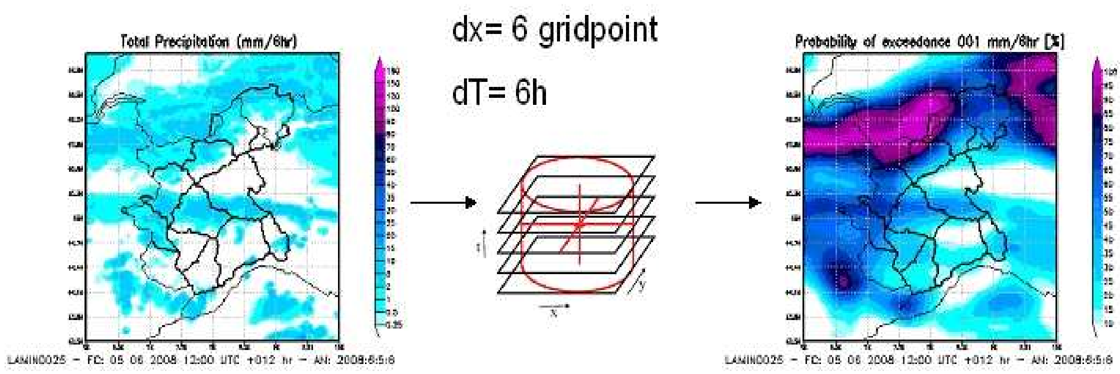

Figure 1: Example of the procedure: for each neighborhood (red cylinder) of every grid point, a probability distribution is calculated and it is converted in a map of probability of exceeding a given threshold.

The key assumption about this procedure is: QPF at the grid points within the neighborhood are assumed to be independent and identically distributed according to the probability density function of the precipitation forecast at the central grid point. Fig. 1 shows a schematic view of the procedure with the cylindrical neighborhood in the space-time plane (x, y and t), with the base in the space plane, and the vertical height in the time plane. The probability of exceeding a certain threshold is the number of the grid points within the neighborhood which are greater than the given threshold, divided by the total number of grid points within the neighborhood. In detail, a spatial radius of six grid cells is used (DX=6) and the spatial radius is kept constant within the temporal radius of 6 hours (DT=1). A crucial issue for these methods is the determination of an optimal size and shape of the neighborhood: in the following paragraph three space-time neighborhood methods have been tested in order to define the optimal one.

3 Verification

In the present study, we would like to answer to these two questions in an objective way:

- How does PPPF depend on the space and time windows ?
- Does this approach improve the DMO ?

In order to answer to these questions the verification procedure compares each observation in Piemonte (about 350 stations) with the nearest grid point, and calculates these scores (see Wilks, 1995 for more details):

- Reliability diagram (or attribute diagram): how well the predicted probabilities of an event correspond to their observed frequencies ?
- Brier Score (BS) & Brier Skill Score (BSS): what is the magnitude of the probability forecast errors ?
- ROC diagram: what is the ability of the forecast to discriminate between events and non-events ?
- Value diagram: given a cost/loss ratio C/L for taking action based on a forecast, what is the relative improvement in economic value between climatological and perfect information ? (see for instance Richardson, 2000)

The verification period is 1 year (June 2007 - May 2008) and 3 kinds of neighborhood are considered:

- DX=6 and DT=1 (i.e.: a circle of 6 grid points plus next and previous time steps, where the neighborhood has a cylindrical shape)
- DX=12 and DT=1 (cylindrical shape again)
- DX=6 and DT=0 (circular shape)

Since we have many combination of time steps (from +12 to + 42 hours), thresholds (1, 5, 15 and 30 mm/6h) and scores, in the following we will show only the results that are representative for deriving some general conclusion.

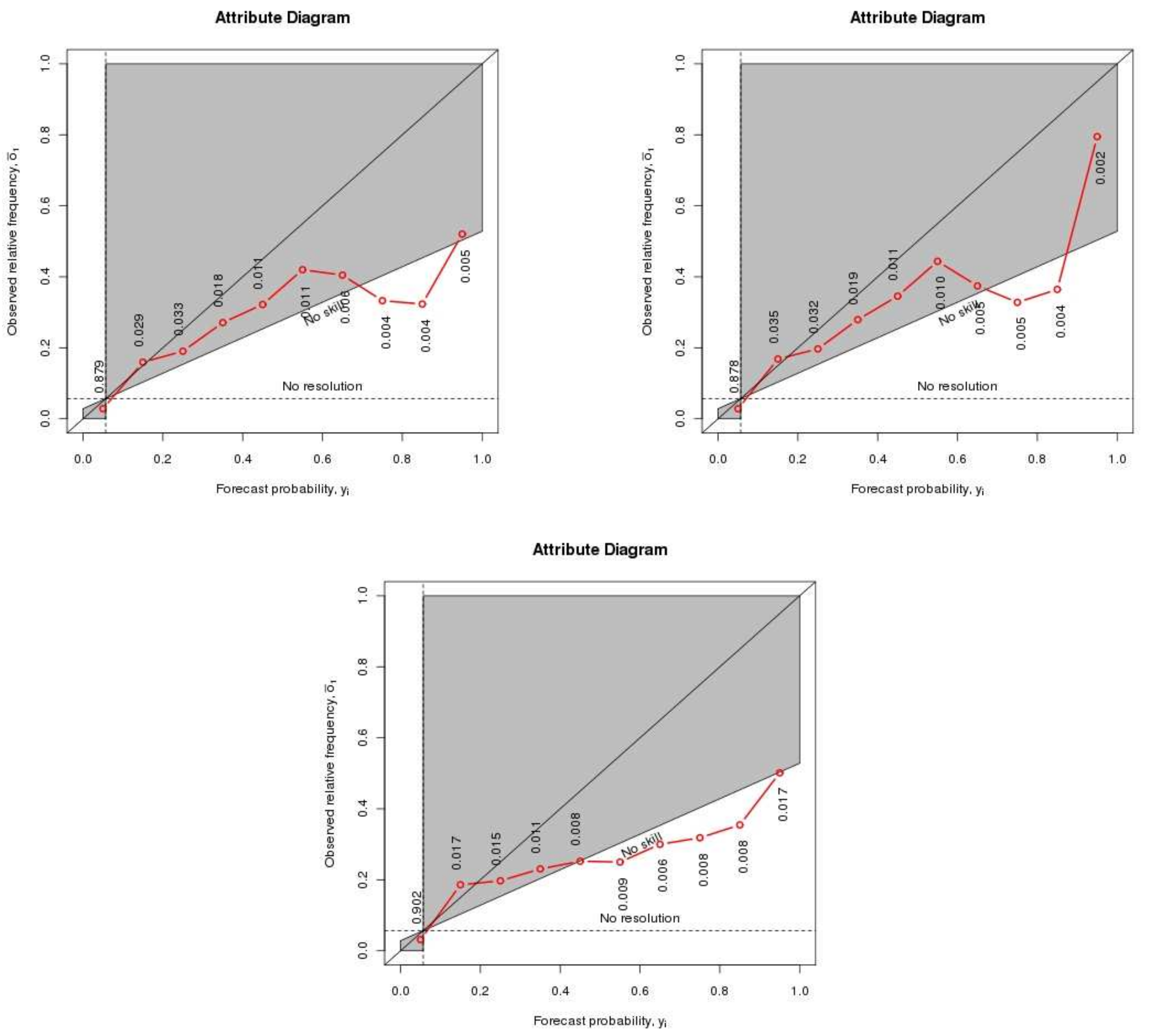

Figure 2: Reliability diagram (or attribute diagram) for the three different configurations: DX=6 and DT=1 (top left), DX=12 and DT=1 (top right), DX=6 and DT=0 (bottom) respectively with forecast time +42h and threshold 5 mm/6h.

Regarding the first question it is important to note that, as showed in Theis et al. (2005), an optimal universal neighborhood size cannot be determined: the effect of neighborhood size does not only depend on the precipitation amount, but also on the user's needs. The degree to which the forecast probabilities match the observed frequencies is shown in Fig. 2. The location of all the reliability curves (referred to the three neighborhoods for time step +42h and threshold 5 mm/6h) to the right of the diagonal indicates that the probabilities were always overestimated except for the 0.1 probability. For DX=12 and DT=1 and DX=6 and DT=1, only for 0.8 and 0.9 probabilities the PPPF have no skill whereas for the case with DX=6 and DT=0 all the higher probabilities have no skill.

Considering the accuracy (measured by the Brier Score BS, see Fig. 3), we show two examples: for the lowest (1 mm/6h) and the highest (30 mm/6h) thresholds. The best accuracy is achieved with DX=12 and DT=1 (small differences with respect to DX=6 and DT=1) and for the highest threshold, but this fact is not surprising since the Brier score is sensible to the climatological frequency of the event: if an event is rare, it is easier to get a good BS without having any real skill. In order to verify the ability of the forecast to discriminate between two

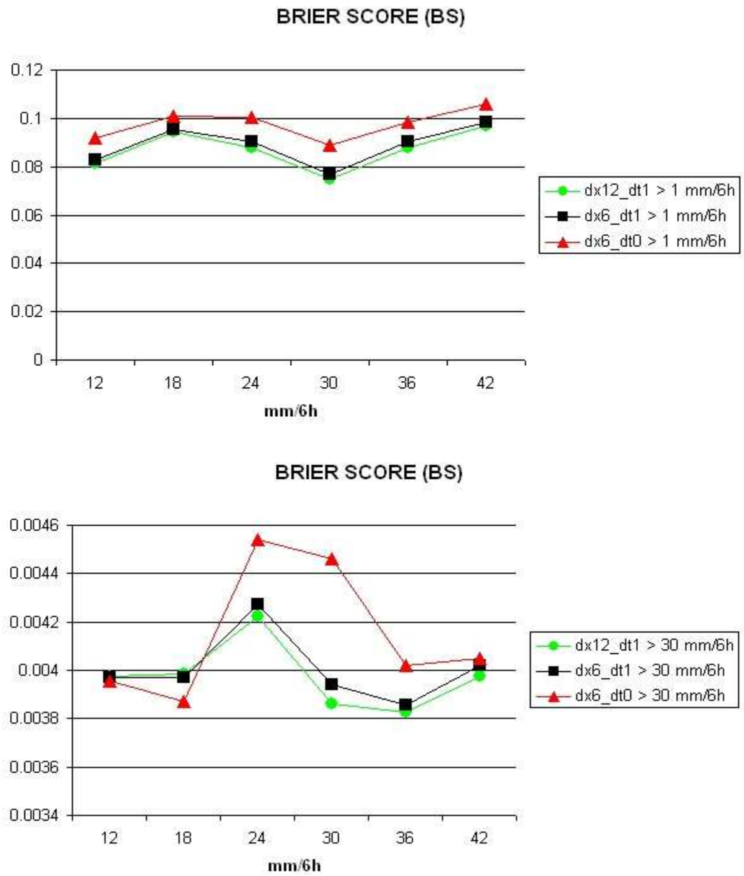

Figure 3: Brier Score for the lowest (1 mm/6h, top) and the highest threshold (30 mm/6h, bottom), for the three different configurations.

alternative outcomes, we consider the "Relative Operating Characteristic (ROC) Diagram", shown in Fig. 4. A perfect forecast would have a ROC curve starting in the lower left corner following the y-axis (false alarm rate=0) up to the top left corner, then following the x-axis (hit rate=1) until the upper right corner. The ROC curve for the three different PPPFs demonstrated that the best neighborhood is with DX=12 e DT=1, with small differences with respect to DX=6 and DT=1.

Summarizing, the results of the sensitivity study on DX and DT suggest that the best performances are obtained with DX=12 and DT=1, but they are not far from the DX=6 and DT=1 results, as it was also evident from all the other indices shown before. Therefore, since the DX=12 and DT=1 neighborhood has much more CPU costs (at least on our operational UNIX machines, where the DX=12 and DT=1 method requires 27 minutes of CPU and the DX=6 and DT=1 one only 10 minutes), we have chosen to use the latter in the operational setting.

In order to verify the ability of the forecast to discriminate between two alternative outcomes, we consider the "Relative Operating Characteristic (ROC) Diagram", shown in Fig. 4. A perfect forecast would have a ROC curve starting in the lower left corner following the y-axis (false alarm rate=0) up to the top left corner, then following the x-axis (hit rate=1) until the upper right corner. The ROC curve for the three different PPPFs demonstrated that the best neighborhood is with DX=12 e DT=1, with small differences with respect to DX=6

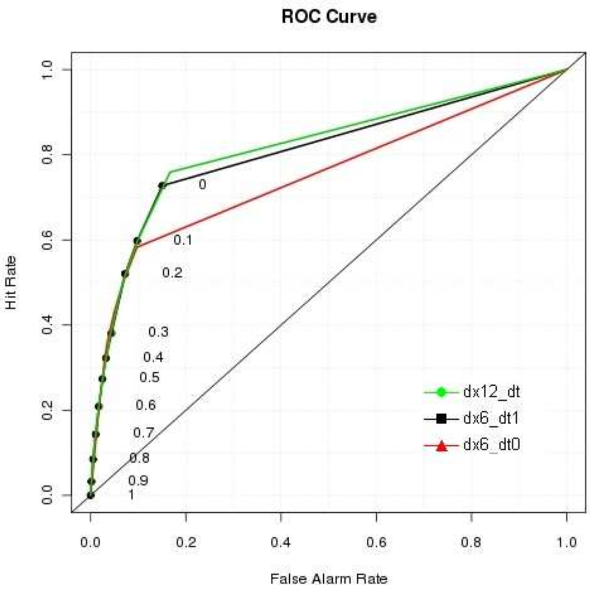

Figure 4: ROC diagram for the three neighborhoods fixed at forecast time +36h and threshold 5 mm/6h.

and DT=1. It can be noticed that considering the threshold of 5 mm/6h it is not evident the worsening with the forecast time. The best performances are found with DX=12 and DT=1, but they are not far from the DX=6 and DT=1 results, as it was also evident from all the other indices shown before. Therefore, since the DX=12 and DT=1 neighborhood has much more CPU costs (at least on our operational UNIX machines, where the DX=12 and DT=1 method requires 27 minutes of CPU and the DX=6 and DT=1 one only 10 minutes), we have chosen to use the latter in the operational setting.

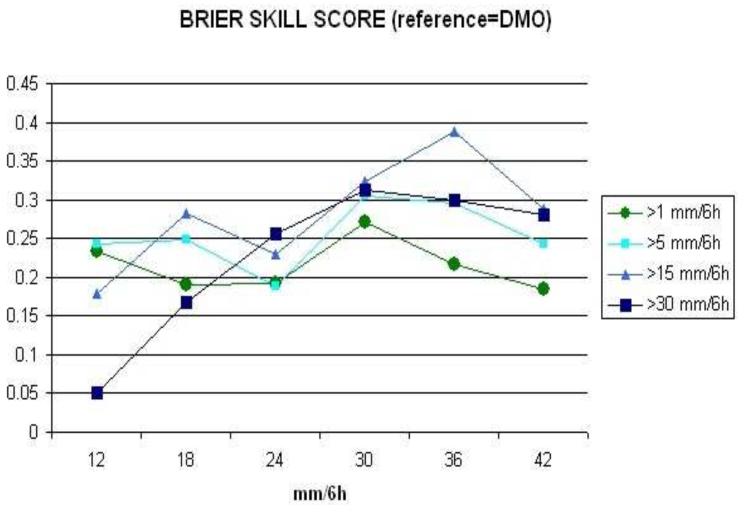

Figure 5: BSS with DMO used as a reference forecast.

The second question is "Does this approach improve the DMO ?". An answer is provided by the Brier Skill Score (BSS) that measures the improvement of the PPPF relative to a reference forecast (in this case the DMO). The BSS is constructed so that perfect forecast

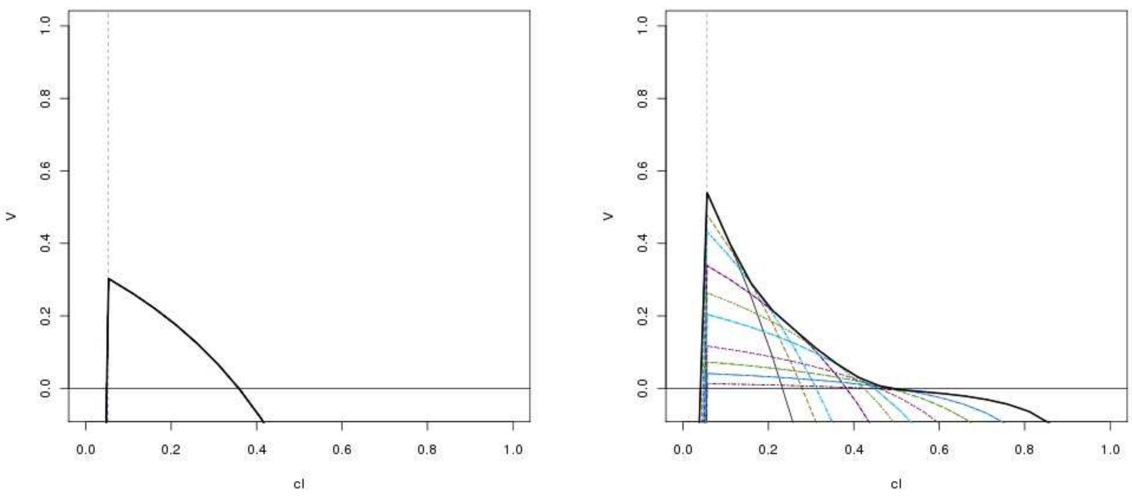

Figure 6: Relative value for DMO (left) and PPPF (right) at Forecast Time +18h and Thresholds=5mm/6h.

takes value 1 and reference 0, so it is positive (negative) if the forecast is better (worse) than reference. The BSS of the system with DX=6 and DT=1 is always positive, for all the thresholds and all the time steps and this means that the PPPF has more accuracy than the DMO (see Fig. 5).

Another diagnostic to measure the possible add value of the PPPF in respect to the DMO is the relative value (Richardson, 2000). This score is related to forecast resolution, but inserts the performance into a decision-making framework. The relative value V quantifies the usefulness of a forecast in minimizing the economic costs associated with protecting against the effects of bad weather and the losses incurred when bad weather occurs but the user did not take protective action. The improvement in economic value of the forecast is measured relatively to a climatology forecast and it is plotted as a function of the cost-loss ratio C/L. The relative value curves shown here are relative to forecast time +18h and to threshold 5 mm/6h (Fig. 6). For the PPPF box the lighter curves represent the relative value as a function of C/L using each of the probabilities (in this case, 0.0, 0.1, 0.2, ... 1.0) as a yes/no threshold for the forecast, while the heavy curve is the envelope representing the maximum relative value possible. The maximum PPPF relative value of 0.55 occurred for C/L close to 0.08, which is the climatological frequency of rain in the sample. These plots show:

- the PPPF have an added-value with respect to the DMO;
- the PPPF have value for all decision makers except those with very low C/L ratios or C/L ratios greater then 0.4.

4 Conclusions

The main conclusions could be here summarized:

- this method of post-processing improves the DMO;

- the best performances are found with DX=12 and DT=1, but the differences with respect to DX=6 and DT=1 are not sufficiently large to justify more CPU costs. Therefore we use operationally DX=6 and DT=1.

There are also some general observations:

- probabilistic forecasts provide an estimate of uncertainty that may be very useful for forecasters and end users;
- a direct link exists between probabilities and C/L;
- this approach should be complementary with an ensemble prediction system since they answer to different questions. For example:
  - EPS: will there be convection/front in a general area ?
  - PPPF: if there is convection/front, what will the peak precipitation be and where is it most likely to be located ?

References

Kaufmann, P., 2007: Post-processing of the aLMo Precipitation with the Neighborhood Method. COSMO Newsletter No. 7.

Richardson, D.S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. Quart. J. Royal Meteorol. Soc., 126, 649-667.

Theis, S., Hense, A., Damrath, U., and Renner, V., 2002: Ensemble Prediction and Statistical Post-processing of Weather Parameters. COSMO Newsletter No. 2.

Theis, S., Hense, A., Damrath, U., and Renner, V., 2003: Deriving Probabilistic Precipitation Forecasts from LM Simulations. COSMO Newsletter No. 3.

Theis, S., 2005: Deriving probabilistic short-range forecasts from a deterministic high-resolution model. Ph. D. thesis, University Bonn, Germany.

Theis, S., Hense, A., and Damrath, U., 2005: Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. Meteorological Applications, 12, 3, 257-268.

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences. An Introduction. Academic Press, San Diego, 467 pp.