

# The Soverato flood in Southern Italy: performance of global and limited-area ensemble forecasts

A. Montani<sup>1</sup>, C. Marsigli<sup>2</sup>, F. Nerozzi<sup>1</sup>, T. Paccagnella<sup>1</sup>, S. Tibaldi<sup>1</sup>, and R. Buizza<sup>3</sup>

<sup>1</sup>Regional Meteorological Service ARPA-SMR, Bologna, Italy

<sup>2</sup>Department of Earth Sciences, University of Bologna, Bologna, Italy

<sup>3</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK

Received: 19 October 2001 – Revised: 23 December 2002 – Accepted: 5 March 2003

**Abstract.** The predictability of the flood event affecting Soverato (Southern Italy) in September 2000 is investigated by considering three different configurations of ECMWF ensemble: the operational Ensemble Prediction System (EPS), the targeted EPS and a high-resolution version of EPS. For each configuration, three successive runs of ECMWF ensemble with the same verification time are grouped together so as to generate a highly-populated “super-ensemble”. Then, five members are selected from the super-ensemble and used to provide initial and boundary conditions for the integrations with a limited-area model, whose runs generate a Limited-area Ensemble Prediction System (LEPS). The relative impact of targeting the initial perturbations against increasing the horizontal resolution is assessed for the global ensembles as well as for the properties transferred to LEPS integrations, the attention being focussed on the probabilistic prediction of rainfall over a localised area. At the 108, 84 and 60-hour forecast ranges, the overall performance of the global ensembles is not particularly accurate and the best results are obtained by the high-resolution version of EPS. The LEPS performance is very satisfactory in all configurations and the rainfall maps show probability peaks in the correct regions. LEPS products would have been of great assistance to issue flood risk alerts on the basis of limited-area ensemble forecasts. For the 60-hour forecast range, the sensitivity of the results to the LEPS ensemble size is discussed by comparing a 5-member against a 51-member LEPS, where the limited-area model is nested on all EPS members. Little sensitivity is found as concerns the detection of the regions most likely affected by heavy precipitation, the probability peaks being approximately the same in both configurations.

## 1 Introduction

The increase of computer power resources has recently enabled a marked development of numerical weather prediction (NWP) models as concerns not only the possibility to afford higher horizontal and vertical resolution but also a more detailed description of the physical processes related to atmospheric instability. As a consequence, the average performance of NWP models has noticeably improved and the operational use of sophisticated global-model ensemble prediction systems is widespread in several weather centres (Toth and Kalnay, 1995; Houtemaker et al., 1996; Molteni et al., 1996). Nevertheless, the forecast of localised and severe weather events (e.g. heavy rainfall, strong winds, cold temperature anomalies) is still nowadays a challenging problem, despite the more and more careful detection of precursors, developments and mature phase of this kind of events. The key role played by mesoscale and orographic-related processes can seriously limit the predictability of intense and localised events. Although the use of high-resolution limited-area models (LAMs), nested on the fields predicted by the global runs, has improved the short-range prediction of locally intense events, it is sometimes difficult to forecast accurately their spatio-temporal evolution for ranges longer than 48 h. Therefore, several methodologies have recently been implemented and experimented to tackle the problem and improve upon the short to medium-range prediction of those surface fields heavily affected by local processes, the attention being often focused on quantitative precipitation forecast (QPF).

Thanks to the generation of ensemble systems, many weather centres have given more and more emphasis to the probabilistic approach, which enable to estimate the predictability of the atmospheric flow and assess the reliability of the deterministic forecast beyond the very short range.

As for the use of global models, the European Centre for Medium-Range Weather Forecasts (ECMWF) has recently increased the horizontal resolution of the operational ensemble

ble prediction system (EPS) from  $T_L159L40$  to  $T_L255L40$  (that is, approximately from 120 to 80 km in grid spacing at mid-latitudes) in order to enhance the description of orographic-related processes (Buizza and Hollingsworth, 2001). The introduction of a stochastic forcing in the physical parameterisation of ECMWF model has been also introduced to account for model errors and to increase EPS spread (Buizza et al., 1999). In addition to this, ECMWF has recently developed a targeted ensemble prediction system (TEPS), where the initial perturbations for the global-model integrations are constructed so as to maximise the 48 h total-energy perturbation growth over the European area. Hersbach et al. (2000) showed that TEPS performance in the probabilistic prediction of heavy precipitation events over Europe (rainfall exceeding  $20 \text{ mm day}^{-1}$ ) was slightly better than the EPS for ranges between 72 and 96 hours. The validity of running TEPS so as to provide initial and boundary conditions for limited-area integrations was tested by Frogner and Iversen (2000a, b) and by Montani et al. (2001), the results being encouraging in both cases. As an alternative approach using global models, Atger (1999) and Ziehmann (2000) propose the “poor man’s EPS”, obtained by joining together forecasts from different weather centres. Model uncertainties are implicitly considered in this approach, which turns out to provide reliable probabilistic forecasts both in the early and in the medium range.

If the attention is focused on the use of limited-area models, it is worth mentioning the development of the short-range ensemble forecasting (SREF) at the National Center for Environmental Prediction (NCEP). This prediction system is generated by high-resolution versions of Eta and Regional Spectral Model (Tracton et al., 1998) and turns out to be very accurate to assess the QPF for forecast ranges up to 63 h. Alternatively, Stensrud et al. (1999a, b) propose the generation of ensembles by either perturbing model physics or by constructing initial perturbations within the limited-area model itself, the more encouraging results being achieved with the former methodology. As a further approach, the Regional Meteorological Office of Emilia-Romagna ARPA-SMR (in Bologna, Italy) developed the Limited-area Ensemble Prediction System (LEPS). By means of a clustering-selection technique, ECMWF members are first grouped into five clusters, then a representative member (RM) is selected within each cluster (Molteni et al., 2001; Marsigli et al., 2001). The RMs provide both initial and boundary conditions for the integrations with a limited-area model, which is run five times (once per RM), so generating a small-size high-resolution ensemble for forecast ranges up to 120 h. Hence, the probabilistic products typically generated at ECMWF (e.g. probability maps for rainfall rates or wind intensity exceeding particular thresholds) can be produced on the basis of the information provided by LAM integrations, each LEPS run being weighted according to the population of the cluster where the RM is selected. The results presented by Marsigli et al. (2001) and Montani et al. (2001) indicate that LEPS methodology allows to combine the benefits gained by the probabilistic approach (a set of different evolution scenarios

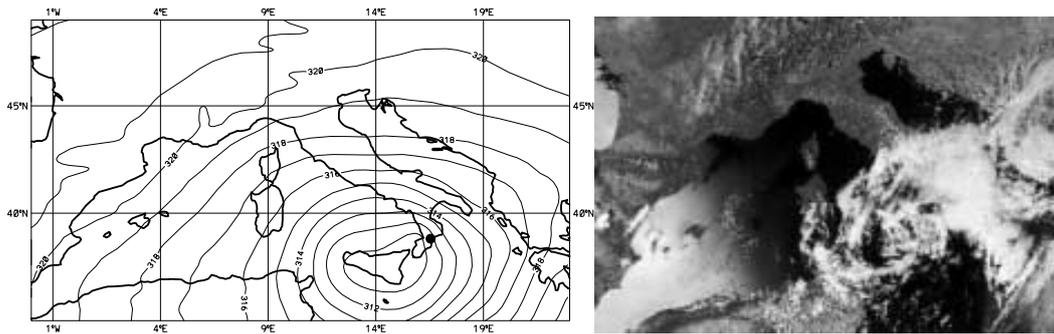
is highlighted) with the high-resolution detail of LAM integrations, without having to pay too much in terms of computer power. It has been shown that, over a number of test cases and for several forecast ranges (48–120 h), LEPS performs better than EPS and TEPS as concerns the estimate of precipitation intensity as well as the detection of the regions most likely affected by heavy rain. An objective verification against observed value of precipitation confirms the accuracy of LEPS forecasts.

In this work, different configurations of ECMWF global ensemble and of LEPS system are tested for QPF purposes. For one case study, relative to a disastrous flood which affected Southern Italy in September 2000, the impact of increased horizontal resolution and the use of European-targeted perturbations in ECMWF ensemble is evaluated for forecast ranges between 108 and 60 h. A new clustering-selection technique (referred to as super-ensemble technique in the remaining of the paper) is proposed and its impact on the performance of LEPS system for each of the global-model configurations is tested. It is aimed to investigate the extent to which flood risk alerts could have been issued a few days ahead the rainfall event and, then, either confirmed or dismissed on the basis of either ECMWF or LEPS forecasts. For one particular range, the performance of a 51-member LEPS is also evaluated (that is, the LAM is nested on each of EPS members) against a small-size LEPS. Although this methodology (referred to as “brute-force” approach) is not affordable on an operational basis with present computing facilities, it enables to assess the impact (if any) brought about by the ensemble-size reduction of the clustering-selection technique. It is immediately worth pointing out that the results presented in this work are relative to a single-case evaluation, since the attention is focussed on the potential benefits of LEPS system rather than on its objective evaluation based on many cases.

The paper is organised as follows: in Sect. 2, the synoptic description of the case study is reported, while the super-ensemble methodology followed to generate the limited-area ensembles is presented in Sect. 3. The results from the global model and the limited-area integrations are presented in Sects. 4 and 5, respectively. In Sect. 6, the results obtained by the “brute-force” approach are presented. Finally, conclusions are drawn in Sect. 7.

## 2 Case study description

At 12:00 UTC on 7 September 2000, the synoptic situation over the Mediterranean Sea was characterised, at upper levels, by a large-scale trough, northwesterly winds blowing over the Italian peninsula. At the surface, a mean-sea-level pressure minimum developed off the Gulf of Genoa and was then advected southeastwards, reaching a position at about  $40^\circ \text{ N } 13^\circ \text{ E}$  by 00:00 UTC 8 September. In the following hours, the cyclonic circulation intensified (minimum of 1006 hPa) and slowly moved over Sicily. The system remained almost stationary over southern Italy for two more



**Fig. 1.** Left panel: ECMWF analysis of 700 hPa geopotential height at 00:00 UTC 10 September 2000; contour interval: 10 dam. The black dot denotes the approximate location of Soverato (38.68° N 16.55° E). Right panel: meteosat image (visible channel) valid at 15:00 UTC 9 September 2000.

days, mainly because of a 500 hPa anticyclonic circulation over eastern Mediterranean and the Arabian peninsula. The signature of the low-pressure system was noticeable at both lower and upper levels, a cut-off low being evident at 700, 500 and 300 hPa (not shown). Moist southeasterly winds impinged over the Southern Italian Apennines and heavy precipitation was observed over a large part of Calabria (the western-tip of Southern Italy). Figure 1 shows the geopotential height pattern at 700 hPa valid at 00:00 UTC 10 September (left panel) as well as the visible satellite image from METEOSAT (right panel), where the cyclonic system, covering a good deal of the Italian peninsula, is well evident. Between 9 and 10 September, rainfall peaks above 300 mm in 24 h were recorded close to the village of Soverato (on the eastern coast of Calabria at 38.69° N 16.55° E), causing landslides, great disruption and losses of life. The four panels of Fig. 2 show the observed values of precipitation, accumulated every 6 h in four stations in the vicinity of the flooded areas, while Fig. 3 reports the locations of the four stations. Observations clearly indicate that the most rainy period includes the first hours of 10 September, with the exceptionally high values at the station of Chiaravalle Centrale (top-left panel), with 185 mm in 6 h and about 350 mm in 24 h.

### 3 Methodology

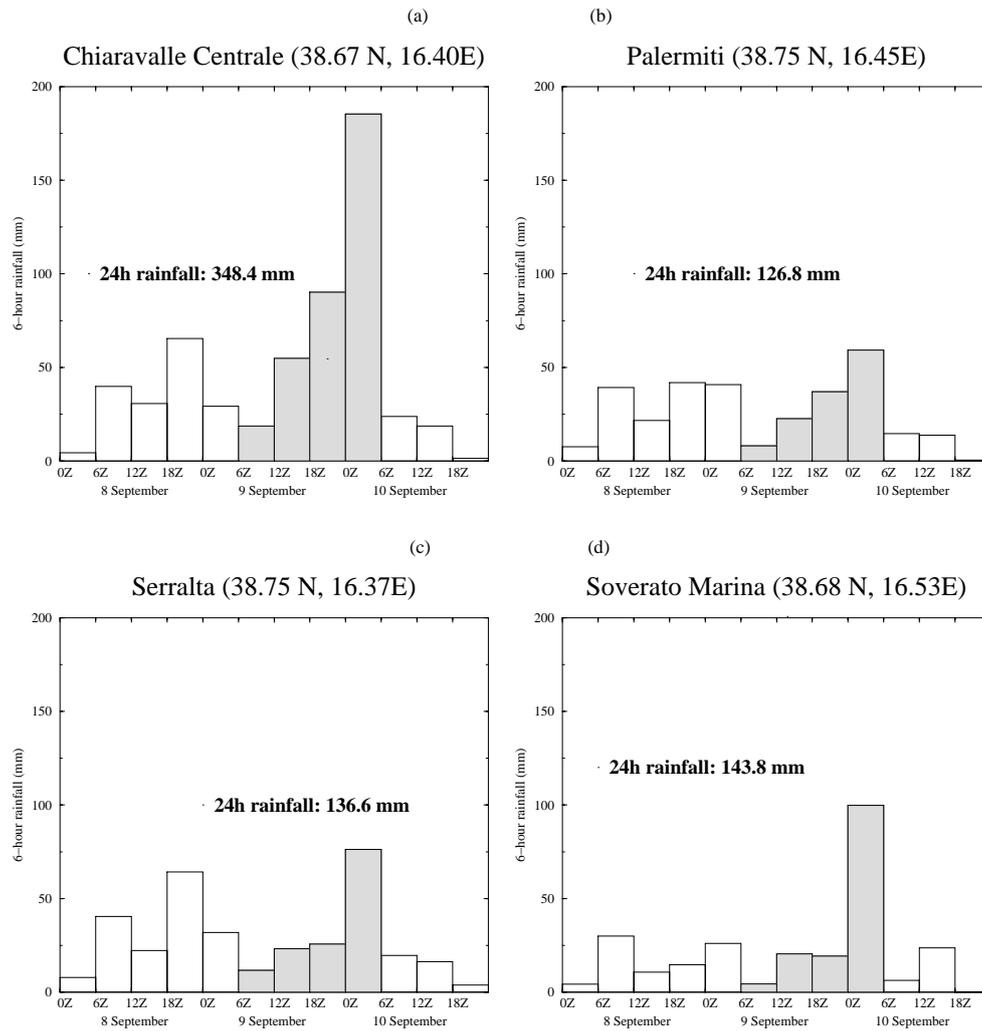
As stated in the introduction, three different configurations of ECMWF global ensemble are considered, namely the operational EPS, the targeted EPS and a high-resolution version of EPS (hereafter, referred to as “OPE”, “TEPS” and “HRES”, respectively). Each ensemble consists of 51 runs of ECMWF global model, one integration (the control) starting from the unperturbed analysis, the other fifty ones from different analyses generated by perturbing the control with linear combinations of scaled singular vectors, fast-growing initial perturbations which maximise total energy over a 48 h time interval (Molteni et al., 1996; Buizza et al., 1998). The main features of the three configurations can be summarised as follows:

- “OPE”: it is run at  $T_L159$  L40 (about 120 km of horizontal resolution with 40 vertical levels) and was the operational set-up of EPS at the time of the flood;
- “TEPS”: it is run at  $T_L159$  L40 (Hersbach et al., 2000) and the initial perturbations are targeted to maximise total energy growth over the European area (about 35° N–75° N, 40° W–30° E) instead of the Extra-Tropical Northern and Southern Hemisphere, like in “OPE”;
- “HRES”: the same as “OPE” except for the higher horizontal resolution ( $T_L255$  L40, that is about 80 km in the horizontal)<sup>1</sup>;

For each configuration, the performance of three successive ECMWF ensembles is examined, starting at 12:00 UTC of 5, 6 and 7 September 2000. Therefore, three forecast ranges (108, 84 and 60 h, respectively) are considered for the performance of each of the “OPE”, “TEPS” and “HRES” global ensembles. As concerns the generation of LEPS system, the following procedure, referred to as “super-ensemble methodology”, is followed:

1. the three successive ensembles are joined together into a unique super-ensemble of 153 members;
2. at verification time, all members of the super-ensemble are grouped into 5 clusters, the discriminating variable being a combination of four variables at three pressure levels: the two horizontal wind components, the geopotential height and the specific humidity at 500, 700 and 850 hPa (Montani et al., 2001);
3. the clustering variables are standardised (for each variable, the mean over the clustering area is calculated and subtracted from any grid-point value; then, the result is divided by the standard deviation, so as to obtain a non-dimensional quantity);

<sup>1</sup> On 21 November 2000, the “HRES” configuration became the operational suite of ECMWF EPS.

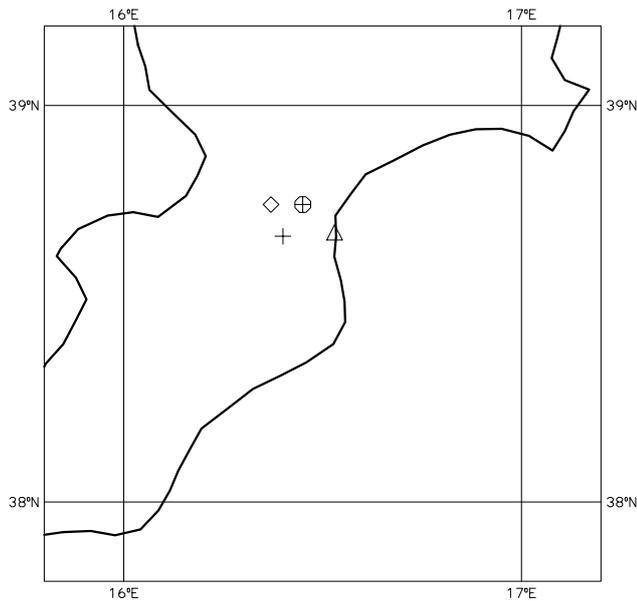


**Fig. 2.** 6-hour rainfall observed in the vicinity of the flooded region from 8 to 10 September 2000: Chiaravalle Centrale (top-left panel), Palermiti (top-right), Serralta (bottom-left) and Soverato Marina (bottom-right). The location of each station and the precipitation cumulated over the 24-hour period ending at 6:00 UTC of 10 September 2000 (gray areas in the panels) are also reported (data kindly provided by “Servizio Idrografico e Mareografico Nazionali; Dipartimento di Catanzaro”).

4. the space-averaged quadratic distances between the super-ensemble members are computed for all variables at all levels;
5. the cluster analysis is performed over Southern Europe (approximately  $53^{\circ}\text{N}$ – $35^{\circ}\text{N}$   $5^{\circ}\text{W}$ – $25^{\circ}\text{E}$ ) and the complete-linkage algorithm<sup>2</sup> (Wilks, 1995) is used to construct the five clusters;
6. one representative member (RM) is chosen from each cluster among the members within the same cluster, the discriminating variables being the same as before;
7. the RM of a cluster is that element which minimises the ratio between its distance from the other members of its own cluster and its distance from the members of the other clusters;
8. the so-selected RMs provide both initial and boundary conditions for the integrations with the limited-area model;
9. the 5 limited-area integrations generate the (small-size) high-resolution LEPS;
10. the above steps are repeated for the “OPE”, “TEPS” and “HRES” configurations, so as to generate three LEPS ensembles.

If compared to the methodologies presented in Molteni et al. (2001), Marsigli et al. (2001) and Montani et al. (2001), the innovative element of the super-ensemble technique is the grouping of different ensembles run on consecutive days (like in a lagged-average technique) so as to get a highly-populated ensemble (Hoffman and Kalnay, 1983), which will

<sup>2</sup>The complete-linkage algorithm is preferred, because, unlike other algorithms, it has the ability to generate few-member clusters (possibly predicting extreme events)



**Fig. 3.** Geographical location of the stations of Fig. 2: the symbol + denotes Chiaravalle station, ⊕ denotes Palermiti, ◇ denotes Serralta, Δ denotes “Soverato Marina” (a few hundred metres from the village of Soverato).

provide the RMs for the limited-area integrations. It is clear that the ensembles generating the super-ensemble start from different analyses and are considered at different forecast times. Preliminary results, based on 38 cases during 1999, confirmed that the skill of the most recent ensemble forecasts is better than that of older ensembles and of the whole super-ensemble. Despite this, mixing ensembles provides higher quality probabilistic forecasts, possibly because the super-ensemble explores regions of the phase-space of the system otherwise left unexplored by the most recent ensemble. Results have indicated, for example, that the Brier Score for the probabilistic prediction of rainfall events is higher for the super-ensemble than for an ensemble based on only the most recent members (that is, the most recent ensemble has a lower Brier Score, hence it performs better). On the other hand, the super-ensemble turns out to be a more reliable ensemble than the most recent one (“more reliable” means that a higher match between forecast probability and verified occurrence for a particular event is encountered for the super-ensemble). In fact, the percentage of outliers decreases and a stronger relationship is found between each cluster population and the probability of the analysis being inside the cluster. The benefits provided by the better reliability of the super-ensemble are preferred to the loss of skill highlighted by the Brier Score, because the former allows a proper weighting of each nested run. The reliability of the super-ensemble forecast system, tested over a larger number of cases and different EPS configurations, seems to be confirmed by the most recent studies performed at ARPA-SMR, which will be published in forthcoming publications. Following these results, it is assumed, as a working hypothe-

sis, that the cluster population is indicative of the possibility of occurrence for the weather scenario highlighted by the RM of that cluster and that this property is transferred to the limited-area integration. Hence, a LAM run nested on a RM selected within a highly populated cluster will be given a greater weight than a run where boundary and initial conditions were provided by a RM in a cluster with fewer members.

The limited-area model used to generate LEPS is the Limited Area Model Bologna (LAMBO), operational at ARPA-SMR since 1993. LAMBO is based on an early version of the NCEP Eta Model (Mesinger et al., 1998) and, in these experiments, it is run up to 108 h, at 20 km of horizontal resolution and with 32 vertical levels. As already described, the initial and boundary conditions of LAMBO runs are provided by the fields of ECMWF RMs. More details about LAMBO can be found in Janjic (1990) and Marsigli et al. (2001).

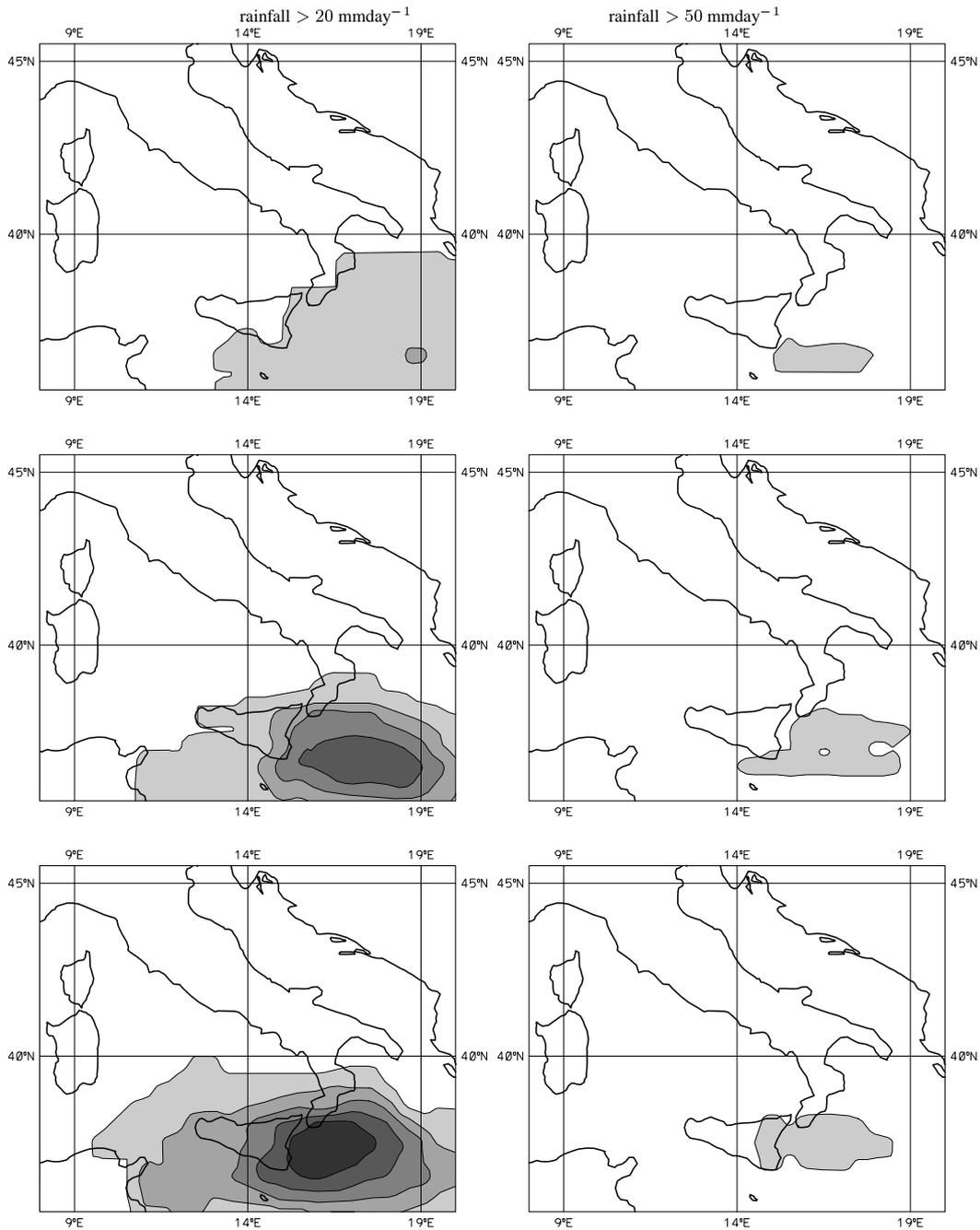
#### 4 Global-model ensembles

The probabilistic forecasts of the operational EPS, of the targeted EPS and of the high-resolution EPS are shown, for different forecast ranges, in Sects. 4.1, 4.2 and 4.3, respectively. The attention is focused on the prediction of rainfall rates exceeding the 20 and 50 mm day<sup>-1</sup> thresholds in order to assess merits and shortcomings of each configuration as regards the potential guidance provided to issue flood alerts before the weather event.

##### 4.1 Operational EPS

The performance of the EPS operational at the time of the flood (hence, the only one available to the bench forecaster), is presented for the three different forecast ranges in Fig. 4. The panels show the probability maps for rainfall exceeding 20 and 50 mm day<sup>-1</sup> (left and right-column panels, respectively) and the results can be summarised as follows:

- at the longest forecast range (108 h, top-row panels), the probability of rainfall is very low for either thresholds, because the amount of expected rainfall is underestimated in most of ECMWF runs. The locations affected by precipitation are misplaced to the southeast, because the final-time position of the cyclone is not over Sicily, like in Fig. 1, but rather in middle-east Mediterranean. The probability of rainfall exceeding 20 mm day<sup>-1</sup> is about 20% over the sea and 2% over some parts of the Calabria region, the signal at the higher threshold being negligible everywhere in the area (top-left and top-right panels, respectively).
- The previous considerations apply also in substance to the 84-hour forecasts (middle-row panels). For this range, the rainfall probability maps indicate higher probabilities of precipitation. In fact, a peak can be noticed for the lower threshold (middle-left panel), but the displacement problem of the cyclone location is still



**Fig. 4.** ECMWF “OPE” experiments: probability maps (valid at 00:00 UTC 10 September 2000) for 24 h cumulated rainfall exceeding 20 (left column) and 50 mm day<sup>-1</sup> (right column) for the ensembles starting at 12:00 UTC 5 September 2000 (top row; 108 h forecast), 12:00 UTC 6 September 2000 (middle row; 84 h forecast), 12:00 UTC 7 September 2000 (bottom row; 60 h forecast). Contour intervals: 2%, 20%, 40%, 60%, 80%.

well evident. In the region actually affected by heavy precipitation, the probability is below 10% and little assistance to a forecaster could be given to issue flood alerts in the Calabria region.

- At the 60-hour range, the predicted location of the cyclone is partly more accurate in some ECMWF runs. Nevertheless, the predicted probability maxima could not give a useful guidance to forecasters because of the

misplacement of the cyclone location in most EPS forecasts. The bottom-left panel of Fig. 1 shows a 30–40% of probability of rainfall rate exceeding 20 mm day<sup>-1</sup> in the correct region, with a peak above 80% evident between Calabria and Sicily. At the higher threshold (bottom-right panel), the performance of the operational EPS is still poor, since only one ECMWF forecast predicts precipitation above 50 mm day<sup>-1</sup> over Calabria

(hence only the 2% contouring is plotted).

Therefore, the overall performance of the operational ensemble is not particularly accurate at any considered forecast range. Because of the displacement of the cyclone position and of the poor representation of the Apennines' forcing on the impinging flow, the amount of deployed rainfall is highly underestimated. Hence, both intensity and location of the rainfall event are not properly captured. Only at the shortest forecast range, the possibility of rainfall above 20 mm is highlighted in the correct region, although with low probability of occurrence.

#### 4.2 Targeted EPS

The performance of the targeted EPS turns out to be very similar to that of the "OPE" configuration. At all forecast ranges, the regions highlighted as the most likely to be affected by heavy rainfall are the same in both configurations (hence, the maps relative to the TEPS experiments are not shown). Very small differences can be noticed for the 60-hour forecasts. At the 50 mm threshold, TEPS runs predict higher chances of heavy precipitation. Nevertheless, the peak is located over the sea and the probability is still below 10% in the vicinity of the Soverato area (not shown).

Therefore, the performance of the TEPS ensembles is approximately as poor as that of the operational EPS and the impact of targeted perturbations is not particularly noticeable in terms of rainfall probabilistic forecast. Only at shortest range, the signal for heavy precipitation is slightly clearer than in EPS runs, although the predicted probability is low over the actually flooded region.

#### 4.3 High-resolution EPS

Unlike the two previously-examined global prediction systems, the "HRES" configuration provides more accurate probabilistic forecasts of the Soverato flood. Figure 5 shows the impact of increased horizontal resolution on the rainfall probability maps of the global ensemble. The top-left panel of the figure indicates that, already at the longest range, there is a non-zero probability of rainfall exceeding 20 mm day<sup>-1</sup> over the Soverato area. As the prediction range decreases, the signal becomes clearer. At the lower threshold, both 84 and 60-hour forecasts indicate high probabilities of rainfall over the Southern tip of the Italian Peninsula as well as over Sicily, thus including the flooded regions (middle-left and bottom-left panels of Fig. 5). The probability is about 20% also for precipitation exceeding the 50 mm day<sup>-1</sup> threshold (middle-right and bottom-right panels), although the southward misplacement of the cyclone location is still present in most ECMWF runs.

Hence, the impact of increased horizontal resolution is definitely noticeable and greater than that of targeting the initial perturbations. About one day in predictability is gained in comparison with the previous configurations, presumably thanks to the more accurate modelling of the orography-related processes. On the basis of these "HRES" forecasts,

**Table 1.** Cluster population and RMs of the super-ensembles for the "OPE", "TEPS" and "HRES" configurations (columns 2, 3 and 4, respectively). The clustering time is set at 00:00 UTC on 10 September 2000. The RMs (in parenthesis) are expressed as MMDD"p"nn, where nn indicates the perturbed element of the ensemble started on day DD, p0 denoting the control run. See text for more details

Cluster number	Cluster population and RM		
	"OPE"	"TEPS"	"HRES"
cl 1	41 (0906p46)	17 (0905p0)	72 (0907p0)
cl 2	16 (0905p24)	25 (0905p18)	19 (0905p33)
cl 3	58 (0907p47)	38 (0906p30)	17 (0907p40)
cl 4	3 (0905p34)	22 (0907p0)	23 (0906p19)
cl 5	35 (0907p7)	51 (0905p33)	22 (0906p24)

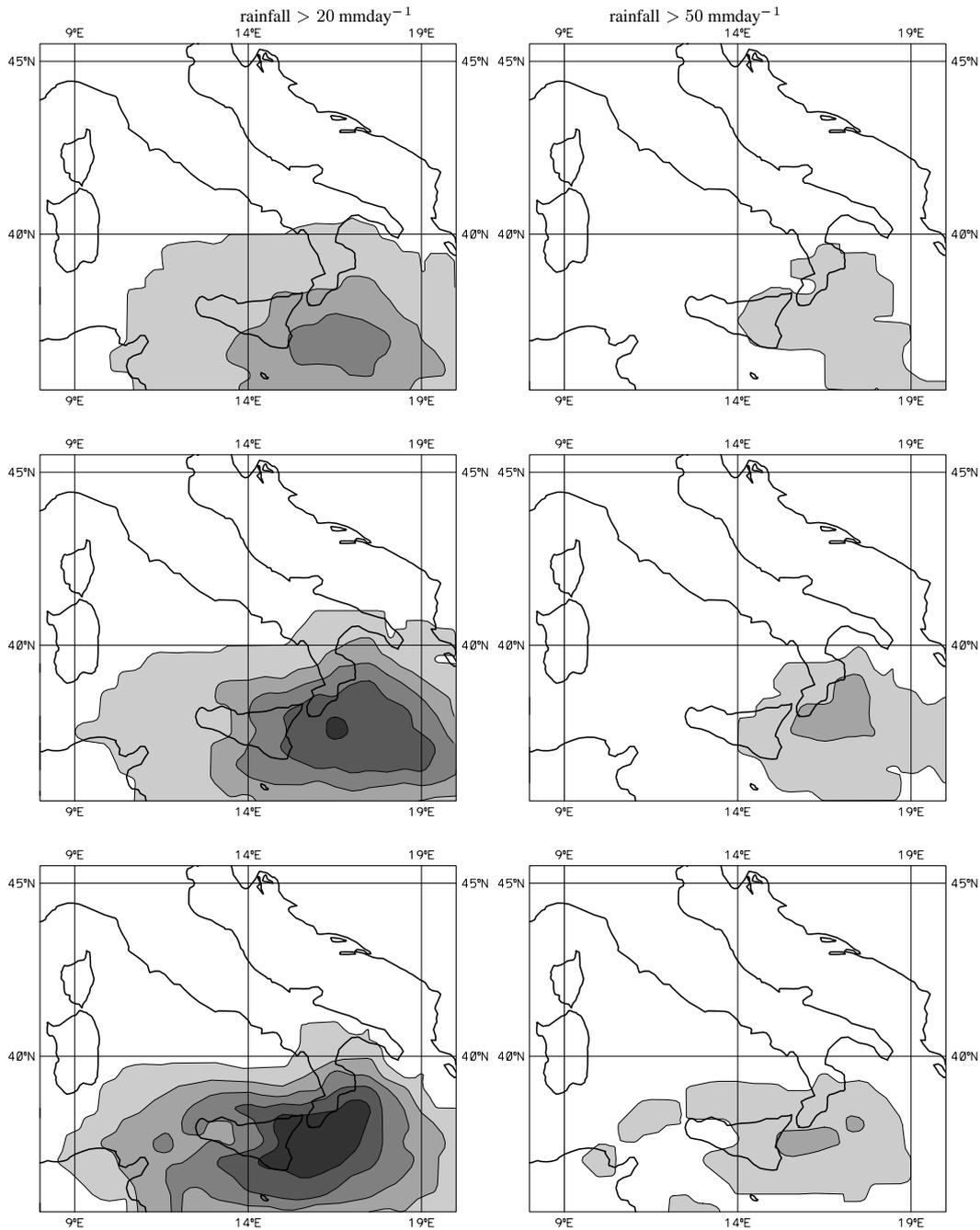
more guidance could have been given to a forecaster to issue a "weak" and widespread flood alert over Southern Italy, although it would have been difficult to detect with sufficient spatial detail the locations most heavily affected by heavy precipitation.

## 5 Limited-area model ensembles

The extent to which the properties of the global-model systems described in the previous section are transferred to the limited-area integrations is assessed by considering the probabilistic rainfall forecasts based on the five LAMBO integrations nested on the RMs calculated in the "OPE", "TEPS" and "HRES" configurations.

As described in Sect. 3, the "super-ensemble" method enables to select 5 RMs out of the 153 elements which compose the super-ensemble. Table 1 reports cluster populations and RMs for each of the three configurations. It can be noticed that, in any case, the method selects members from different ensembles which compose the super-ensemble. As an example, relative to the "OPE" experiments (column 2 of the table), the first RM is selected from the ensemble starting on the 6 September (more precisely, the perturbed element "p46" is selected within a cluster with 41 elements), while the second and the fourth RMs are chosen from those of the oldest ensemble (elements "p24" and "p34", relative to clusters with 16 and 3 elements, respectively). The two other RMs belong to the most recent ensemble, starting at 12:00 UTC of 7 September, members "p47" and "p7" being selected in the third and fifth clusters (58 and 35 elements, respectively). Since the clustering time is fixed at 00:00 UTC on 10 September 2000, different prediction times (108, 84 and 60 h) are considered for each RMs, depending on the initial time of the run.

In Sects. 5.1 and 5.2, the performance of LEPS probabilistic forecasts is evaluated for LAMBO nested on the RMs calculated in the "OPE" and "HRES" configurations, respectively. The "TEPS" configuration is not reported, because



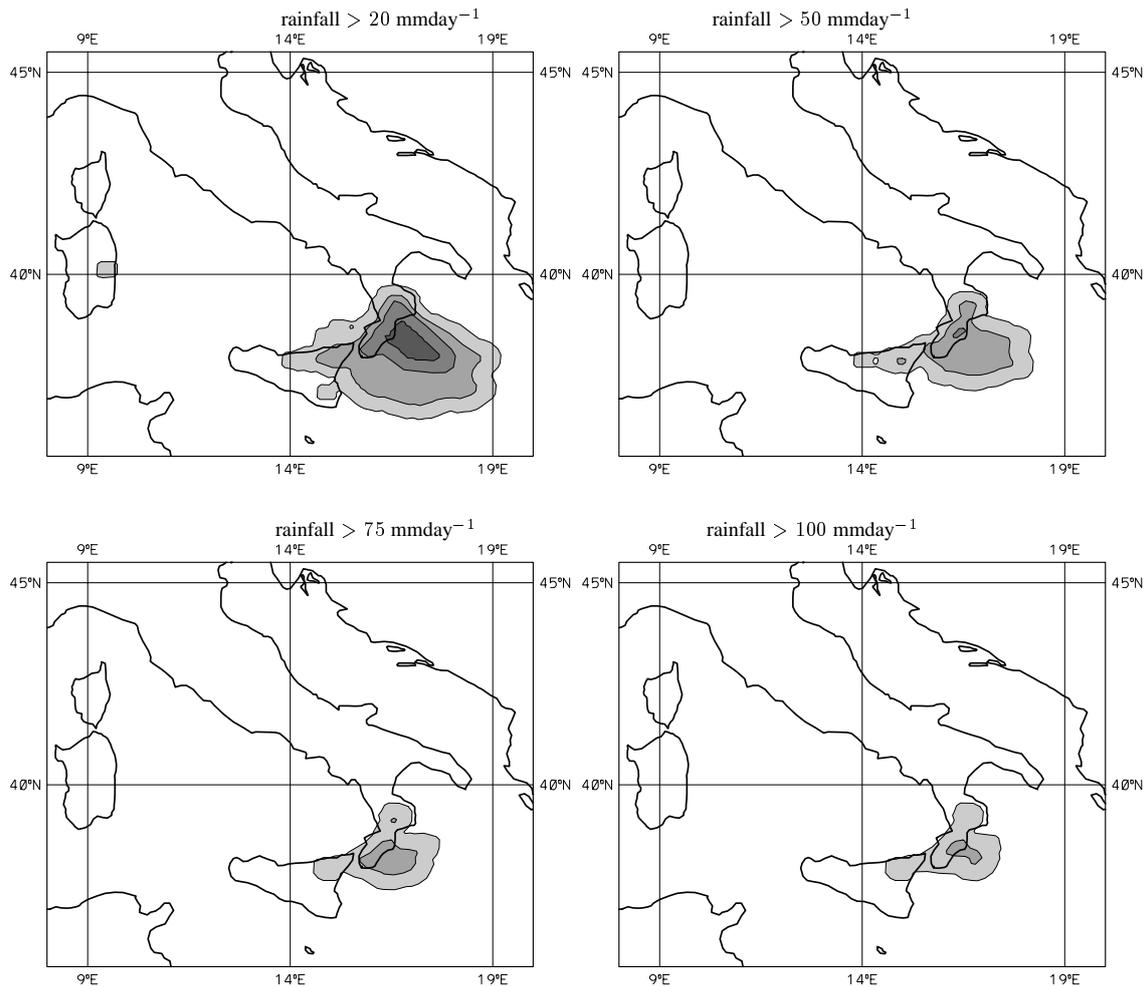
**Fig. 5.** The same as Fig. 4 but for ECMWF “HRES” experiments.

the results are substantially identical to those obtained when LEPS is generated from the “OPE” ensembles.

### 5.1 LEPS on operational EPS

Figure 6 shows the LEPS rainfall probability maps, once the RMs providing initial and boundary conditions for LAMBO runs are selected among those of the “OPE” super-ensemble. In order to describe in great detail the performance of the high-resolution forecasts for heavy rainfall scenarios, two more thresholds ( $75$  and  $100$  mm day $^{-1}$ ; bottom-row panels)

are used to evaluate the prediction skill of the system. Since the most recent LAMBO forecasts are nested on the RMs starting at 12:00 UTC 7 September 2000, LEPS maps have to be compared to those obtained by the global ensemble at the shortest range, namely the 60-hour forecasts (as reported in the bottom-row panels of Fig. 4). A comparison between the panels of Figs. 4 and 6 shows that LAMBO precipitation patterns are sharpened than in the ECMWF runs (this might come from the fact that the spread of a 5-member LEPS is probably smaller than in a 51-member EPS). At the  $20$  mm day $^{-1}$  threshold (top-left panel of Fig. 6), the geographical exten-

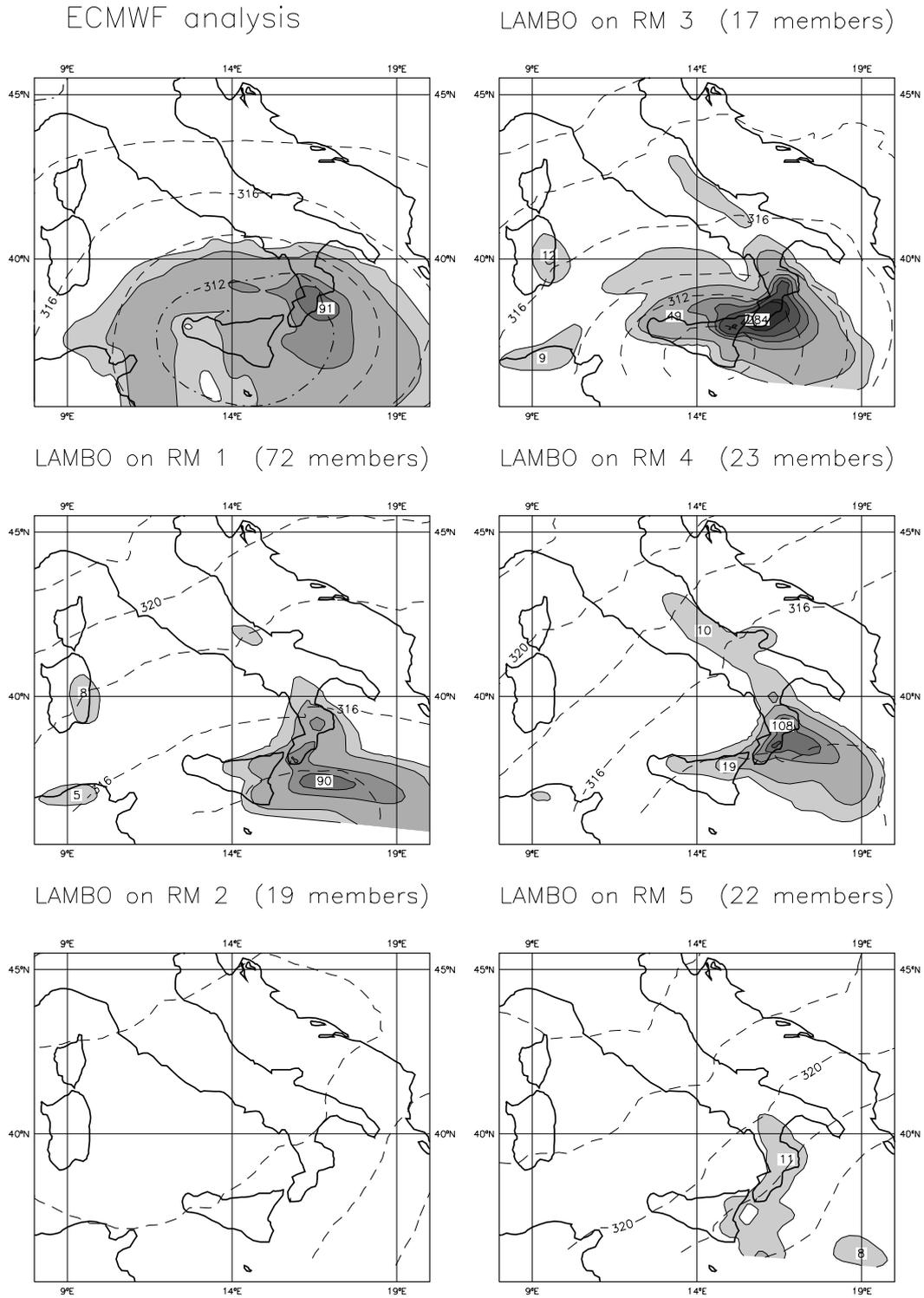


**Fig. 6.** LEPS nested on the RMs of ECMWF “OPE” super-ensemble: probability maps (valid at 00:00 UTC 10 September 2000; 60 h forecast) for 24 h cumulated rainfall exceeding 20 (top-left panel), 50 (top-right), 75 (bottom-left) and 100 mm day<sup>-1</sup> (bottom-right panel). Contour intervals: 2%, 20%, 40%, 60%, 80%.

sion of the region likely to be affected by precipitation is approximately centred around the Soverato area and appears to be heavily influenced by orography. The forcing of the Southern Apennines appears to be well captured in the LEPS runs, since heavy precipitation is much more likely on the up-wind rather than on the lee side of Southern Apennines, the wind blowing from the southeast. Unlike in the ECMWF-model probability maps, the top-right panel indicates a fair chance, above 40%, of rainfall exceeding 50 mm day<sup>-1</sup> in the correct region. At the higher thresholds, the impact of using a high-resolution limited-area model is more evident. The localisation of the regions most heavily affected by the flood is more accurately predicted. In fact, the bottom-row panels of Fig. 6 indicate probabilities about 30% for rainfall exceeding 75 and 100 mm day<sup>-1</sup>, the peaks being located in the Soverato area. Therefore, the possibility of the flood scenario is clearly highlighted 60 hours before the event. More than two days would have been available to issue hazardous-weather warnings and alert the local population, in order to limit damages and loss of life.

## 5.2 LEPS on high-resolution EPS

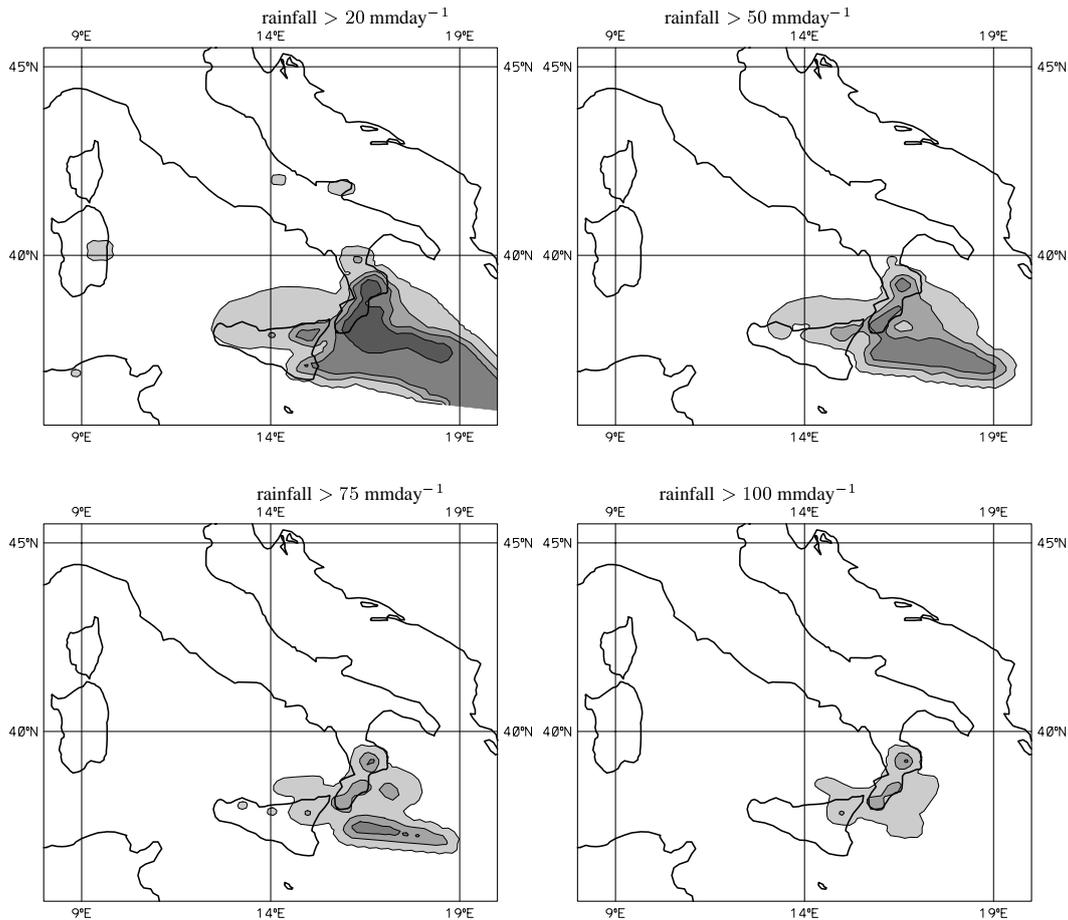
Figure 7 shows the performance of the 5 LEPS integrations in terms of geopotential height at 700 hPa and total precipitation. From a subjective point of view, only the forecast provided by LAMBO nested on the RM of the second cluster (bottom-left panel) is very poor in terms of both circulation field and quantitative rainfall deployed in the flooded area, when compared to the verifying ECMWF “analyses” (top-left panel). This LEPS integration is nested on the RM of a little-populated cluster (19 members, according to Table 1); hence, its weight on the rainfall probability maps will be relatively low. On the other hand, the remaining LEPS runs all provide accurate predictions; in particular, LAMBO nested on the RM of the third cluster, as reported in the top-right panel of the figure, is very good in terms of geopotential height pattern. Both this run and the one nested on the RM of the fourth cluster (middle-right panel) predict rainfall amounts above 100 mm in the actually flooded areas. The LAMBO integration nested on the RM of the most populated



**Fig. 7.** 700hPa geopotential height (dashed contours) and 24-hour cumulated precipitation (shaded areas) predicted by LAMBO nested on RMs 1–5 of “HRES” super-ensemble (middle-left, bottom-left, top-right, middle-right and bottom-right panels, respectively), valid at 00:00 UTC of 10 September 2000. The top-left panel reports the ECMWF verifying analyses. Contour intervals: 20 dam; 2, 10, 50, 75, 100, 150 and 300 mm.

cluster (the first one with 72 members) is relatively accurate in terms of geopotential height, the precipitation amount being slightly underestimated over the Soverato area. Once

each of the LEPS integrations is given a weight proportional to the population of the cluster where the RM was selected, it is possible to generate, like in Sect. 5.1, the high-resolution



**Fig. 8.** The same as Fig. 6, but for LEPS nested on the RMs of ECMWF “HRES” super-ensemble.

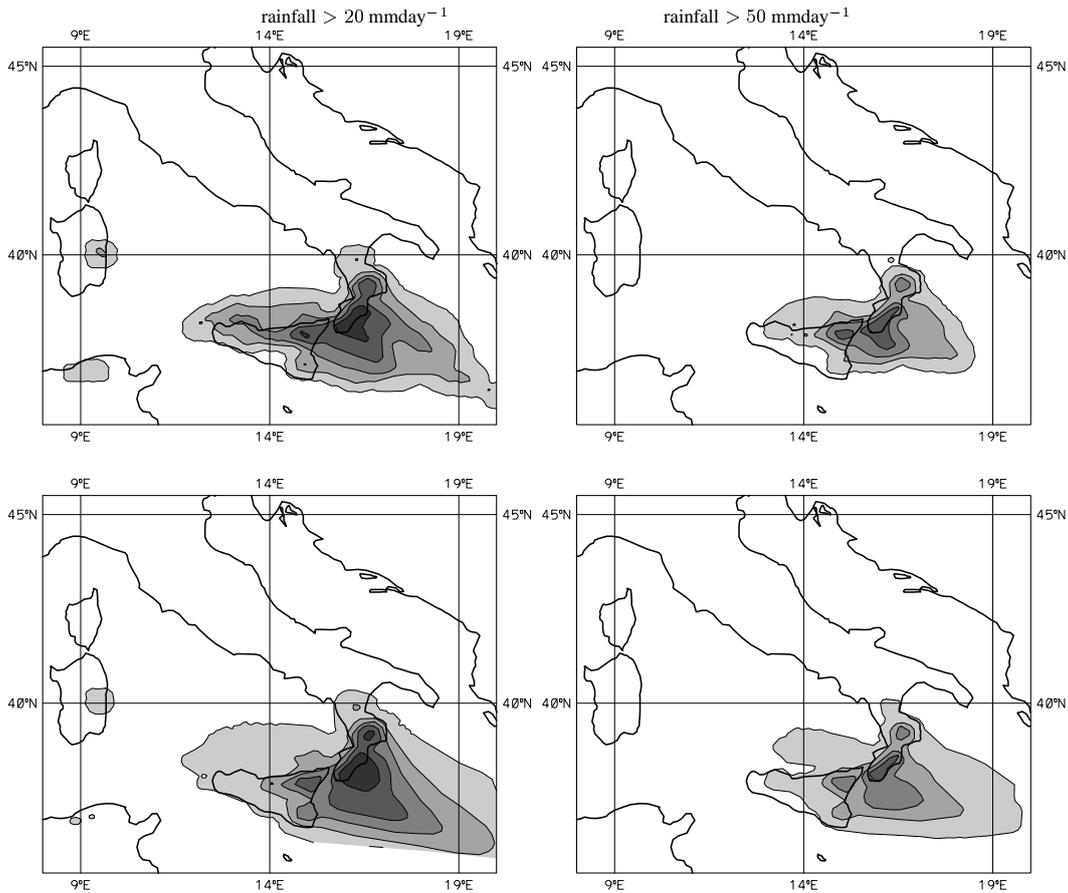
rainfall probability predictions. Figure 8 shows that the precipitation forecasts are particularly accurate in this configuration. The improvement with respect to the “OPE” LEPS is noticeable at the highest thresholds, the probability peaks being located in the vicinity of the Soverato area. It looks as if the forecast improvement gained in the global-model integrations between “OPE” and “HRES” configurations (Sects. 4.1 and 4.3) is partly transferred to the LAM runs. The probability of rainfall exceeding  $20 \text{ mm day}^{-1}$  is above 60% in the flooded region and still about 30% at the  $100 \text{ mm day}^{-1}$  threshold (top-left and middle-right panels, respectively). Therefore, the combination of accurate boundary and initial conditions provided by the “HRES” global ensembles and the spatial detail gained in the LAM integrations enables to detect precisely the areas likely to be flooded. This would have given valuable guidance for flood alert purposes, since 60 hours would have been left to take measures of protection.

## 6 “The brute-force” approach

As described in Marsigli et al. (2001) and Molteni et al. (2001), the main motivations behind the ensemble-size reduction of LEPS methodology are dictated by the comput-

ing power limitations as well as by the (supposed) possibility to capture, thanks to the clustering-selection technique, alternative weather patterns, with higher or lower probability of occurrence. This latter aspect is critical and deserves further investigation. In fact, it is not fully clear the extent to which the ensemble-size reduction may induce a loss of information about the spectrum of possible atmospheric flows and thus affect the overall performance of LEPS. Hence, the idea is to assess, for this case study, the impact of LEPS ensemble size reduction on the rainfall probabilistic prediction of this severe weather event.

Only the most recent EPS (initial conditions at 12:00 UTC 7 February) is considered and the probability maps of the 5-member LEPS (weighted according to the cluster population) are compared to those where LAMBO is nested on each EPS run and all runs are given the same weight. The two configurations are referred to as “5-recent” and “brute-force”, respectively. The 60-hour forecast range is commented for brevity and the “HRES” global ensemble provides initial and boundary conditions to the limited-area runs. The four panels of Fig. 9 report the results for the two different LEPS set-ups, the attention being focused on the probabilistic prediction of rainfall exceeding the  $20$  and  $50 \text{ mm day}^{-1}$  thresholds (left and right-column panels, respectively). It can be noticed that

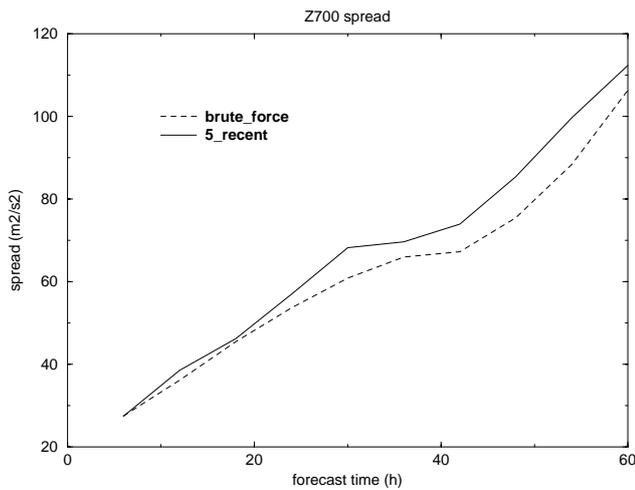


**Fig. 9.** Probability maps of 24 h cumulated rainfall exceeding 20 (left column) and 50 mm day<sup>-1</sup> (right column) for LEPS nested on the RMs of “HRES” ensemble starting at 12:00 UTC 7 September 2000 (60 h forecast). Top-row panels: “5-recent” LEPS; bottom-row panels: “brute-force” LEPS (51 members). Contour intervals: 2%, 20%, 40%, 60%, 80%.

the results obtained by the two configurations are very similar. At the lower threshold, both LEPS ensembles show a probability peak above 80% (left-column panels), while at the 50 mm day<sup>-1</sup> threshold the results are almost identical. A southward displacement problem is evident in both configuration and does not seem to be due to the ensemble-size reduction, but rather to the global-model runs forcing the limited-area integrations.

Hence, the impact of the ensemble-size reduction on the forecast accuracy does not appear to be so important, since the heavy rainfall scenario is properly captured in the “5-recent” configurations as well as in the “brute-force” approach. This result is very encouraging: for this particular case study, the clustering-selection technique highlights the most important evolution scenarios and enables the generation of reliable and accurate probability maps. The location of the regions most likely to be hit by the severe weather event do not change considerably from the 5-member to the 51-member LEPS, the probability maxima being approximately the same in both configurations. If the indications of this case study should be borne out by a larger number of test cases, then it could be concluded that, since a large amount of computer time is spent to perform 51 limited-area

integrations and the accuracy of the probabilistic forecast is not noticeably improved by the increase of LEPS size, the advantages of the clustering-selection methodology are well evident. In particular, the operational use of LEPS methodology would enable the preparation of daily high-resolution probabilistic forecasts of surface parameters (e.g. total precipitation, surface temperature, surface winds). The routine generation of a 5-member LEPS is affordable with present computing facilities, while the “brute-force” approach can be tested, at the moment, only for case study analysis and for relatively short forecast ranges. It is also interesting to compare the performance of LEPS nested on the most recent ensemble against LEPS nested on the super-ensemble (already presented in the previous section). A comparison of the top-row panels of Fig. 8 (relative to the 20 and 50 mm day<sup>-1</sup> thresholds) with those of Fig. 9 shows that the probability of precipitation is higher when the limited-area model is nested on the most recent EPS. This is particularly evident in the Soverato area, where the probability peaks are about 20% higher in the “5-recent” and in the “brute-force” configurations. On the other hand, the 5-member LEPS nested on the super-ensemble is preferable to the previous configurations because of the lesser need of computer-time resources



**Fig. 10.** Average spread of the “5-recent” LEPS (solid line) and of the “brute-force” LEPS (dashed line) in terms of geopotential height at 700 hPa ( $m^2/s^2$ ).

(in comparison with the “brute-force” approach) and of the higher reliability of the limited-area integrations (in comparison with the “5-recent” one).

As a final test, it was decided to compare the spread of the two limited-area ensembles. Figure 10 shows the results averaged over the Italian domain in terms of geopotential height at 700 hPa. It can be noticed that the spread of the “5-recent” configuration is slightly higher than in the “brute-force” one. Anyway, differences are quite limited, being more marked as forecast time increases, and the impact of reducing ensemble size does not seem to affect heavily LEPS features also in this case.

## 7 Conclusions

The performance of the ECMWF global-model and of the ARPA-SMR high-resolution limited-area ensembles is assessed for a flood case occurred in September 2000 over Southern Italy. Three different configurations of ECMWF EPS are tested in order to evaluate the impact of targeting the initial perturbations and of increasing horizontal resolution against the operational ensemble (“TEPS”, “HRES” and “OPE” configurations, respectively). The attention is focused on the probabilistic forecast of rainfall exceeding given thresholds at three different prediction ranges (108, 84 and 60 h). The performances of “TEPS” and “OPE” ensembles are almost identical, while the “HRES” configuration gives better results, with approximately a 1-day gain in predictability. This result is in line with the 12–36 h gain in predictability detected by Mullen and Buizza (2001) when comparing 87-cases quantitative predictions based on the old  $T_L159$  and the new  $T_L255$  ensemble systems. In any case, the performance of the global ensembles is not very accurate, since the amount of precipitation is underestimated and the orographic forcing of the Southern Apennines is not properly

represented. At all forecast ranges, little assistance could be given to a forecaster to issue flood alerts in the correct location, because of the low probability values as well as of the southward displacement of the rainfall peaks in most of the global-model integrations.

A new methodology, referred to as “super-ensemble”, is introduced to generate the Limited-area Ensemble Prediction System (LEPS). Three successive runs of the global-ensemble are grouped together so as to form a highly-populated (153 members) ensemble. Then, 5 representative members (RMs) are selected by a clustering-selection technique using 4 variables at 3 pressure levels to construct the clusters and to choose a RM within each cluster. Each of the 5 RMs provides initial and boundary conditions for the integrations with a limited-area model, thus generating LEPS. The performance of LEPS is assessed for each of the global-model configurations, that is when the RMs are selected within the “OPE”, “TEPS” and “HRES” ensembles. It turns out that the performance of the limited-area integrations is always more accurate than in the global runs. The possibility of a heavy precipitation event is clearly indicated and localised in the area actually affected by the flood. In the “OPE” and “HRES” LEPS, the probability maps indicate high probabilities of precipitation above  $50 \text{ mm day}^{-1}$  (40% and 60%, respectively), with probabilities peaks above 20% for rainfall exceeding  $100 \text{ mm day}^{-1}$  (“HRES” configuration). The performance of the “TEPS” LEPS is similar to the “OPE” one. In all cases, the information of a highly possible flood is available 60 h before the event; hence, enough time is left to take some protection measures and hopefully mitigate the flood consequences.

As concerns the “HRES” configuration, the sensitivity of forecast accuracy to the LEPS size is also assessed.

For the 60-hour forecast range, only one EPS is considered and the probability maps of the 5-member LEPS are compared to those obtained by nesting the limited-area model on each of the 51 EPS members (“brute-force” approach). The results indicate a limited sensitivity to the LEPS size as regards the localisation of the regions most likely affected by heavy precipitation. At both 20 and  $50 \text{ mm day}^{-1}$  thresholds, the probability rainfall maps are very similar. Therefore, for this case study, the overall information which can be extracted by both configurations is approximately the same. It is worth pointing out that this result is not in agreement with Mullen and Buizza (2001), who concluded that an ensemble size reduction can have a very strong negative effect on the prediction of rare events associated with intense precipitation. The contradiction may be due to the fact that this study was based on one case only, while Mullen and Buizza worked on a larger data-set (87 cases, with 16 characterised by very intense precipitation). The fact that the 5 LEPS integrations are given weights based on the cluster populations may also be another reason for this result. Indeed, a comparison on a larger data-set is needed to draw statistically significant conclusions on this issue. The practical advantages of the ensemble-size reduction have also to be mentioned for operational implementation of LEPS system. The “brute-force”

approach is not operationally affordable because of the large computer power required to run many high-resolution integrations. Therefore, the generation of a 5-member LEPS as in the super-ensemble methodology looks a suitable compromise between the need of a probabilistic approach with limited-area models and present computing facilities. In the near future, it is envisaged to start the generation of limited-area ensemble forecasts on an experimental-operational basis (with a set-up similar to that presented in this work). The limited-area domain will cover most of Europe and the attention will be focused on the real-time probabilistic prediction (from day 2 on) of severe weather situations: rainfall exceeding high thresholds, very warm and very cold 2-metre temperature anomalies and intense surface winds. Probabilistic products will be issued to several weather centres so as to improve the present-day probabilistic prediction of localised and severe events. The results concerning the quality of this new product will be presented in forthcoming studies.

*Acknowledgements.* The authors thank Davide Cesari and Carlo Cacciamani for valuable discussions and Paolo Patrino for technical assistance. We are also grateful to Dr. Iiritano for providing us with the data of Fig. 2. This work was partly sponsored by a grant from GNDCI-CNR (Gruppo Nazionale Difesa Catastrofi Idrogeologiche-Consiglio Nazionale delle Ricerche).

## References

- Atger, F.: The Skill of Ensemble Prediction System, *Mon. Weather Rev.*, 127, 1941–1953, 1999.
- Buizza, R., and Palmer, T. N.: The singular-vector structure of the atmospheric global circulation, *J. Atmos. Sci.*, 52, 1434–1456, 1995.
- Buizza, R., Petroliaigis, T., Palmer, T. N., Barkmeijer, J., Hamrud, M., Hollingsworth, A., Simmons, A., and Wedi, N.: Impact of model resolution and ensemble size on the performance of an ensemble prediction system, *Quart. J. Roy. Meteor. Soc.*, 124, 1935–1960, 1998.
- Buizza, R., Miller, M., and Palmer, T. N.: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, ECMWF Technical Memorandum, 279, 1999.
- Buizza, R. and Hollingsworth, A.: Storm prediction over Europe using the ECMWF Ensemble Prediction System, *Met. Appl.*, in press, 2001.
- Frogner, I-L. and Iversen, T.: Targeted ensemble prediction for northern Europe and parts of the north Atlantic Ocean, *Tellus*, 52A, in print, 2000a.
- Frogner, I-L. and Iversen, T.: High resolution limited area ensemble predictions based on low resolution targeted singular vectors, submitted to *Quart. J. Roy. Meteor. Soc.*, 2000b.
- Hoffman, R. H. and Kalnay, E.: Lagged average forecasting, an alternative to Monte Carlo forecasting, *Tellus*, 35A, 100–118, 1983.
- Houtemaker, P. L., Derome, J., Ritchie, H., and Mitchell, H. L.: A systematic simulation approach to ensemble prediction, *Mon. Weather Rev.*, 124, 1225–1242, 1996.
- Janić, J.: The step-mountain coordinate: physical package, *Mon. Weather Rev.*, 118, 1429–1443, 1990.
- Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., Tibaldi, S., Molteni, F., and Buizza, R.: A strategy for high-resolution ensemble prediction. Part II: limited-area experiments in four Alpine flood events, *Quart. J. Roy. Meteor. Soc.*, 127, 2095–2115, 2001.
- Molteni, F., Buizza, R., Palmer, T. N., and Petroliaigis, T.: The ECMWF Ensemble Prediction System: Methodology and validation, *Quart. J. Roy. Meteor. Soc.*, 122, 73–119, 1996.
- Molteni, F., Buizza, R., Marsigli, C., Montani, A., Nerozzi, F., and Paccagnella, T.: A strategy for high-resolution ensemble prediction. Part I: definition of representative members and global-model experiments, *Quart. J. Roy. Meteor. Soc.*, 127, 2069–2094, 2001.
- Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T. and Buizza, R.: Performance of the ARPA-SMR Limited-area Ensemble Prediction System: two flood cases, *Nonlin. Proc. Geophys.*, 8, 387–399, 2001.
- Mullen, S. and Buizza, R.: Quantitative Precipitation Forecasts over the United States by the ECMWF Ensemble Prediction, *Mon. Weather Rev.*, 129, 638–663, 2001.
- Palmer, T. N., Gelaro, R., Barkmeijer, J., and Buizza, R.: Singular vectors, metrics and adaptive observations, *J. Atmos. Sci.*, 55, 633–653, 1998.
- Stensrud, D. J., Brooks, H. E., Du, J., Tracton, S., and Rogers, S.: Using Ensembles for Short-Range Forecasting, *Mon. Weather Rev.*, 127, 433–446, 1999.
- Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: the generation of perturbations, *Bull. Am. Met. Soc.*, 125, 2317–2330, 1993.
- Tracton, M. S., Du, J., Toth, Z., and Juang, H.: Short-range ensemble forecasting (SREF) at NCEP/EMC, Proceedings of the 12th AMS/NWP Conference, Phoenix AZ, 269–272, 1998.
- Wilks, D. S.: Statistical atmospheric sciences, Academic Press, San Diego, CA, USA, 1995.
- Ziehmann, C.: Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models, *Tellus*, 52A, 280–299, 2000.