Preliminary results of LAMEPS experiments at the Hungarian Meteorological Service

Edit Hágel and Gabriella Szépszó

hagel.e@met.hu, *szepszo.g@met.hu*

1. Introduction

The ensemble technique is based on the fact that small errors in the initial condition of any numerical weather prediction model (or errors in the model itself) can cause big errors in the forecast. When making an ensemble forecast the model is integrated not only once (starting from the original initial condition), but forecasts are also made using little bit different (perturbed) initial conditions. This ensemble of initial conditions consists of equally likely analyses of the atmospheric initial state and, in an ideal case, encompasses the unknown "true" state of the atmosphere. This technique is capable to predict rare or extreme events and has the advantage of predicting also the probability of future weather events or conditions. Despite its success, at the moment the ensemble method is mainly used for medium range forecasting and on global scales, though nowadays the emphasis is more and more moving towards the short ranges and smaller scales. However methods used in the medium range cannot be directly applied to short-range forecasting. Research has already been done in this field and there are some operational short-range ensemble systems (e.g. at NCEP, or the COSMO-LEPS). We also wish to develop a short-range ensemble system with as main goal the better understanding and prediction of local extreme events like heavy precipitation, wind storms, big temperature-anomalies and also to have a high resolution probabilistic forecast for 2 meter temperature, 10 meter wind and precipitation in the 12-48 h time-range.

For making an ensemble forecast lots of methods can be used (e.g. multi-model, multi-analysis, perturbation of observations, singular-vector method, breeding etc.). It is not known yet (especially at mesoscale) which method would provide the best forecasts. Therefore the following methods will be tried:

- ALADIN EPS coupled with global (ARPEGE based) ensemble members. This would include the investigation of the impact of the target domain and target time-window of the global singular-vector computation.
- ALADIN EPS coupled with representative members of clusters formed from ARPEGE based ensemble forecasts (the so called "super ensemble")
- ALADIN EPS based on ALADIN native singular-vector perturbations

Hereafter the first activities and results of this LAMEPS project will be briefly described.

2. Verification and visualization

The first task was to implement and develop the special verification and visualization tools needed for an ensemble system. The tools are mainly based on the softwares MAGICS and METVIEW (both are ECMWF visualization softwares).

Our verification package includes the most important scores and methods:

- ROC diagram
- Talagrand diagram
- Brier score, Brier skill score and reliability diagram

In the case of wind speed, temperature and geopotential the models are verified against SYNOP data in grid points. In the case of precipitation it was decided to do it in a different way because of the following reason. The forecast model predicts precipitation fluxes over areas of the order of about 10 km×10 km while SYNOP stations report values representing less than a m^2. Because of this inconsistency it was decided to use a special verification method in the case of precipitation. Forecasts are verified not in grid points but instead the average values computed over bigger areas (such as watersheds) are verified.

Our visualization package includes the usual plots, such as:

- Spaghetti diagrams
- Plume diagrams
- Ensemble mean

- Members (together or one by one)
- Probabilities

3. LAMEPS runs – Experiments

It was decided to start our experiments with the downscaling of the global (ARPEGE based) ensemble. This work can be divided into two parts:

- Downscaling the ARPEGE/PEACE[1] members
- Investigation of the impact of the target domain and target time-window and downscaling the ARPEGE ensemble members (the integration of the global ensemble is performed locally)

3.1 Downscaling of the PEACE ensemble members

We started with running ALADIN EPS coupled with PEACE ensemble members. The PEACE system is now run at Météo-France operationally once a day (at 18 UTC). It has 11 members (10 perturbed and a control one). It is based on the global spectral model ARPEGE. The initial perturbations of this global ensemble system are based on targeted singular vectors, the target domain covering Western Europe and the North Atlantic region. The target time-window is 12 h. We performed the ALADIN EPS integrations coupled with PEACE members for a 4 day period in October 2003 (the time interval was short because of the heavy computational cost). Both the ALADIN EPS and the PEACE members have been verified over the LACE domain (resolution 12 km, domain covering Central Europe) and the following results were obtained.

3.1.1 Talagrand diagram

The Talagrand diagram is a very useful measure of the spread. If the spread in the ensemble is big enough the histograms should be flat. A U shape indicates lack of spread (the verifying analysis lies outside the ensemble lots of times), a L shape indicates overestimation, a J shape means underestimation. For 10 meter wind speed it can be seen that the histograms have a U shape both in the case of PEACE (fig. 1) and ALADIN EPS (fig. 2). For geopotential (fig. 3) the situation is better especially if we go ahead in time. At +36 h the histogram is nearly flat.

It can be seen that the diagrams for the two models are very similar, which means that in this situation no extra information came from the integration of the limited area model ALADIN.

3.1.2 ROC diagram

In this method the bigger the area under the curve, the better the forecast is. The diagonal line represents the climate. If our curve lies below this line (so the area under the curve is less than 0.5) then our forecast gives less information than the use of the climate. ROC diagrams were made for many different parameters. For example the following events were examined: 10 m wind speed exceeds 5 m/s, and 2 m/s. In the first case it was found that the area under the curve at analysis time is smaller than at later stages of the forecast (fig. 4). The reason of this might be that the forecast starts at 00 UTC, when the wind is usually not so strong, therefore the number of cases is quite small. As we go ahead in time we get better results because of the growing perturbations. If we look to the event that 10 m wind speed exceeds 2 m/s (fig. 5), this problem can not bee seen, which can be explained by the fact that there are more cases for this lower wind speed.

3.1.3 Reliability diagram

In this case the forecast probability (x axis) and the observed probability (y axis) is plotted. In an ideal situation the points lie on the diagonal which means that the event is forecasted as many

1 Prévision d'Ensemble A Courte Echéance

times as it was observed. If the points lie above the diagonal it means underestimation, if they are under the diagonal, then it is overestimation. Because of the short time-interval (only 4 days) the number of cases is quite small, that is why the curve has a zigzag shape in the early stages of the forecast (fig. 6). The curves get smoother as we go ahead in time.

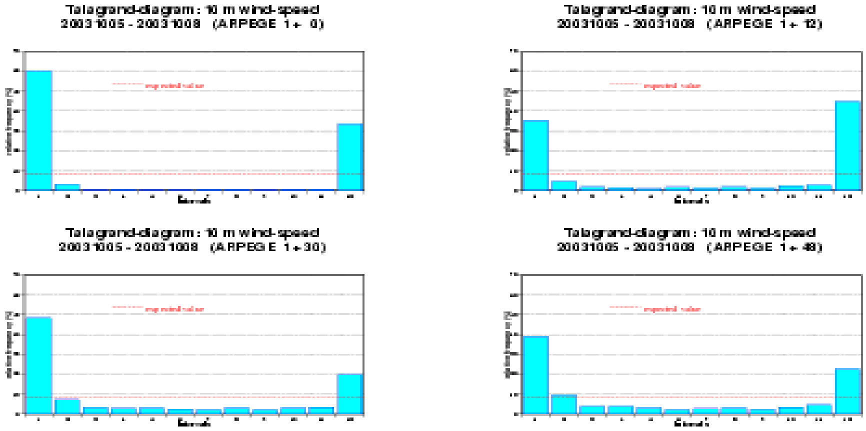

Figure 1. Talagrand diagram for the model ARPEGE, for 10 m wind speed (time steps: 00, 12, 30, 48)

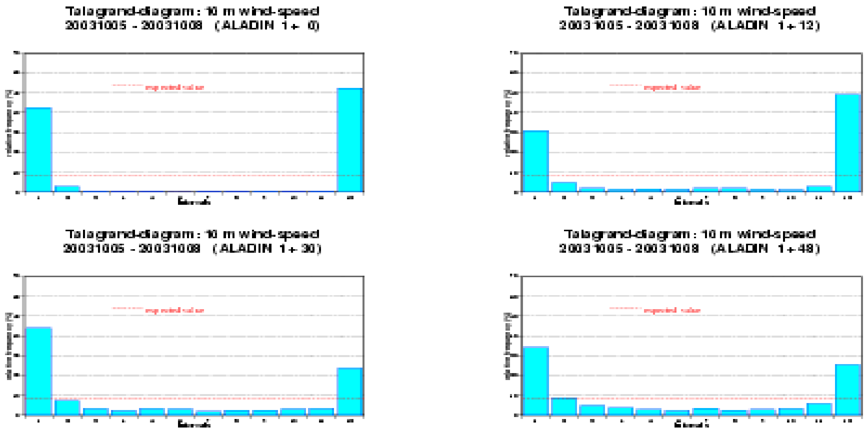

Figure 2**.** Talagrand diagram for the model ALADIN, for 10 m wind speed (time steps: 00, 12, 30, 48)

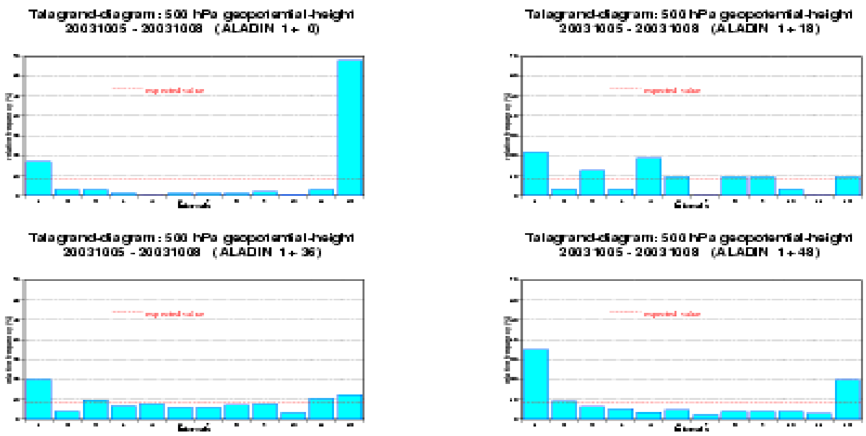

Figure 3. Talagrand diagram for the model ALADIN, for 500 hPa geopotential height (time steps: 00, 18, 36, 48)

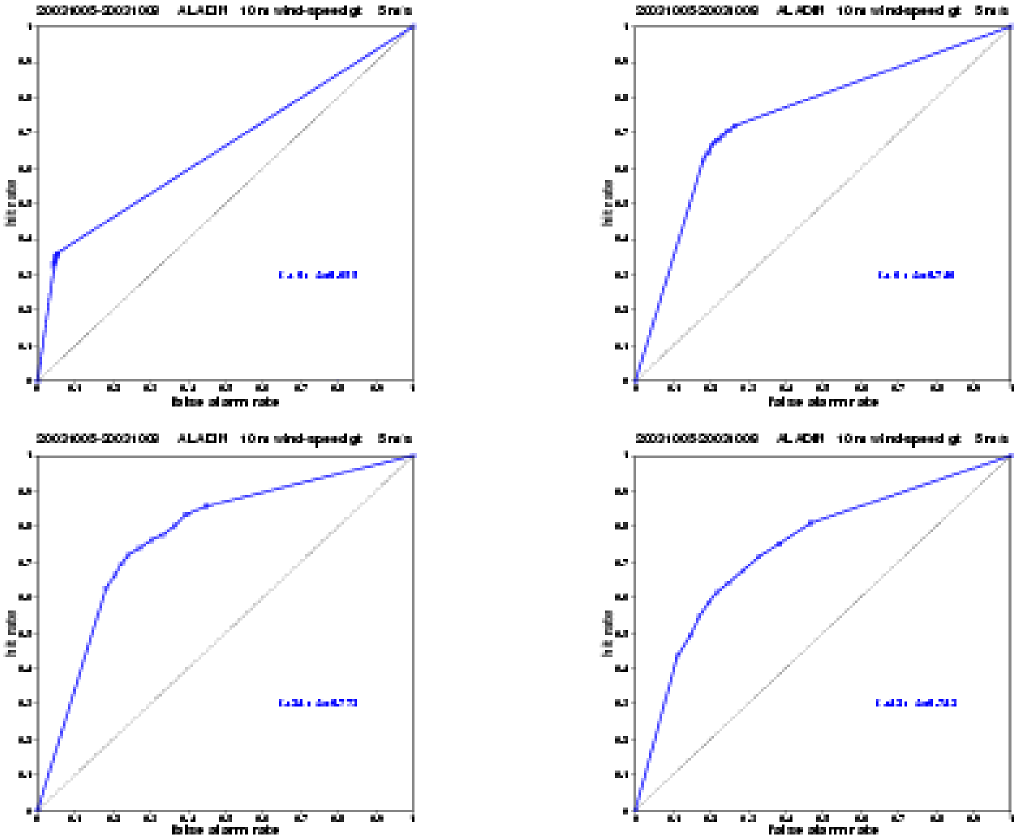

Figure 4. ROC diagram for the model ALADIN, event: 10 m wind speed exceeds 5 m/s (time steps: 00, 06, 24, 42)

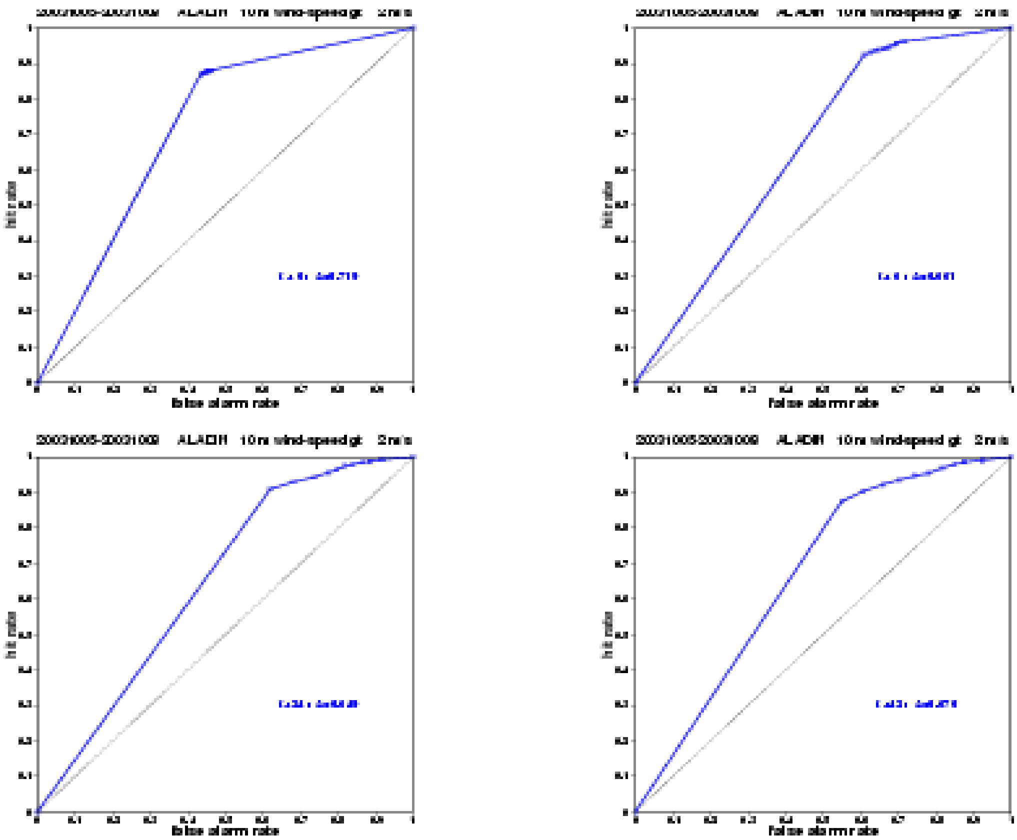

Figure 5. ROC diagram for the model ALADIN, event: 10 m wind speed exceeds 2 m/s (time steps: 00, 06, 24, 42)

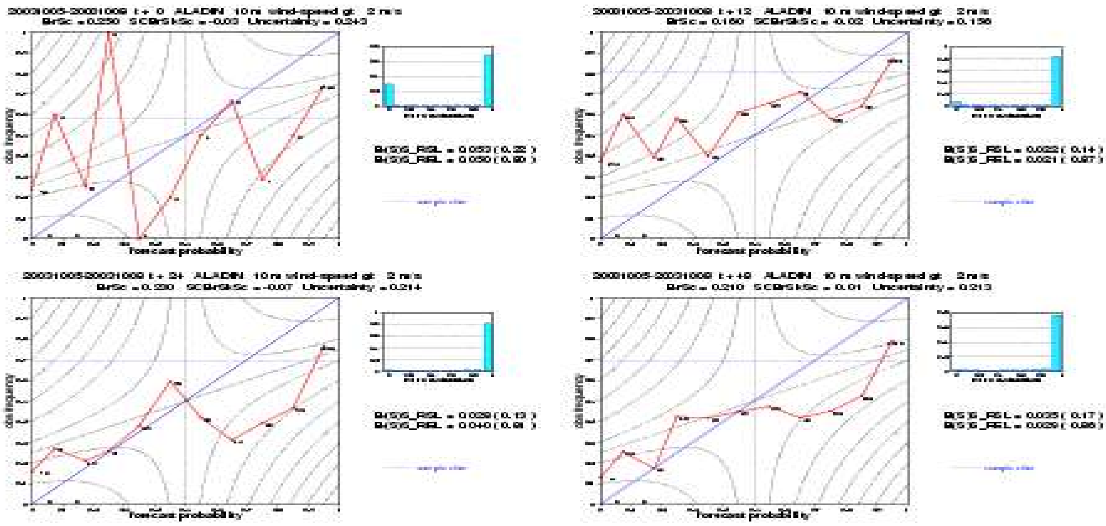

Figure 6**.** Reliability diagram for the model ALADIN, event: 10 m wind speed exceeds 2 m/s (time steps: 00, 12, 24, 48)

3.1.4 Results of the downscaling

From these first experiments with downscaling the PEACE members it seems that the spread is not big enough in our area of interest (Central Europe, especially Hungary). It seems reasonable if we consider that the PEACE system was calibrated in order to get enough spread over Western Europe between 24 and 72 h steps, for wind speed, 500 hPa geopotential and mean-sea-level pressure. The aim of the PEACE system is to detect strong storms. This raises some questions :

- Are the PEACE provided initial and boundary conditions convenient for the local EPS run, for a Central European application?
- What is the impact of different target domains and target times?

To answer these questions it was decided to make some case studies.

3.2 Experiments with different target domains

In our experiments an ARPEGE ensemble system was used, based on PEACE. The main difference is that the target domain was not fixed (for the target time 12h was used). Four different target domains were defined (fig. 7):

- Domain 1: Atlantic Ocean and Western Europe (the same as in PEACE)
- Domain 2: Europe and some of the Atlantic
- Domain 3: covering nearly whole Europe
- Domain 4: slightly bigger than Hungary

We expect that in different meteorological situations the use of different target domains would provide the better results and a compromise should be found to choose the best domain. So far three different meteorological situations were examined. One of them was a convective event in 2002. In this situation large quantity of precipitation (40-70 mm during 24 h) was measured at some places along the river Danube and all the models (ALADIN, ARPEGE, ECMWF) failed to forecast the event. The second case (from 2001) was a situation with a fast moving cold front coming from the west. This time the models overestimated the precipitation. The third situation (from 2004) was one with a quite big temperature overestimation. This error in the forecast of temperature caused a big problem : the models predicted rain, but in reality it was sleet.

Every time the ARPEGE ensemble runs were performed locally with the use of the above mentioned singular-vector target domains, and the ALADIN model was coupled with these ensemble members. In all three cases domains 1, 2 and 3 were used. In the convective situation target domain 4 was also tried. Every time the average standard deviation over Hungary was computed (for 850 hPa temperature, 10 meter wind speed, mean-sea-level pressure and 500 hPa geopotential) and we also looked at different meteorological parameters.

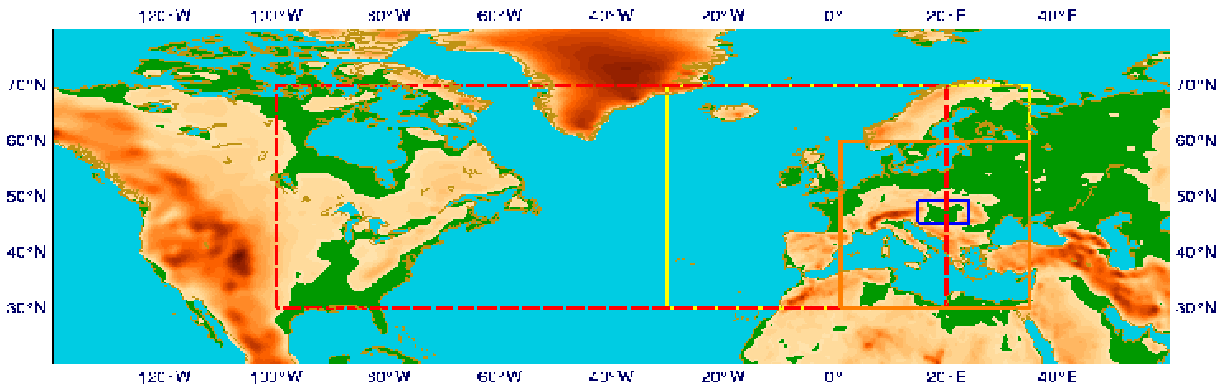

Figure 7. The defined target domains (red: domain 1, yellow: domain 2, orange: domain 3, blue: domain 4)

3.2.1 Results – Spread

In every situation it was found that with the use of the first singular-vector target domain (this is the one used in the PEACE system) the average standard deviation was quite small in the beginning of the forecast and it increased quite slowly. Around the end of the forecast range it usually reached the values obtained by the use of the other domains, but we do not want to concentrate only on the last few hours of the forecast. Instead we would like to find an optimal target domain for the singular-vector computation which guarantees sufficient spread in the 12-48 h time-range.

When target domains 2 and 3 were used the (average) standard deviation was bigger and quite similar both times. In the convective situation the fourth domain (its size being a bit larger than Hungary) was also tried. Doing so the spread over Hungary was quite big in the beginning of the forecast but started to decrease as we went ahead in time. The second case (fast moving cold front) was the only one when standard deviations were nearly the same with the use of domain 1, 2 and 3. The reason of this might be that in this case the examined phenomenon was a large scale one.

It seems that for our purposes the first domain in not convenient in every meteorological situation because the area of biggest spread is usually far from our area of interest (which is Hungary and Central Europe).

3.2.2 Results - Meteorological parameters

Not only the standard deviation was examined but also we looked at different meteorological parameters each time. In the first case (convective case) we got nearly no precipitation at all when we used target domain 1 in the global singular-vector computation. The best results were obtained with the use of domain 4 : some of the members predicted big amount of precipitation at right position (fig. 8). The second case (fast moving cold front) was the only one where standard deviations were nearly the same with the use of domain 1, 2 and 3, and also the predicted amount of precipitation was quite similar. In the third case (sleet) the result was not so good. In reality the temperature was around or below 0 °C all day, but the models predicted much more. A sufficient spread was obtained when domain 2 and 3 was used, but still the values for the temperature were very high. At least some of the members were colder than the control one, but they were not cold enough (fig. 9).

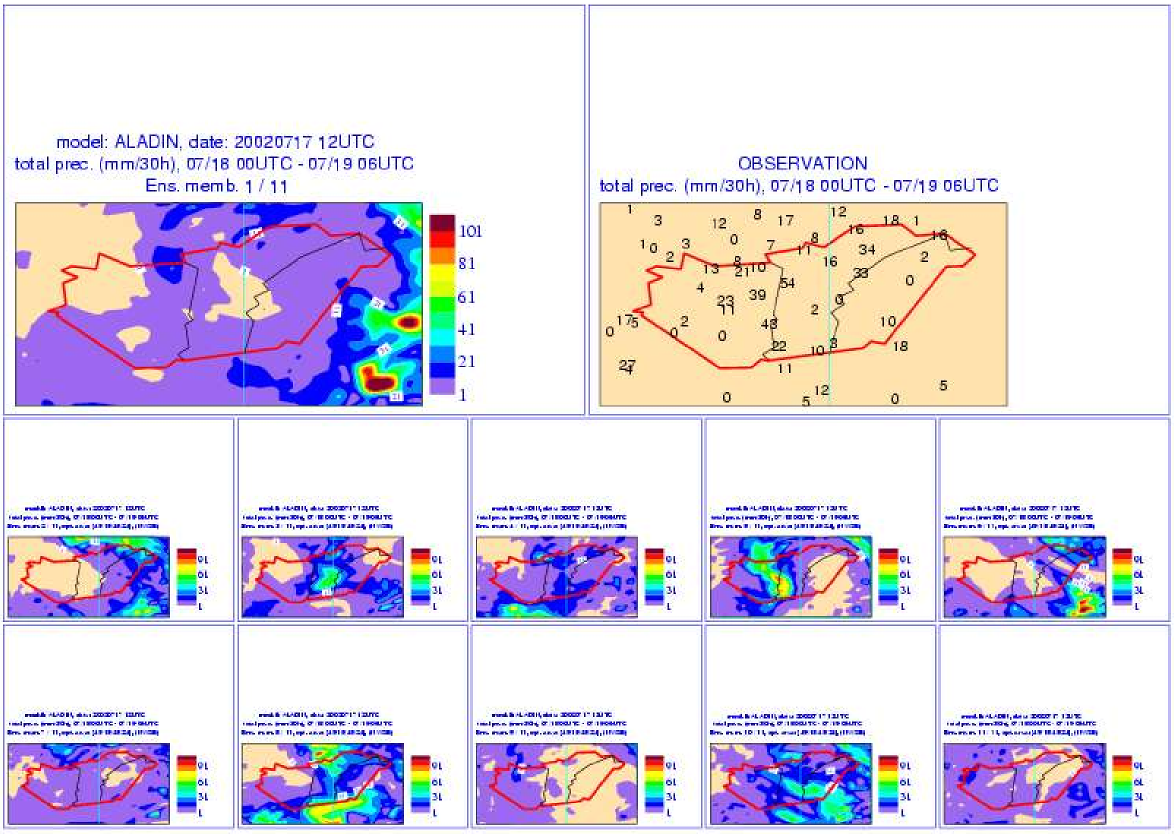

Figure 8. Ensemble members for the 2002 July case. The plotted parameter is total precipitation (mm/30 h) from 18 July 00 UTC until 19 July 06 UTC. The control forecast is at the top left corner, observations at the top right corner. The 10 ensemble members are also plotted. Some of the ensemble members forecasted big amount of precipitation at right position.

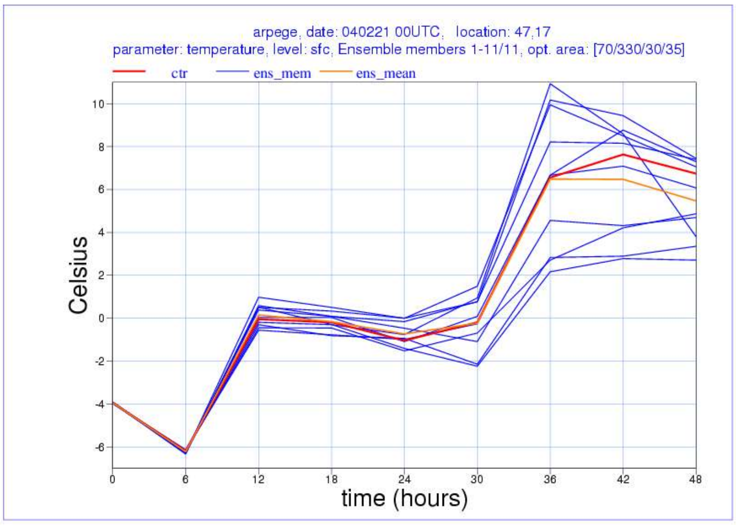

Figure 9**.** Plume diagram for the 2004 February case. The plotted parameter is 2m temperature. Forecast started at 21 February 2004 00 UTC; target domain 2 was used in the global singular-vector computation. On 22 July the highest observed temperature in the country was around three celsius. At +36 h (which is 22 July, 12 UTC) the spread is quite big, but all members are above zero, which means overestimation.

4. Preliminary conclusions

From the case studies and the experiment with downscaling the PEACE members it seems that the PEACE provided initial and boundary conditions are not really optimal for the local ensemble run, for a Central European application. It can be understood if we consider that it was calibrated to Western Europe. Our aim is to find an optimal target domain which fits our purposes, but some case studies still have to be done to find out which domain is the better to use.

5. Future plans

We would like to continue with further case studies to investigate the sensitivity with respect to target domain and also start experiments with different target times. Scores obtained by using the first singular-vector target domain and a different one (domain 2 or 3) will be compared for a longer period (one week - 10 days). It would be interesting to try what would happen when using more perturbations (e.g. to integrate 20 members instead of 10). Also it is planned to start the experiments with other methods especially with ALADIN native singular-vector perturbations.

CONTENTS