Toward an ensemble nowcasting system: describing the steering field’s uncertainty in an advection scheme for radar images

Virginia Poli[1,2], **Pier Paolo Alberoni**[1,3], **Frank S. Marzano**[4,5]

[1] ARPA - SIM, Bologna (Italy).

[2] CIMA - Università di Genova e della Basilicata, Savona (Italy).

[3] ISAC - CNR, Bologna (Italy).

[4] Dipartimento di Ingegneria Elettrica – Centro di Eccellenza CETEMPS - Università dell'Aquila, L'Aquila (Italy).

[5] Dipartimento di Ingegneria Elettronica – Università “La Sapienza” di Roma, Roma (Italy).

1 Introduction

Advection nowcasting techniques of rainfall propagate ahead in time the structure observed at the initial forecasting time. This is usually done using a numerical advection scheme and, as initial condition, a steering vector(s) or field.

As for any numerical prediction technique, forecast depends on the knowledge of initial conditions. A direct consequence is that a lack and/or error in the initial condition propagate on the predicted fields. For advection schemes this is mainly due to uncertainty embedded in the motion field.

The aim of this work is to explore how this uncertainty is originated by the motion field generation mechanism and to understand how it propagates in the prediction.

In the used algorithm the steering field is evaluated using a multi-scale recursive cross-correlation analysis of radar reflectivity pattern. This operation involves areas bounded by values subjectively chosen and leads up to the definition of a three-level motion vectors merged to provide a motion field for the entire radar coverage. The areas are defined as a geometrical segmentation of the radar map as a function of reflectivity level analyzed.

It is evident that the final motion field is strongly influenced by values assumed by parameters of the cross-correlation (i.e. reflectivity, area dimension, how to merge different levels).

To evaluate the uncertainty in the motion field, we generate and analyze an ensemble of fields based on a set of parameters randomly chosen.

This technique will be used with a semi-lagrangian advection scheme to create a probabilistic precipitation forecast. As a term of reference this way to express the uncertainty will be compared with a white-noise directly added to the “standard” steering field.

2 Determination of motion field

Motion field used as initial condition in semi-lagrangian advection scheme is determined in two different steps: cross-correlation and analysis.

Cross-correlation allows to retrieve steering vectors analyzing two subsequent radar reflectivity fields. In the standard algorithm the methodology used is based on a multi-scale recursive analysis. This means that both of radar

---

Correspondence to: Virginia Poli.

vpoli@arpa.emr.it

patterns are divided in three different layers lower bounded by 0, 30 and 50 dBZ.

For the first layer and for each of the two data fields, an area of reflectivity centered on radar image is considered. The one from field at time t_0 is shifted around the one at time $t_1=t_0+15$ minutes in order to find in what position and in what direction cross-correlation coefficient is maximum. Best cross-correlation coefficient determines components of steering vector for this layer which is assumed to represent the average motion of the whole reflectivity pattern.

For the identification of steering vector for the second and the third layers, radar scans are divided in regular grids, with increasing resolution corresponding to an increasing reflectivity. Each box of the grid of the first scan is compared to all possible box in the second scan until maximum cross-correlation coefficient is derived. Also in this case values calculated are used to retrieve components of motion vectors (Poli et al., 2005; Poli et al., 2006).

The second step assumes that single vectors associated at each reflectivity level are merged using a method based on successive correction analysis. In this step retrieved vectors are associated to a radius that circumscribes an area in which they are supposed to have a certain influence. This weight is an exponential function of the relative displacement between vectors and analysis gridpoint and depends on the reflectivity level considered. In particular higher reflectivity are supposed to be representative of smaller structures and hence smaller influence area.

In the standard algorithm mentioned parameters, such as reflectivity values, the dimension and position of the research area for the definition of the first level motion field in cross-correlation step and influence radius in analysis step, are fixed.

The idea is to create an ensemble changing each one of these variables and try to understand what is the one that have the better impact on the forecast. These modifications are performed by a random selection of parameters themselves in order to generate a great set of motion fields. Running the semi-lagrangian advection algorithm with these different configurations generates three ensembles.

3 Case analysis

Case study deals with a convective episode in Po valley characterized by high reflectivity peaks and quick passage of rainfall structures. Data are from S.Pietro Capofiume radar with a spatial resolution of 1 Km. Forecasts are made for lead times ranging from 15 minutes up to 1 hour at intervals of 15 minutes.

Probabilistic forecasts associated to each ensemble show different behaviors (Figures 1, 2 and 3). For example probabilistic field generated changing influence radius in the analysis step is more stretched. In the second and in the third images, areas related to great probabilities are bigger especially for the northward structure.

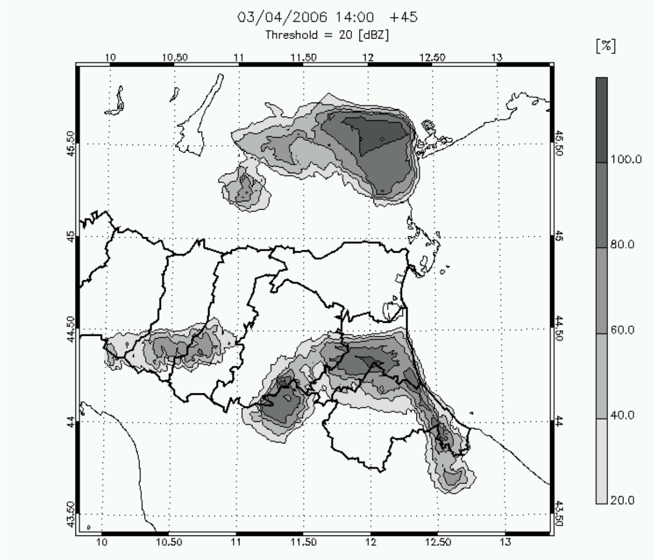

Fig. 1. Probabilistic forecast for a random choice of reflectivity levels. Lead time is equal to 45 minutes and threshold is 20 dBZ.

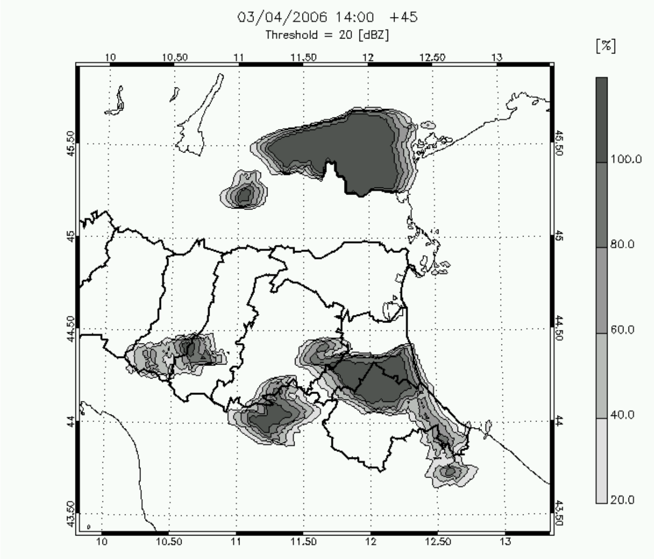

Fig. 2. Same as for Figure 1, but with a random choice of research area.

Single ensembles are merged in order to produce a larger set of data. Figures 4 and 5 display probabilistic forecast for the same lead time of the other images, but with increasing reflectivity threshold.

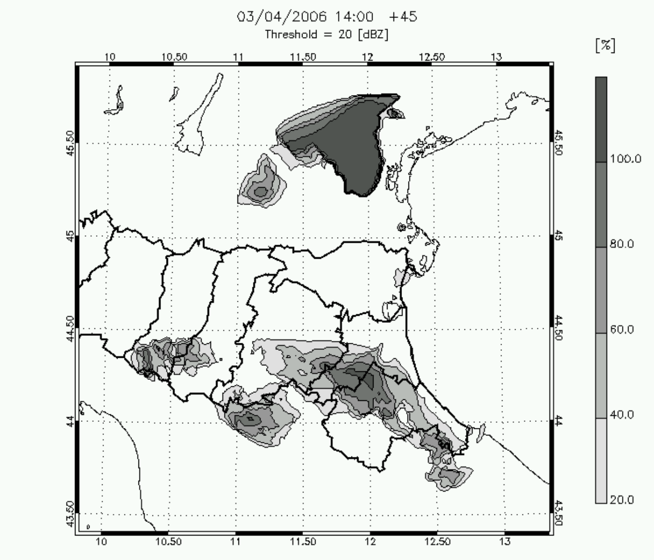

Fig. 3. Same as for Figure 1, but with a random choice of influence radius.

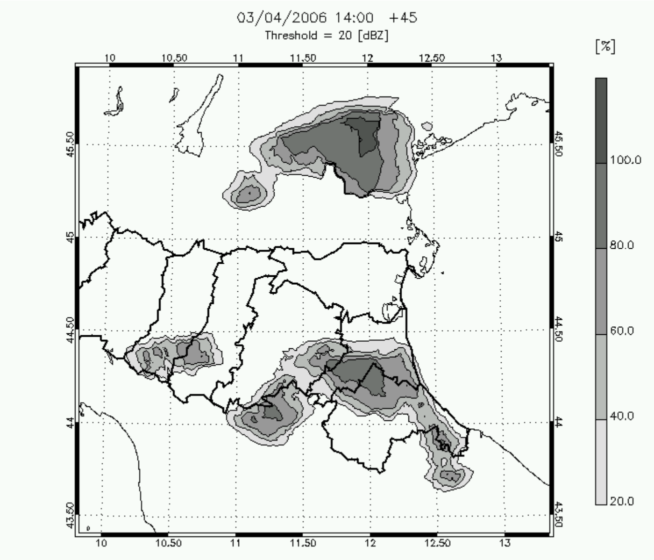

Fig. 4. Probabilistic forecast for 45 minutes lead time and with a threshold of 20 dBZ for the total ensemble.

Through this outcomes visualization an estimate of uncertainty is supplied. As a term of reference these results are compared to forecasts obtained adding a white-noise directly to the "standard" steering field.

Currently white-noise field is generated by a random function that add a value included in the interval of ±1 m/s.

From a numerical point of view standard ensemble scores, such as Brier score (BS) and Brier skill score (BSS), are calculated. They are introduced because of their plain and intuitive nature. BS measures the mean square probability error, while BSS estimates the improvements in accuracy of the probabilistic forecast over the reference forecast. Values for each lead time are investigated. Numerical results are presented in the form of a graphic (Figures 6 and 7).

In Table 1 dimensions of studied samples are summarized.

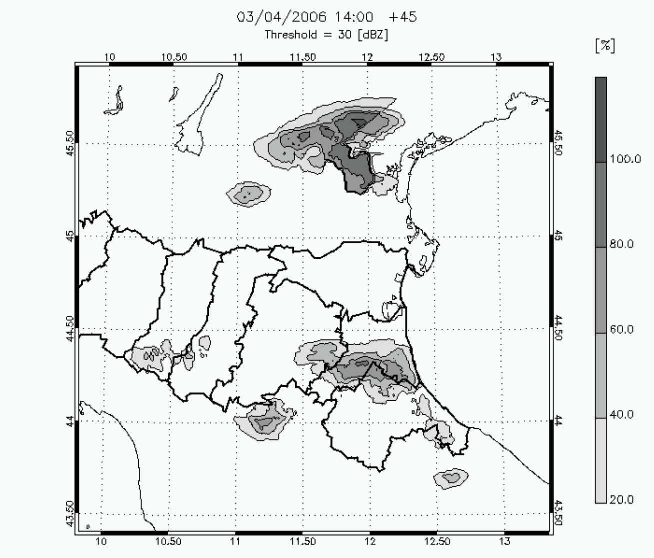

Fig. 5. Same as for Figure 4, but with a threshold of 30 dBZ.

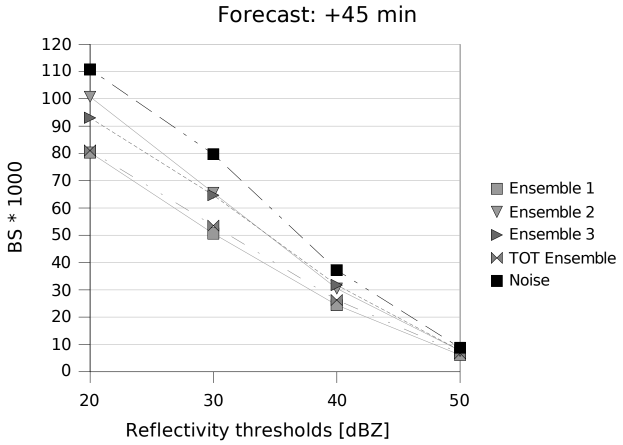

Fig. 6. Comparison between BS for the different ensembles in function of increasing thresholds for forecast with 45 minutes lead time

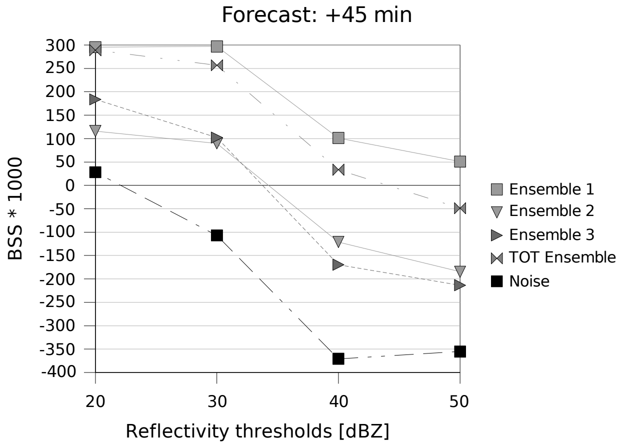

Fig. 7. Same as for Figure 7, but shown parameter is BSS.

Table 1. Number of observations greater than the different thresholds for forecast with 45 minutes lead time.

Thresholds [dBZ]	Number of observations > threshold
20	8597
30	5120
40	1831
50	426

4 Conclusions

The issue of this work is to understand if an ensemble generated changing parameters which determine motion field gives more information in spite of an ensemble created adding only white-noise to motion field.

Examined forecast represents the peak of the event. It is obvious that, because of the absence of physical equations in semi-lagrangian algorithm, neither growth nor dissipation are represented.

Outcomes tendency is similar for all of the forecast made.

Brier score trend is rapidly decreasing, otherwise the set of analysed data is dramatically poor for high thresholds. One of the causes of these short results resides in structures characterized by high reflectivity. They are very localized and following their motion becomes very difficult.

Best score is obtained by noise ensemble which is, on the contrary, associated to the worst brier skill score. Between the other three single ensembles, the one that use random reflectivity levels has the better impact on forecasts. It weighs a lot on the general ensemble and for thresholds lower than 30 dBZ, it is useful in forecast improving.

An effort will do in the future to draw a more complete and deep verification of results.

Acknowledgements: This activity is carried on in the framework of PROSCENIO and is partially supported by the European Union under the INTERREG IIIB CADSES programme “RISK-AWARE”, contract number 3B064. The authors thank Chiara Marsigli for her assistance in the use of validation tools.

References

Daley R., 1991: Atmospheric data analysis. Cambridge University Press, 457 pp.

Poli V., P. P. Alberoni, T. Paccagnella and D. Cesari, 2005: Preliminary results of the comparison of two advection methods. *WWRP Symposium no Nowcasting and Very Short Range Forecasting*, Toulouse, France, 5-9 September 2005.

Poli V., P. P. Alberoni and D. Cesari, 2006: Preliminary analysis of the intercomparison of two nowcasting methods. *MAP*, Risk-AWARE special issue, Submitted.