

A Strategy for High-Resolution Ensemble Prediction Part II: Limited-area experiments in four Alpine flood events

C Marsigli¹, A Montani¹, F Nerozzi¹,
T Paccagnella¹, S. Tibaldi¹,
F Molteni^{2,3} and R Buizza³

Research Department

¹ ARPA-Servizio Meteorologico Regionale Emilia Romagna, Bologna, Italy.

² CINECA, Centro di Calcolo Inter-Universitario,
Casalecchio di Reno (Bo), Italy

³ ECMWF, Reading, UK

May 2001

Also published in Q. J. R. Meteorol. Soc.

For additional copies please contact

The Library
ECMWF
Shinfield Park
Reading, Berks RG2 9AX

library@ecmwf.int

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/pressroom/publications.html>

© Copyright 2001

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.



ABSTRACT

A high-resolution ensemble system, based on five runs of a Limited Area Model (LAM), is described. The initial and boundary conditions for the LAM integrations are provided by the representative members (RMs) selected from the European Centre for Medium-Range Weather Forecasts Ensemble Prediction System (EPS). EPS members are grouped in five clusters; then, from each cluster, a RM is selected, according to the methodology described in the companion paper (Molteni et al., 2001). The ability of the high-resolution ensemble system to predict the occurrence of heavy rainfall events (either five or six days ahead) is tested for four cases of floods over the Alpine region. Results show that, in two case studies, the LAM integration corresponding to the RM of highly populated cluster predicts the observed rainfall with a very good degree of time and spatial accuracy. In the other two cases, the extreme events are captured by at least one of the runs nested on the members of the less populated clusters. Probability maps constructed from LAM integrations provide great detail on the location of the regions affected by heavy precipitation and the information gained with respect to EPS probability maps and LAM deterministic forecasts is highlighted. The probabilistic estimates based on the LAM ensembles are also shown to be of valuable assistance to forecasters in order to issue early flood alerts, contributing to the definition of a flood risk alarm system.

1. INTRODUCTION

Quantitative precipitation prediction is still a very difficult task, despite recent model developments and the implementation of ensemble prediction systems in some of the major meteorological centres (Tracton and Kalnay, 1993; Molteni et al., 1996). This is particularly true for extreme events (see Buizza et al., 1999; Mullen and Buizza, 2000), such as flash floods in regions with complex orography. One of the reasons of the poor predictive skill of the global ensemble systems, as far as quantitative precipitation forecasts are concerned, is that they are still characterised by a relatively coarse horizontal resolution. At the time of writing this paper and running the experiments, the European Centre for Medium-Range Weather Forecasts (ECMWF) global Ensemble Prediction System (EPS) has a grid scale of approximately 120 km at mid-latitudes* (Buizza et al., 1998) and the National Center for Environmental Prediction Short Range Ensemble Forecasting System has a grid spacing of about 80 km (Tracton et al., 1998). With such a resolution, these ensemble systems are not able to represent in a realistic way the orographic modulation of mesoscale systems. This shortcoming reduces the possibility of predicting few days ahead the risk of extreme flooding events over a localized region and to alert the population of possible serious damages.

Due to the very high cost of running an ensemble system with a large number of high resolution members, the Regional Meteorological Service of Emilia-Romagna (ARPA-SMR, in Bologna, Italy) developed a Limited-area Ensemble Prediction System (LEPS), based on 5 members, with initial and boundary conditions provided by the representative members (hereafter, RMs) of ECMWF EPS. The RMs are selected first by applying a cluster analysis to the 51-member EPS to define 5 clusters and, then, by identifying the RM of each cluster. Clusters are defined by considering the atmospheric flow at 700 hPa and by using the wind vector as clustering variable. Once the 5 clusters have been constructed, for each cluster the RM is defined as the member closest to all members of its own cluster and most distant from the members of the other clusters, with distances computed using an L1 norm applied to the precipitation field. The reader is referred to the companion paper, Molteni et al. (2001), and to Marsigli (1998) for a detailed description of the selection methodology. The LEPS is based on integrations of the limited area model LAMBO (Limited

* On November 2000, the horizontal resolution of the operational EPS was increased to T_{L255}, corresponding to a grid scale of approximately 80 km



Area Model BOlogna), operational at ARPA-SMR since 1993. LAMBO runs are performed at high horizontal resolution (about 20 km) in order to resolve those orographic and mesoscale processes responsible for heavy precipitation events. A probability of occurrence is assigned to each scenario, based on the population of the corresponding EPS cluster. In this way, it is possible to combine the ability of the EPS to highlight a set of possible evolution scenarios (keeping into account the intrinsic predictability of a particular synoptic situation) with the high spatial detail of the predictions obtained using a LAM. Ideally, an operational use of LEPS technique would provide a quantitative evaluation of the probability of occurrence for a particular extreme event, with enough a few days left to alert the population. It is worth pointing out that, because of the size of LAMBO domain (about 2000 km; see also section 3) and of the length of the forecast runs (120–144 hours), we expect the spread obtained by LAM integrations to depend mainly on differences in the boundary rather than in the initial conditions provided by the RMs (Vukicevic and Paegle, 1989). In fact, it is believed that the information provided by the initial conditions within the LAM domain will soon be advected outside the region of integration, while boundary conditions will carry the signal of upstream developments inside the same domain, heavily influencing LAM predictions after a couple of days of integrations (Vukicevic and Errico, 1990).

This paper investigates the predictability issues discussed above in four cases, chosen among flood events occurred in recent years in the southern side of the Alps. Since an accurate precipitation prediction is a crucial factor in the set-up of an appropriate flood alarm system by the agencies involved in the protection from natural hazards, attention will be focussed on quantitative precipitation prediction. Therefore, the aim of this work is to evaluate to what extent the LAM ensemble, nested on five members of the ECMWF EPS, can predict heavy rainfall events and their localisation in time and space a few days ahead.

The present paper is organised as follows. In section 2, a synoptic description of the case studies investigated in this work is presented. In section 3, the main features of LAMBO are described, while section 4 presents the verification tools used to evaluate LEPS performance. Results are shown in section 5 and conclusions are presented in section 6.

2. CASE STUDY DESCRIPTION

The four selected cases are chosen among the periods investigated during the field campaign of the Mesoscale Alpine Programme (MAP). The observed precipitation values were retrieved from the MAP database (web-site <http://www.map.ethz.ch>), where precipitation data from the participating countries are archived as 24h-cumulated values over each station point. Since the database (stored in Zurich, Switzerland) includes a number of stations ranging between 4000 and 6000, depending on the case study, a large number of observations could be used to evaluate the performance of LEPS in terms of quantitative rainfall. The initial dates and the clusters' population for the four experiments are reported in Table 1 and 5 of the companion paper, respectively. The attention is focussed on the 120-h forecast range for 3 cases and on the 120-h and 144-h forecast range for the fourth case (when the period of maximum rainfall was split between the fifth and the sixth forecast day). For each case study, a total of 6 runs is performed with LAMBO. In all cases, mesoscale forcing plays a key role at intensifying the precipitation which falls in a relatively short time interval; nevertheless, each flood event has its own synoptic features. The main characteristics of the synoptic situation for each case as well as the associated precipitation patterns are described hereafter.



2.1 Case 1 (Vaison-la-Romaine, 1992)

At 12Z UTC 21 September 1992, the mean-sea-level pressure pattern over western Europe (not shown) was characterised by a south-eastward shift of the Azores anticyclone; at the same time, a surface low, centred off the western coast of France, was moving eastwards. In the afternoon of 21 September, a cold front reached the French Atlantic coast, with mesoscale convective systems developing ahead of the front (not shown). A low-pressure system, deepening over the French Mediterranean coast, was associated with the steady flow of warm moist air from the south-west. The surface flow, enhanced by the mountains' channelling effect, was characterised by an intense, low-level jet in the Gulf of Lion (Senesi et. al., 1996). On 22 September, the axis of the trough tilted and a mesoscale convective system developed to the north of the French Mediterranean coast. Intense precipitation was recorded in many locations of the region, with rainfall rates of about 200 mm in few hours. The village of Vaison-la-Romaine* suffered the heaviest precipitation with a maximum of 300 mm in 24 hours observed during 23 September at about 44N 5E (top-left panel of fig.1 and bottom-right panel of fig.6), a secondary rainfall maximum being observed about 100 km to the north-west of the village.

2.2 Case 2 (Brig, 1993)

At 12 UTC 22 September 1993, the 500 hPa flow was characterised by an upper-level trough centred over the Biscay gulf and extending over the British Isles and Iberia (not shown). At the surface, an elongated cold front crossed Spain in the north-east south-west direction. Ahead of it, warm moist air was advected over western Mediterranean, producing an unstable atmospheric state (Benoit and Desgagne, 1995). In the following days, the upper-level trough deepened and slowly moved south-eastwards, while a cut-off cyclone developed over the Gulf of Valencia. The low-pressure system migrated to the northeast of the Balearic Islands, causing an intense meridional flow impinging on the Alps. Both the cut-off cyclone and the eastward propagation of the cold front contributed to sustain the warm advection which fed several mesoscale systems over the south-west of the Alps. Rainfall was found to affect several locations on both the Italian and the Swiss sides of the Alps (bottom-left panel of fig. 1). Although heavy precipitation was more widespread over the southern rather than the northern side of the Alps, a flash flood actually occurred over the Swiss town of Brig (at about 46N 8E), with a rainfall peak of about 150mm/day, during 23 and 24 September 1993.

2.3 Case 3 (Piemonte, 1994)

At 12 UTC 3 November 1994, a deep upper-level trough was localised over western Spain, while a ridge extended its influence over central and eastern Europe (not shown). In the following days, the ridge remained stationary, while the axis of the trough slowly rotated anti-clockwise, intensifying the geopotential gradient over western Mediterranean. At the surface, a southerly flow of warm moist air dominated the circulation in the same region, especially over Sardinia and Corsica. Synoptic charts also indicate that the wind direction was deflected to the west when approaching western Alps (not shown).

* This case study is referred to as the "Vaison case" in the remaining part of this work.

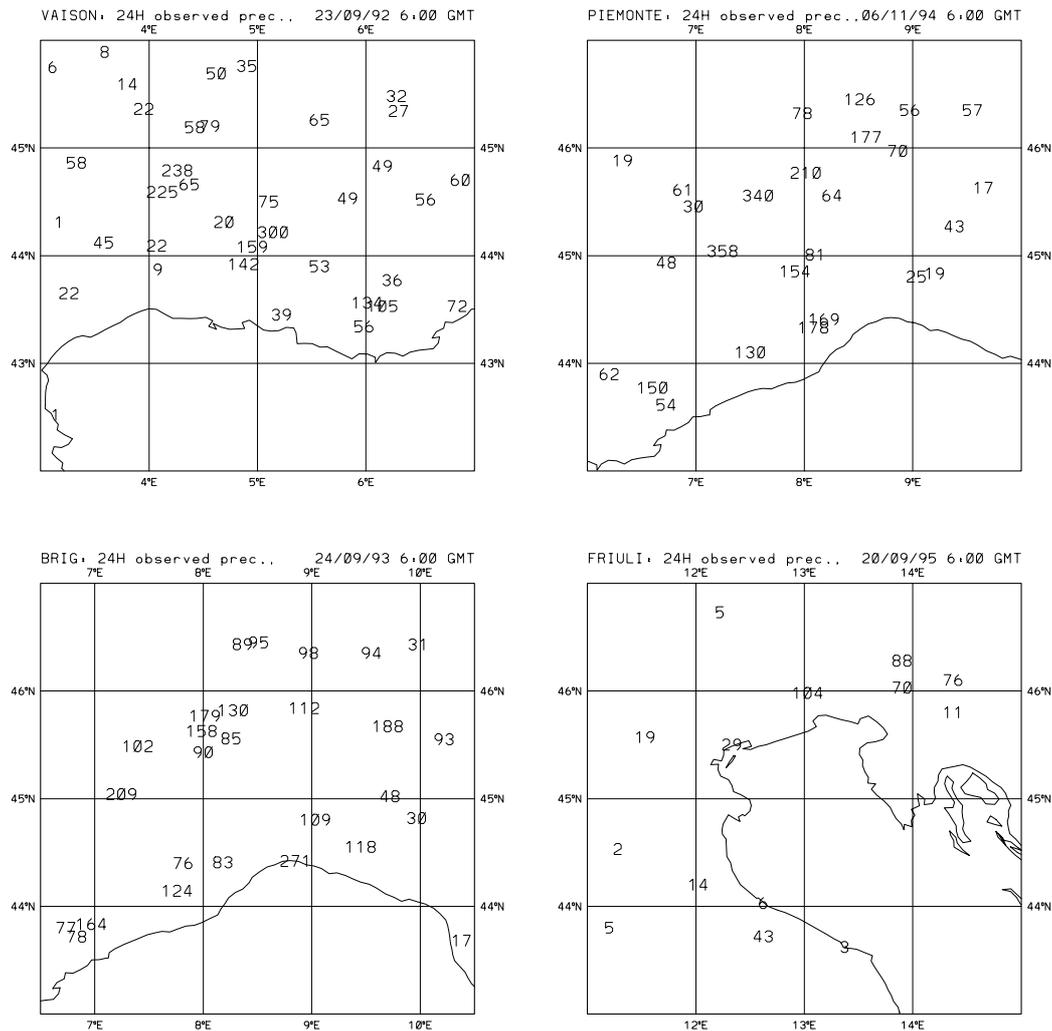


Fig 1 Cumulated rainfall observed over the 24-hour periods ending at 6 UTC of 23 September 92 (top-left panel), 6 UTC of 24 September 93 (bottom-left), 6 UTC of 6 November 94 (top-right) and 6 UTC of 20 September 95 (bottom-right). All data are in mm (from MAP web site). To make the plot more readable, not all the observations have been reported.

A weaker flow along the Adriatic Sea was also deflected, converging towards the same region through the Po Valley. This convergence was associated with a substantial uplift of warm moist air along the mountain slope (Buzzi et al., 1998). Heavy precipitation mainly affected two areas in north-west Italy with a double-peak structure over Piemonte and between the Piemonte and Liguria regions. Precipitation rates exceeding 200 mm/day were recorded during 5 and 6 November, with a maximum of about 358 mm/day at about 45N 7E (top-right panel of fig.1 and bottom-right panel of fig. 3). This was probably the most disastrous flood over the Alpine region in the last few decades, since it caused 70 casualties and damages for about 12 billion euros.

2.4 Case 4 (Friuli, 1995)

The synoptic situation relative to 00UTC 19 September 1995 (not shown) was characterised by a main low-pressure system centred over the French Mediterranean coast with a surface minimum value of 1005 hPa located at about 42N 6E. At this time, a south-westerly flow associated with the cyclonic circulation centred



over France affected Friuli (north-east Italy, not shown). Thickness charts indicate that, at low levels, cold air masses were moving from the north-east over Friuli, leading to relatively stable conditions. The frontal system, associated with the surface low, moved north-eastwards and, at 12UTC, the surface minimum was at about 48N 10E. The cold front was elongated in the meridional direction and crossed north-east Italy, warm humid air being advected over the Friuli region from south and south-east (Kerkmann, 1996). The confluence region between cold and warm air masses, localised over Friuli, lead to highly unstable atmospheric conditions, which caused heavy precipitation in the afternoon hours as well as in the first hours of the following day. Rainfall rates greater than 100 mm/day were observed over the Friuli region, causing a flash flood in that area (bottom-right panel of fig. 1).

3. DESCRIPTION OF THE LIMITED AREA MODEL

LAMBO, the ARPA-SMR limited area model, is a grid-point, split-explicit, primitive equation model, based on an early version of the NCEP Eta Model (Mesinger et al., 1988). The prognostic variables are zonal and meridional wind components, specific humidity, temperature and surface pressure. As vertical coordinates, the model can use either the terrain-following σ or the quasi-horizontal η (Mesinger et al., 1988), with a variable spacing in the vertical so as to give higher resolution near the surface rather than at the tropopause (in this work, all runs are performed using σ coordinate). Lorenz-type discretisation is used, the vertical velocity being defined at intermediate levels between those where the prognostic variables are integrated (Janjic, 1990). Horizontal discretisation of the momentum equation is performed on an Arakawa E-type grid with latitude and longitude as independent variables (Mesinger and Arakawa, 1976). Physical parameterisations include vertical turbulent exchanges, horizontal diffusion, precipitation, radiation and surface exchanges of moisture, heat and momentum. Vertical turbulent exchanges of heat, moisture and momentum are parameterised using the Mellor-Yamada closure schemes level 2.5 and 2 for the boundary and surface layer, respectively (Mellor and Yamada, 1974). The model also includes a non-linear fourth-order horizontal diffusion, the diffusion coefficient depending on flow deformation and on the turbulent kinetic energy. A two-layer bare soil model is also included. The precipitation is parameterised as dynamic (large-scale) and convective. Both deep (precipitating) and shallow (non-precipitating) convection are based on Betts and Miller scheme (Janjic, 1990). The Meteo-France radiation scheme (Ritter and Geleyn, 1992) is used to parameterise radiative processes.

At ARPA-SMR, the operational suite is based on two consecutive LAMBO runs:

- the former one, referred to as “father” run, has a horizontal resolution of 0.25° (about 40 km), 21 vertical levels and the model top at 100 hPa. The initial conditions are provided by ECMWF operational analysis, interpolated to LAMBO resolution; the boundary conditions are provided by ECMWF operational forecast, available every 6 hours throughout all integration time. The integration region covers approximately the area 4W–29E, 33N–52N (corresponding to 93x81 grid points);
- in the latter run, referred to as “son” integration, the model top is still set at 100 hPa, while the horizontal resolution is 0.125° (about 20 km), with 32 vertical levels. Boundary and initial conditions are now provided by the “father” run, the boundary conditions being updated every 3 hours. The “son”



integration area is slightly smaller than before, still covering the Italian peninsula and the Alpine region (1E–25E 36N–50N, corresponding to 137x113 grid points).

LAMBO is operationally integrated twice a day, nested on ECMWF operational runs of 00 and 12 UTC, the forecast length being 72 and 84 hours, respectively. In this study, the initial and boundary conditions for the “father” integration are provided by the representative members selected from ECMWF EPS, the integration time being extended to either 5 (Vaison, Brig and Piemonte cases) or 6 (Friuli case) days. Only results from the “son” runs will be shown.

4. VERIFICATION SCORES

An ensemble system can provide both deterministic and probabilistic forecasts. While each ensemble member represents a deterministic forecast, the whole ensemble can be used to compute probabilities of various kinds of event and it can be validated in terms of probabilistic scores. In the following two subsections, we introduce the deterministic and probabilistic verification scores, which will be used in the following sections to evaluate LEPS performance. In the calculation of both deterministic and probabilistic scores, the predicted 24-hour cumulated precipitations after 5 days (6 days for the Friuli case) are interpolated at the station point and then verified against the available observations (from the MAP database). All scores are calculated over the domain of the “son” runs (approximately 1E–25E, 36N–50N).

4.1 Deterministic scores

The accuracy of the single, deterministic prediction is evaluated using both the Bias (B) and the Threat Score (TS), defined respectively as (Wilks, 1995):

$$B(P) = \frac{F}{O}, \quad TS(P) = \frac{C}{F + O - C} \quad (1)$$

In the above equations, F is the total number of forecasts with 24-hour cumulated precipitation exceeding a certain threshold value P , O is the total number of observations with precipitation exceeding the same threshold and C is the number of coincidences, that is events which appear both in F and O . A perfect forecast has a Bias equal to 1, a greater (lower) value indicating that the forecast overestimates (underestimates) the occurrence of the event. The best value of TS is obtained when $F = O = C$, yielding $TS = 1$, while the worst possible value is 0, when there are no coincidences. B is a global indicator over the spatial domain, which does not require any spatial coherence, while TS is a better measure to indicate the ability of the model to forecast precipitation amounts in the correct location. A comparison is also performed between the model precipitation scores and the scores obtained from a “no-skill forecast” where C is defined as the product of the observed and forecast probabilities (Buzzi et al., 1994).

4.2 Probabilistic scores

The accuracy of probabilistic forecasts can be evaluated using the Relative Operating Characteristic (ROC) curves (Mason and Graham, 1999) and the Brier Scores (Stanski et al., 1989). The ROC curve is a x - y diagram where the Hit Rate is plotted against the False Alarm Rate.



Contingency table		Forecast	
		Yes	No
Observed	Yes	a	b
	No	c	d

Table 1 Contingency table.

According to the schematic contingency Table 1, the Hit Rate (HR) and the False Alarm Rate (FAR) can respectively be defined as:

$$HR = \frac{a}{a+b}, \quad FAR = \frac{c}{c+d}. \quad (2)$$

The two scores indicate, respectively, the proportion of events which were predicted and actually happened, and the proportion of forecast events which did not occur in reality. In this study, "event" means rainfall rate exceeding a threshold (e.g. 20 mm/day). Within a probabilistic forecast system, a warning can be issued when the forecast probability of an event exceeds a particular value (e.g. the probability of rainfall rate exceeding 20 mm/day has a value greater than 50%). Therefore, if several warning values are used for the event, a set of hit and false alarm rates may be determined for the same threshold. The set of HR, plotted against the corresponding FAR, generates the ROC curve. The area under the curve is commonly used as a probabilistic score, its maximum value being 1, and a value of 0.5 indicating a no-skill forecast system (for further explanation, see Mason and Graham, 1999).

The Brier Score (BS) is defined as the average of squared differences between forecast probabilities and the corresponding binary values representing the occurrence/non-occurrence of the event in observations (for a definition, see the companion paper in section 2.3). BS can take on values only in the range [0,1], the perfect forecast having $BS = 0$ (Stanski et al., 1989; Wilks, 1995).

In this work, the values for the ROC area and the Brier Score are calculated in two different configurations: in the former one, each member of LEPS is weighted according to the population of the represented cluster (that is, of the cluster in which a given RM was selected); in the latter one, the same weight is assigned to any high resolution run, irrespective of the cluster population. The results obtained in these two configurations should clarify to what extent the information on cluster population is useful for a definition of a probabilistic forecast based on a limited number of high-resolution integrations. In the remaining part of the paper, the two configurations will be referred to as "weight" and "no-weight", respectively.

5. PRESENTATION OF THE RESULTS

For each case study, LAMBO is nested both on the 5 RMs, selected within EPS members, and on EPS control forecast, considered as a reference forecast. An objective verification is performed using both the deterministic and the probabilistic scores introduced in the previous section. In addition to this, probability maps can be computed from LEPS integrations. Since in this study the attention is focussed on quantitative rainfall prediction, probabilities of rainfall exceeding thresholds of 20 and 50 mm/day have been calculated as the percentage of LEPS forecasts which predict total precipitation greater than the prescribed threshold. As for the verification scores, the probability maps are constructed both in the "weight" and "no-weight" configurations.



In the following subsections, the performance of LEPS is assessed, with particular emphasis on the amount of information gained with respect to both ECMWF EPS and the deterministic approach (represented by the control forecast), and on the sensitivity of the results to the configuration used.

5.1 Performance of LAMBO nested on the control forecast

The 24h cumulated rainfall and the 700 hPa geopotential height (hereafter, Z700) patterns predicted by LAMBO nested on EPS control for the four case studies are reported in fig. 2. A direct comparison against the observed precipitation patterns of fig. 1 shows that a clear distinction is possible between Piemonte and Brig cases, on the one side, and Vaison and Friuli cases, on the other one.

In the first two cases, the forecast obtained nesting LAMBO on the EPS control member gives a precipitation pattern similar to the observed. The top-right panel of fig. 2 (Piemonte case) indicates that the two rainfall maxima of 196 mm and 123 mm are correctly localised over Piemonte and between Piemonte and Liguria, respectively. In the bottom-left panel of fig. 2 (Brig case), a precipitation maximum of 142 mm is predicted at about 46N 10E, close to the region actually affected by the flood (and reported in fig. 1). In addition to this, the reference forecast of Z700 is quite accurate, with a minimum correctly placed at about 43N 6E.

In the other two cases, the LAMBO simulations nested on the EPS control are less satisfactory, with no clear indication of the possibility of a flood event. In the Vaison case (top-left panel of fig. 2), the predicted rainfall pattern is not consistent with the observations: the forecast values are too small over the region affected by heavy precipitation. The same is true also for the Friuli case (bottom-right panel), the precipitation field predicted after 138 hours by the reference forecast being considerably different from the observed one. In particular, a large-scale maximum of about 20 mm is predicted over northern Italy and no peaks in rainfall intensity are forecast anywhere in LAMBO domain. Nevertheless, the Z700 forecast is relatively accurate in the prediction of the cyclonic curvature of the flow over Italy.

For reason of brevity, the results obtained by LEPS are extensively presented only for the Piemonte and Vaison cases, in order to test the efficacy of the methodology for the two opposite situations of “accurate” and “non-accurate” reference forecast, respectively. With regard to the other two cases (Brig and Friuli), the obtained results are qualitatively similar to those presented for the Piemonte and Vaison floods respectively, and are briefly discussed in subsection 5.4. A global evaluation is also presented in subsection 5.5.

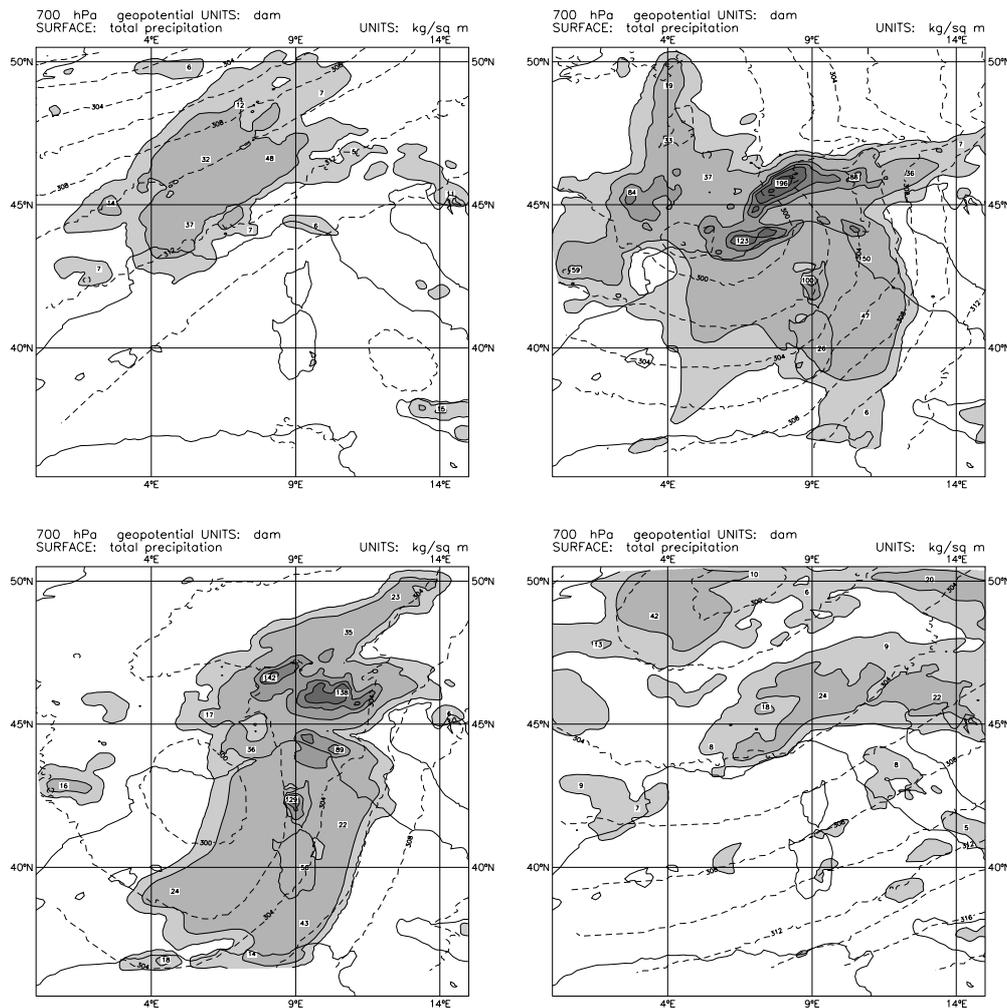


Fig 2 Z700 (dashed contours) and 24-hour cumulated precipitation (shaded areas) predicted by LAMBO nested on EPS control and valid at: 6 UTC of 23 September 92 (VAISON case), (t+114, top-left panel), 6 UTC of 24 September 93 (BRIG case), (t+114, bottom-left), 6 UTC of 6 November 94 (PIEMONTE case), (t+114, top-right), 6 UTC of 20 September 95 (FRIULI case), (t+138, bottom-right). Contour intervals: 20 dam and 2, 10, 50, 75, 100, 150, 200, 250 mm.

5.2 LEPS performance during Piemonte flood

The bottom-right panel of fig. 3 shows the verifying Z700 analysis valid at 6UTC of 6 November 1994 while, in the other panels, the 5 LEPS forecasts of total precipitation and Z700 are shown. Above each panel, the population of the represented cluster is reported. It can be noticed that the top-right panel of fig. 2 is identical to the top-left panel in fig. 3, because it happens that the EPS control is also the RM of the first cluster (the most populated one, with 26 elements; see tables 1 and 5 of the companion paper).

If the cluster population is taken into account, the precipitation pattern forecast by this member is given the highest probability of occurrence and, as already discussed, the prediction turns out to be very accurate. A great improvement in the amount of information given to the forecaster is evident when comparing this map with the corresponding ECMWF RMs prediction (reported in the companion paper, Fig. 9a). Both the low



and high resolution ECMWF global runs are able to forecast high precipitation values over a broad region covering the flooded area, but none of them is able to represent the double-peak structure which characterised the event. The same is true also for the LAMBO forecast nested on the RM of cluster 2 (9-element population). In fact, the middle-left panel of the figure reports two well-localised maxima over northern and southern Piemonte, although another region affected by heavy precipitation (maximum of 124 mm/day) is wrongly predicted over southern France and the low-level geopotential height pattern fails in representing properly the trough which extends over north-western Italy. With regard to the LAMBO run nested on the RM of the 3rd cluster (6-element population, bottom-left panel of fig. 3), even if the forecast patterns are quite different from both rainfall observations and Z700 analysis, a precipitation maximum of 82 mm is still evident over north-western Italy. The last two LEPS members, when LAMBO is nested on the RMs of the 4th and 5th cluster (respectively, 6 and 4-element population, reported in the top-right and middle-right panels of fig. 3), predict heavy rainfall mainly localised over southern France; the Z700 patterns indicate that the trough is located too much to the west, as if the system had not progressed enough during the LAM integrations. Nevertheless, in the 4th LAMBO run, a local maximum is correctly placed over southern Piemonte.

In order to have an overall measure of forecast accuracy, the total precipitation absolute error[†] has been computed for every LAMBO run, the numerical values being reported above each panel of fig. 3. This measure confirms to a great extent the above considerations on the forecast accuracy of the single runs, with lower errors for those LAMBO runs nested on the RMs of the most populated clusters.

The top row of fig. 4 shows the Bias (B) and the Threat Score (TS) for each LAMBO run (top-left and top-right panels, respectively) as a function of the rainfall rate. The values of B for the LAMBO integrations nested on the RMs of the first, second and fourth cluster (thin-solid, dotted and dot-dashed curves, respectively) are close to the ideal value of 1, for any threshold[‡] Also the TS for these three members is quite good; in particular, the accuracy of the first LAMBO integration is evident, the TS being the highest for all thresholds. For the LAMBO runs relative to the second and fourth clusters, the TS is always greater than that of the no-skill forecast (thick-dashed profile), indicating some correlation between predicted and observed rainfall values. On the other hand, the values of both B and TS for LEPS integrations nested on the third and the fifth RM confirm the poor performance of these two predictions. Overall, these results suggest that the uncertainty of the deterministic forecast is rather low, since the very intense rain predicted over the flooded region is confirmed by the RMs of the most populated clusters.

[†] The total precipitation absolute error (TPE) is defined as

$$TPE = \sum_{k=1}^{N_k} |f_k - o_k|,$$

where o_k is the observed precipitation, f_k is the predicted one (interpolated at the station point) and N_k is the number of station points.

[‡] For this particular case, since the control run is also the RM of the first cluster, the thin-solid line is “hidden” by the thick-solid one relative to LAMBO nested on EPS control.

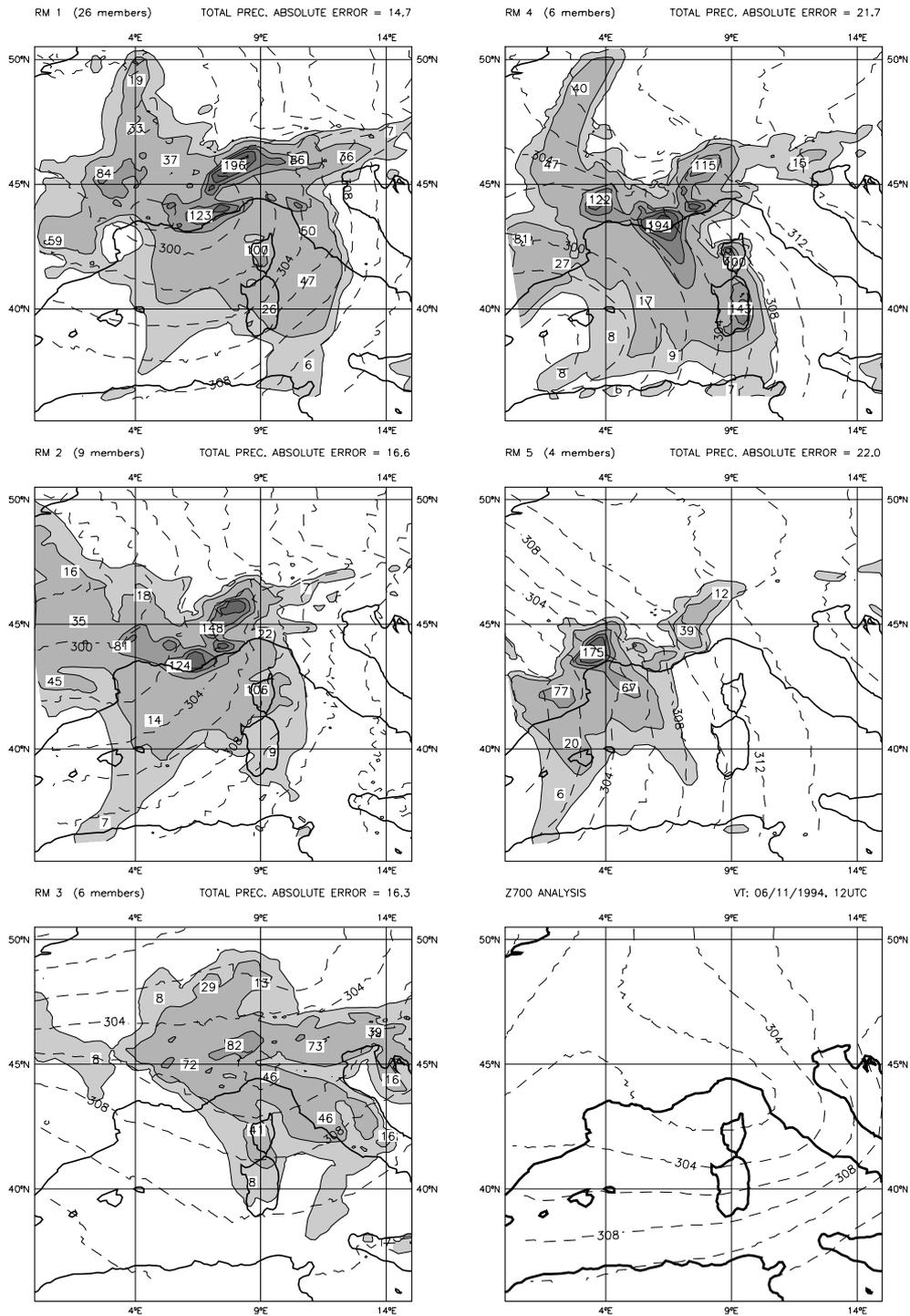


Fig 3 Piemonte case: Z700 (dashed contours) and 24-hour cumulated precipitation (shaded areas) predicted by LAMBO nested on RMs 1–5 (top-left, middle-left, bottom-left, top-right and middle-right panels, respectively), valid at 6 UTC of 6 November 94 ($t+114$). The bottom-right panel reports Z700 verifying analysis. The total precipitation absolute errors (in mm) are also reported for each LAMBO forecast. Contour intervals: 20 dam and 2, 10, 50, 75, 100, 150, 200, 250 mm.

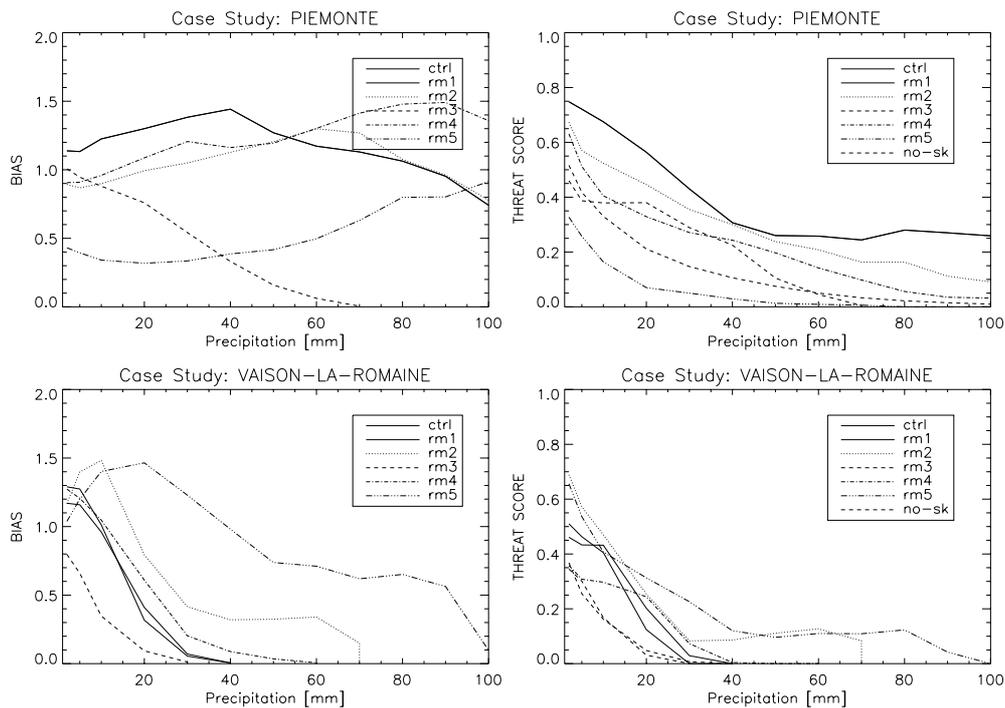


Fig 4 Piemonte (top row) and Vaison (bottom row) cases: Bias (left column) and Threat Score (right column) for LAMBO runs nested on EPS control (thick–solid line) and on RMs 1–5 (thin–solid, dotted, dashed, dot–dashed and 3–dot–dashed, respectively). The thick–dashed line in the right–column panels refers to the no–skill forecast (labelled with “no–sk”)

Rainfall probability maps obtained from LEPS integrations are presented in fig. 5. The top–row panels of the figure are obtained in the “weight” configuration, and indicate that LEPS assigns high values to the probability of precipitation exceeding 20 and 50mm/day over the two different regions affected by the flood. In particular, for the 50 mm/day threshold (top–right panel), a double–peak structure in the probability values (80% and 89% over northern and southern Piemonte, respectively) is evident, indicating the ability of the high resolution runs to represent small–scale features which cannot be simulated in EPS members. Very similar results are obtained in the “no–weight” configuration, as shown by the middle–row panels of fig. 5. The probability values are slightly lower than before, due to the smaller weight given to the LAMBO runs which predict the heaviest rainfall over Piemonte. Nevertheless, it is important to underline that in either configuration the signal indicating high chance of rain is very intense, and LEPS probability maps of precipitation exceeding 100mm/day (not shown) present a 60% maximum over Piemonte, the secondary maximum of 50% being located slightly to the west of the flooded region.

It is of interest to compare the LEPS probabilities with the probabilities predicted by the lower resolution ECMWF EPS (LEPS probabilities are based on forecasts with a 20km grid spacing, while the EPS probabilities are based on forecasts with 120km grid spacing). For the 20 mm/day threshold, the region where high probability is predicted covers a good deal of northern Italy, while for the higher threshold (bottom–right panel of fig. 5) the EPS indicates a maximum of 64% over a broad area in north–west Italy.

Due to the coarser resolution, no evidence of a double peak structure in precipitation maxima is evident, unlike in LEPS maps. It is also worth pointing out that the probability of rainfall exceeding higher



precipitation values (e.g. 100 mm/day) is below 10%, a substantially lower value than in the high resolution ensemble.

5.3 LEPS performance during Vaison flood

Fig. 6 reports the Z700 analysis and the predictions of rainfall and Z700 from the 5 LEPS runs for the Vaison case. It is evident that the low-level trough over southern France (bottom-right panel) is badly predicted by both the 1st and the 3rd LAMBO runs, corresponding to the two most populated clusters (respectively, 18 and 20 elements, top-left and bottom-left panels). In addition, both integrations strongly under-predict precipitation, and give no indication of a possible flood in the Vaison area. Also the LEPS run relative to the 4th RM (2-element population, top-right panel) predicts a rainfall pattern very different (and much weaker) from the observed one, corresponding to a weaker flow at 700 hPa. In the 2nd LAMBO integration (relative to a 6-element cluster, middle-left panel), the predicted rainfall maximum of 80 mm has a relatively good location with respect to observations, but it still underestimates the actual value, and the cluster population is low. The most accurate precipitation prediction is given by the member nested to the RM of cluster 5 (middle-right panel of fig. 6). This integration shows a precipitation maximum of 108 mm, well localised in the flooded region, although the associated Z700 pattern is somehow different from the one actually observed.

In particular, the secondary rainfall maximum is well captured, while the intensity of the main peak is partly underestimated. A comparison with the rainfall pattern forecast by this ECMWF RM (reported in the companion paper, Fig. 3e) underlines the advantage of the LAMBO run, which predicts a more detailed and better localised rainfall distribution, with more realistic precipitation values. Despite the low predictability of this rainfall event (evident in the poor forecasts of the LAMBO runs nested either on the control or on the RMs of the most populated clusters), at least one LEPS integration indicates the possibility (although with low probability) of a flood scenario five days ahead.

The bottom-row panels of fig. 4 show the values of B and TS for the Vaison case. At first glance, the poor performance of LAMBO runs in this case (when compared to the Piemonte one) is well evident in the lower values of both scores. Only the 3-dot-dashed profile, and to a lesser extent the dotted one (relative to the fifth and second LEPS runs, respectively), show a reasonable prediction of the weather event, while the scores obtained by LAMBO nested on the control are very poor (the values of B and TS decreasing very quickly with increasing threshold).

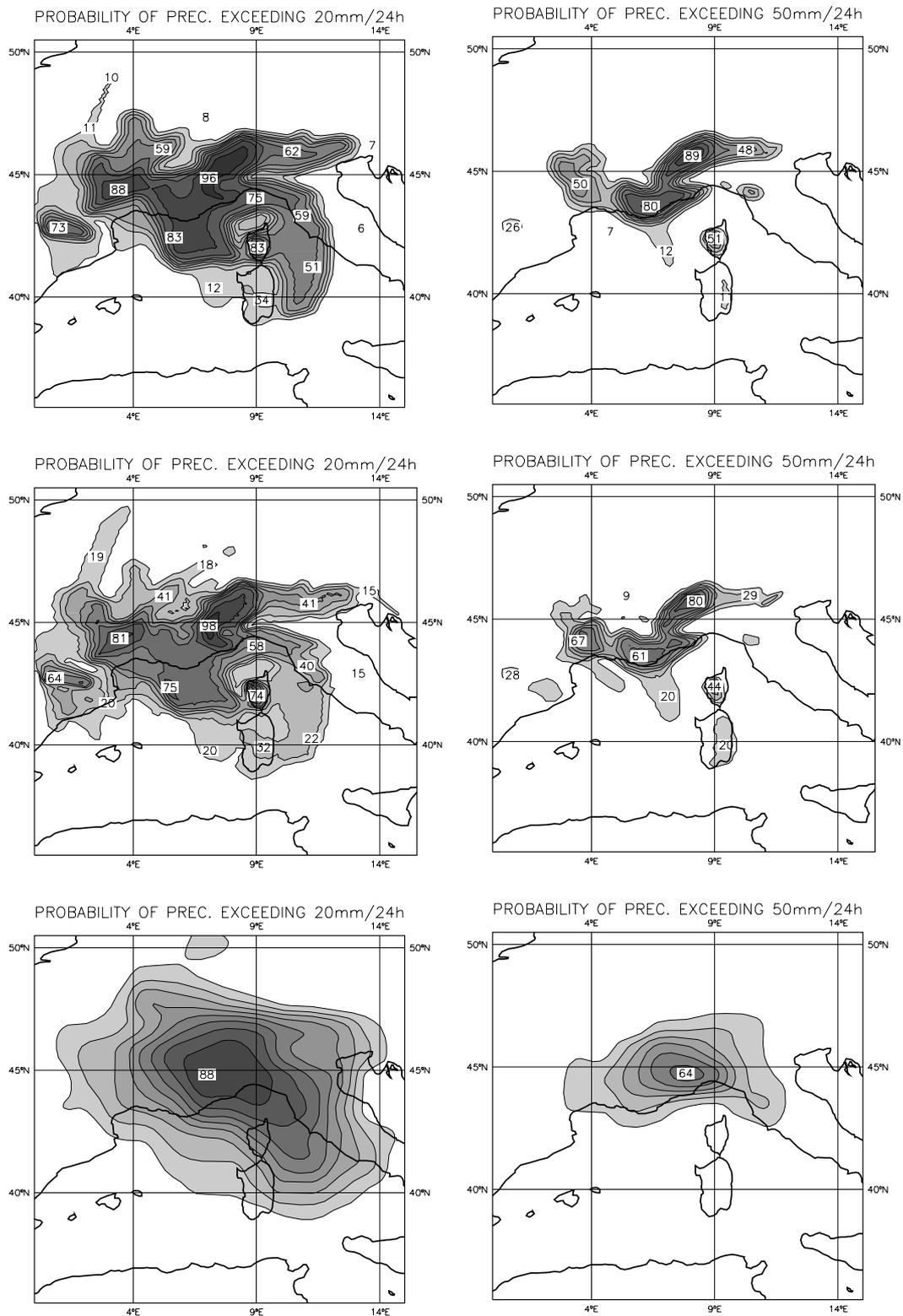


Fig 5 Piemonte case: probability maps valid at 6 UTC of 6 November 94 ($t+114$) for predicted rainfall rates exceeding 20 mm/day (left column) and 50 mm/day (right column). In the top-row panels, LAMBO probability maps in the “weight” configuration; in the middle-row, LAMBO probability maps in the “no-weight” configuration; in the bottom-row, ECMWF probability maps. Contour intervals: every 10%.

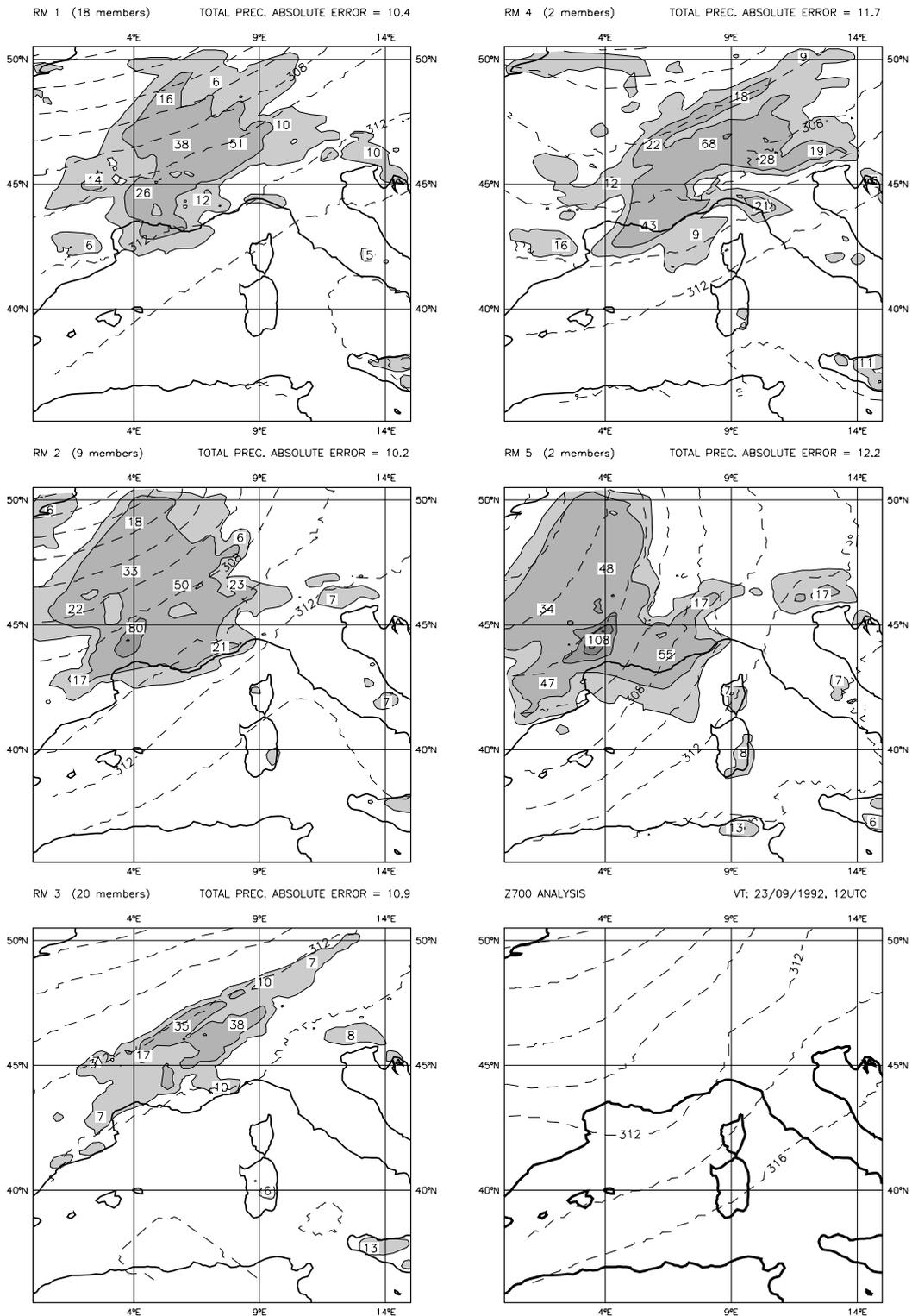


Fig 6 Vaison case: the same as fig. 3, but valid at 6 UTC of 23 September 92.



The small population of the cluster related to the most skilful high resolution forecast has a direct impact on LEPS rainfall probability maps in the “weight” configuration, as reported in the top-row panels of fig. 7. LAMBO runs relative to the two most populated clusters do not predict heavy rainfall in the Vaison region, but rather on the western side of the Alps. Therefore, both top-row panels of fig. 7 show low probabilities for rainfall exceeding 20 and 50 mm/day in the region of interest, the peak values ranging between 41% and 19% (top-left and top-right panel, respectively). The situation is different in the “no-weight” configuration, the relative importance of the best integration being higher. In fact, the middle-row panels have higher probability values close to the correct location for both thresholds and a weaker signal, about 20%, appears also for rainfall rate exceeding 100 mm/day (not shown). The panel relative to the 20 mm threshold well indicates a double peak in rainfall probability, with peaks of 40% and 60% close to the two areas where heavy rainfall was actually observed. The performance of ECMWF EPS is very poor for this case. The EPS probability prediction of precipitation exceeding 20mm/day is misplaced (bottom-row panel of fig. 7) and a null probability is predicted for precipitation exceeding 50mm/day (not shown). The highest probability of heavy precipitation is located on the Italian side of the Alps, at about 46N 8E, several hundred kms to the east of the correct location. Therefore, also in this case of lower predictability, the gain in information with respect both to ECMWF EPS and LAMBO reference run is still evident. Although the flood is not forecast at the medium range by the reference run, a non-zero probability of flood in the correct location is highlighted in both LEPS configurations, enabling a hypothetical forecaster to pay particular attention to the evolution of this weather event in the following days, and either to confirm or to dismiss flood alerts.

5.4 LEPS performance during Brig and Friuli floods

As already mentioned, the performance of LEPS in the two left cases (Brig and Friuli) is relatively similar to that obtained for the Piemonte and Vaison floods, respectively. Also in these cases, discussed in a more cursory way for brevity reason, the flood event is always predicted by at least one LEPS integration, which provides detailed information about the location of the region affected by heavy precipitation.

For the Brig case (left panel of fig.8), the forecast of LAMBO nested on the RM of the first cluster (15-element population) is quite accurate, predicting heavy precipitation over a region centred around Brig. Also the LAMBO runs nested on the RMs of the third and fourth clusters (not shown) provide realistic precipitation patterns. The high peaks of precipitation predicted by these LEPS integrations provide more reliable information to the forecaster, in comparison with the rainfall patterns forecast by EPS RMs (reported in fig. 6 of the companion paper). The area where the flood is more likely to occur according to LEPS maps, is less broad than in ECMWF runs and the intensity of the event is better captured.

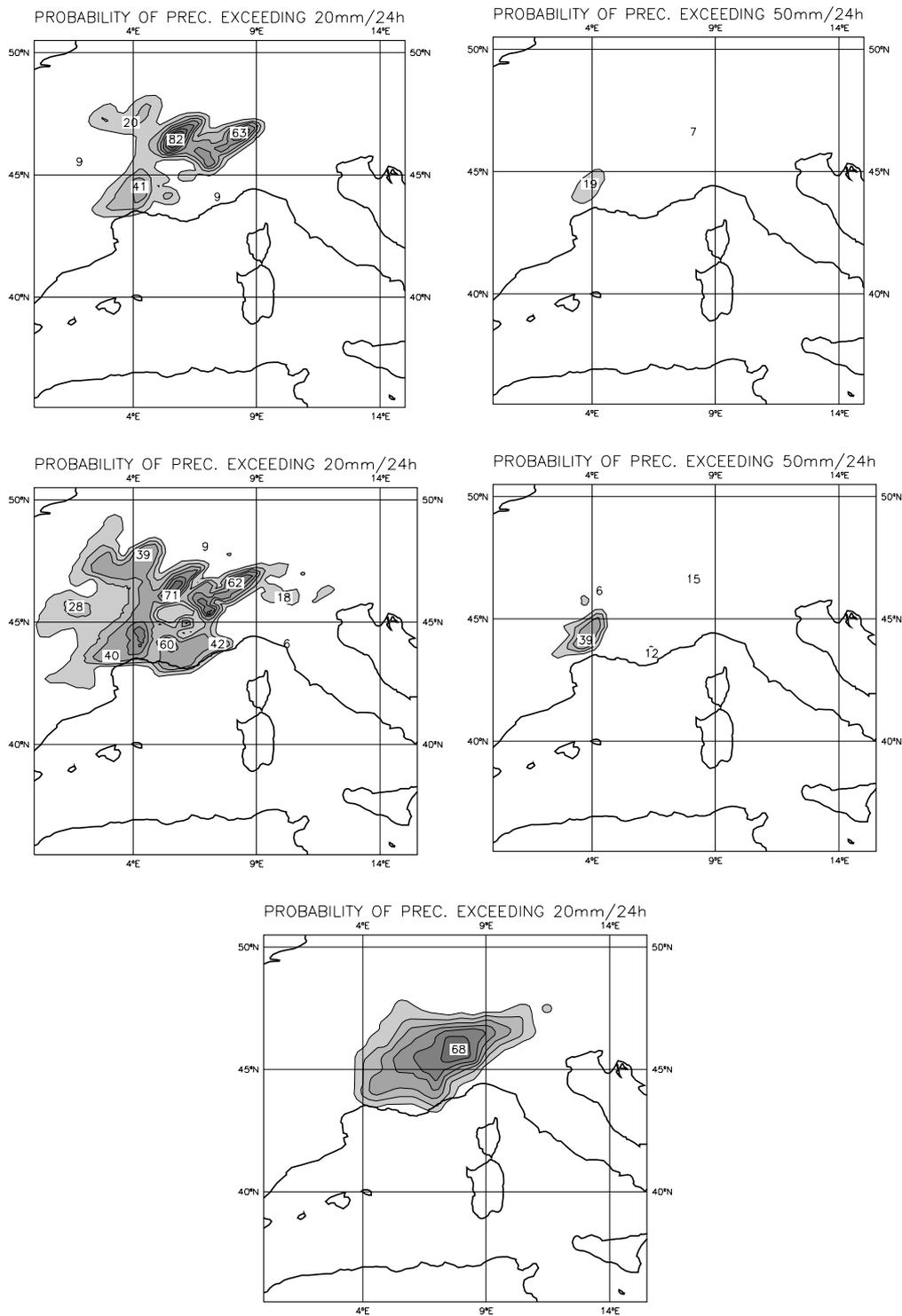


Fig 7 Vaison case: the same as fig. 5, but valid at 6 UTC of 23 September 92. The bottom-row panel reports the ECMWF probability map for rainfall exceeding the 20 mm/day threshold only.

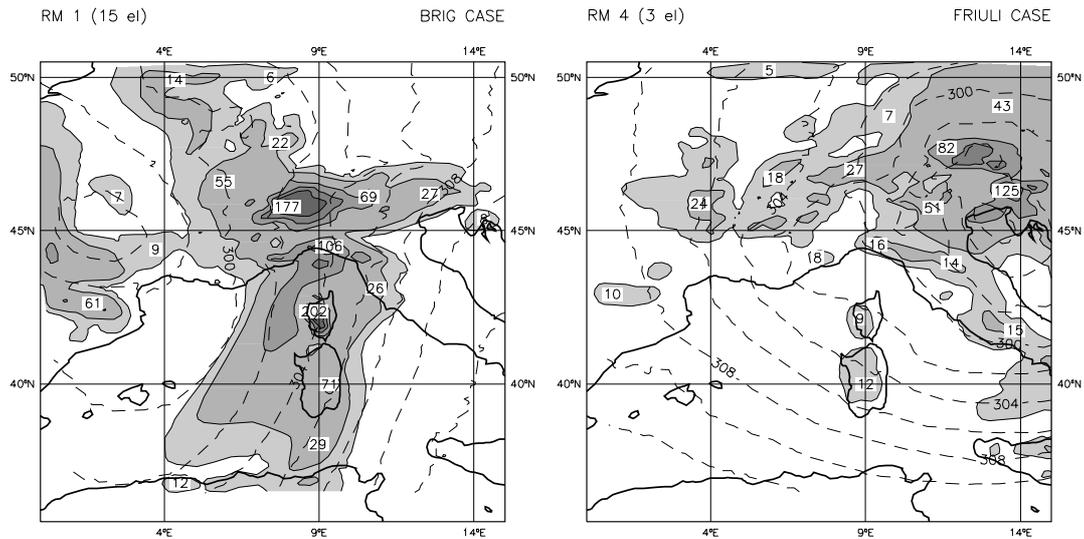


Fig 8 Left panel: Z700 (dashed contours) and 24-hour cumulated precipitation (shaded areas) predicted by LAMBO nested on the RM of the first cluster and valid at 6 UTC of 24 September 93 ($t+114h$; Brig case). Right panel: the same as the right one, but valid at 6 UTC of 20 September 95 ($t+138h$; Friuli case) and LAMBO nested on the RM of the fourth cluster.

In the Friuli case, the positive impact of the high-resolution ensemble is more evident, because of the poor performance of LAMBO integration nested on EPS control and on most of EPS RMs. In fact, the possibility of heavy precipitation over the Friuli region is highlighted by only one LEPS integration, nested on the RM of the fourth cluster (right panel of fig. 8). Despite the low population of the cluster (3 elements), this LEPS run predicts 6 days ahead heavy precipitation in the right location and provides the forecaster with the correct evolution scenario, so enabling the issue of flood alerts. The improvement with respect to EPS runs is noticeable, since none of ECMWF RMs properly predicts either the precise localisation of the event or the observed rainfall rates (see fig. 11 of the companion paper).

5.5 Global evaluation of LEPS results

It is not straightforward to summarize the performance of LEPS for the four case studies on the basis of the objective scores introduced in section 4. The whole experimental set is not large enough for a global statistical evaluation and, as already discussed, the set includes cases which span situations of both high and low predictability. A statistical evaluation of the method based on a larger experimental set will be the subject of future investigation. Nevertheless, the probabilistic scores can be used as a tool to compare the performance of the methodology in the two configurations tested (“weight” and “no-weight”).

The left and right panels of fig. 9 report, respectively, the values of the ROC area and of the BS for different thresholds, averaged over the four case studies; the solid (dotted) line represent the results for the (“weight” (“no-weight”) configuration. The ROC area values are, for any threshold, greater than 0.6 in both configurations, confirming the results of the previous subsections about the validity of LEPS methodology. It can also be noticed that slightly higher values are achieved in the “no-weight” curve. This is especially true for low rainfall rates, while the difference becomes gradually less marked as the 50 mm/day threshold is



approached. The better performance of the “no-weight” configuration is confirmed by the BS curves, for all rainfall rates, although also for this score the difference is hardly noticeable at high rainfall rates.

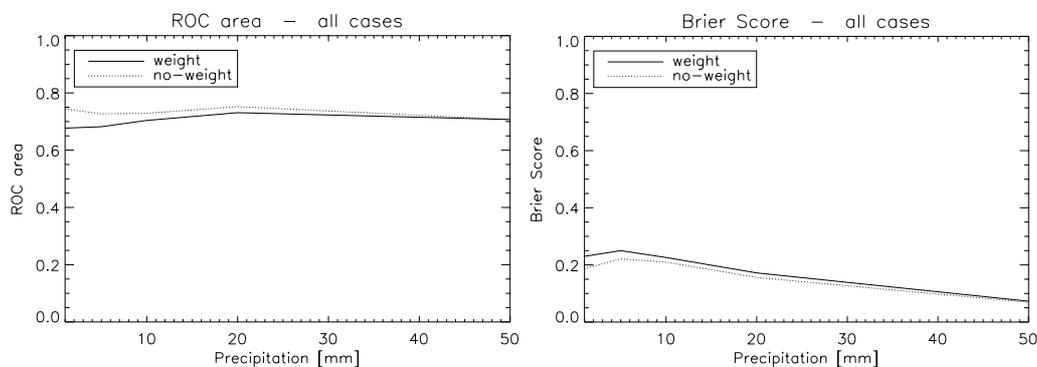


Fig 9 ROC area (left panel) and Brier Score (right panel) as a function of the rainfall threshold, averaged over the four case studies, in the “weight” (solid lines) and “no-weight” (dotted lines) configurations.

If these scores are calculated case by case (not shown), it turns out that only for the Piemonte case the “weight” configuration performs definitely better than the “no-weight” one. For the Brig case, although the reference forecast is quite accurate, the LEPS integration relative to the second cluster (the most populated one with 17 elements) is rather poor, and this penalises the weighing procedure in the calculation of both the ROC area and of the BS. On the other hand, in the Vaison and Friuli cases, the best LEPS integrations (in terms of precipitation forecast) are related to the least populated clusters, and their impact on the calculation of the probabilistic scores is enhanced in the “no-weight” configuration. In other words, the LEPS run nested on the RM of the most populated cluster is not necessarily the most accurate as far as quantitative precipitation forecasts are concerned. This can be partly due to the particular experimental set, which is made up only of heavy precipitation cases. A natural extension of this work would be to test the impact of the weighing procedure on a more representative experimental set, which includes little-rain and no-rain cases. For this set the weighing procedure could possibly lead to a reduction of the false alarms and to an improvement of the objective scores. Despite this shortcoming, it is crucial to underline that the cluster analysis could detect a forecast scenario close to the actual evolution of events, in both the high and low predictability case studies. This can be considered a very good result, since it should not be expected a probability forecast system always to predict what happens as the most probable event; what really matters is that the LEPS did successfully predict the event, albeit with a low probability.

6. CONCLUSIONS

The possibility of predicting extreme weather events over localised regions a few days ahead has been the focus of this study. Results obtained using LEPS, a high resolution ensemble prediction system based on a limited area model (LAMBO), have been presented, by investigating the medium-range predictability of four flood cases over the Alpine region. LEPS rainfall predictions have been verified against direct observations of precipitation over a large number of stations within the LAMBO domain and available at the MAP web-site. The verification has been performed on the basis of both subjective (direct comparison between forecast and measured rainfall) and objective scores (e.g. Bias and Threat Score; ROC area, Brier



Score), with very encouraging results obtained with either methodology. In two case studies (Piemonte and Brig), the high-resolution runs nested on the RMs of highly populated clusters predict, to a very good degree of spatial and temporal accuracy, the heavy rainfall which actually occurred.

In the other cases (Vaison and Friuli), the spread provided by LAMBO integrations is large enough that the extreme events are predicted by at least one of the LEPS runs (although by those nested on the RMs of the least populated clusters.) Both subjective and objective evaluations indicate that this new methodology can provide forecasters with information unavailable from either the ECMWF EPS or the deterministic LAMBO run. The precipitation values predicted by LAMBO are generally higher than in ECMWF RMs (shown in Molteni et al., 2001). Because of the higher spatial resolution, LEPS rainfall patterns are more detailed and closer to the observed ones, the small-scale orographic and mesoscale forcing being more accurately described. The better performance of the high resolution ensemble system is also evident when the attention is focussed on LEPS rainfall probability maps: the values for high rainfall thresholds are higher than in ECMWF probability maps. In cases of very heavy rainfall, a region of non-zero probability not evident or misplaced in EPS maps, is clear and reasonably well located (shown for the Piemonte and Vaison cases). This is true regardless of the specific methodology used to construct LEPS probability maps, that is whether or not each LAMBO integration is weighted according to the population of the cluster where the RM was selected (“weight” and “no-weight” configuration, respectively).

The overall results suggest that higher ROC area values and lower BS are obtained when every LEPS run has the same weight. This particular result, however, is strongly influenced by the presence of two rather unpredictable cases in our (limited) sample, and seems to be in contrast with that of the companion paper, where the average Brier Score for the “weight” configuration is slightly better than that of the “no-weight” configuration. In view of this result, further investigation is probably needed on the clustering methodology, and on the way the probability of occurrence of the different scenarios is assessed.

In any case, it is important to use the information provided by the probability maps (as in figs. 5 and 7) in the evaluation of the different members of LEPS (as presented in figs. 3 and 6), in order to maximise the amount of information extracted from the high-resolution integrations. This is particularly true for those cases of low predictability, since the LEPS methodology can still assign a non-zero probability of occurrence to the scenario which provides a realistic representation of the actual extreme event.

As previously discussed, only a small number of case studies could be investigated, due to the computational cost of the involved experimentation. Although no statistical evaluation concerning the validity of this methodology could be done, the results indicate that the limited-area ensemble prediction system can improve medium-range weather prediction of extreme precipitation events. A systematic experimentation of LEPS is being undertaken to give solid statistical grounds to this methodology, which may provide a substantial contribution to the definition of an early flood risk alarm system.

A natural extension of this work would be to test this methodology on ECMWF ensemble members initially perturbed using targeted perturbations. The ensemble so generated, usually referred to as “Targeted Ensemble Prediction System” (TEPS), is being developed at ECMWF to improve the performance of EPS over Europe in the early-medium range (Hersbach et al., 1999). TEPS performance is under investigation by a joint research group including the Royal Netherlands Meteorological Institute, Oslo University and ARPA-



SMR, in order to test the extent to which a joint use of LEPS and TEPS methodology can improve the predictions of extreme rainfall events for forecast ranges between 2 and 4 days. Research will also be developed by applying this methodology to the predictability of other particular weather events (e.g., heavy snowfall over large towns, strong winds, \dots) whose impacts on the local population could be highly reduced by a more advanced warning system than that available at the moment.

Finally, further investigation is needed on clustering algorithm, and on the methodology to select the RMs. Work is in progress to test new clustering algorithms which enable the grouping of EPS members and the selection of the RMs on the basis of more than one variable at more than one level. The results of these studies will be presented in forthcoming papers.

Acknowledgements

The authors thank Carlo Cacciamani for useful discussions and Paolo Patrino for helpful assistance throughout this work. We are also grateful to an anonymous referee whose detailed comments helped to improve the quality of the manuscript. This work was partly sponsored by a grant from GNDCI-CNR (Gruppo Nazionale Difesa Catastrofi Idrogeologiche – Consiglio Nazionale delle Ricerche).

References

- Benoit, R. and Desgagne, M., 1995: Application of the non hydrostatic MC2-LAM to the Brig 1993 flash flood event and the PYREX IOP3 foehn case. *MAP Newsletter*, 3, 24–26.
- Buizza, R., Hollingsworth, A., Lalauette, F. and Ghelli, A., 1999: Probabilistic predictions of precipitation using the ECMWF Ensemble Prediction System. *Wea. and Forecasting*, 14, 168–189.
- Buzzi, A., Fantini, M., Malguzzi, P. and Nerozzi, F., 1994: Validation of a limited area model in cases of Mediterranean cyclogenesis: surface fields and precipitation scores. *Meteorol. Atmos. Phys.*, 53, 137–153.
- Buzzi, A., Tartaglione, N. and Malguzzi P., 1998: Numerical simulation of the 1994 Piedmont Flood: Role of Orography and Moist Processes. *Mon. Wea. Rev.*, 126, 2369–2383.
- Hersbach, H., Mureau, R., Opsteegh, J. D. and Barkmeijer, J., 2000: A short-range to early-medium-range Ensemble Prediction System for the European Area. *Mon. Wea. Rev.*, 128, 3501–3519.
- Janjic, Z.I., 1990: The step-mountain coordinate: physical package. *Mon. Wea. Rev.*, 118, 1429–1443.
- Kerkmann, J.K.H., 1996: Case studies of severe convective storms in the Friuli area. *MAP Newsletter*, 4, 4–5.
- Marsigli, C., 1998: Previsioni di ensemble della circolazione atmosferica con modelli numerici ad area limitata. Thesis Degree. University of Bologna, Bologna, Italy, pp. 117.
- Mason, S.J. and Graham, N.E., 1999: Conditional probabilities, relative operating characteristics and relative operating levels. *Wea. and Forecasting*, 14, 713–725



- Mellor, G. L. and Yamada, T., 1974: A hierarchy of turbulence closure models for Planetary Boundary Layer. *J. Atmos. Sc.*, 31, 1791–1806.
- Mesinger, F. and Arakawa, A., 1976: Numerical Methods Used in Atmospheric Models, Vol. 1. GARP Publication Series, 17, WMO-ICSU Joint Organizing Committee, Geneva, pp. 64.
- Mesinger, F., Janjic, Z.I., Nickovic, S., Gavrilov, D. and Deaven, D.G., 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of Appalachian redevelopment. *Mon. Wea. Rev.*, 116, 1493–1518.
- Molteni, F., Buizza, R., Palmer, T.N. and Petroliagis, T., 1996: The ECMWF Ensemble Prediction System: methodology and validation. *Quart. J. Roy. Meteor. Soc.*, 122, 73–119.
- Molteni, F., Buizza, R., Marsigli, C., Montani, A., Nerozzi, F. and Paccagnella, T., 2001: A Strategy for High-Resolution Ensemble Prediction. Part I: Definition of Representative Members and Global-Model Experiments. Accepted at *Quart. J. Roy. Meteor. Soc.*
- Mullen, S. L. and Buizza, R., 2000: Quantitative precipitation forecasts over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, in press.
- Ritter, B. and Geleyn, J. F., 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.*, 112, 303–325.
- Senesi, S., Bougeault, P., Cheze, J. L., Cosentino, P. and Thepenier, R. M., 1996: The Vaison la Romaine flash flood: meso-scale analysis and predictability issues. *Wea. and Forecasting*, 11(4), 417–442.
- Stanski, H. R., Wilson, L. J. and Burrows, W. R., 1989: Survey of common verification methods in meteorology. WMO World Weather Watch Tech. Rep., 8, pp 144.
- Tracton, M. S. and Kalnay, E., 1993: Operational ensemble prediction at the National Meteorological Centre: Practical Aspects. *Wea. and Forecasting*, 8, 379–398.
- Tracton, M. S., Du, J., Toth, Z. and Juang, H., 1998: Short-range ensemble forecasting (SREF) at NCEP/EMC. Proceedings of the 12th AMS/NWP Conference, Phoenix AZ, 269–272.
- Vukicevic, T. and Paegle, J., 1989: The influence of one-way interacting lateral boundary conditions upon predictability of flow in bounded numerical models. *Mon. Wea. Rev.*, 117, 340–350.
- Vukicevic, T. and Errico, R., 1990: The influence of Artificial and Physical Factors upon Predictability Estimates Using a Complex Limited-Area Model. *Mon. Wea. Rev.*, 118, 1460–1482.
- Wilks, D.S., 1995: Statistical atmospheric sciences. Academic Press, San Diego, CA, USA.