Short-to-Medium-Range Limited-Area Ensemble Prediction: the LEPS system

Stefano Tibaldi, Tiziana Paccagnella, Chiara Marsigli, Andrea Montani and Fabrizio Nerozzi

ARPA-SMR, Bologna, Italy

1. Introduction

The main purpose of developing the LEPS system is to be able to provide the forecaster with probabilistic guidance (e.g. estimates of the probability of overcoming given thresholds for a number of meteorological variables as a function of space and time) to identify the possible/probable occurrence of severe weather conditions in the time range "late-short-range (72h) - early-medium-range (120h)" and with a higher spatial detail, typical of Limited-Area prediction Models (LAMs). The meteorological variables one has mostly in mind in such an exercise are first and foremost quantitative precipitation, but also temperature and wind. Let us consider very briefly how the problem came about, with no pretence of completeness.

Indeed the improvement of Quantitative Precipitation Forecasting (QPF) is in itself still one of the major challenges in numerical weather prediction (NWP). Despite the constant increase of computer power resources, which has allowed the development of more and more sophisticated and resolute NWP models, accurate forecasts of extreme weather conditions, especially when related to intense and localised precipitation structures, are still difficult beyond day 2 (Mullen and Buizza, 2001a) and, in rare and selected cases, even at 24 hours. This limitation is due, among other reasons, to the inherently low degree of predictability typical of the relevant physical phenomena. The probabilistic approach has been recently more and more explored to try to come to terms with the chaotic behaviour of the atmosphere and to help forecasting phenomena with low predictability.

In addition to this, almost twenty years ago Henk Tennekes (at the time member of the ECMWF Scientific Advisory Committee) raised the question of the opportunity of producing apriori estimates of forecast skill stating that "no forecast is complete without a forecast of the forecast skill". It is not an overstatement to say that his bold assertion contributed greatly to the development of forecast skill studies, estimates and prediction techniques (e.g. Palmer and Tibaldi, 1988) and to the related development of statistical-dynamical prediction methods like ensemble forecasting, at least at ECMWF.

In fact, global ensemble prediction systems, implemented operationally for several years now by some of the major meteorological centres (Tracton and Kalnay, 1993; Molteni et al., 1996; Houtekamer et al., 1996) have become extremely important tools to tackle the problem of predictions beyond day 3-to-4 and are becoming more and more the bread-and-butter of operational forecasters all around the world. For a probabilistic approach to be useful, that is to predict within a certain detail the evolution of the probability density function (PDF) in the atmosphere's phase space, the population of the ensemble should be of the same order of magnitude as the dimensionality of the unstable subspace of the phase space itself (Montani, 1998). This implies a number of integrations probably much higher than what can be achieved by present-day operational ensemble populations (50 at most). Furthermore, probabilistic global ensemble systems are usually run at a coarser resolution with respect to (single) deterministic global predictions. Hence, ensemble skill in forecasting intense and localised events in the short- and medium-range is still nowadays limited. In order to enhance the present-day prediction capabilities of ensemble systems, several approaches have been attempted.

As for the global model, the European Centre for Medium-Range Weather Forecasts (ECMWF) has recently increased the resolution of the operational Ensemble Prediction System (EPS), which is now based on a T255L40 model (spectral model with triangular truncation at wavenumber 255 and 40 vertical levels), corresponding to a horizontal resolution of about 80 km (Molteni et al., 1996; Buizza et al., 1999a; Buizza et al., 1999b, see also the paper by Buizza in this volume). This improved the description of orography-related processes, with a 12-36 hour gain in predictability detected by Mullen and Buizza (2001a) when comparing 87 cases of quantitative precipitation predictions based on the old T159L40 (horizontal resolution of about 120 km) and the new (T255L40) ensemble systems.

During 1999, ECMWF also developed a Targeted Ensemble Prediction System (TEPS), where the perturbations applied to the analysis to obtain the different initial conditions were selected in order to maximise the 48-hour total energy perturbation growth over the European Area (about 35N-75N, 40W-30E), instead of the whole extratropical Northern and Southern Hemisphere as in the operational EPS. This TEPS system was based on a T159L40 model (the same configuration as the operational EPS at that time) and was developed within an ECMWF special project as a collaboration amongst KNMI (Royal Meteorological Institute of the Netherlands), ARPA-SMR (Regional Meteorological Service of Emilia-Romagna, Italy) and DNMI (Norwegian Meteorological Institute). The project aim was to increase the ensemble spread over the European Area in the short-range and early medium-range. First results based on a limited number of cases suggest that TEPS performs marginally better than EPS in the 72-to-96 forecast range in terms of probabilistic prediction of severe events over Europe (Hersbach et al, 2000; Frogner and Iversen, 2001).

As an alternative approach (trading ensemble numerosity for model horizontal resolution) a smaller-ensemble-size global ensemble, but at a higher resolution, was tested by Molteni et al. (2001). Although the system turns out to be very expensive from a computational point of view, it gives promising results for the probabilistic prediction of heavy precipitation.

As for the use of LAMs, NCEP (the National Center for Environmental Prediction) developed the Short-Range Ensemble Forecasting system (SREF; Tracton et al., 1998), composed of 25 integrations of the ETA model and of the Regional Spectral Model. This system provides accurate QPFs for ranges up to 72 hours (see the paper by Toth in this volume). At the DMNI, a limited-area ensemble was generated by nesting the operational limited-area model in each element of a 21-member TEPS set (Frogner and Iversen, 2002). This common approach (sometimes referred to as the “brute-force” approach, BFA) appears to provide better results than global TEPS for the prediction of heavy rainfall events, but at the cost of such an increased computational burden as to make it unaffordable on an operational basis, at least for the time being.

2. The LEPS approach

In this short work, we describe the main features and some provisional results obtained by the Limited-area Ensemble Prediction System (LEPS), described in its essential principles in Molteni et al. (2001) and Marsigli et al. (2001) and developed at ARPA-SMR. LEPS is designed to combine the advantages of a global-ensemble prediction system with the ability, typical of LAMs, to detail atmospheric phenomena on more local scales, particularly in those regions dominated by the effects of complex orography. It must be understood from the outset that the goals set out for the LEPS system could, in principle, be achieved in a very straightforward way by what was referred to in the previous section as the “brute force” approach' (BFA), that is by nesting a LAM on each member of a global ensemble. This, however, while being conceptually very simple (a desirable feature in itself), is today still extremely expensive (possibly prohibitive from an operational point of view), due to the very large amount of computer resources required, as well as of data to be transferred, if the BFA system is to be integrated in a different meteorological centre from the one that carries out the global integrations. A BFA set of integrations performed on the ECMWF

EPS output on an European domain at a horizontal resolution of order 10 km requires the transfer of approximately 5 Gigabytes per day.

The LEPS methodology is based on the idea of reducing the number of LAM integrations and of the overall resources) needed by an order of magnitude by retaining a hopefully large amount of the global ensemble information while decreasing the number of ensemble elements subjected to LAM runs. This is achieved by, first, grouping the global ensemble members into 5 clusters and then choosing a “Representative Member” (RM) within each cluster. Each RM is considered to be representative of the possible evolution scenario associated to each particular cluster and provides both initial and boundary conditions for a high-resolution LAM integration. LEPS is therefore the 5-member (only!) ensemble generated by these LAM integrations.

In order to be able to compute (hopefully) meaningful probabilities, each of the five LAM integrations carries a weight proportional to the population of the cluster from which the RM (providing initial and boundary conditions to the LAM) was selected. Probabilistic forecast products over the LAM domain (e.g. probability of 24-hour rainfall, or temperature and wind extrema, exceeding given thresholds) are then computed for forecast times up to 120 hours.

Both Marsigli et al. (2001) and Montani et al. (2001) reported about the performance of a 5-member LEPS on a number of selected case studies involving heavy rainfall in or close to mountain areas. At all forecast ranges, and concerning probabilistic quantitative prediction of heavy precipitation events, the LEPS system is shown in these works to perform better than the global model ensembles (either operational ECMWF EPS or targeted EPS), both in terms of better geographical localisation and more realistic intensity of the rainfall patterns.

Since diagnostic work had shown that the efficacy of the sampling of the phase space by the global ensemble was still critical (Mullen and Buizza, 2001b), Montani et al. (2003) introduced the super-ensemble methodology. A number of global ensemble sets (either two or three), starting at different (lagged) initial times, are grouped together so as to generate a more populated (either 102 or 153 members) lagged-super-ensemble.

The same ensemble-size reduction technique is then applied to the super-ensemble, grouping all members (properly shifted in time so as to verify accordingly) into five clusters by applying a multivariate hierarchical cluster analysis. Then the subsequent selection of a RM from each of the clusters takes place (Molteni et al., 2001; Marsigli et al., 2001). The number of clusters is kept fixed at 5 and the clustering algorithm is based on a Complete Linkage method (Wilks, 1995): the two horizontal wind components, the geopotential height and the specific humidity at 3 pressure levels in the middle-to-low troposphere (500, 700 and 850 hPa) are used as clustering variables (Montani et al., 2001). A RM is selected within each cluster by minimising the ratio between intra- and inter-cluster distances: the RM is that member closest to the other members of its own cluster and most distant from the members of the other clusters. This ratio is calculated on the basis of the same variables and metrics used for clustering. The identified RMs (global EPS integrations) provide both initial and boundary conditions for the 5 integrations of the LAM.

For practically the entire preliminary phase of the experimentation of LEPS, the LAM used has been LAMBO (Limited-Area Model BOlogna), the limited-area model running operationally at the time at ARPA-SMR. LAMBO is a hydrostatic LAM based on an early version of the NCEP ETA model (Janjic, 1990; Mesinger et al., 1988) and, for this work, was integrated at a horizontal resolution of approximately 20 km and with 32 vertical levels (domain: 1-25E, 36-50N); consequently most case-studies reported below were produced in such configuration. However, following the encouraging results of this first experimental phase, a COSMO-LEPS project was launched recently under the auspices of the COSMO consortium. COSMO (Consortium for Small-scale MOdelling, www.cosmo-model.org) is a consortium involving

Germany, Italy, Switzerland, Greece and Poland which has the purpose of continuously developing a limited-area non-hydrostatic model named Lokal Modell (LM). COSMO-LEPS aims therefore at the development and pre-operational test of a “short to medium-range” (48-120 hours) probabilistic forecasting system for severe weather prediction. It uses 10km mesh, 32 levels Lokal Modell as the LAM of the LEPS system, Tl255L40 ECMWF EPS as the operational global ensemble providing the basis integrations and runs on the ECMWF supercomputing facility, so to further minimize the need for large data transfer. COSMO-LEPS is integrated on a comparatively larger domain (see fig. 1) covering the geographical territory of all countries participating to COSMO.

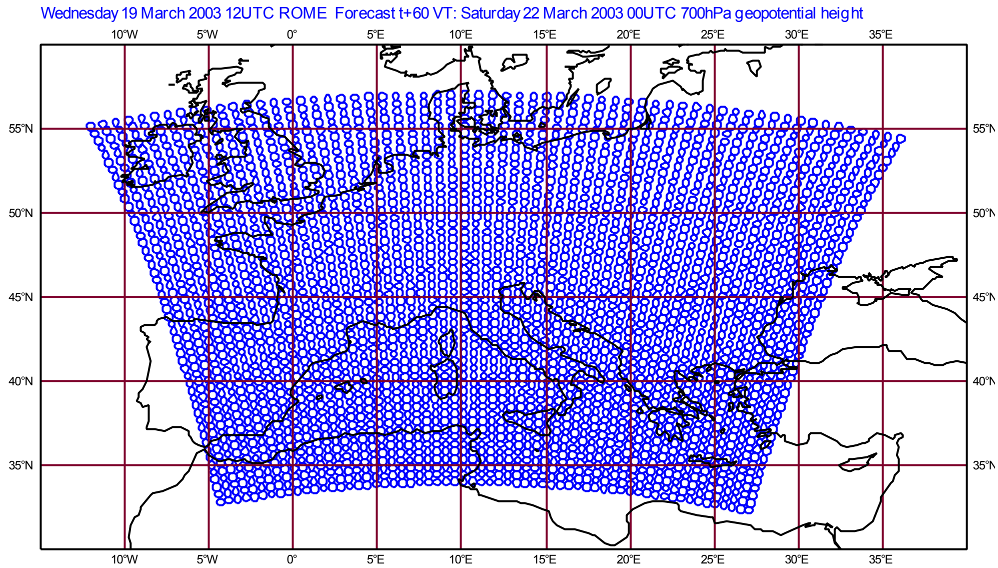

Figure 1 The COSMO-LEPS domain

3. Results from selected case studies

3.1. The Soverato case

Between 9 and 10 September 2000, rainfall maxima above 300 mm in 24 hours were recorded close to the small town of Soverato (on the eastern coast of Calabria at 38.69N 16.55E), this causing landslides, great disruption and loss of life. The four panels of fig. 2 show the observed values of precipitation, accumulated every 6 hours in four stations in the vicinity of the flooded areas. Observations clearly indicate that the most of the observed rain fell during the first hours of September 10, with exceptionally high values at the station of Chiaravalle Centrale, with 185 mm in 6 hours and about 350 mm in 24 hours.

For this case-study, three successive, 24h-lagged, EPS ensembles (starting at 12UTC of 5, 6 and 7 September 2000) are joined together into a unique super-ensemble of 153 members.

The performance of the most recent ensemble (at 60-hour forecast range) is presented in terms of rainfall probability maps (fig. 3). For the 20mm/day threshold (left panel), high probabilities of rainfall over the southern tip of the Italian Peninsula as well as over Sicily are evident, including the flooded regions. The probability is about 20% also for precipitation exceeding the 50 mm/day threshold (right panel), although the main probability peak is located probably too much to the south. At higher thresholds, no indication of a possible truly intense event is given by the probability maps.

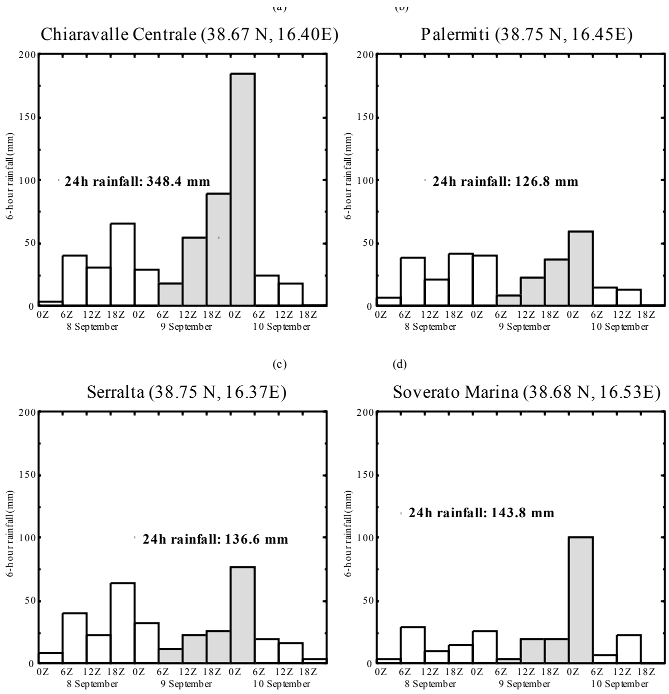

Figure 2 Soverato case. Observed values of precipitation, accumulated every 6 hours in four stations in the vicinity of the areas flooded during the "Soverato flood". The position of the station is indicated on the top of each panel, while the 24-hour cumulated precipitation relative to the 9th of September (grey bars) is reported in the panel.

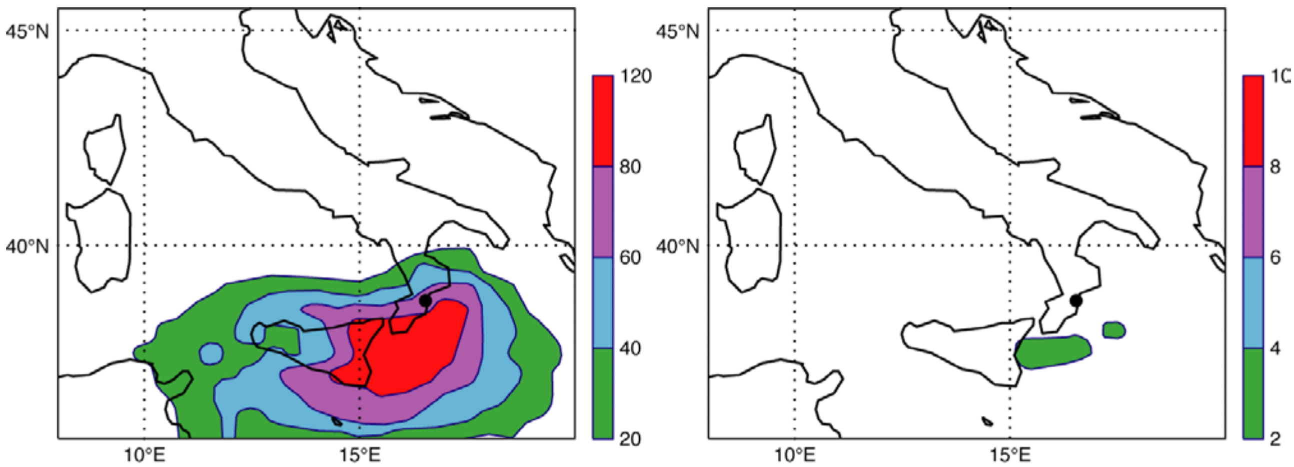

Figure 3 - Soverato case. Probability maps of precipitation exceeding 20mm/day (left panel) and 50mm/day (right panel) relative to the most recent global ensemble EPS (60-hour forecast range, model used T255). The black dot denotes the approximate location of Soverato (38.68°N, 16.55°E).

On the other hand, the performance of the limited-area integrations is more accurate than the one of the global runs. The possibility of a heavy precipitation event is clearly indicated and localised in the area actually affected by the flood. LEPS probability maps indicate about 60% of probability of precipitation above 50 mm/day (fig. 4, left panel), with peaks above 20% for rainfall exceeding 100 mm/day (fig. 4, right panel). The information about a possible (or even probable) intense rainfall event, likely to lead to flash floods, would have then been available 60 hours before the event; hence, enough time would have been available for preventive measures to be taken and hopefully mitigate the flood consequences.

The sensitivity of forecast accuracy to the limited area integrations ensemble size is also assessed. At the 60-hour forecast range, probability maps of 5-member LEPS (weighted according to the cluster population) are compared to those obtained by nesting the limited-area model in each of the 51 EPS members (the BFA). The results (fig. 5 in comparison with fig. 4) indicate no substantial degradation of the LAM forecast information in going from the BFA to the LEPS technique, at least in terms of the localisation of the regions most likely affected by heavy precipitation. At both 50 and 100 mm/day thresholds, the probability rainfall maps are very similar. The peaks in probability are slightly higher in the 5-member LEPS, since a large weight happens to be given to the more rainy scenarios. In conclusion, for this case study, the overall information which can be extracted by both configurations (BFA and LEPS) is approximately the same.

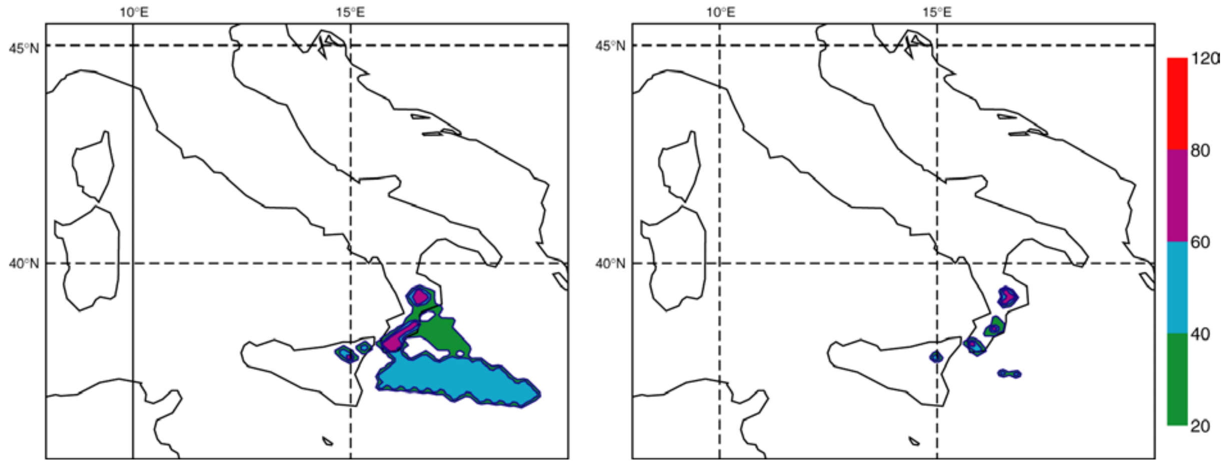

Figure 4 Soverato case. Probability maps of precipitation exceeding 50mm/day (left panel) and 100mm/day (right panel) relative to the limited-area ensemble LEPS, weighting each member according to the cluster population; the forecast range 60 hours.

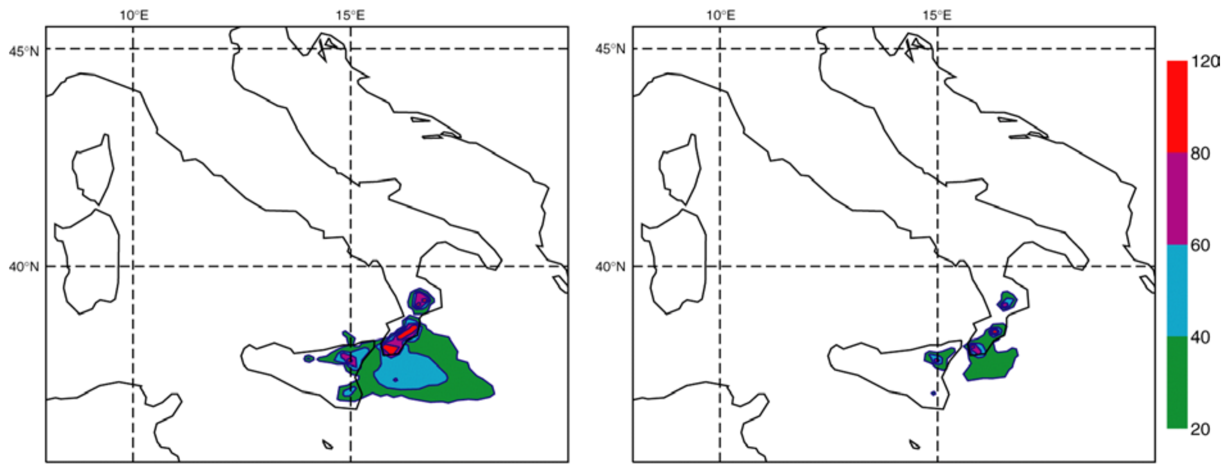

Figure 5 Soverato case. The same as in Fig. 4 but for the BFA system, obtained by nesting the limited-area model on each of the 51 EPS members.

3.2. The MAP IOP2b case

The MAP IOP2b case study was characterised by a deep trough reaching as far south as the northern African coast and moving, on 19 September 1999, from Southern France to the Alpine area. The episode had its maximum intensity, for the Western Alpine area, on September 20, when the trough was associated with a strong southerly flow and a cold frontal passage over the area. The flow remained southerly for many hours on the Alpine area and the strong precipitation area propagated towards the eastern Alps on September 20, while the Lago Maggiore area was almost continuously affected by considerably intense precipitation (see also MAP Science Director's notes on IOP 02 by Philippe Bougeault and POC Science coordinator Notes: IOP 02 by Robert Houze, available on the MAP web site: http://www.map.ethz.ch). Observed precipitation, cumulated over 24 hour from 20/09 06 GMT, is shown in Fig. 6. Precipitation exceeding 200 mm were observed in the Lago Maggiore area and in the eastern part of the Alps.

For this case-study, two TEPS, starting at 12UTC of 16 and 18 September 1999, are joined together into a unique super-ensemble (SE) of 102 members. The performance of this SE is presented in terms of rainfall probability maps (fig. 7). At the 20mm/day threshold there is a low probability of precipitation in the Lago Maggiore area; higher values are present in the eastern part of the alpine basin. No signal is evident in the 50 mm/day probability map. The correspondent LEPS probabilities are shown in Fig. 8. It is evident how high precipitating areas are well identified with high values of probability, over 70%, in the 50 mm/day threshold map.

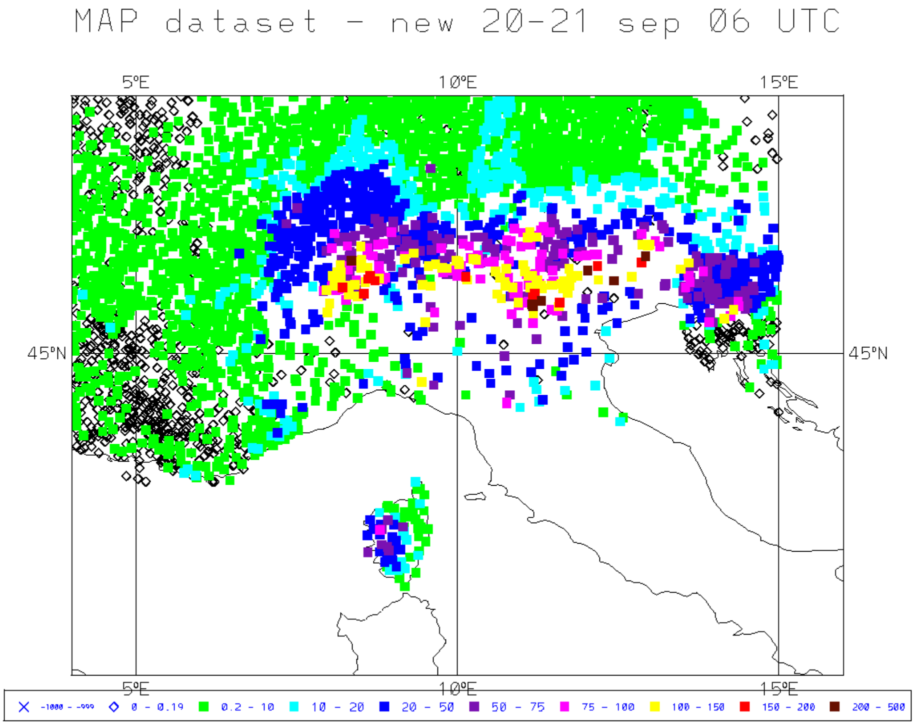

Figure 6 MAP IOP2b case. Observed precipitation, cumulated over 24 hour from 20/09/1999 06 GMT (MAP database).

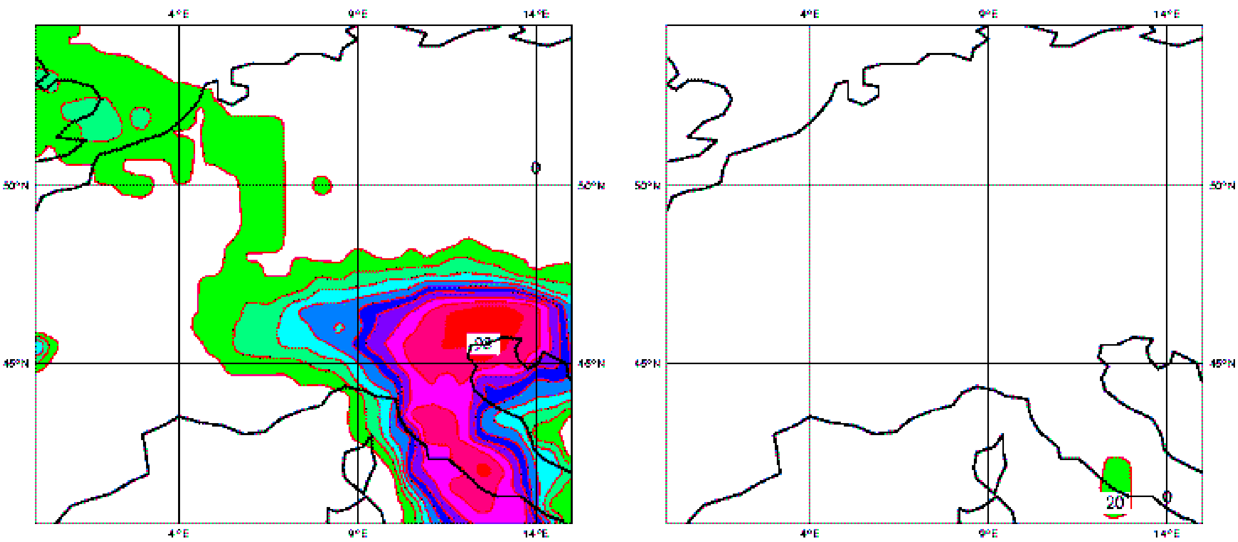

Figure 7 MAP IOP2b case. Probability maps of precipitation exceeding 20mm/day (left) and 50mm/day (right) relative to the 102 members of the TEPS Super-Ensemble. Contour interval every 10%.

The same experiment was also carried out using as LAM the non-hydrostatic model Lokal Modell (LM) at 7 Km. resolution in place of the 20 km mesh hydrostatic LAMBO. Even if the very large gap in horizontal resolution between TEPS and LM is unsuitable for a correct nesting cascade (and generates some evident negative boundary effects), the good results of LEPS employing LAMBO are confirmed (Fig. 9) and at high precipitation threshold high probability values are obtained over the two heavy precipitation areas.

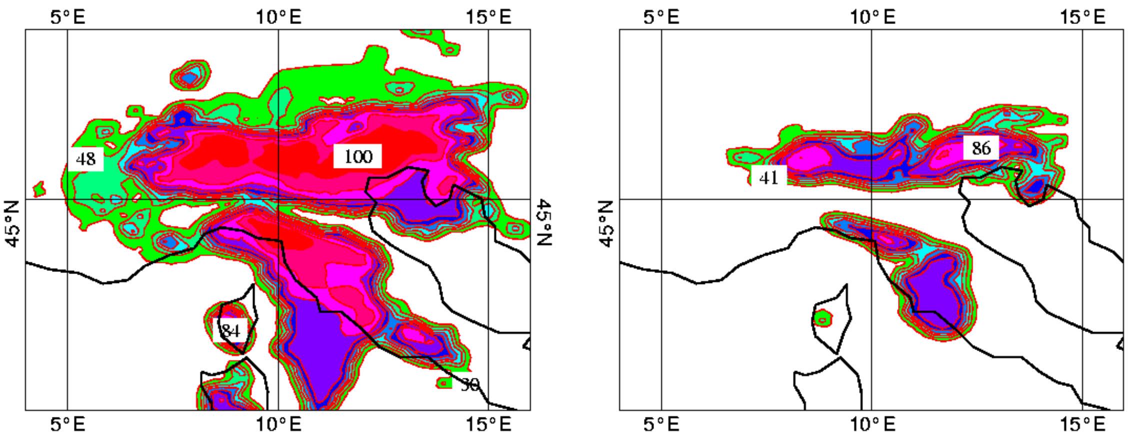

Figure 8 MAP IOP2b case. Probability maps of precipitation exceeding 20mm/day (left) and 50mm/day (right) relative to the 5-member LEPS using the hydrostatic model LAMBO at 20 Km horizontal resolution. Contour interval every 10%.

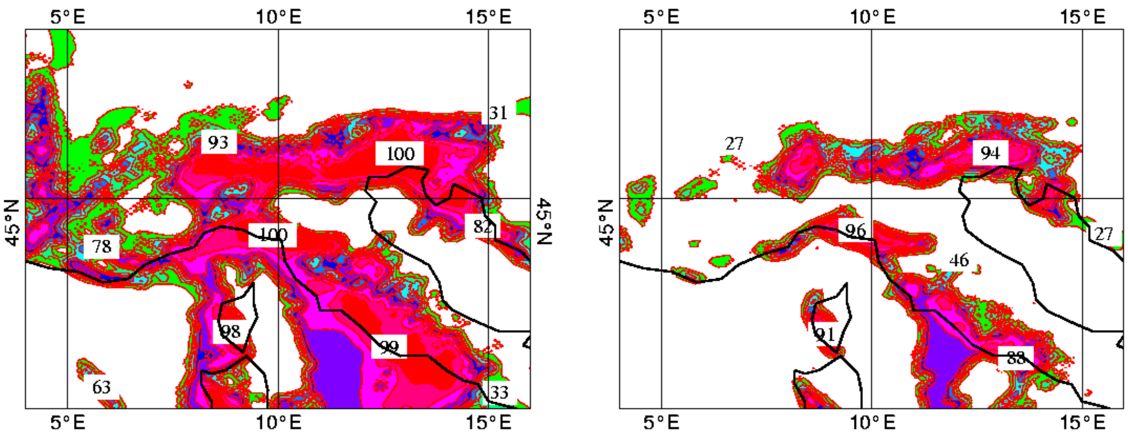

Figure 9 MAP IOP2b case. Probability maps of precipitation exceeding 20mm/day (left) and 50mm/day (right) relative to the 5-member LEPS using the non-hydrostatic model Lokal Modell at 7 Km horizontal resolution. Contour interval every 10%.

For this case, the impact of the ensemble reduction technique has been also tested. A brute force approach was performed by nesting LAMBO in each of the 51 members of the most recent TEPS of the SE (fig.10). As for the previous BFA-LEPS comparison, the LEPS system does not show any appreciable deterioration in forecast quality compared to the BFA system. The BFA forecast shows a slight positive increase in probabilities values in the correct areas of the alpine region but without essentially changing the already correct LEPS forecast.

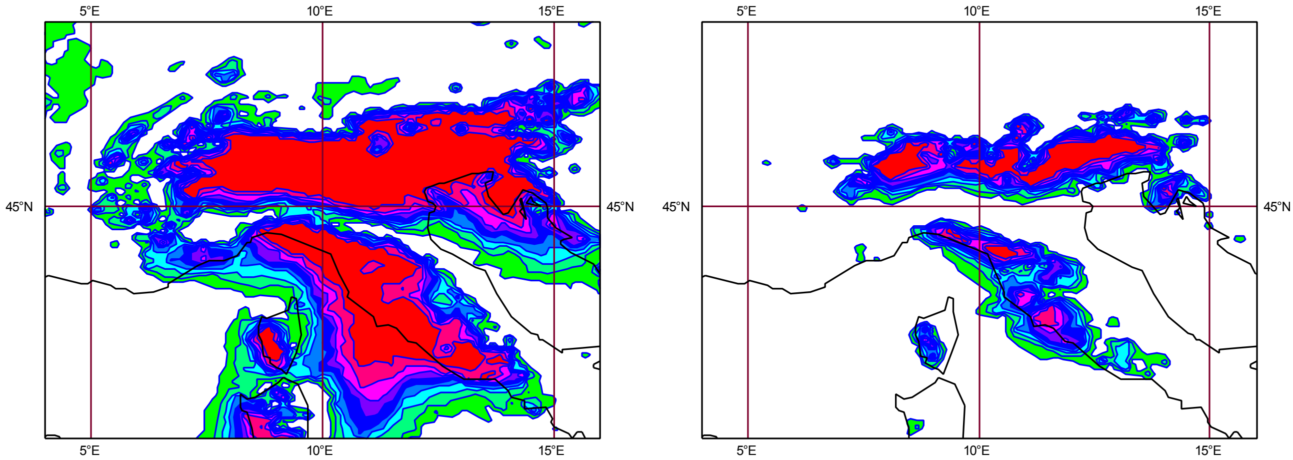

Figure 10 MAP IOP2b case. Probability maps of precipitation exceeding 20mm/day (left) and 50mm/day (right) relative to the brute force approach experiment performed by nesting LAMBO in each of the 51 members of the most recent TEPS of the SE. Contour interval every 10%.

4. Statistical evaluation

In this section, a preliminary statistical evaluation of the LEPS performance is attempted, so as to assess average abilities and shortcomings of this prediction system. LEPS forecasts are verified on the basis of the usual probabilistic scores (Brier Skill Score, ROC area, Cost-loss Curve) over a set of 15 cases selected during the field campaign of the Mesoscale Alpine Programme (MAP; Binder and Schar, 1996, Bougeault et al., 2001), between 7 September and 16 November 1999. All forecasts are compared against observations by a (brutal) bilinear interpolation of the model fields from surrounding grid points to station locations.

In this study, TEPS (T_L159L40 resolution) is used as the global ensemble providing the RMs for the limited-area integrations. During the TEPS experimentation, the ensemble was run twice a week (on Fridays and Sundays at 12UTC) for 120 hours, so as to generate a probabilistic forecast system for the early-to-medium range (the two ensembles are referred here to as "5day-ensemble" and "3day-ensemble", respectively). The system is then verified on Wednesdays. In our experiments, the 5day-ensemble and the 3day-ensemble are grouped together, so as to produce a super-ensemble of 102 members. The clustering is performed at 12UTC of Wednesdays, this corresponding to a 72-hour forecast for the 3day-ensemble.

The experimentation is aimed at investigating the extent to which LEPS precipitation forecasts are more skilful than those produced by ECMWF global ensembles. The appropriateness of the weighting procedure is also investigated, bearing in mind the assumptions underlying this approach:

- cluster population is proportional to the probability of occurrence of the scenario forecast by the cluster representative member and
- this proportionality carries over when LAM integrations nested in the global representative members are considered.

In particular, the performance of (ECMWF or LEPS) ensembles where each of the five members is weighted according to the cluster population is compared with that of ensembles where each of the five members is given the same weight. These two configurations are referred to as "weighted" and "no-weighted" respectively.

4.1. Comparison of global and limited-area ensembles

The evaluation of LEPS performance with respect to the global ensembles is carried out on the basis of the precipitation forecast only. Only scores relative to two precipitation thresholds are reported: 10mm/day (table 1) and 50mm/day (table 2).

Different ensemble configurations are compared:

- T102: global-model super-ensemble (102 members)
- T5w: weighted global-model ensemble made up of the 5 RMs
- T5nw: no-weighted global-model ensemble made up of the 5 RMs
- L5w: weighted limited-area ensemble
- L5nw: no-weighted limited-area ensemble

Index	**T102**	**T5w**	**T5nw**	**L5w**	**L5nw**
BSS	213	103	143	21	107
ROC area	834	777	762	753	750

Table 1 - Values (multiplied by 1000) of BSS and ROC area (second and third row, respectively) for the 10mm/day threshold for the different ensembles. T102: global-model super-ensemble (102 members), T5w: weighted global-model ensemble made up of the 5 RMs, T5nw: no-weighted global-model ensemble made up of the 5 RMs, L5w: weighted limited-area ensemble, L5nw: no-weighted limited-area ensemble

Index	**T102**	**T5w**	**T5nw**	**L5w**	**L5nw**
BSS	-18	-24	-31	105	129
ROC area	742	524	524	751	749

Table 2 - The same as table 1 but for the 50mm/day threshold.

At the lower precipitation threshold (table 1), the global ensemble T102 obtains the best results, both in terms of BSS and ROC area. Furthermore, also the ensemble made up of the 5 TEPS RMs (T5w and T5nw) performs better than LEPS (L5w and L5nw), both in the weighted and no-weighted configurations. On the other hand, LEPS scores are better than those of TEPS super-ensemble (both 102-member and 5-member) when high precipitation thresholds are considered (table 2), confirming the usefulness of this approach for extreme events.

The reduction of ensemble size seems to be more than compensated by the increase of resolution from TEPS to LEPS (approximately, from 120 to 20 km in the horizontal) and by the better representation of mesoscale and orographic-related processes in the limited-area model runs. LEPS forecasts appear to be particularly accurate when high precipitation thresholds and intense precipitation cases are considered (not shown), without a noticeable increase in false alarms.

4.2. Evaluation of the impact of the weighting procedure

It is here examined the extent to which the RM of a highly-populated cluster is more likely to be “closer” to the verifying analysis than the RM of a cluster with fewer members. This feature is also assessed for those LAM integrations nested in the RMs of such clusters.

When 5-member TEPS integrations are considered (fig. 11), the probability of occurrence of a certain synoptic scenario, expressed in terms of 700 hPa geopotential height (black-dashed line), is well related to the population of the cluster represented by that ensemble member. The correspondence is also good when

TEPS precipitation forecasts are verified against either 24-hour precipitation forecast or observed precipitation. Hence, a TEPS member selected in a highly-populated cluster is more likely to be closer to the verifying analysis rather than a member taken from a cluster with fewer elements.

This property is still evident, although to a lesser extent, when the geopotential height LEPS predictions are considered (fig. 12, black-dashed line). On the other hand, a confirmation of this result in terms of LEPS precipitation is less clear, since the cluster population seems to be only weakly related to the probability of occurrence. Among other reasons, this could be due to the higher model error in forecasting precipitation rather geopotential height, this hiding the ability of the member with the highest weight to represent correctly the large-scale structure. As an alternative hypothesis, this could be attributed to the fact that the verification procedure and indices do not enable to fully understand which member gives the best forecast in terms of precipitation.

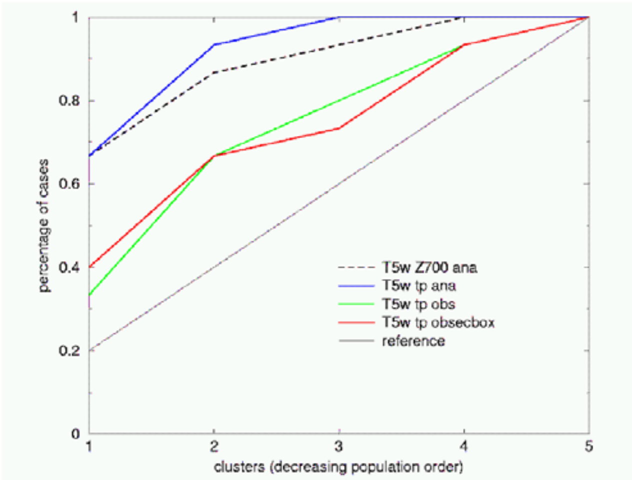

Figure 11 Percentage of times the best member of the TEPS 5 RM belongs to the most populated cluster, to the two most populated clusters, to the i most populated clusters as a function of i. Black dashed line: geopotential height at 700 hPa against ECMWF analysis, blue line: forecast precipitation against ECMWF operational +24h forecast ("analysis"), green line: forecast precipitation interpolated on station points against MAP observations, red line: forecast precipitation against MAP observation averaged over ECMWF TEPS grid boxes.

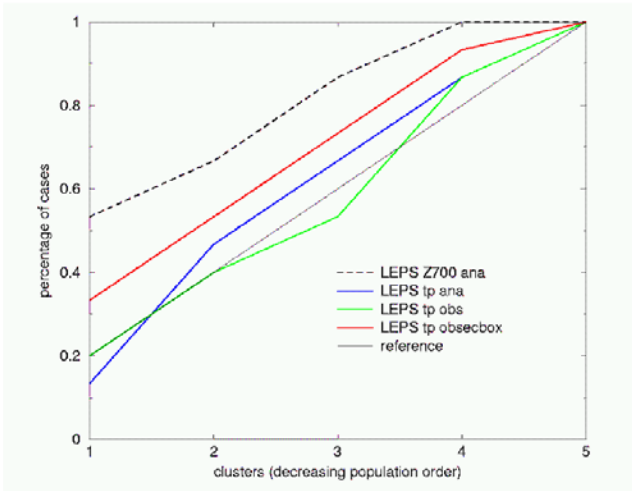

Figure 12 The same as in Fig. 11 but for LEPS. Black dashed line: forecast geopotential height at 700 hPa against ECMWF analysis, blue line: forecast precipitation against MAP analysed precipitation (Frei and Haller, 2001), green line: forecast precipitation interpolated on station points against MAP observations, red line: forecast precipitation averaged over ECMWF TEPS grid boxes against observation averaged over the same boxes.

5. Future plans: the COSMO-LEPS experiment

The results obtained so far have encouraged the operational implementation of LEPS, which appears to be a useful probabilistic tool to support forecasters in the prediction of severe precipitation events. This implementation will make use of a more sophisticated non-hydrostatic limited-area model (Lokal Modell), this enabling a better representation of orographic-related processes. This operational implementation will allow to test the system on a regular, instead of a case-to-case, basis and to answer the open questions. Following the encouraging results of the experimental phase, the generation of an ``experimental-operational" limited-area ensemble prediction system, the **COSMO-LEPS project**, has recently started on the ECMWF computer system under the auspices of COSMO consortium. COSMO-LEPS aims therefore at the development and pre-operational test of a "short to medium--range" (48-120 hours) probabilistic forecasting system using a LAM over a comparatively large domain (see Fig.1), covering all countries involved in COSMO. The COSMO-LEPS suite (Fig. 13) is running every day applying the procedure on a 153-element super ensemble from three 12-hour lagged operational EPS ensembles (today midday, yesterday midnight, yesterday midday) and integrating five times a 10-km mesh version of LM once a day for 120 hours. Products are usually available by 4UTC the next day (tomorrow morning), well in time to be evaluated by operational forecasters.

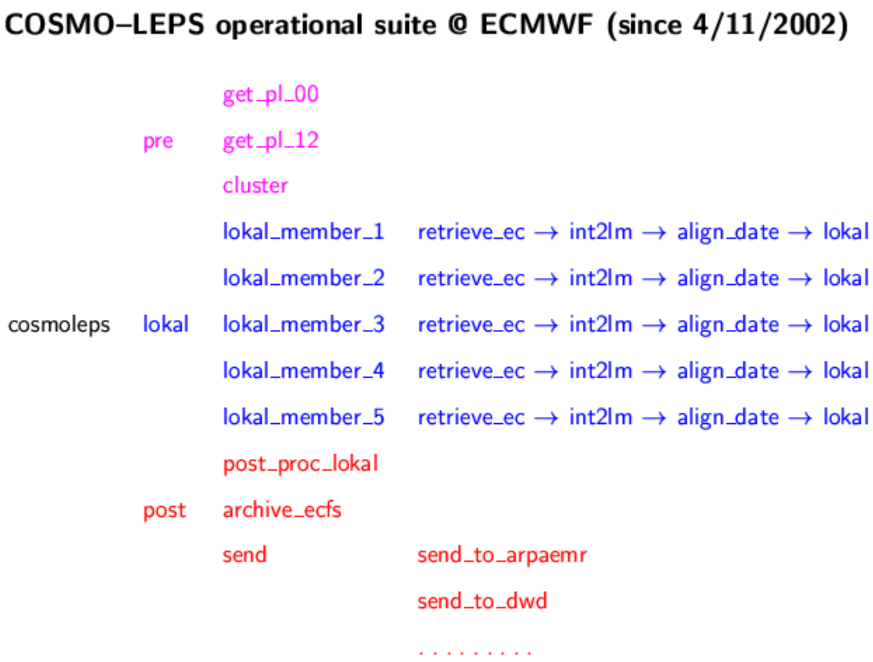

Figure 13 Schematic of the COSMO-LEPS suite.

COSMO-LEPS pre-operational dissemination started during November 2002 and, at the time of writing (April 2003), the system is still under test to assess its usefulness in met-ops rooms, particularly in terms of the assistance given to forecasters in cases of extreme events. An objective probabilistic verification of the system is also being carried out. The COSMO-LEPS products are transfered to Cosmo Partners and other interested Member States. The deterministic products for each of the 5 LM runs are: Precipitation, Mean Sea level pressure, 700 hPa Geopotential, 850 hPa Temperature. The probabilistic products are: probability of 24h rainfall exceeding 20,50,100,150 mm, probability of 24h Tmax exceeding 20,30,35,40 C, probability of 24h Tmin below -10,-5,0,+5 C, probability of 24h Vmax exceeding 10,15,20,25 m/s, probability of 24h snowfall exceeding 1,5,10,20 mm of equivalent water. In the probability maps, each LEPS member is weighted according to the population of the cluster from which the representative member has been chosen.

6. Summary and concluding remarks

Since the day, almost 20 years ago, in which Henk Tennekes stated, during a meeting of the Scientific Advisory Committee of the European Centre for Medium-Range Weather Forecasts, that "no forecast is complete without a forecast of the forecast skill", the demand for numerical forecasting tools which should have, at the same time, the capability of providing quantitative estimates of forecast reliability and of casting quantitative forecasts of various meteorological parameters in terms of probabilities, has been ever increasing. Consistently with this, among ECMWF's operational users, output from the now twice-daily EPS system is eroding more and more the historical primacy detained by the single, daily, purely deterministic, highest possible resolution, "main", 12:00GMT, 10-day forecast run, which has for a long time represented by far the main source of reliable, good-quality, numerical short- and medium-range forecast information.

EPS products, however, suffer from a number of drawbacks, most of them related to the comparatively low resolution of the global model used to produce them. The size of the ensemble, i.e. the number of the forecasts which compose it, is a very important factor in producing a sufficiently adequate "exploration" of the phase-space of all possible future atmospheric states which are compatible with our imperfect knowledge of the atmospheric initial conditions. This imposes a compromise between model resolution and the total number of elements needed in the ensemble, i.e. the model resolution has to be kept low enough to make the total computational requirements affordable. This has as a consequence that in the EPS model a number of local (mainly orography-related) atmospheric features/phenomena/processes are still misrepresented or underestimated. Orographically-related local and/or intense precipitation is one of the best possible such examples, but not the only one. This carries with it a reduced capability of the EPS system to provide useful guidance in case of intense meteorological events (often catastrophical in terms of consequences on life and property), notably those characterized by large rainfall amounts leading to floods.

In the attempt, therefore, to push the point of compromise between resolution and number-of-elements further toward a better representation of local effects (such as the one typical of high-resolution Limited-Area-Models), without having to sacrifice too much to completeness of exploration of the atmospheric phase-space, the LEPS (Limited-are Ensemble Prediction System) system was developed at ARPA-SMR. The basic goal is to get most of the advantages of a brute-force approach, which can hardly be afforded on an operational basis (run a high-resolution LAM integration from each and every element of an ECMWF EPS, for example), without having to pay the full computational.

The LEPS idea is in fact very simple: once all EPS elements have been grouped in clusters (in our case five, a fixed number), from all the elements of each cluster a most "Representative Member" (RM) is defined, which carries with it a weight proportional to the population of the cluster it represents and from which a high-resolution LAM is integrated. These five RM LAM integrations can be combined, using the cluster-population weights, to produce all statistically-based predictions typical of a complete EPS system, like for example the probability for a meteorological variable to overcome given thresholds as function of space and time.

A number of case-study tests on intense precipitation episodes (some of which produced catastrophic floods) conducted combining ECMWF EPS output with the hydrostatic, 20km mesh, LAMBO limited-area model operational at ARPA-SMR, showed that LEPS seems to succeed in combining the ability typical of LAMs to produce more realistic intense precipitation (mostly in connection with orographic forcing) without losing large amounts of meteorologically important information due to the drastic reduction in the number of cluster elements which are actually integrated at high resolution. Further, the (limited) experimentation carried out does not suggest that the number and size of false alarms produced by the LEPS system should be of particular concern, although a much more complete statistical evaluation on a much larger set of cases is needed before this tentative conclusion is put on firmer ground.

In addition to this, it was found that aggregating together two or three immediately subsequent EPS outputs into a super-ensemble of 102 or 153 elements (a kind of time-lagged-super-ensemble), and then defining the five clusters within such super-ensemble, was yielding considerably better results. The system was also tested, in a few cases, using the COSMO-Consortium non-hydrostatic, 7-km mesh size, Lokal Modell (LM) as the LAM, in place of the older hydrostatic LAMBO, providing equally good, if better, results.

Such overall preliminary results were considered promising enough to launch a COSMO-LEPS pre-operational, regular experimentation exercise, carried out nesting, in LEPS mode, LM within the ECMWF EPS model output. The experimentation is carried out daily, constructing a 153-element super ensemble from three 12-hour lagged operational EPS ensembles (today midday, yesterday midnight, yesterday midday) and integrating five times a 10-km mesh version of LM once a day for 120 hours. The limited-area model integrations using Lokal Modell are carried out on a domain covering the geographical territory of all countries participating to COSMO. The experiment is carried out on the ECMWF computer system, to minimize file transfer requirements. This extended pre-operational test started during the first week of November 2002 on computer resources coming from the ECMWF COSMO partners (that is Germany, Greece, Italy and Switzerland).

At the end of this experimental period, it is hoped that enough cases will be available to provide, at the same time, a good synoptic variety in terms of meteorological situations conducive to intense events and a good basis for statistical reliability in objective verification of model/system performance. These will allow to assess the usefulness of the COSMO-LEPS system in providing forecasts of intense meteorological events in the short-to-medium range. The problem of a possible truly operational implementation within the COSMO community will then be addressed on the basis of quantitative evidence based on firmer ground.

References

Binder P. and Schar C., 1996. MAP design proposal. MeteoSwiss, 75 pp. Available at the MAP Office, MeteoSwiss, CH-8044, Zurich, Switzerland.

Bougeault P., Binder P., Buzzi A., Dirks R., Houze R., Kuettner J., Smith R. B., Steinacker R. and Volkert H., 2001. The MAP Special Observing Period. *Bull. Am. Met. Soc.,* **82**, 433-462.

Buizza R., Miller M., Palmer T. N., 1999a. Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.,* **125**, 2887-2908.

Buizza R., Barkmeijer J., Palmer T. N. and Richardson D. S., 1999b. Current status and future developments of the ECMWF Ensemble Prediction System. *Meteorol. Appl.,* **6**, 1-14.

Frei C. and Haller E., 2001. Mesoscale precipitation analysis from MAP SOP rain-gauge data. MAP Newsletter Nr. 15.

Frogner I. and Iversen T., 2001. Targeted ensemble prediction for northern Europe and parts of the North-Atlantic Ocean. *Tellus*, **53A**, 35-55.

Frogner I. and Iversen T., 2002. High-resolution limited-area ensemble predictions based on low-resolution targeted singular vectors. *Quart. J. Roy. Meteor. Soc.,* **128** ,1321-1341.

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.

Houtekamer P. L., Derome J., Ritchie H. and Mitchell H. L., 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

Janjic Z. I., 1990. The step-mountain co-ordinate: physical package. *Mon. Wea. Rev.*, **118**, 1429-1443.

Marsigli C., Montani A., Nerozzi F., Paccagnella T., Tibaldi S., Molteni F. and Buizza R., 2001. A strategy for high-resolution ensemble prediction. II: Limited-area experiments in four Alpine flood events. *Quart. J. Roy. Meteor. Soc.*, **127**, pp. 2095-2115.

Mesinger F., Janjic Z. I., Nickovic S., Gavrilov D. and Deaven D. G., 1988. The step-mountain co-ordinate: 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.,* **127**, 2069-2094.

Molteni F., Buizza R., Marsigli C., Montani A., Nerozzi F. and Paccagnella T., 2001. A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments. *Quart. J. Roy. Meteor. Soc.*, **127**, pp. 2069-2094.

Montani A., 1998. Targeting of observations to improve forecasts of cyclogenesis. PhD thesis. University of Reading, Reading, UK.

Montani A., Marsigli C., Nerozzi F., Paccagnella T. and Buizza R., 2001. Performance of ARPA-SMR Limited-area Ensemble Prediction System: two flood cases. *Nonlinear Processes in Geophysics,* **127**, pp. 2095-2115.

Montani A., Marsigli C., Nerozzi F., Paccagnella T., Tibaldi S. and Buizza R., 2003. The Soverato flood in Southern Italy: performance of global and limited-area ensemble forecasts. In print on *Nonlinear Processes in Geophysics*.

Mullen S. L. and Buizza R., 2001a. Quantitative precipitation forecast over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **129**, 638-663.

Mullen S.L. and Buizza R., 2001b. The Impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF Ensemble Prediction System. Mon. Wea. Rev., in press.

Palmer T. N. and Tibaldi S., 1988. On the prediction of forecast skill. *Mon. Wea. Rev.,* **116**, 2453-2480.

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. and Juang H., 1998. Short-range ensemble forecasting (SREF) at NCEP/EMC. Proceedings of the 12^{th} AMS/NWP Conference, Phoenix AZ, 269-272.

Wilks D. S., 1995. Statistical methods in atmospheric sciences. Academic Press, New York, 467 pp.