

Dynamical downscaling of  
ECMWF EPS forecasts applied to  
cases of severe weather  
in Croatia

Čedo Branković\*, Blaženka Matjačić\*,  
Stjepan Ivatek-Šahdan\* and  
Roberto Buizza

Research Department

\* Croatian Meteorological and Hydrological Service  
Grič 3, 10000 Zagreb, Croatia

January 2007

Submitted for publication in Meteor. Applications

This paper has not been published and should be regarded as an Internal Report from ECMWF.  
Permission to quote from it should be obtained from the ECMWF.



**Series: ECMWF Technical Memoranda**

A full list of ECMWF Publications can be found on our web site under:  
<http://www.ecmwf.int/publications.html>

Contact: [library@ecmwf.int](mailto:library@ecmwf.int)

**© Copyright 2007**

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

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

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

## Abstract

Dynamical downscaling has been applied to global ensemble forecasts to assess its potential impact during cases of severe weather (precipitation and wind) over various parts of Croatia. It was based on the Croatian 12.2-km version of the Aladin limited area model, nested in the ECMWF TL255 (approximately 80 km) global Ensemble Prediction System (EPS). The 3-hourly EPS output was used to force the Aladin model over the central European/northern Mediterranean domain. Four synoptic cases are considered, for which both global EPS and regional Aladin 51-member ensembles were run.

From synoptic cases studied in more detail, downscaling brings improvement in one case where a well-resolved small-scale orography was important. In the second case, no improvement in the downscaled ensemble is seen. This was the case with severe weather defined at very small scales over an area of Croatia with no major orographic differences between the global and regional models. In the third case, severe weather covered a large area over the Adriatic Sea and the coastal Croatia and it was well captured by the global model. For this case, the limited area model yielded only a relatively little improvement in both synoptic development and in terms of probabilistic forecasts for precipitation and wind.

Our results indicate that downscaling can have a large impact on clustering: when applied to both global and downscaled sets of ensembles the same clustering algorithm may yield differing results. For example, even in the case when clusters in both global and regional models are made of the same individual members, cluster means could differ considerably. This implies that downscaling may affect dynamical and physical properties of the global ensemble forecasts. It has been argued that this is due to explicitly resolved small spatial scales in the downscaled ensemble, in particular to those related to orography. Therefore, in the process of downscaling, it may not always be feasible to make a selection (or a subset) of global ensemble members that might be representative of all possible evolution scenarios. This notion is supported by a larger spread in the Aladin than in the EPS ensembles for the fields closely related to small-scale variations in orography (for example, wind and vertical velocity).

Although this paper discusses too few cases for any statistically significant conclusions to be drawn, it provides useful indications on the accuracy of the ECMWF EPS and on the Aladin ensembles in cases of severe weather. Since from the cases considered no clear-cut difference between the global and the downscaled ensembles was detected, additional work with more synoptic cases and more detailed analysis is required to be able to draw some comprehensive conclusions.

## 1. Introduction

Synoptic case studies have always been useful for analysis and better understanding of the capabilities of numerical models to simulate physical and dynamical processes related to severe weather like, for example, strong precipitation, flash floods, gale force winds, hail storms, etc. Although these studies are normally based on a limited amount of readily available observational data, they provide very valuable indications of the strengths and weaknesses of the models under assessment.

The focus of a case-study analysis is usually closely linked with the characteristics of the models used in the investigation - whether they are relatively low resolution atmospheric global circulation models (AGCMs) or relatively high resolution regional models (RMs). For an AGCM, the main concern is to establish whether it is able to reproduce large-scale circulation that preceded and ultimately led to a given severe weather event. For example, using the European Centre for Medium Range Weather Forecasts (ECMWF) deterministic AGCM, Jung et al. (2004) studied several European storms: the Dutch storm of 1953, the Hamburg storm of 1962 and the October 1987 storm. The latter severely affected the northern France and the southern United Kingdom. They concluded that despite some underestimation of the severity of the storms, reliable predictions of severe weather by ECMWF global model were possible several days in advance.

In the second group of case studies, attention is focused on smaller-scale phenomena, since RMs enable a better spatial and temporal resolution, normally required to accurately analyse synoptic and meso-scale atmospheric features related to severe weather. For example Shutts (1990) claims that the UK Met Office fine-mesh model successfully predicted the details of the above-mentioned October 1987 storm. Several synoptic-scale studies have been made in the Croatian Meteorological and Hydrological Service (CMHS) with a limited area model, normally employed for short-range operational forecasting. Various local phenomena have been studied, many of them related to severe weather (e.g. Tudor and Ivatek–Šahdan 2002, Ivančan–Picek et al. 2003, Strelec–Mahović and Drvar 2005, Ivatek–Šahdan and Ivančan–Picek 2006). Although in each of the above CMHS studies a single RM run was made, the results enabled a detailed insight into various aspects of processes defined on relatively small scales.

In the past decade, ensembles of forecasts brought new, both qualitative and quantitative, improvements to the process of operational forecasts and have become an indivisible part of operational forecasting practice in many weather services. Thus, the 1990s pioneering work at ECMWF (e.g. Molteni et al. 1996, Buizza 1997, Palmer et al. 1997) has led to a widespread application of the atmospheric predictability theory in the medium-range (e.g. Lorenz 1982). Global ensemble forecasts have been used to revisit interesting cases: for example, Buizza and Hollingsworth (2002) used the ECMWF ensemble prediction system (EPS) to study the December 1999 European storms. They found that the EPS was a valuable tool for assessing the risk of severe weather and issuing early warnings. Similarly, Buizza and Chessa (2002) studied the impact of a stochastic parameterisation (Buizza et al. 1999) on the EPS forecasts of a severe storm that affected the US in January 2000, and concluded that the stochastic parameterisation improved the capability of the EPS to simulate the storm's development. Jung et al. (2005) showed that the ensemble forecasts could have provided extremely valuable supplementary information to that generated by a single, high-resolution (deterministic) forecast in the case of the October 1987 storm. They concluded that the EPS was capable of predicting large uncertainties associated with the timing of the storm.

Because of their relatively coarse resolutions, AGCMs cannot resolve small-scale features and their usefulness to study phenomena at sub-synoptic scales is limited. Severe weather events are often associated with a relatively large variability of atmospheric small-scale parameters, which in turn might be related to variations in local topography. Thus, in order to better evaluate and understand local processes, and to be able to verify them against observations, there appears a need to carry out dynamical downscaling of global ensembles. Such downscaled ensembles might be useful for further applications in e.g. hydrological and crop-yield models.

The purpose of this study is twofold. First, we try to identify the differences in the statistics between forecast ensembles made by an AGCM and the downscaled forecast ensembles made by a RM. Such statistics is extended to clusters that were derived in identical way from both global and regional ensembles. Second, the differences in the (ensemble) analysis of selected synoptic cases between the two types of models are shown and discussed. In this paper, multiple model realisations are linked, through the downscaling of the AGCM results, to a limited domain of interest. Of course, such attempts have already been made (e.g. Stensrud et al. 1999, Montani et al. 2001, Marsigli et al. 2001, Frogner and Iversen 2002). There is no unique approach to dynamical downscaling of ensemble forecasts. Some authors, for example, applied the so-called “representative members” approach (e.g. Molteni et al., 2001; Montani et al., 2001). This implies various intermediate steps in order to define a reduced number of members for dynamical downscaling that characterise all possible evolution scenarios of the global model EPS. It is important to emphasise that in this

study dynamical downscaling is applied to all global model ensemble members. Whilst some pieces of information are inevitably lost in the selection of representative members, no such loss in our approach is possible. Thus, dynamical downscaling of the whole global model ensemble represents an ultimate way of applying ensembles to smaller spatial scales. Due to a relatively large computational demand such an approach might not, at present, be viable in an operational practice.

In the next section experiments and methodology are described. Section 3 deals with a general (statistical) comparison between the results from the global model and regional model ensembles, irrespective of how a particular model performed during the given synoptic developments. The details of synoptic situations and various aspects of the results from the two sets of ensembles (global and regional), with the emphasis on precipitation probabilities and possible gains that may be attained by dynamical downscaling, are shown and discussed in section 4. Summary and some conclusions are given in section 5.

## **2. Experimental design**

### **2.1. Models and dynamical downscaling**

For this study, the global ensembles were generated by the ECMWF model cycle 28R4 that was in ECMWF operations from 18 October 2004 till 5 April 2005. The model was integrated at the TL255 spectral resolution (approximate horizontal resolution 80 km) with 40 levels in the vertical and a 45-minute timestep. ECMWF EPS initial conditions (ICs) were made as in ECMWF operations. The global model ensembles contained 51 members each.

In order to create lateral boundary conditions (LBCs) for dynamical downscaling by the Aladin (Aire Limitée Adaptation dynamique Développement InterNational) RM, the 3-hourly output from ECMWF EPS (hereafter referred to as ECEPS) was used. LBCs were defined separately from ECMWF model levels and surface data. The upper-air LBCs were interpolated from ECMWF model levels to Aladin model levels, and also converted to the appropriate format. For surface fields, a selected output from the ECMWF 4-layer land-surface model (van den Hurk et al. 2000) was scaled to accommodate the Aladin 2-layer surface scheme (Giard and Bazile 2000). The ECMWF surface parameters soil temperature, soil water and snow depth were converted to Aladin's soil temperature, soil water and ice and snow depth. No other modification or manipulation with ECEPS output data has been carried out.

The full set of such defined ICs and LBCs from ECEPS was then applied to Aladin, i.e. Aladin ensembles (hereafter referred to as ALEPS) also contained 51 members. The Aladin RM was run at the regular 12.2-km grid, at the Lambert conformal projection, with 37 levels in the vertical and with the timestep of 514 s (8.6 min). The central point of the integration domain was positioned at (17°E, 46.2°N) with 229 grid points in the x direction and 205 grid points in the y direction, thus covering central Europe and the northern Mediterranean. The Aladin dynamical core is similar to that of ECMWF EPS (the integrated forecasting system - IFS); however, the physical parameterisations in the two models are different. The references that describe the Aladin RM in more details are given in, for example, Ivatek-Šahdan and Tudor (2004).

In addition to physical parameterisations, one of the most important differences between ECEPS and ALEPS is the definition of orography field. Fig. 1a,b shows orography in the Aladin integration domain for both models interpolated, for the comparison purpose, to the regular  $0.5 \times 0.5^\circ$  latitude/longitude grid. The 200-m

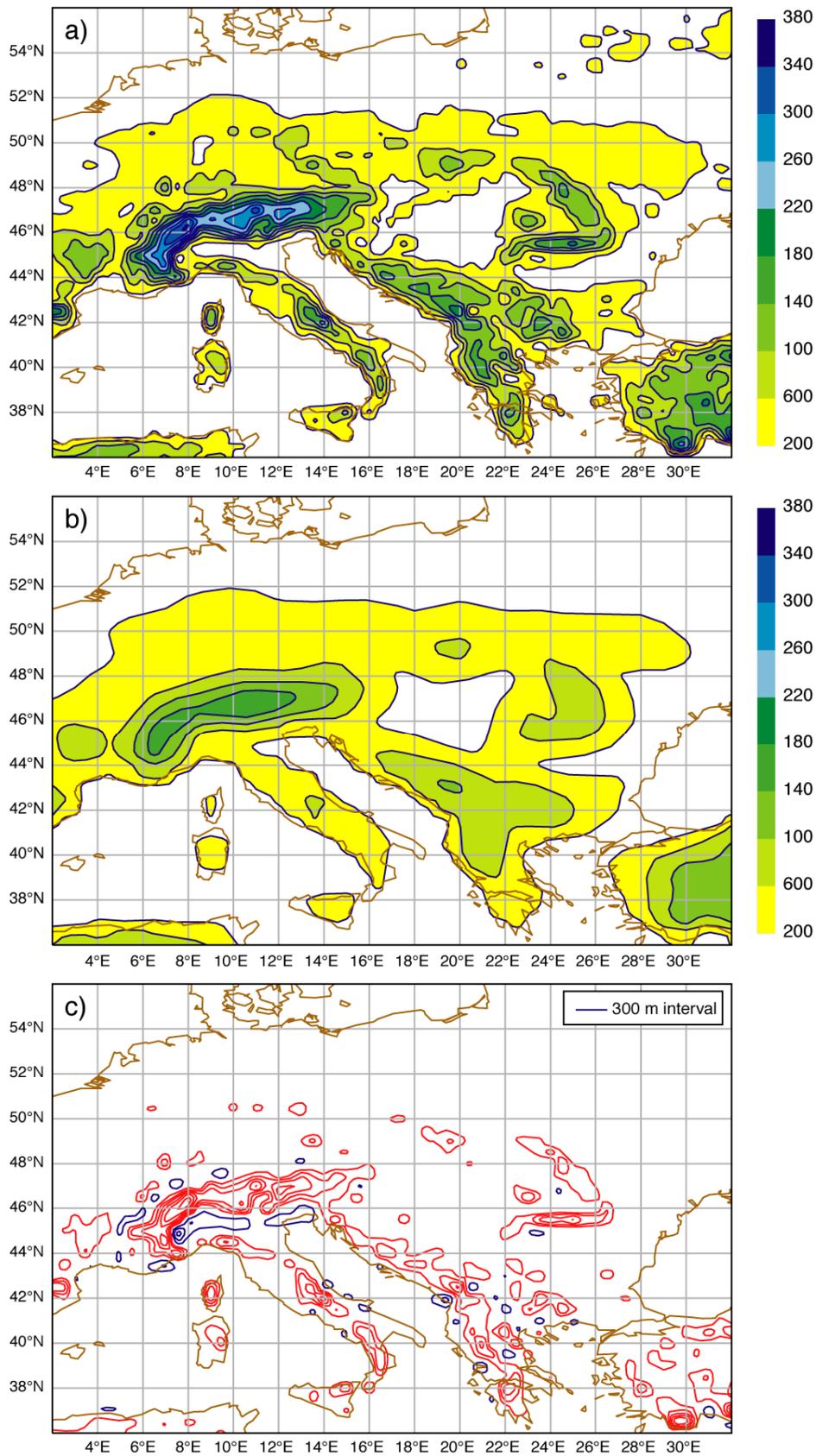


Figure 1 Model orography at 0.5°x0.5° resolution in a) Aladin, b) ECMWF EPS, and c) the difference Aladin minus ECMWF EPS. Contours in a) and b) 400 m starting from 200 m. Contours in c) 300 m with positive contours in red and negative contours in blue.

contour (yellow shading) encircles approximately the same area in both models. However, whilst the maximum height of the Alps in ECEPS is below 1800 m, in the Aladin model it exceeds 3400 m. The maximum orographic difference between ECEPS and ALEPS models, located in the western Alps at about (8°E, 46°N), exceeds 1800 m (Fig. 1c). Other large differences are seen in the southern Turkey (over 1500 m), central Italy (1200 m) and in the Carpathian Mountains, the Pyrenees and in the southern Greece (over 900 m). In addition to the Alps, the improvement in the orographic representation in ALEPS is seen over the mountains of the Balkan and Italian peninsulas and large Mediterranean islands. When compared to ECEPS, the orography in Croatia has been “raised” in ALEPS between 300 and 600 m in the eastern Adriatic coastal region and its hinterland. Fig. 1c also indicates an increase in the horizontal orography gradient in Aladin near steep and high mountains. For example, the north-south section near 8°E shows the “deepening” of the river Po Valley in the northern Italy by about 900 m. It is expected that such differences in orography would contribute to the differences in orography-related fields, like, for example, precipitation or surface temperature.

## 2.2. Data and methods

Four ECMWF ensembles were run with a 3-hourly output for dynamical downscaling with the Aladin model. The results from these four cases were used to compare various statistical properties of ECEPS and ALEPS. These four synoptic cases, two summer (SU) and two autumn (AU) cases, are summarised in Table 1. Whereas ensemble statistics is based on all four cases (section 3), only three cases are studied in more detail in section 4. They are the cases of severe weather that occurred over various parts of Croatia in the summer of 2003 and autumn 2004. Based on their intensity and inflicted damage, these three cases may by no means be considered as extreme weather events when compared with some other Mediterranean storms discussed by, for example, Montani et al. (2001, 2003) or by Tripoli et al. (2005). However, they are typical of severe weather that occasionally hits Croatia and therefore are appropriate to be studied by an ensemble prediction system. The fourth synoptic case has not, in terms of severe weather, affected Croatia. Some other parts of central Europe, in particular the lower Austria and the Tatre Mountains region in Slovakia and south Poland experienced gale force winds and excessive precipitation (both rain and snow). For this case, the ECMWF EPS indicated a possible increase in precipitation and wind in the westernmost part of Croatia, and therefore it could be characterised as a sort of a “mild” false alarm.

Case	Initial time and date	Target forecast period for case study
SU1	00UTC 2 July 2003	T+60 to T+72
SU2	00UTC 26 July 2003	T+66 to T+78
AU1	00UTC 12 November 2004	T+36 to T+72
AU2	00UTC 17 November 2004	

*Table 1 The four ensembles and three synoptic cases described and discussed in the main body of the paper with respective initial dates and target forecasting periods.*

Since this study derives from a presumed capability of the RM to better simulate small scale features, and since a limited number of cases was considered, the verification is focused mainly on the synoptic analysis of the quality of cluster and of probabilistic forecasts. Such an approach could be justified because probability density functions (PDFs) might have been changed due to dynamical downscaling in a model with higher

resolution and different physical parameterisations. Thus, the clustering and probabilities in (downscaled) ensembles would help to identify details in forecast development and would hopefully enable more accurate operational prediction of severe weather events.

### 3. Global model vs. limited area model

#### 3.1. Comparison of full ensembles

We first discuss and compare some general properties (statistics) for the full ECEPS and ALEPS ensembles over the domain of interest. A relatively simple comparison of the two sets of ensembles is made against the ECMWF operational analyses. The latter appears a natural choice, because a common reference might seem desirable for the comparison of the two, to a certain extent, correlated ensembles. However, it also might be assumed that due to the difference in spatial scales between the two ensembles, the choice of ECMWF analysis as the reference for the comparison would favour ECEPS. As it is demonstrated below, this proves not to be the case. In the following, all the fields from both ensembles and from ECMWF analysis were interpolated to the regular  $0.5 \times 0.5^\circ$  latitude/longitude grid. In section 4, a more detailed comparison between ECEPS and ALEPS based on case studies is given.

In Table 2, the following three simple and relatively crude measures are shown: the ensemble mean absolute difference with respect to ECMWF operational analysis  $\delta_e$ , the ensemble mean deviation  $\sigma_e$  (also measured with respect to ECMWF operational analysis), and the ensemble spread  $S_e$  measured with respect to ensemble mean. For an ensemble of  $M$  individual forecast  $F_i$ , ensemble mean absolute difference  $\delta_e$  with respect to verifying analysis  $A$  is given by

$$\delta_e = \frac{1}{M} \sum_i^M |F_i - A|,$$

and likewise, ensemble mean deviation  $\sigma_e$  could be defined as the following

$$\sigma_e = \left[ \frac{1}{M-1} \sum_i^M (F_i - A)^2 \right]^{1/2}$$

$\delta_e$  and  $\sigma_e$  measure overall departures of the two sets of ensembles from ECMWF verifying analysis and represent some sort of aggregated modelling error. Ensemble spread  $S_e$  is computed as the distance between each individual ensemble member  $F_i$  and the mean of the ensemble  $\bar{F}_e$ , and averaged over all ensemble members. The expression is essentially similar to that for  $\sigma_e$  where the verifying analysis  $A$  is replaced by the ensemble mean  $\bar{F}_e$

$$S_e = \left[ \frac{1}{M-1} \sum_i^M (F_i - \bar{F}_e)^2 \right]^{1/2}.$$

In terms of probability distribution, spread of an ensemble measures the dispersion of forecast states. In the presence of small model errors, small spread usually indicates that the ensemble mean is relatively skilful. All statistical quantities discussed here,  $\delta_e$ ,  $\sigma_e$  and  $S_e$ , are computed for various upper-air parameters at the T+48 hr forecast step over the domain ( $36^\circ\text{N}$ – $56^\circ\text{N}$ ,  $2^\circ\text{E}$ – $32^\circ\text{E}$ ) and for all synoptic cases considered.

Parameter	Synoptic case	Mean absolute difference $\delta_e$		Mean deviation $\sigma_e$		Ensemble spread $S_e$	
		ECEPS	ALEPS	ECEPS	ALEPS	ECEPS	ALEPS
<b>Z500</b> (dam)	SU1	<b>0.65</b>	0.57	<b>1.64</b>	1.56	<b>1.39</b>	1.35
	SU2	<b>0.63</b>	0.59	1.45	<b>1.46</b>	1.23	<b>1.27</b>
	AU1	1.97	<b>2.01</b>	<b>4.91</b>	4.83	<b>4.29</b>	4.18
	AU2	<b>1.51</b>	1.44	<b>4.10</b>	3.96	<b>3.45</b>	3.37
<b>Z700</b> (dam)	SU1	<b>0.75</b>	0.63	<b>1.50</b>	1.36	<b>1.14</b>	1.07
	SU2	<b>0.78</b>	0.74	<b>1.39</b>	1.35	<b>1.07</b>	1.03
	AU1	1.45	<b>1.49</b>	<b>3.40</b>	3.34	<b>2.92</b>	2.82
	AU2	<b>1.05</b>	0.97	<b>2.64</b>	2.59	<b>2.24</b>	2.21
<b>T850</b> (°)	SU1	0.70	<b>0.78</b>	<b>1.44</b>	1.35	<b>1.16</b>	0.91
	SU2	0.97	<b>1.14</b>	1.70	<b>1.79</b>	<b>1.17</b>	1.10
	AU1	1.17	<b>1.18</b>	<b>2.43</b>	2.35	<b>1.91</b>	1.82
	AU2	<b>1.36</b>	1.15	<b>2.39</b>	2.04	<b>1.64</b>	1.41
<b>Wind 850</b> (ms <sup>-1</sup> )	SU1	1.85	1.85	3.10	<b>3.18</b>	2.00	<b>2.10</b>
	SU2	<b>1.50</b>	1.49	2.83	<b>2.85</b>	2.01	<b>2.07</b>
	AU1	<b>2.63</b>	2.60	4.87	<b>4.97</b>	3.45	<b>3.57</b>
	AU2	3.08	<b>3.12</b>	<b>5.47</b>	5.45	3.50	<b>3.60</b>
<b>RT 500/1000</b> (dam)	SU1	0.92	0.92	<b>1.76</b>	1.73	<b>1.36</b>	1.27
	SU2	<b>1.00</b>	0.91	<b>2.00</b>	1.91	1.49	<b>1.53</b>
	AU1	<b>1.69</b>	1.57	<b>4.16</b>	4.07	<b>3.61</b>	3.52
	AU2	<b>2.45</b>	2.05	<b>4.92</b>	4.48	<b>3.59</b>	3.38
<b>RT 700/1000</b> (dam)	SU1	0.57	<b>0.63</b>	<b>1.17</b>	1.14	<b>0.94</b>	0.80
	SU2	0.70	<b>0.89</b>	1.39	<b>1.49</b>	<b>1.07</b>	1.03
	AU1	<b>0.98</b>	0.95	<b>2.23</b>	2.16	<b>1.84</b>	1.77
	AU2	<b>1.26</b>	0.97	<b>2.32</b>	2.00	<b>1.61</b>	1.45
<b><math>\omega</math>500</b> (Pa)	SU1	0.17	<b>0.18</b>	0.30	<b>0.39</b>	0.13	<b>0.26</b>
	SU2	0.12	<b>0.13</b>	0.26	<b>0.33</b>	0.18	<b>0.26</b>
	AU1	0.29	<b>0.30</b>	0.52	<b>0.67</b>	0.24	<b>0.47</b>
	AU2	0.20	0.20	0.35	<b>0.44</b>	0.16	<b>0.30</b>
<b><math>\omega</math>700</b> (Pa)	SU1	0.18	<b>0.19</b>	0.30	<b>0.42</b>	0.15	<b>0.29</b>
	SU2	0.14	<b>0.15</b>	0.28	<b>0.36</b>	0.19	<b>0.28</b>
	AU1	0.31	<b>0.32</b>	0.58	<b>0.71</b>	0.28	<b>0.46</b>
	AU2	0.20	<b>0.22</b>	0.35	<b>0.48</b>	0.19	<b>0.34</b>

Table 2 The T+48 hr absolute mean difference with respect to ECMWF operational analysis (two left columns), mean deviation (two middle columns) and ensemble spread for ECEPS and ALEPS ensembles computed over the domain (36°N–56°N, 2°E–32°E). Bold print indicates the larger value of the two ensembles.

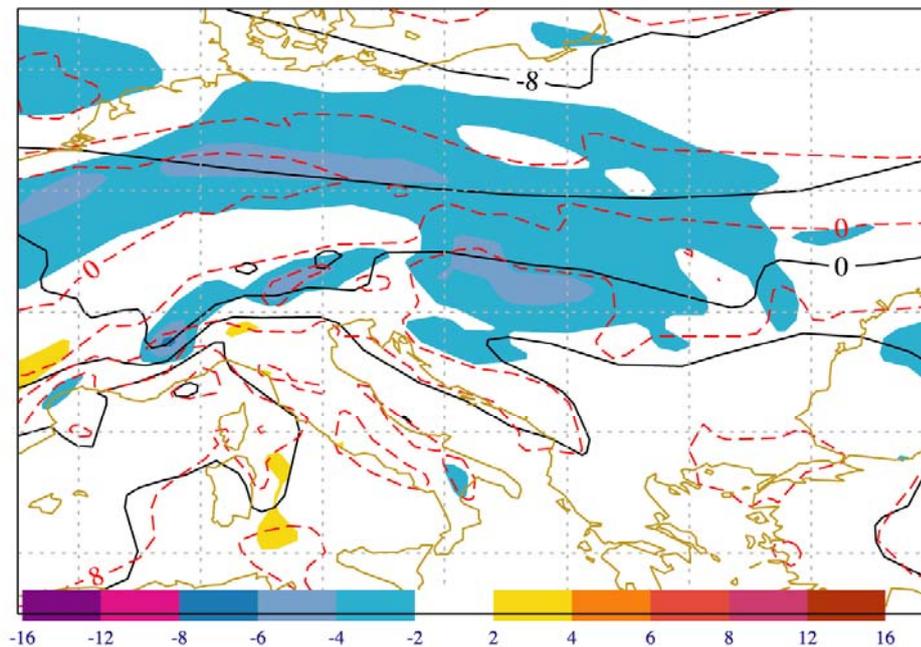
Overall, the mean absolute difference  $\delta_e$  for ALEPS does not differ substantially from that for ECEPS. In 15 out of the total 32 synoptic case/parameter combinations (8 parameters times 4 synoptic cases) the ALEPS  $\delta_e$  was larger than the ECEPS  $\delta_e$  (indicated by the bold typeset in Table 2). In 14 combinations, the opposite is seen, i.e. the ECEPS  $\delta_e$  was larger than the ALEPS  $\delta_e$ , and in 3 combinations the  $\delta_e$  values were identical in both models. When the values for vertical velocity  $\omega$  are excluded from the consideration the total number of the ALEPS combinations being larger than ECEPS drops dramatically from 15 to 8. Such a result broadly implies that, in our four synoptic cases, the dynamical downscaling reduces the modelling error for most parameters considered except for  $\omega$ . In other words, a better horizontal resolution may introduce errors in parameters closely dependent on orography if they are “verified” against a relatively coarse resolution reference field.

This is further confirmed for ensemble mean deviation  $\sigma_e$ . When  $\sigma_e$  values for  $\omega$  parameters are excluded from Table 2, a fairly regular pattern of higher  $\sigma_e$  values for ECEPS than for ALEPS could be clearly seen (18 out of 24 values). This may seem somewhat surprising bearing in mind a potentially higher variability in ALEPS that would have stemmed from better-resolved smaller spatial scales. Part of the explanation for such results could be that most values for the non-omega parameters in Table 2 are related to upper-air fields that usually have relatively smooth features above model orography. This is confirmed by overall higher ALEPS than ECEPS  $\sigma_e$  values for wind at the 850-hPa level. The  $\omega$  field from Table 2, though defined at upper-air levels, is by definition and its physical characteristics susceptible to a much stronger orographic influence than the other parameters.

The higher ALEPS than ECEPS  $\sigma_e$  values for wind magnitude at 850 hPa could be indicative of an increased influence of a high-resolution orography on the low-level circulation variability. Clearly, the higher ALEPS orography “interferes” with the 850 hPa atmospheric flow in more grid points than in ECEPS. Also, a higher absolute mean difference  $\delta_e$  for ALEPS T850 could be possibly attributed to the discrepancy in orographic heights between ECEPS and ALEPS (cf. Fig. 1). On the other hand, consistently larger  $\delta_e$  and  $\sigma_e$  values for the autumn than for summer cases, irrespective of the model, could be associated with an increased natural atmospheric variability due to seasonal cycle.

From Table 2, the largest differences in  $\delta_e$  and  $\sigma_e$  values between ECEPS and ALEPS are found for T850 in the AU2 synoptic case (1.36 vs. 1.15 for  $\delta_e$ , and 2.39 vs. 2.04 for  $\sigma_e$ ) and again for relative topography. For T850, Fig. 2 illustrates that both models largely underestimate analysed temperature, i.e. the erroneous cooling is found over much of the western and central Europe. This cooling in ECEPS exceeds  $-4^\circ\text{C}$  over the Hungarian Plain, central Germany and the Czech Republic (Fig. 2a), the regions with a relatively inconspicuous orography (Fig. 1). Similar error pattern, but with somewhat reduced amplitude, is seen in ALEPS (Fig. 2b). For relative topography 1000/500, which in Table 2 shows the largest difference between ECEPS and ALEPS, the reduction in geopotential extends farther north, over the northern Germany and Poland (not shown). These results indicate that the largest differences between the two models are not always associated with high orography. For example, a part of the erroneous warming between  $+2$  and  $+4^\circ\text{C}$  seen in ALEPS over the river Po Valley (Fig. 2b) is in the region of relatively large model orographic differences (more than 300 m, cf. Fig. 1c) though the orography itself is not very high there.

(a) ECMWF 20041117 +48h t850 em (black) & an (red) & (em - an) shaded  
 cl. area: [ 56N, 2E, 36N, 32E, ], ensemble\_std = 2.39 abs(em-an) = 1.36



(b) ALADIN 20041117 +48h t850 em (black) & an (red) & (em - an) shaded  
 cl. area: [ 56N, 2E, 36N, 32E, ], ensemble\_std = 2.04 abs(em-an) = 1.15

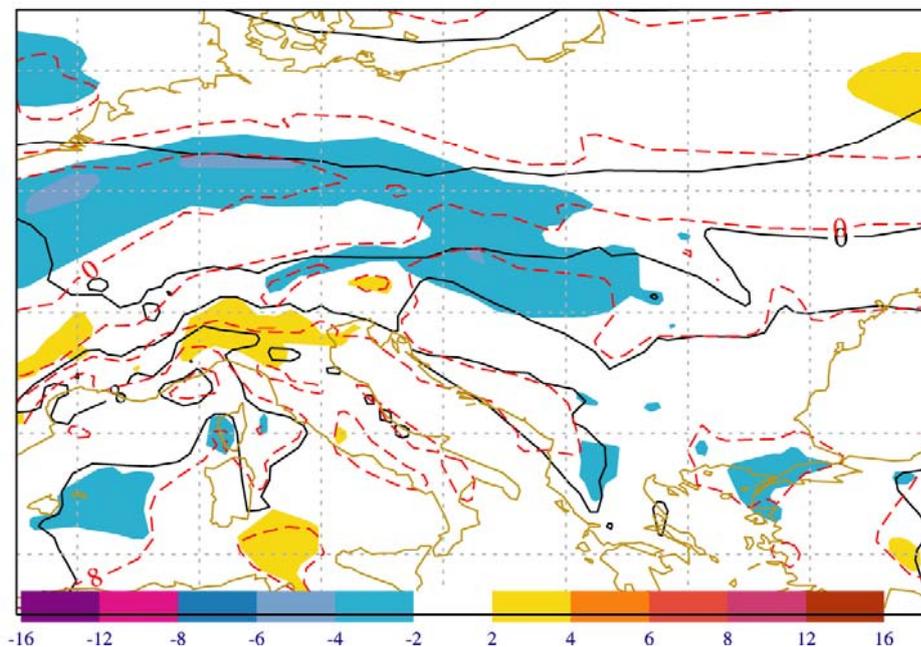


Figure 2 The 850-hPa temperature difference between ensemble mean and ECMWF analysis for the AU2 synoptic case at T+48 hr for a) ECEPS and b) ALEPS. Ensemble mean in black solid, ECMWF analysis in red dashed, and the difference shaded. Contouring interval 4° for full fields and 2° for differences.

Ensemble spread  $S_e$  (the last two columns in Table 2) is constantly higher in ALEPS than in ECEPS for wind at 850 hPa and both  $\omega$  fields. It could be argued that small spatial scales in Aladin cause a larger ensemble

spread in those fields that are in general more influenced by orography. This statistic is consistent with ensemble mean deviation with respect to operational analysis  $\sigma_e$ .

### 3.2. Clustering

Clustering is the method by which individual forecasts from an ensemble that are close to each other are grouped together. The “closeness” of individual forecasts is based upon some objective criterion, for example, the root-mean squared (RMS) difference among individual members. The most common algorithm for clustering is the so-called Ward hierarchical clustering algorithm (Anderberg 1973). Clustering enables to identify some atmospheric features (flows, synoptic regimes) that might attain an increased probability for the occurrence. For example, in a forecasting system with relatively small systematic errors the most populated cluster would be normally accepted as the one having the highest probability (chance) of realisation.

After manipulating the data from both ensembles in identical way, the comparison of the ECEPS and ALEPS results would enable a better insight and a more detailed assessment of potential benefits of dynamical downscaling. Thus, the standard ECMWF clustering algorithm (see for example, Atger 1999) has been applied to both ECEPS and ALEPS ensembles for all four cases considered. The smaller spatial scales contained in ALEPS may affect some ensemble properties, and eventually clustering results for ALEPS could be different from those for ECEPS, even if clusters were made of the same members. As before (see section 3.1), both ECEPS and ALEPS data were interpolated to the regular  $0.5^\circ \times 0.5^\circ$  latitude/longitude grid. The clustering was performed over the domain ( $36^\circ\text{N}$ – $56^\circ\text{N}$ ,  $2^\circ\text{E}$ – $32^\circ\text{E}$ ) for all parameters from Table 2 at the T+48 hr forecast step. In addition, it has been decided beforehand that the clustering algorithm would generate no more than three clusters.

#### 3.2.1. Cluster size and common members

First we discuss some basic properties of clustering for the two sets of ensembles and the differences between them. Table 3 shows the cluster size (number of individual members) for all three clusters in both ECEPS and ALEPS. The clustering algorithm has defined the ordering of clusters in Table 3, i.e. they have not been sorted out according to their size. Overall, in terms of this simple statistics, there are more differences between ECEPS and ALEPS than similarities. Only in 2 out of the total 32 parameter/synoptic case combinations the clustering algorithm yields identical result: for Z700 and RT 500/1000, both in the AU2 synoptic case. For both cases, not only the size of ECEPS and ALEPS clusters is identical, but also are individual members within each cluster. However, as mentioned above, the resulting ensemble mean fields for corresponding clusters may not be necessarily identical, although in our case they are very similar (not shown).

There are several very similar groupings of individual forecasts for ECEPS and ALEPS, differing only slightly in cluster size - for example, in AU1 for Z700, T850, wind 850 and RT 1000/700, in SU2 for RT 1000/500 and in SU1 for RT 1000/700. However, for most combinations in Table 3 the differences in size between ECEPS and ALEPS clusters are relatively large. The largest differences in the number of cluster members are found for Z500 and  $\omega$  fields in all four synoptic cases. Fig. 3 illustrates these differences for Z500 in the AU1 synoptic case – the shaded areas are the differences between cluster means and ECMWF analysis (cluster “errors”). One may argue that there is some similarity between the most populated clusters: 32 members in ECEPS (Fig. 3 top right) and 24 members in ALEPS (Fig. 3 middle left). The error pattern

and amplitude for the second most populated clusters (13 members in ECEPS, Fig. 3 middle right and 20 members in ALEPS, Fig. 3 top left) look very different. The third ALEPS cluster (7 members, Fig. 3 bottom left) shows some similarity with the ECEPS cluster number 2, as does the second ALEPS cluster with the third ECEPS clusters (6 members). It is clear from Fig. 3 that similar error patterns between the ECEPS and ALEPS clusters arise from identical individual ensemble members being included in both clusters. For example, all seven members from the ALEPS third cluster (Fig. 3 bottom left) are included in the second ECEPS cluster (with 13 members; Fig. 3 middle right).

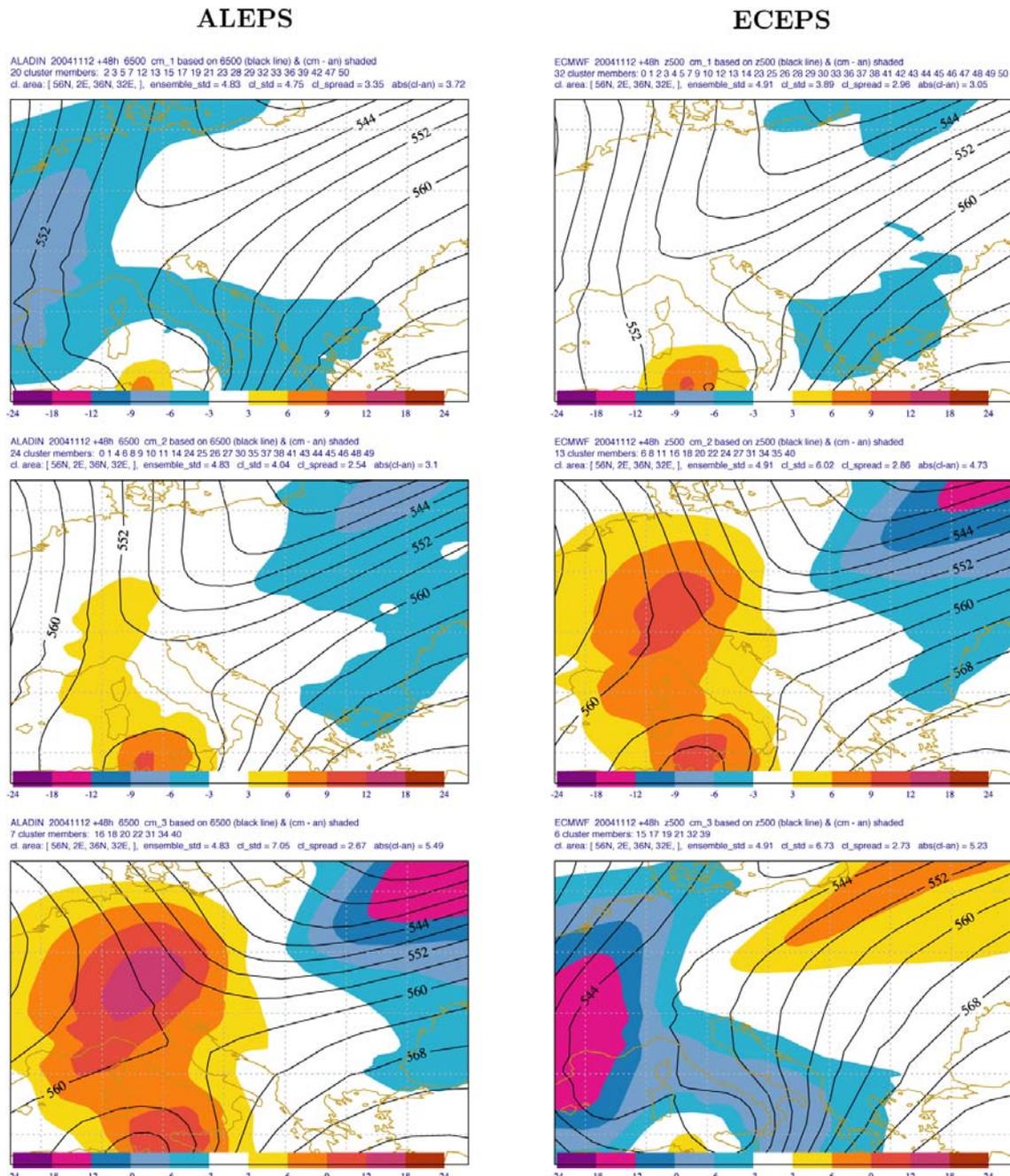


Figure 3 The 500-hPa geopotential height clusters for the AUI synoptic case at T+48 hr for ALEPS (left) and ECEPS (right). Contouring interval 4 dam for cluster means (solid black), and 3 dam for errors with respect to ECMWF operational analysis (shaded).

Parameter	Synoptic case	ECEPS clusters			C	ALEPS clusters		
		1	2	3		1	2	3
<b>Z500</b>	SU1	22	23	6	18	33	12	6
	SU2	30	12	9	21	11	21	19
	AU1	32	13	6	18	20	24	7
	AU2	16	23	12	20	32	14	5
<b>Z700</b>	SU1	31	14	6	23	24	21	6
	SU2	19	20	12	17	17	18	16
	AU1	26	11	14	24	24	12	15
	AU2	17	19	15	19	17	19	15
<b>T850</b>	SU1	19	<u>18</u>	14	16	30	10	11
	SU2	14	18	19	19	24	18	9
	AU1	34	10	7	33	34	8	9
	AU2	19	24	8	23	26	16	9
<b>Wind 850</b>	SU1	22	<u>22</u>	7	21	27	10	14
	SU2	23	12	16	15	24	<u>23</u>	4
	AU1	25	14	12	24	26	14	11
	AU2	21	18	12	21	26	18	7
<b>RT 500/1000</b>	SU1	28	11	12	21	32	13	6
	SU2	21	<u>19</u>	11	19	20	<u>20</u>	11
	AU1	32	9	10	17	27	21	3
	AU2	29	16	6	29	29	16	6
<b>RT 700/1000</b>	SU1	27	11	13	14	28	13	10
	SU2	31	14	6	18	15	18	18
	AU1	33	9	9	30	30	9	12
	AU2	24	19	8	20	<u>20</u>	24	7
<b>ω500</b>	SU1	24	14	13	22	27	16	8
	SU2	11	28	12	9	16	21	14
	AU1	10	30	11	12	18	16	<u>17</u>
	AU2	<u>18</u>	20	13	14	24	17	10
<b>ω700</b>	SU1	22	10		18	25	15	11
	SU2	34	10	7	19	20	21	10
	AU1	32	8	11	23	12	24	15
	AU2	19	<u>19</u>	13	10	24	15	12

Table 3 Cluster size of each cluster (1, 2, 3) for ECEPS and ALEPS. In the shaded middle column is the number of common members in both ECEPS and ALEPS from the most populated or the second most populated (underlined) clusters. The base clustering time is  $T+48$  hr.

The number of common members, i.e. ensemble members residing in both ECEPS and ALEPS most populated or second most populated clusters is shown in the shaded middle column in Table 3 (denoted C). The highest number of common members, 33, is found for T850 in AU1. Even in this case when ECEPS and ALEPS clusters have the highest number of common members, errors in cluster means may not look identical or even similar (not shown). Thus, many common members do not necessarily guarantee similarity between clusters from the two different populations. It is through dynamical downscaling that physical attributes of clustering might be changed. Since for both ECEPS and ALEPS the clustering algorithm is applied at the same grid, this example indicates how different results may be reached after dynamical downscaling. The lowest number of common members, 9 in Table 3, is found for  $\omega 500$  in SU2.

From Table 3, the number of the occurrences when the most populated cluster contains more members than a given (arbitrary) threshold could be also found. If such a threshold is set, for example, to 26 members, indicating that the most populated cluster contains more than one half of all ensemble members, the result yields 15 (out of 32) and 14 occurrences for ECEPS and ALEPS respectively. This implies that in nearly one half of all (parameter/synoptic case) combinations from Table 3, the clustering algorithm will sort out all the ensemble members into one major cluster and two (much) less populated clusters. The experience from the operational forecasting practice at CMHS indicates that clustering results of operational ECMWF EPS (based on the ECMWF algorithm) often show inadequate number of clusters for southern Europe. The extreme situation is that all ensemble members are grouped into a single cluster. Our results and the discussion of Table 3 indicate that possibly no major gain in this respect would be attained with ALEPS - the clustering algorithm essentially yields, in a crude statistical sense, similar results for both ECMWF and Aladin ensembles.

The above “skewed” distribution of cluster populations may not be true if a different clustering criterion is applied. When the clustering is based on the proportion of explained variance, then the number of clusters may vary. If for example the explained variance is set to 50%, for most parameters considered the clustering algorithm yields four or more clusters with more “uniform” distribution of cluster populations.

### 3.2.2. *Most populated clusters*

Table 4 shows the statistics similar to that from Table 2 but for ECEPS and ALEPS most populated clusters. Cluster mean deviation with respect to ECMWF operational analysis  $\sigma_c$  is computed in the same way as ensemble mean deviation  $\sigma_e$  discussed in section 3.1, the only difference being that for  $\sigma_c$  the summation runs over the total number of cluster members. The comparison of the cluster mean deviation  $\sigma_c$  in Table 4 (the two middle columns) with the ensemble mean deviation  $\sigma_e$  in Table 2 shows that the number of combinations when the ECEPS deviation is larger than the ALEPS deviation is reduced from 18 in Table 2 to 14 in Table 4. This would essentially imply that, in terms of mean deviation, the relationship between ECEPS and ALEPS most populated clusters does not substantially differ from that for ECEPS and ALEPS full ensembles. However,  $\sigma_c$  values are generally lower than  $\sigma_e$  values irrespective of the model, thus pointing to a reduction in mean variation of the most populated cluster with respect to operational analysis when compared to mean variation of the full ensemble. Bearing in mind that clusters contain members that are much more similar to each other than normally found in an ensemble (i.e. RMS differences among cluster members are smaller than among ensemble members), such a reduced variability of clusters when compared to full ensembles should be expected.

Parameter	Synoptic case	Mean absolute difference $\bar{\delta}_c$		Mean deviation $\sigma_c$		Cluster spread $S_c$	
		ECEPS	ALEPS	ECEPS	ALEPS	ECEPS	ALEPS
<b>Z500 (dam)</b>	SU1	<b>1.04</b>	1.03	<b>1.39</b>	1.35	1.05	<b>1.07</b>
	SU2	0.94	<b>0.98</b>	1.30	<b>1.36</b>	<b>0.99</b>	0.94
	AU1	3.05	<b>3.10</b>	3.89	<b>4.04</b>	<b>2.96</b>	2.54
	AU2	2.06	<b>2.30</b>	2.80	<b>3.32</b>	2.16	<b>2.51</b>
<b>Z700 (dam)</b>	SU1	<b>0.97</b>	0.91	<b>1.27</b>	1.24	<b>0.99</b>	0.90
	SU2	<b>0.88</b>	0.85	<b>1.12</b>	1.11	0.75	0.75
	AU1	<b>2.18</b>	2.15	<b>2.71</b>	2.68	<b>2.01</b>	1.97
	AU2	<b>1.59</b>	1.46	<b>2.04</b>	1.93	1.57	<b>1.59</b>
<b>T850 (°)</b>	SU1	<b>1.07</b>	0.99	<b>1.46</b>	1.28	<b>0.93</b>	0.74
	SU2	1.31	<b>1.37</b>	1.73	<b>1.74</b>	<b>0.85</b>	0.82
	AU1	<b>1.54</b>	1.42	<b>2.07</b>	1.92	<b>1.22</b>	1.18
	AU2	<b>1.66</b>	1.51	<b>2.18</b>	1.98	<b>1.20</b>	1.19
<b>Wind 850 (ms-1)</b>	SU1	<b>2.57</b> (2.16)	2.32	<b>3.34</b> (2.79)	3.01	1.78 (1.68)	<b>1.85</b>
	SU2	<b>2.09</b>	1.99	<b>2.92</b>	2.66	<b>1.87</b>	1.78
	AU1	3.24	<b>3.29</b>	4.37	4.53	2.58	<b>2.93</b>
	AU2	<b>3.98</b>	3.93	<b>5.76</b>	5.58	2.57	<b>2.78</b>
<b>RT 500/1000 (dam)</b>	SU1	1.16	<b>1.24</b>	1.50	<b>1.63</b>	<b>1.07</b>	1.01
	SU2	1.34	<b>1.50</b> (1.34)	1.96	<b>2.19</b> (1.82)	<b>1.26</b>	1.23 (1.14)
	AU1	2.26	<b>3.15</b>	2.99	<b>4.46</b>	2.08	<b>2.73</b>
	AU2	<b>2.19</b>	1.99	<b>3.27</b>	2.99	<b>2.60</b>	2.51
<b>RT 700/1000 (dam)</b>	SU1	0.80	<b>0.84</b>	1.06	<b>1.11</b>	<b>0.75</b>	0.64
	SU2	0.97	<b>1.16</b> (1.19)	1.40	<b>1.49</b> (1.53)	<b>0.86</b>	0.70 (0.75)
	AU1	<b>1.32</b>	1.16	<b>1.76</b>	1.60	<b>1.14</b>	1.08
	AU2	<b>1.98</b>	1.16	<b>2.85</b>	1.63	<b>1.28</b>	1.00
<b><math>\omega</math>500 (Pa)</b>	SU1	0.18	<b>0.23</b>	0.29	<b>0.38</b>	0.12	<b>0.24</b>
	SU2	0.14	<b>0.18</b>	0.22	<b>0.33</b>	0.13	<b>0.24</b>
	AU1	0.33	<b>0.41</b>	0.52	<b>0.66</b>	0.20	<b>0.47</b>
	AU2	0.24	<b>0.25</b>	0.37	<b>0.40</b>	0.13	<b>0.25</b>
<b><math>\omega</math>700 (Pa)</b>	SU1	0.20	<b>0.26</b>	0.31	<b>0.41</b>	0.14	<b>0.26</b>
	SU2	0.16	<b>0.19</b>	0.27	<b>0.33</b>	0.16	<b>0.25</b>
	AU1	0.35	<b>0.42</b>	0.57	<b>0.70</b>	0.25	<b>0.42</b>
	AU2	0.25 (0.21)	<b>0.30</b>	0.39 (0.32)	<b>0.47</b>	0.15 (0.17)	<b>0.31</b>

Table 4 As Table 2 but for ECEPS and ALEPS most populated clusters. Values in parenthesis are for the second cluster when two most populated clusters are equal in size (cf. Table 3)

For cluster mean absolute difference  $\delta_c$  (the first two columns in Table 4) almost the opposite of  $\sigma_c$  is found. When compared with ensemble mean absolute difference  $\sigma_e$  (Table 2), it is clear that  $\delta_c$  values are almost always larger than  $\delta_e$  values. In other words, irrespective of the model the mean distance between cluster members and analysis is larger than the mean distance for the full ensemble. Such a smaller ensemble mean difference is the consequence of the fact that in the full ensemble some members outside the most populated cluster are being closer to operational analysis, thereby reducing the overall ensemble departure from analysis. Thus, when compared with the full ensemble, the clustering will reduce the mean variability among cluster members, but will not guarantee that the mean (cluster) error is also reduced.

In addition to the above two statistical measures, cluster spread  $S_c$  is also included in Table 4 (the last two columns). Similar to ensemble spread  $S_e$  discussed in section 3.1, it is computed as the distance between each individual cluster member and the mean of that cluster, and averaged over all cluster members. From Table 4 it could be seen that in 16 out of 32 combinations the ECEPS cluster spread is larger than the ALEPS spread (compare the  $S_c$  values in the bold typeset in Table 4). A consistently lower spread in ALEPS is found for T850 and RT 700/1000 implying that these fields are potentially more predictable with the Aladin RM. Z500 and RT 500/1000 could be also added to relatively more predictable parameters. For wind at 850 hPa and both  $\omega$  fields the ECEPS spread is lower than that in ALEPS. This clearly indicates that a relatively high variability of the ALEPS high-resolution orography (cf. Fig. 1) causes a stronger dispersion of the above fields. In addition, in almost all combinations the ALEPS spread for  $\omega$  is twice as large as the ECEPS spread. From Table 4, the effect of seasonal cycle is also clearly seen for all parameters considered - a generally larger spread is found in autumn than in summer synoptic cases. The comparison of  $S_c$  with  $S_e$  (see Table 2) indicates a reduced spread in most populated clusters when compared with full ensembles. This is consistent with the reduction of cluster mean deviation  $\sigma_c$  in comparison to ensemble mean deviation  $\sigma_e$ .

### 3.2.3. Clustering statistics

The differences between the global and regional most populated clusters, shown and discussed above, rendered by identical clustering algorithm warrant further analysis. In order to quantify these differences in more details the following additional calculations were carried out: (a) the mean distance  $d_i$  of the  $i$ -th regional cluster members from the  $j$ -th global cluster mean (centroid) and *vice versa*, (b) the distance between the  $i$ -th regional and the  $j$ -th global centroids, and (c) the so-called representative members of regional and global clusters (see Molteni et al. 2001 for the definition).

For the  $i$ -th regional cluster, the mean absolute distance  $d_i$  (averaged over all cluster members) from the  $j$ -th global cluster centroid  $C_j^G$  is computed as

$$d_i = \frac{1}{M_i} \sum_k |F_k^R - C_j^G|$$

The summation index  $k$  runs over all  $F_k^R$  members of the  $i$ -th regional cluster (in total  $M_i$  members of the  $i$ -th cluster). Similarly, the mean distance  $d_j$  of the  $j$ -th global cluster from the  $i$ -th regional cluster centroid  $C_i^R$  can also be calculated. If dynamical downscaling does not affect clustering properties, then for the *same* clustering algorithm the mean distance of the  $i$ -th regional cluster with respect to the  $j$ -th global cluster centroid will be the smallest for  $i = j$ . This essentially implies that a pair of global and regional clusters with the same index will be the closest when compared to other pairs of clusters for which  $i \neq j$ .

For our 32 synoptic case/parameter combinations and the given number of three clusters per population, there are in total 96 pairs of regional and global clusters that fulfil the criterion  $i = j$ . In 50 out of these 96 pairs the mean distance  $d_i$  between the ALEPS clusters and ECEPS centroids  $C_j^G$  is the smallest. In other words, there are almost as many pairs (46) for which dynamical downscaling causes the distance  $d_i$  to be the smallest when  $i \neq j$ . The best results are found for Z700, where in all 12 pairs for which  $i = j$  the distance  $d_i$  is the smallest.  $\omega$ 500 and RT 700/1000 follow with 9 and 8 pairs respectively. For wind at 850 hPa the distance is smallest for all 6 pairs with the same index in the autumn seasons, but for none in summer. Somewhat improved statistics emerges when the distances  $D_i = |C_i^R - C_j^G|$  between ALEPS and ECEPS cluster centroids are computed. In this case, in 52 out of 96 pairs the distance  $D_i$  is the smallest when  $i = j$ . As in the case of  $d_i$ , the distance  $D_i$  is the smallest for Z700 in all 12 pairs with identical indices. To summarise, this kind of statistics indicates that dynamical downscaling may induce non-negligible differences between clusters from global and regional ensembles even if the clustering algorithm was identical. The results also demonstrate that the clustering will depend on parameter and/or season considered.

To support the above findings further, Fig. 4 shows examples of scatter diagrams when all pairs of distances  $d_i$  and  $d_j$  are plotted against each other. For each parameter and season, the distances are normalised by ensemble spread in order to obtain comparable values. For three clusters in each population, there are 9 possible pairs (combinations) of  $i$  and  $j$  indices; hence 9 symbols for each season. If there were no impact of dynamical downscaling on clustering properties, the symbols would be positioned along the diagonal. Thus, it could be assumed that the departures from the diagonal measure the impact of dynamical downscaling on regional clusters. Three “types” of scatter diagrams can be identified from Fig. 4. For Z500 (Fig. 4a; and for both RT 500/1000 and RT 700/1000), the dispersion is relatively large, whereas for Z700 (Fig. 4b; and for T850 and wind 850) it is less so. For the latter, there is a tendency of the clusters to group closer to the diagonal, in particular in autumn, confirming a possible importance of seasonal cycle on clustering. For both  $\omega$  fields ( $\omega$ 700 in Fig. 4c), the symbols are arranged closely along the line that is rotated relative to the diagonal. This “tightness” in  $\omega$  fields could be partly explained by a relatively larger ensemble spread for vertical velocity found in ALEPS than in ECEPS (see Table 2) that has an effect on the normalised distance  $d_i$  of the ALEPS clusters from the ECEPS centroids.

Finally, we briefly compare and discuss cluster representative members for both ALEPS and ECEPS defined by following the procedure proposed by Molteni et al. (2001). For the same cluster index, i.e. when  $i = j$ , only in 27 out of total 96 ALEPS/ECEPS pairs (less than 30%) the representative members in both populations are found to be identical. The largest proportion, 9 pairs (out of 12) of all identical members are found for Z700, whereas the least number of identical members (1) was found for Z500 and both  $\omega$  fields - consistent with the cluster distance statistics presented above. The distance  $r_i = |R_i^R - R_j^G|$  between the ALEPS representative member  $R_i^R$  and the ECEPS representative member  $R_j^G$  for identical clusters, i.e. when  $i = j$ , essentially yields the results similar to those for  $d_i$  and  $D_i$  discussed above -  $r_i$  was found to be the smallest in 46 out of 96 combinations. For Z700, this is (again) true for all possible 12 combinations, whereas for Z500 and  $\omega$ 700 this occurs in 2 and 3 cases respectively.

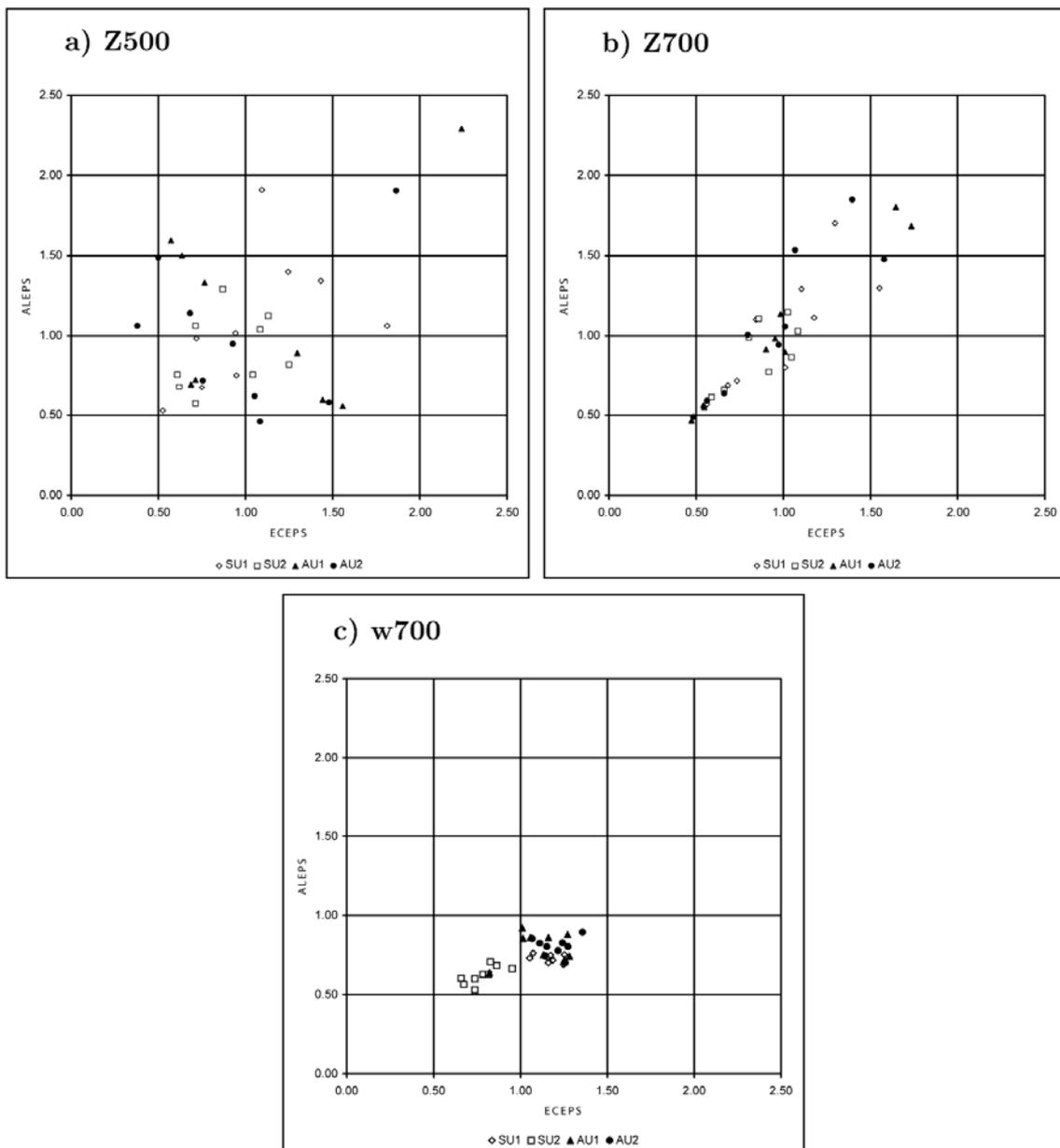


Figure 4 Normalised mean distance  $d_i$  of the  $i$ -th ALEPS cluster from the  $j$ -th ECEPS cluster centroid  $C_j^G$  for a) Z500, b) Z700 and c)  $\omega$ 700.

The results and discussions in this sub-section indicate that different outcomes are likely when the same clustering algorithm is applied to ECEPS and ALEPS. Based on such a consideration, an important implication for the dynamical downscaling could be inferred. It might not always be feasible to make a selection (or a subset) of the global model members for dynamical downscaling that would be based on the global model clustering (although from the computational point of view such a reduction in the number of forecasts for downscaling would be desirable). In other words, similarity (or likeness) found among the members of a given global model cluster might not be necessarily seen or carried over to the subsequent downscaled forecasts.

## 4. Ensemble analysis of synoptic cases

As mentioned above, three cases out of four downscaled ensembles are studied in more detail and their results are compared with ECEPS and against Croatian observation data. In terms of severe weather, the fourth synoptic case (AU2) has not affected Croatia and will not be discussed here. As mentioned earlier, some other parts of central Europe experienced gale force winds and excessive precipitation.

### 4.1. Synoptic cases

#### 4.1.1. Summer case 1 (SUI): 4-5 July 2003

The first synoptic case covers the period between 4 and 5 July. For several days the eastern Adriatic coast was exposed to south-westerly upper-air wind that was bringing a relatively moist and unstable air from the western Mediterranean. At 200 hPa, the jet extended from Italy into the Adriatic and the Balkan Peninsula, reaching according to ECMWF analysis  $42 \text{ ms}^{-1}$  over the central eastern Adriatic (Dalmatian) hinterland at the mid-day on 4 July. In the early hours of 4 July, the two shallow lows, one in the northern and the other in the central Adriatic, were detected by ECMWF operational analysis (Fig. 5a). The pressure in both lows was below 1011 hPa and the south-easterly surface wind was blowing along the Dalmatian coast. In most places 2m temperatures were in excess of  $25^\circ\text{C}$  from early morning hours. These conditions favoured the development of unstable mesoscale convective systems with thunderstorms.

The 24-hour accumulated precipitation from the Dalmatian rain gauges, for the 24-hour period between 06UTC 4 July to 06UTC 5 July, indicates the rain totals between 50 and 56 mm (Fig. 5b; only 24-hr accumulations larger than 20 mm are shown). The hourly reports from the main climatological stations reveal that in fact there were two periods of increased precipitation. In the first period on 4 July, the Dalmatian towns Šibenik and Split (marked by open black circle and black square respectively in Fig. 5b) were hit by heavy rain between 8 and 10 a.m. local time. In Split more than 30 mm fell in 40 minutes. 12 hours later a single shallow low of less than 1008 hPa moved into the southern Adriatic with northerly winds blowing along much of the Dalmatian coast. In the second period, on 5 July between the midnight and 6 a.m. local time, some towns and islands in central Dalmatia were affected by rain of nearly 20 mm in 1 hour.

Both operational deterministic models predicted some precipitation for central Dalmatia. The ECMWF TL511 model (with horizontal resolution of approximately 40 km) predicted between 1 and 5 mm in 12 hr, and Aladin between 1 and 10 mm, however shifted farther inland from the place where maximum precipitation actually occurred (not shown). For this synoptic case, the initial date was chosen to be 2 July 2003 at 00UTC, and the target forecast times were T+60 and T+72 hours, i.e. between 12UTC 4 July and 00UTC 5 July (cf. Table 1).

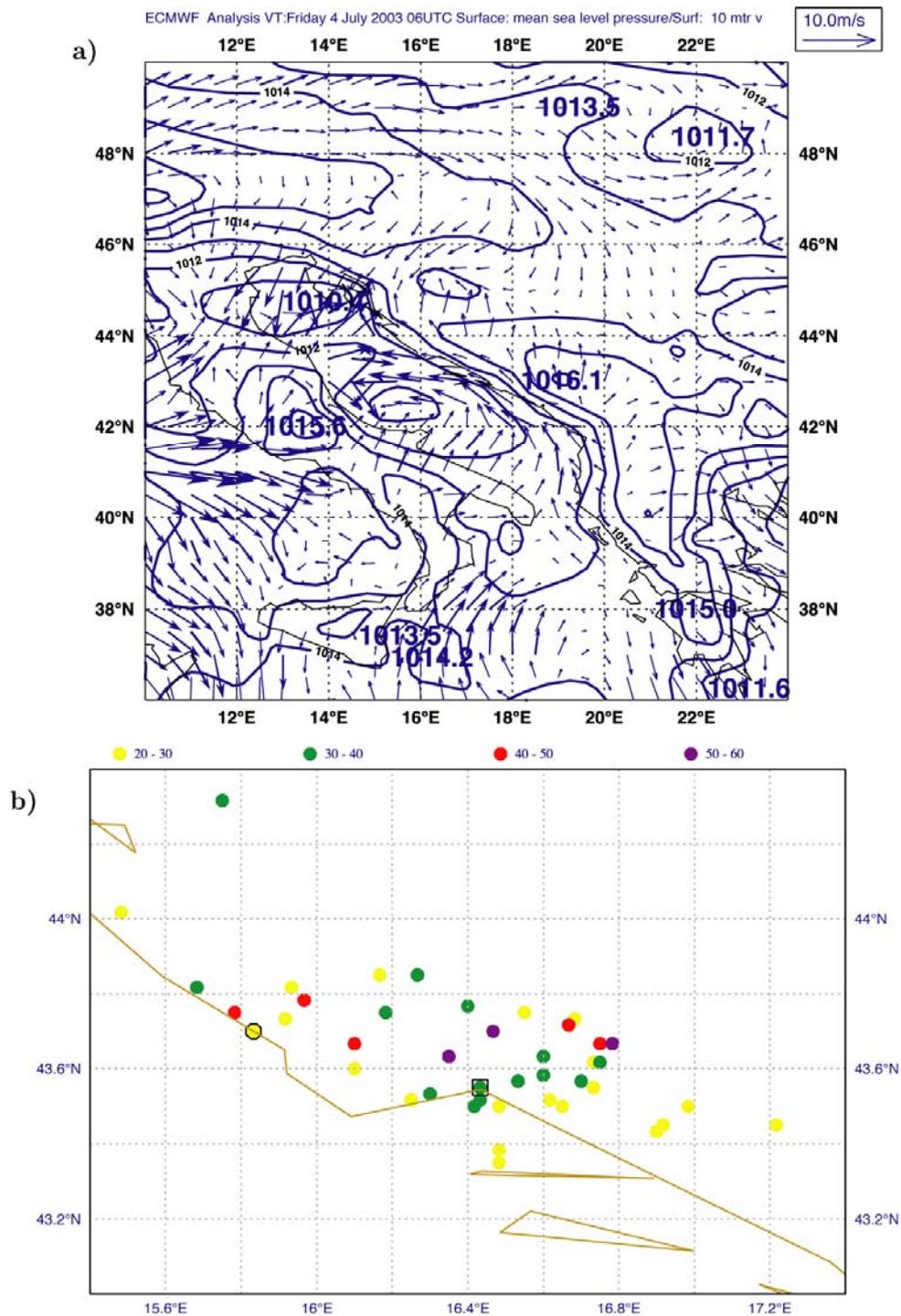


Figure 5 Case SU1: a) ECMWF operational analysis for MSLP and 10 m wind at 06UTC on 4 July 2003, and b) 24-hour accumulated precipitation from the Dalmatian rain gauges between 06UTC 4 July 2003 and 06UTC 5 July 2003. Contours in a) every 1 hPa. In b) only rain gauges with more than 20 mm/24 hr are shown.

#### 4.1.2. Summer case 2 (SU2): 28-29 July 2003

This typical summer storm case occurred over the continental Croatia. A strong thermal ridge in the lower troposphere extended from north Africa into the northern Mediterranean, and a relatively cooler air was covering an area that stretched from the north-western Europe to the north of the Alps. A weak diffluent flow

formed in the mid-troposphere over the western Balkan Peninsula (Fig. 6a), and moderate MSLP and temperature gradients formed between the eastern side of the Alps and the Hungarian Plain (Fig. 6b). The prevailing northerly near-surface winds caused some “spillage” of cold, unstable air around the Alps into the northern Croatia. During the night 28/29 July, between 20UTC and 4UTC many stations in the northwest Croatia reported heavy precipitation with thunder and strong winds (Fig. 6c). In some places precipitation prolonged over several hours, in the others it was a relatively short but intense event. In Fig. 6c Zagreb is marked by a black diamond and the rain gauge at Kapela has measured 60.9 mm in 24 hours (shown as an open square in Fig. 6c).

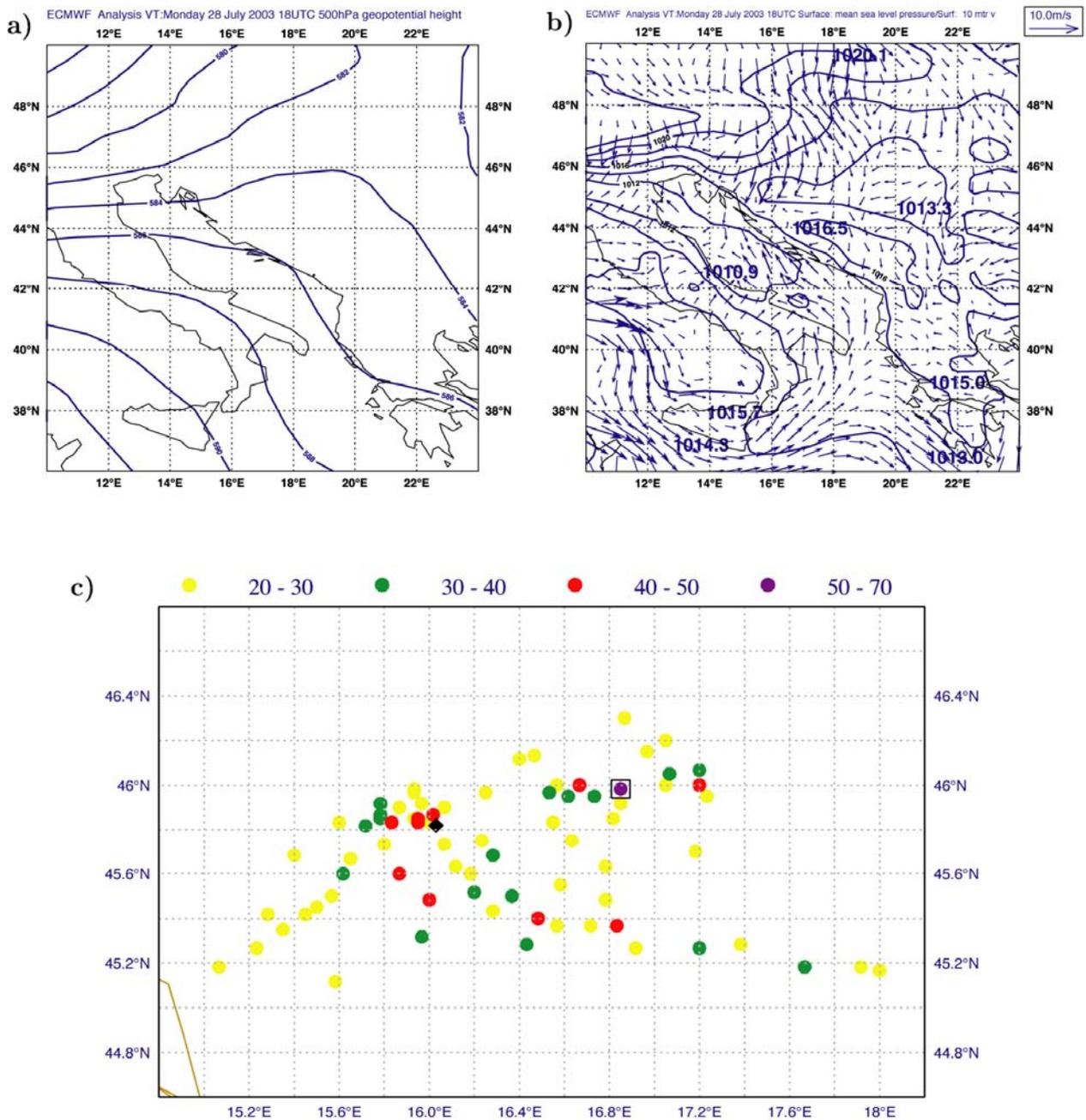


Figure 6 Case SU2: ECMWF operational analysis at 18UTC on 28 July 2003 for a) 500 hPa geopotential height, b) MSLP and 10 m wind, and c) 24-hour accumulated precipitation from the north-western Croatia rain gauges between 06UTC 28 July 2003 and 06UTC 29 July 2003. Contours in a) every 2 dam, in b) every 2 hPa. In c) only rain gauges with more than 20 mm/24 hr are shown.

For the same time interval, the ECMWF operational deterministic model generated between 5 and 10 mm of rain (not shown). Operational deterministic forecasts Aladin predicted more than 10 mm, however slightly displaced to the south of the region of interest. For this case, the initial date for numerical simulations was chosen to be 00UTC 26 July 2003, and the target forecast period is between T+66 and T+78, i.e. between 18UTC 28 July and 06UTC 29 July.

#### 4.1.3. Autumn case 1 (AUI): 13-15 November 2004

This was a synoptic situation that occasionally occurs in the eastern Adriatic and its hinterland during the cold period of the year. In the period 13-15 November a deep low was moving from north Africa, across south Italy into the southern Adriatic, where eventually it became semi-stationary and slowly dissipated over the Ionian Sea and the southern Balkans. The gale force winds and heavy precipitation along the Croatian coast and islands caused loss of life, considerable damage to agriculture and havoc to the air, road and sea ferry traffic. Initially, a strong northerly bora wind (*bura* in Croatian) was blowing along the northern Adriatic coast, and a strong south-easterly wind (*jugo*) was pounding the coast of the southern Adriatic (Fig. 7a,b). In the morning of 14 November, the maximum daily wind gust reached  $33.5 \text{ ms}^{-1}$  in the northern Adriatic town of Senj. Later during the day the automatic weather station at the Island of Krk Bridge recorded a  $206 \text{ kmh}^{-1}$  wind gust (see Figure 4 at [http://klima.hr/priopcenja/bura\\_2004\\_e.html](http://klima.hr/priopcenja/bura_2004_e.html), with the maximum wind speed of  $57 \text{ ms}^{-1}$ ). During the same morning, *bura* in Split peaked at  $36.1 \text{ ms}^{-1}$ . In the southern Adriatic, at the Dubrovnik climatological station, *jugo* was logged at  $30.0 \text{ ms}^{-1}$  in the early morning of 13 November, and in the afternoon of the following day (14 November) it veered to northerly *bura* reaching  $29.6 \text{ ms}^{-1}$ . By 00UTC 15 November *bura* extended and was blowing fiercely along the whole eastern Adriatic coast.

Heavy precipitation was recorded along the southern Adriatic coastal region between 06UTC 13 November to 06UTC 14 November (Fig. 7c). In Split (marked by a square in Fig. 7c), 31.7 mm of rain fell in 24 hours - between the afternoons of 13 November and 14 November. In Ploče (triangle in Fig. 7c) and Dubrovnik, 45.5 and 41.1 mm of rain respectively fell in 6 hours, between 07UTC and 13UTC on 14 November. In 24 hours (18UTC 13 November to 18UTC 14 November) it ultimately amounted to 94.2 and 76.9 mm respectively. The maximum precipitation of 123.4 mm in 24 hours (6UTC 13 November to 6UTC 14 November) has been measured in Gornje Sitno (open diamond in Fig. 7c), and in the same period 87.8 mm was recorded in Blato on the island of Korčula (hexagon in Fig. 7c).

For this synoptic case the operational ECMWF TL511 deterministic model predicted precipitation amounts higher than 50 mm/12 hr only further south, at the Montenegrin and Albanian coastal mountains, and no precipitation in the northern Adriatic. Similar result has been obtained by operational (deterministic) Aladin model (not shown). The initial time for model integrations was set to 00UTC 12 November 2004, and the target period is between T+36 and T+72, depending on the area considered.

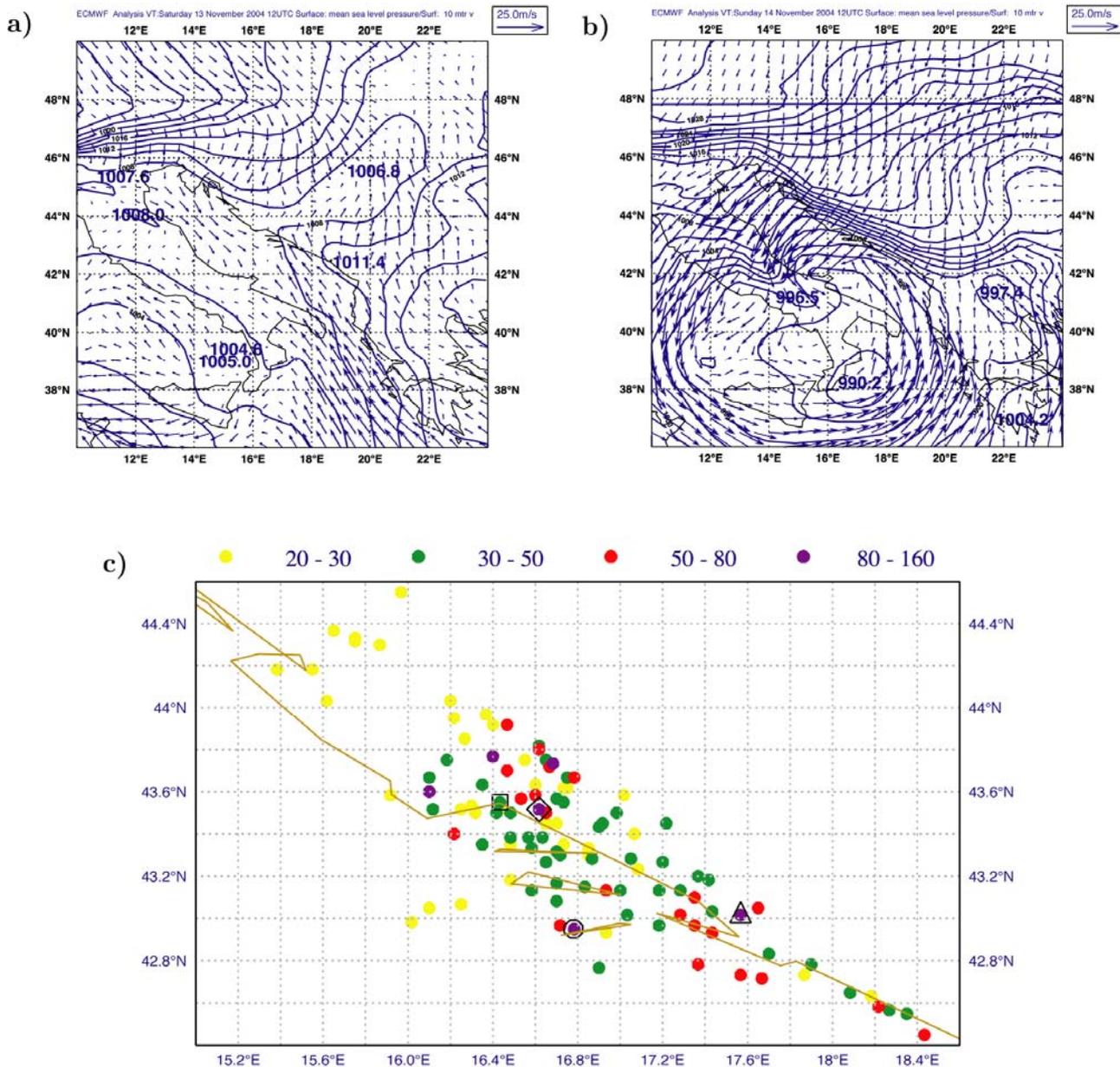


Figure 7 Case AUI: ECMWF operational analysis for MSLP and 10 m wind at a) 12UTC on 13 November 2004, b) 12UTC 14 November 2004, and c) 24-hour accumulated precipitation from the south Croatia rain gauges between 06UTC 13 November 2004 and 06UTC 14 November 2004. Contours in a) and b) every 2 hPa. In c) only rain gauges with more than 20 mm/24 hr are shown.

## 4.2. Clustering proxy for precipitation

The clustering described and discussed in section 3 was performed for upper-air parameters. In view of severe weather in our synoptic cases, it might be desirable to carry out clustering for those parameters that best describe the event itself. Perhaps intuitively one would prefer to base clustering on surface parameters, like for example, precipitation amount or the magnitude of the near-surface wind. However, because of the discrete nature and incoherent structure of many surface fields the clustering algorithm based on, for example, RMS difference between ensemble members may not be a straightforward exercise. In addition, it might be difficult to reconcile the small-scale nature of (often orographic) precipitation with large-scale features that are normally included in the clustering algorithm over the domain of the size of central Europe

and the northern Mediterranean. On the other hand, if upper-air fields are used as the basis for the clustering of surface fields, the question arises which parameter (if any) could be the best-defined proxy for a given surface field? Since no unambiguous relationship between surface and upper-air fields exists, various testing has been performed and discussed below.

Before discussing the results, it is important to note that the clustering time for precipitation proxies is based on the hindsight knowledge of each synoptic case, i.e. we already knew the time intervals when severe weather events (heavy precipitation, gale force winds) occurred. This, of course, would not be possible in an operational forecasting environment. However, our aim is to find out the best possible estimates for the clustering of parameters that constitute severe weather.

#### 4.2.1. Case *SUI*

For the *SUI* case, the most intense precipitation occurred within the forecasting period between T+54 and T+66 hours (cf. Fig. 5). Thus, the proxy clustering for the 12-hr accumulated precipitation is centred on the T+60 forecast time, i.e. at 12UTC 4 July 2003. For the ECEPS clustering based on different upper-air parameters, the most populated clusters yield too little 12-hour precipitation in the area of the central eastern Adriatic (Fig. 8 left panels). Here, for reason of space only selected parameters are shown. Somewhat increased precipitation amounts are found farther north, in the north Adriatic region, indicating a displacement of the precipitation maximum in the ECMWF EPS model when compared with observations. For the central eastern Adriatic almost no difference is found among various parameters. The next 12-hour accumulations (between T+60 and T+72) show no improvement, i.e. the precipitation maxima stay more or less in the same position.

For ALEPS, the same clustering procedure yields an increased precipitation rate relative to ECEPS. When Fig. 8 right-hand panels are compared with Fig. 5b, it could be concluded that such an increase in the ALEPS precipitation means an improvement as well. For RT 700/1000 (Fig. 8 bottom right), the 30-mm contour is seen in the box bounded by 43°–44°N and by 17°–18°E. This maximum is only slightly shifted eastward relative to observations shown in Fig. 5b. The clustering based on the other parameters also indicates increased precipitation in the same area but with somewhat reduced amplitude. For some parameters the maximum precipitation is found further north (e.g.  $\omega 500$ ), similar to ECEPS. In comparison with ECEPS, the precipitation in ALEPS is generally more dispersed over the central part of the integration domain.

It could be argued that the increased precipitation in the southern Dalmatian hinterland in ALEPS, and not seen in ECEPS, is due to the impact of a better-resolved Aladin orography. From Fig. 1, it is clear that, in the area considered, the orography difference between the Aladin and the ECMWF EPS models is more than 600 m in places. For this particular synoptic case, in addition to higher mountains in ALEPS, the orientation of the Dinaric Alps (the main mountain chain in the eastern Adriatic hinterland) must have also played a crucial role to increasing precipitation. As discussed above (section 4.1), in early July 2003 the eastern Adriatic coast was exposed to a strong and moist south-westerly flow. When hitting the orographic obstacle that runs from the northwest to southeast, i.e. perpendicular to the direction of the flow, such a flow is bound to generate some precipitation. Similar increase in the ALEPS precipitation is seen over central Italy where the Apennines represented in Aladin are higher for more than 1000 m than in ECMWF EPS and have identical orientation as the Dinaric Alps (cf. Fig. 1).

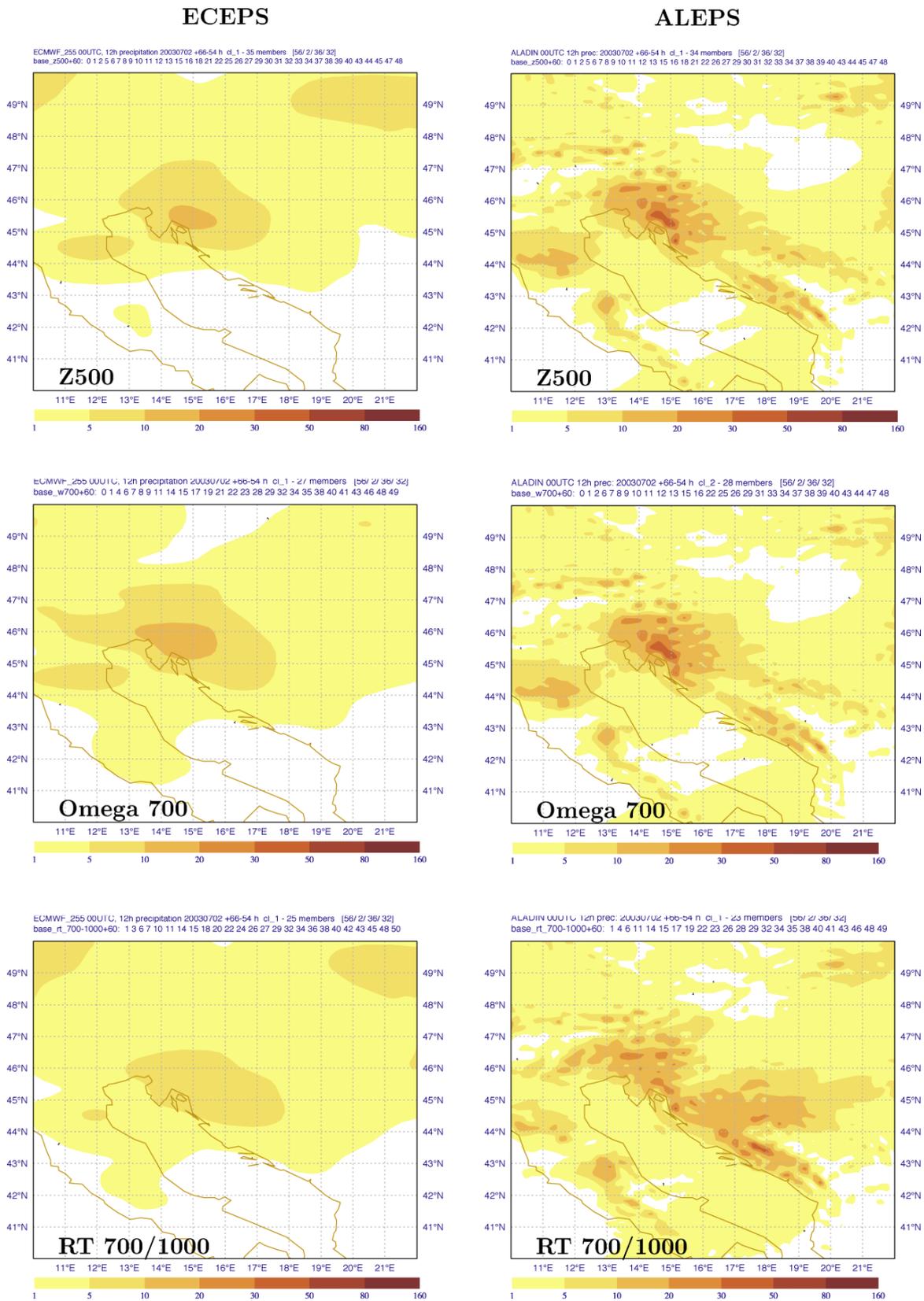


Figure 8 The 12-hour accumulated precipitation between  $T+54$  and  $T+66$  in most populated clusters in the SU1 case and various clustering base parameters for ECEPS (left) and ALEPS (right). Contouring 1, 5, 10, 20, ... mm/12 hr.

It is important to emphasise that the results shown in Fig. 8 and discussed above are related to the most populated clusters. Fig. 9 shows that, for the same clustering base parameters some other, less populated clusters have attained even better results. For example, for ECEPS, the clustering based on RT 1000/500 has brought the 5 mm contour to the region of central Dalmatia (Split) – an increase when compared with Fig. 8 left panels. Likewise, for ALEPS, the clustering based on  $\omega 500$  indicates rainfall of more than 50 mm in 12 hours. Of course, such results might be relevant only in the hindsight analysis. In the real-time forecasting practice less populated clusters would only serve as an indication of a possible alternative development, i.e. they would quantify forecast uncertainty. In the case of ALEPS, the cluster number 2 in Fig. 9b contained 10 members, whereas the most populated cluster contained 32 members. Thus, the development that was closest to the observed precipitation rates was predicted by about 20% of ensemble members, i.e. by a non-negligible margin.

From the above discussion a question arises what upper-air parameter in this synoptic case has served as the best clustering proxy for precipitation. According to Fig. 8 no clear-cut answer could be reached. Some parameters yielded a little more accurate precipitation amount than the others, but overall no dramatic differences among various parameters are found. On the other hand, from Fig. 9b  $\omega 500$  clearly represents a relatively good clustering base for describing well the precipitation maxima.

#### 4.2.2. Case SU2

Generally, in the most populated ECEPS clusters there is much less precipitation over the northwest Croatia than it was actually observed. Fig. 10a shows that for the clustering based on Z500 only between 1 and 5 mm of precipitation in 12 hours was generated in the box bounded by 45°–46°N and 16°–17°E (cf. Fig. 6c). Similar result also holds for  $\omega 500$ , but for other parameters very little or almost no precipitation has accumulated between T+66 and T+78. Interestingly, in the same time interval the then operational ECMWF T<sub>L511</sub> (deterministic) model generated between 5 and 10 mm of rain (not shown), indicating that an improved horizontal resolution might have helped in better representing relatively small spatial synoptic scales. In almost all clusters, a somewhat increased precipitation (5-10 mm) occurred to the northwest, in the neighbouring Slovenia. The ALEPS most populated clusters have not performed better than their ECEPS counterparts. Fig. 10b shows that in the box considered no more than 5 mm of precipitation accumulated over 12 hours for the clustering based on RT 500/1000.

In both ensembles, an increased precipitation is located in the northeast Italy, south Austria and throughout Slovenia. Indeed, heavy showers and thunder have been observed in these areas (not shown) and in both ensembles the advance of a relatively cold air from the northwest has been halted there. Thus, as far as the north-western part of Croatia was concerned, for both ensemble systems this seems to be a difficult case to predict correctly. The additional clustering for other base times and parameters was performed and various forecast time intervals were analysed. However, this brought no improvement to either ensemble. For this particular case both operational deterministic models predicted more accurately the precipitation amounts than their ensemble counterparts (see the discussion in section 4.1.2).

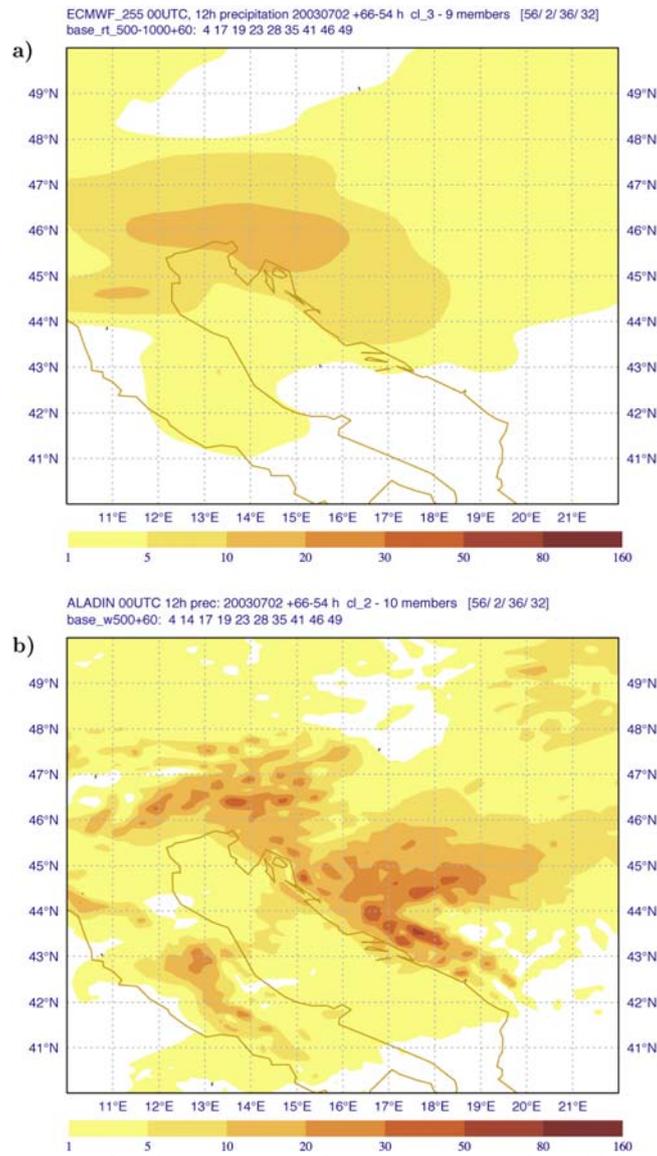


Figure 9 The 12-hour accumulated precipitation between  $T+54$  and  $T+66$  hr in the SU1 case for a) ECEPS cluster no. 3 for the RT 500/1000 clustering base, and for b) ALEPS cluster no. 2 for the  $\omega 500$  clustering base. Contouring is the same as in Fig. 8.

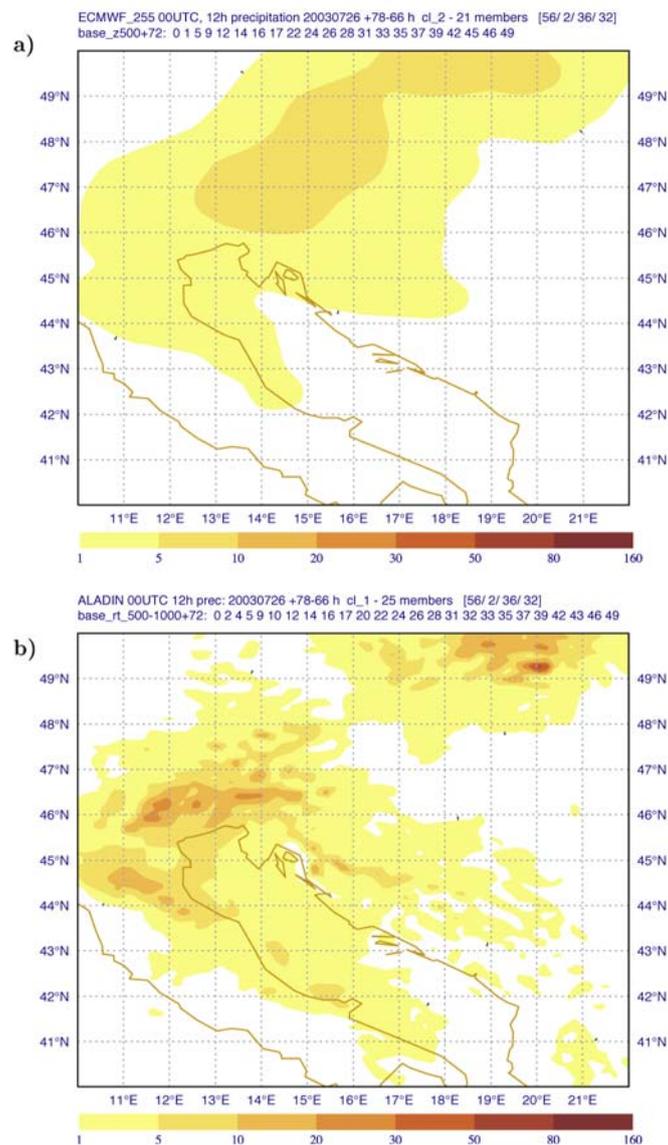


Figure 10 The 12-hour accumulated precipitation between  $T+66$  and  $T+78$  in most populated clusters of the SU2 synoptic case for a) ECEPS for the Z500 clustering base and b) ALEPS for the RT 500/1000 clustering base. Contouring as in Fig. 8.

#### 4.2.3. Case AUI

The bulk of precipitation in the southern Adriatic fell between the evening of 13 November and the early afternoon hours on 14 November. For this case, the clustering proxies for the 12-hour accumulated precipitation are centred on 00UTC 14 November, i.e. at the  $T+48$  forecast time. All of most populated ECEPS clusters based on different upper-air parameters, generate in the south-eastern Adriatic more than 20 mm of precipitation per 12 hours, and in the Dubrovnik area more than 30 mm (see Fig. 11a for the Z500 clustering base). Less populated clusters also yielded similar amounts. Thus, ECEPS was relatively successful in this synoptic situation and a high degree of consensus among individual members was attained. Moreover, a gradual reduction in precipitation amounts towards the northern Adriatic was also correctly predicted: for example, in the northern Adriatic town of Senj (marked as SE in Fig.11) the 12-hour accumulation centred at 00UTC 14 November was only 7 mm.

The ALEPS most populated clusters indicate correctly more than 30 mm/12hr precipitation in the Ploče area and more than 50 mm/12hr in the Dubrovnik area respectively (marked as PL and DU in Fig. 11; cf. section 4.1.3). This is clearly seen in Fig. 11b for the  $\omega 500$  clustering base. The largest precipitation amounts, similar to ECEPS, are found further south down the eastern Adriatic coast. This maximum in precipitation is induced when a strong and moist southerly wind enters the southern Adriatic (Fig. 7b) and impinges the highest mountains of the Balkan Peninsula (more than 1800 m in the Aladin model, cf. Fig. 1a). It could be concluded that the ALEPS most populated clusters improved the detailed distribution of high precipitation amounts in the southern Adriatic relative to ECEPS, although the performance of the latter was very good indeed. When compared to observations, both ECEPS and ALEPS performed better than ECMWF and Aladin deterministic models (see the discussion in section 4.1.3 above).

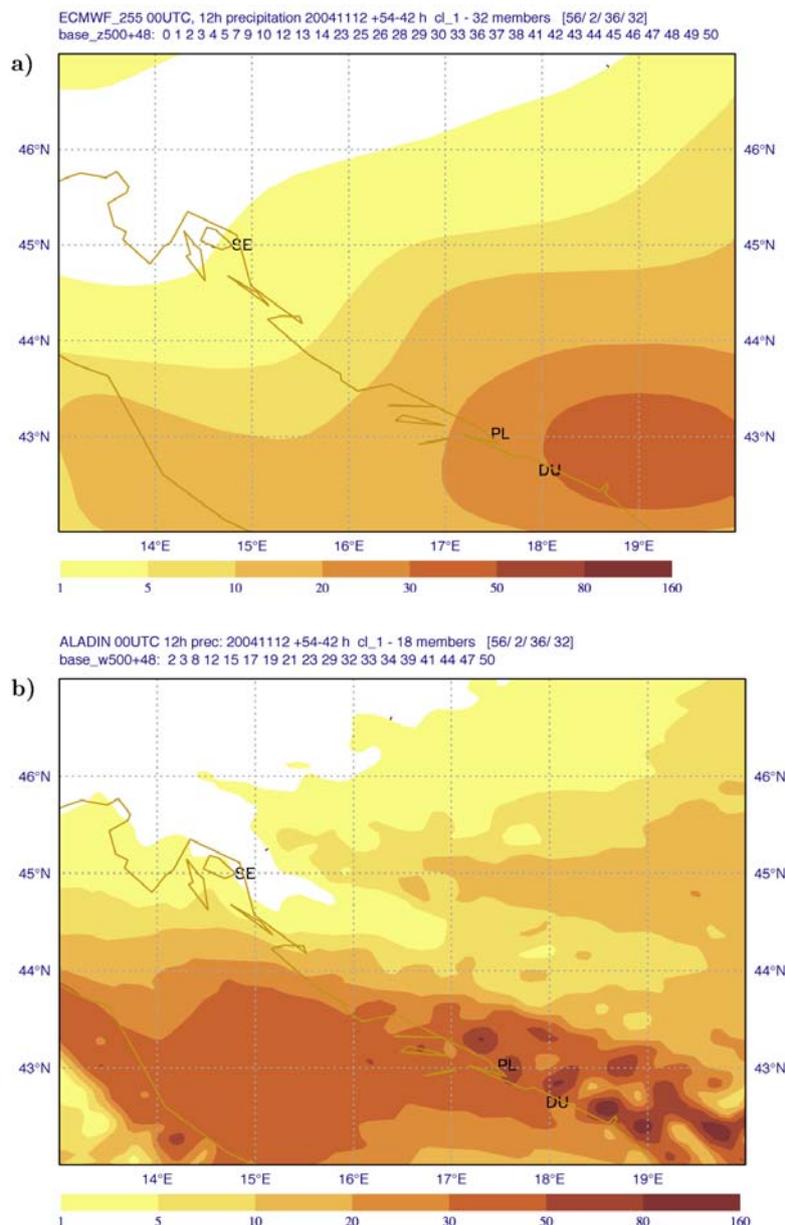


Figure 11 The 12-hour accumulated precipitation between  $T+42$  and  $T+54$  in most populated clusters of the AU1 synoptic case for a) ECEPS for the Z500 clustering base and b) ALEPS for the  $\omega 500$  clustering base. Contouring as in Fig. 8.

For this synoptic case, there is almost no difference among various parameters used as the clustering base for precipitation. This is essentially in agreement with the results for precipitation proxies discussed above for the SU1 case. The fact that no major differences in precipitation amounts have been obtained for the clustering based on various parameters could be useful in an operational forecasting practice, since it would eliminate redundant results.

### 4.3. Probabilities

To complement the results of clustering proxies for precipitation, in this section precipitation probabilities for both ECEPS and ALEPS are analysed. For precipitation, we focus to 12-hour accumulations. In general, the 12-hour (and also 24-hour) accumulations were found to yield higher probabilities than short-period accumulations. A possible explanation could be that (global) models are unable to accurately pinpoint severe weather events - they might be displaced either in space or/and in time. Thus, accumulations over a longer period of time would compensate for model spatial and temporal inaccuracies. In addition to precipitation, probabilities for near-surface wind are also considered, as well as one (non-standard) application of forecast probabilities to the diagnostics based on isentropic analysis developed at CMHS. Most results on probabilities shown and discussed below are defined from the whole ensemble; however, we also briefly discuss precipitation probabilities from clusters that were based on precipitation proxies. In the following, probabilities for only two synoptic cases are discussed because, as shown by cluster analysis above, for the SU2 case both models performed relatively poorly.

#### 4.3.1. Case SU1

For 10-mm and 15-mm thresholds, the 12-hour precipitation probabilities in ALEPS (Fig. 12 middle panels) are generally higher than probabilities in ECEPS (Fig. 12 top panels). The highest probabilities in ALEPS, located over the mountains of the northern Adriatic and reaching more than 95%, coincide with highest probabilities of ECEPS which extend only a little above 35%. Though one could argue that both models misplaced the largest precipitation amounts (cf. section 4.1 (i)), this is not entirely true. Namely, in the region of the northern Adriatic an increased precipitation has also been observed, however, not as much as farther south in central Dalmatia (not shown). On the other hand, though the positioning of the 65% contour in ALEPS over the mountains of the southern Adriatic hinterland does not coincide precisely with the region of the heaviest precipitation (cf. Fig. 5b), it is nevertheless quite close to where it actually occurred.

This case demonstrates the modelling difficulty to capture the details of a heavy precipitation event that was rather confined in time and space. In terms of probabilities, the emphasis in the global model ensemble was on a less important precipitation event – or from pessimistic point of view, one could argue that the global model missed out the event of interest. On the other hand, though the limited area model gave some indications of the event close to the region where it actually occurred, it has not distinguished it with sufficient details.

Precipitation probabilities discussed above were derived from the whole Aladin ensemble. In section 4.2 it has been demonstrated that some upper-air parameters could be used as the basis for the clustering of precipitation. In the bottom panels of Fig. 12, the 12-hour precipitation probabilities from ALEPS most populated clusters for the two parameters chosen as the precipitation proxies (cf. Fig. 8 right) are shown. In both Fig. 12 bottom panels, an increased probability of precipitation greater than the 15-mm threshold is seen

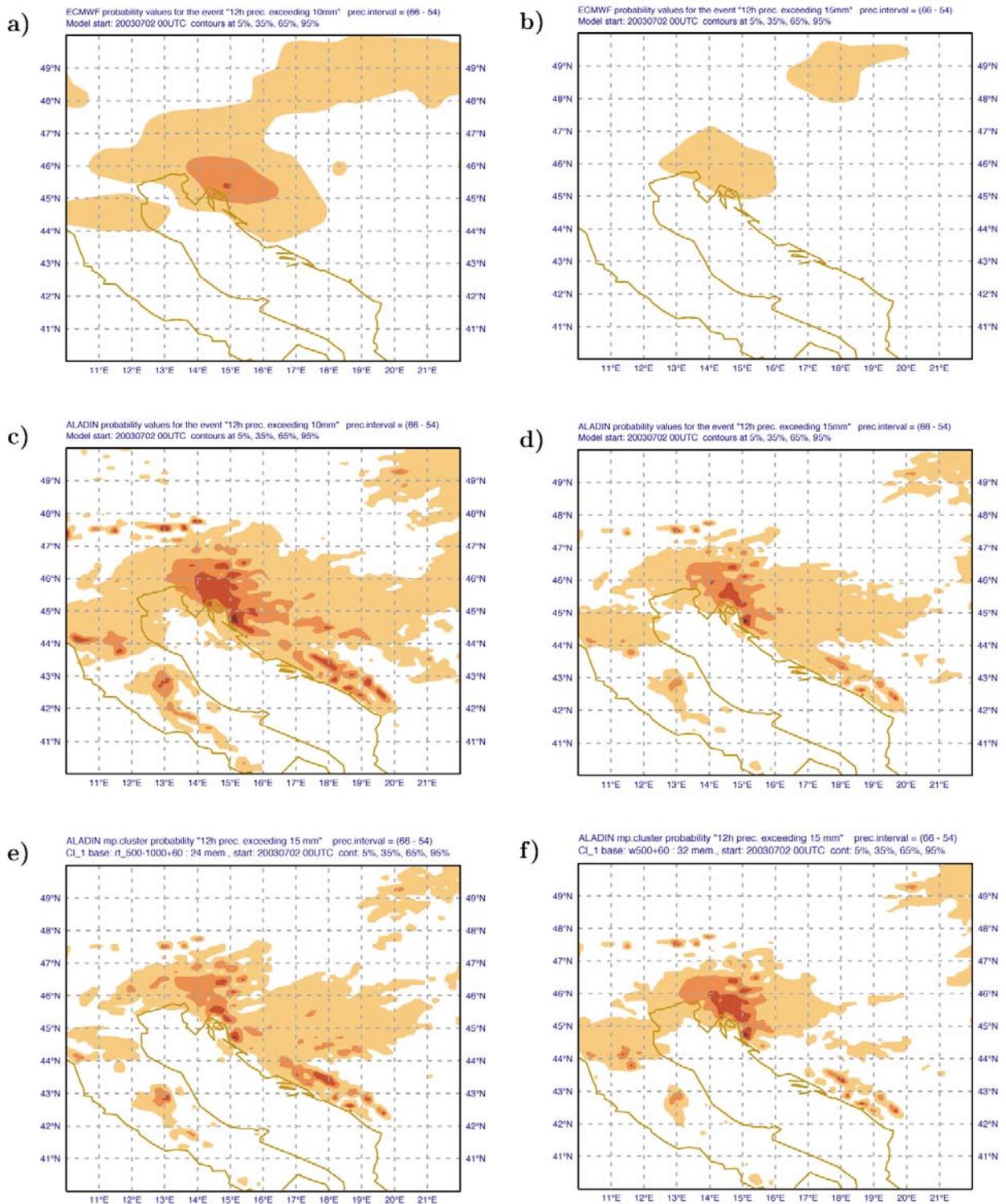


Figure 12 Probability of the 12-hour accumulated precipitation (between T+54 and T+66) in the SU1 case for ECEPS (top) and ALEPS (middle) for the thresholds 10 mm/12 hr (left) and 15 mm/12 hr (right). Same probability for ALEPS and the 15-mm threshold for most populated clusters based on e) RT 500/1000, and f)  $\omega$ 500. Contouring at 5, 35, 65 and 95%.

in the southern Adriatic hinterland, similar to Fig. 12 middle panels. Whilst a correct location of the largest observed precipitation amounts in Fig. 12e,f still remains somewhat elusive, i.e. it is similar to that for the full ensemble (Fig. 13c,d), it nevertheless confirms that both parameters – in this case RT 500/1000 and  $\omega_{500}$  – could be used as precipitation proxies. For ECEPS with 15-mm threshold, no parameter acting as a proxy for precipitation in the most populated clusters has produced any significant precipitation probability in the region of interest (not shown).

Probabilities of meteorological parameters shown and discussed above have been computed directly from model output. However, other applications based on forecast probabilities might prove as being useful prediction tools. In the CMHS forecasting practice the High Resolution Isentropic Diagnosis (HRID) model has been used for a long time to primarily diagnose whether the atmosphere over predefined localities is potentially unstable (see, for example, Ivančan-Picek et al. 2003). For a given location, various surface and upper-air parameters are computed from the Aladin model output and could be displayed at regular time intervals in the form of the height/time cross sections. Here, the capability of HRID has been extended to calculate probabilities of the occurrence for a given event and we demonstrate its potential in a forecast ensemble.

The SU1 storm that occurred in the early hours of 4 July was associated with an increased upper-air wind (see the discussion in the section 4.1 (i)). In the period from 3 July to 5 July, the operational (deterministic) HRID for Split (43.51°N, 16.55°E) has predicted an increase of the upper-air wind above 6 km to the strength of the jet stream (Fig. 13a). The wind peaked at more than 44  $\text{ms}^{-1}$  in the slab between 9 and 10 km at 00UTC 4 July. The ALEPS indicates that probability of the jet stream occurrence (i.e. the wind being stronger than 30  $\text{ms}^{-1}$ ) between 6 and 7 km is about or slightly above 50%, and more than 90% between 8.5 and 10 km (Fig. 13b). This is essentially in agreement with the ECMWF operational analysis for 12UTC 4 July which indicates that the south-westerly wind was gradually increasing from less than 30  $\text{ms}^{-1}$  at 400 hPa to nearly 40  $\text{ms}^{-1}$  at 200 hPa (not shown). Thus, it could be argued that the probabilistic HRID is closer to verifying analysis, whereas its deterministic counterpart overestimates the wind magnitude in the upper portion of the time cross-section in Fig. 13.

The bottom panels of Fig. 13 show the distribution and probabilities of relative humidity in both deterministic and probabilistic HRID. Clearly, in both bottom panels an increase in relative humidity could be seen, at and around the time of heavy precipitation. When compared with the hourly precipitation data (cf. section 4.1 (i)), a higher probability of relative humidity greater than 60% shortly before 12UTC 4 July (more than 80%) is fairly close to the precipitation observations. It has to be pointed out, however, that both deterministic and probabilistic diagnostics have not captured entirely correctly the actual precipitation rate at the given location.



high mountains prevail (see Fig. 1). It might be possible that such an extensive precipitation at a high orographic obstacle in ALEPS is linked to some deficiency in model's parameterisations. In ECEPS (Fig. 14c), on the other hand, a higher probability of precipitation over the western Balkan Peninsula indicates that moist air in the southerly flow penetrated further inland than in ALEPS.

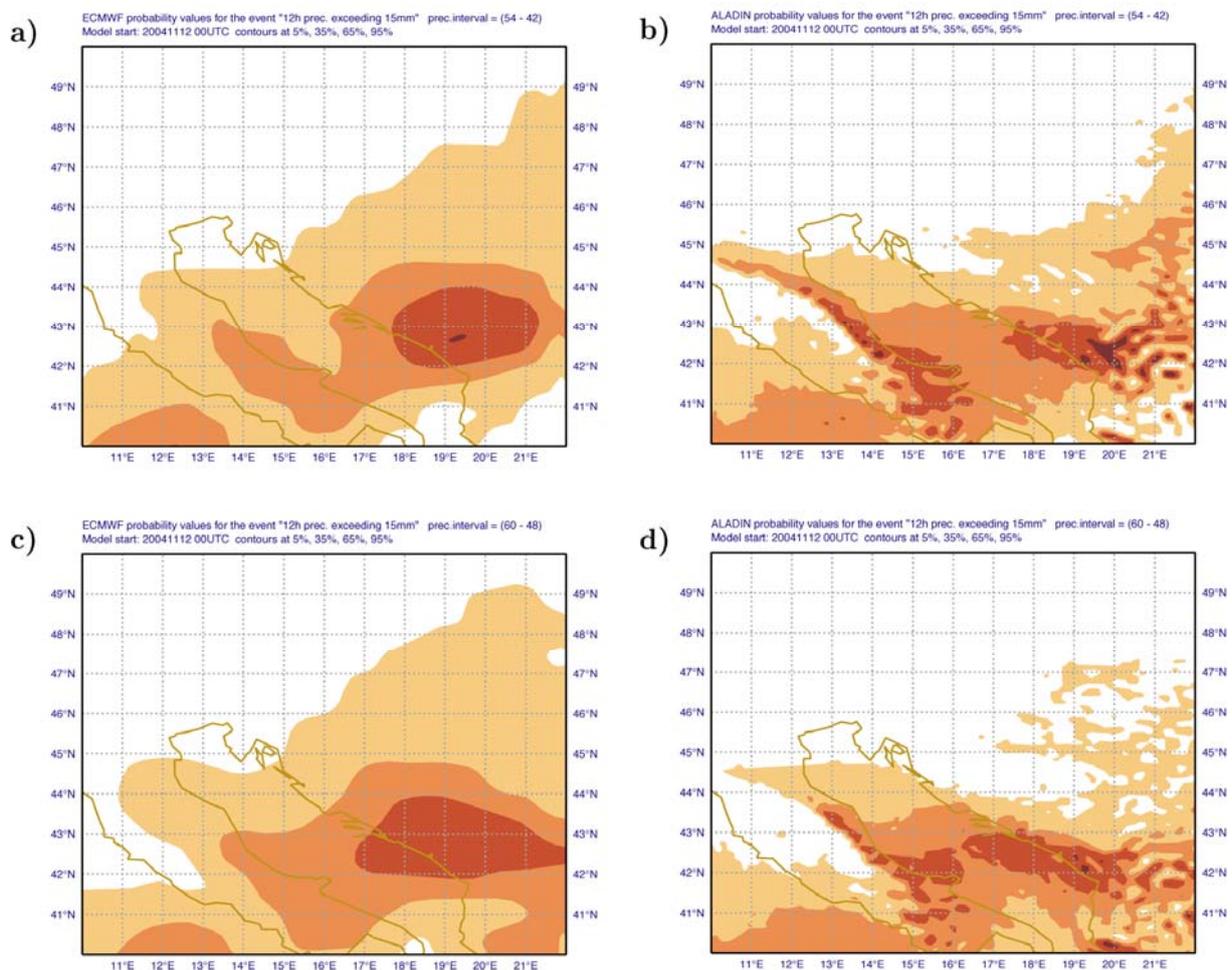


Figure 14 Probability for precipitation exceeding the threshold 15 mm/12 hr accumulated between T+42 and T+54 (top) and between T+48 and T+60 (bottom) in the AUI synoptic case for ECEPS (left) and ALEPS (right). Contouring at 5, 35, 65 and 95%.

In section 4.1.3 it has been discussed that, in addition to heavy precipitation, strong winds at the Croatian Adriatic were also associated with this synoptic case. The gale force *bura* (with wind speed of more than  $20 \text{ ms}^{-1}$ ) in the northern Adriatic has been captured by ALEPS with probability exceeding 95% (Fig. 15d), i.e. nearly all ensemble members predicted such a severe event! In contrast, further south in the central Adriatic (Split), even a somewhat stronger observed wind has been predicted less successfully. Nevertheless, there is some indication of a relatively high wind speed there: from Fig. 15b it could be inferred that probability of wind exceeding  $15 \text{ ms}^{-1}$  is about 65%. The ECEPS has not performed so well in predicting these small-scale features of severe wind at the eastern Adriatic coast. An increased probability is found to the southwest over the open sea, further away from the place where the actual *bura* blows. Further south, in the Dubrovnik area, neither model captured strong winds, almost equally strong to those in the northern and central Adriatic. In ALEPS there is some hint of increased winds (Fig. 15b), however, it is displaced away from the coast, between  $42^\circ$  and  $43^\circ\text{N}$ .

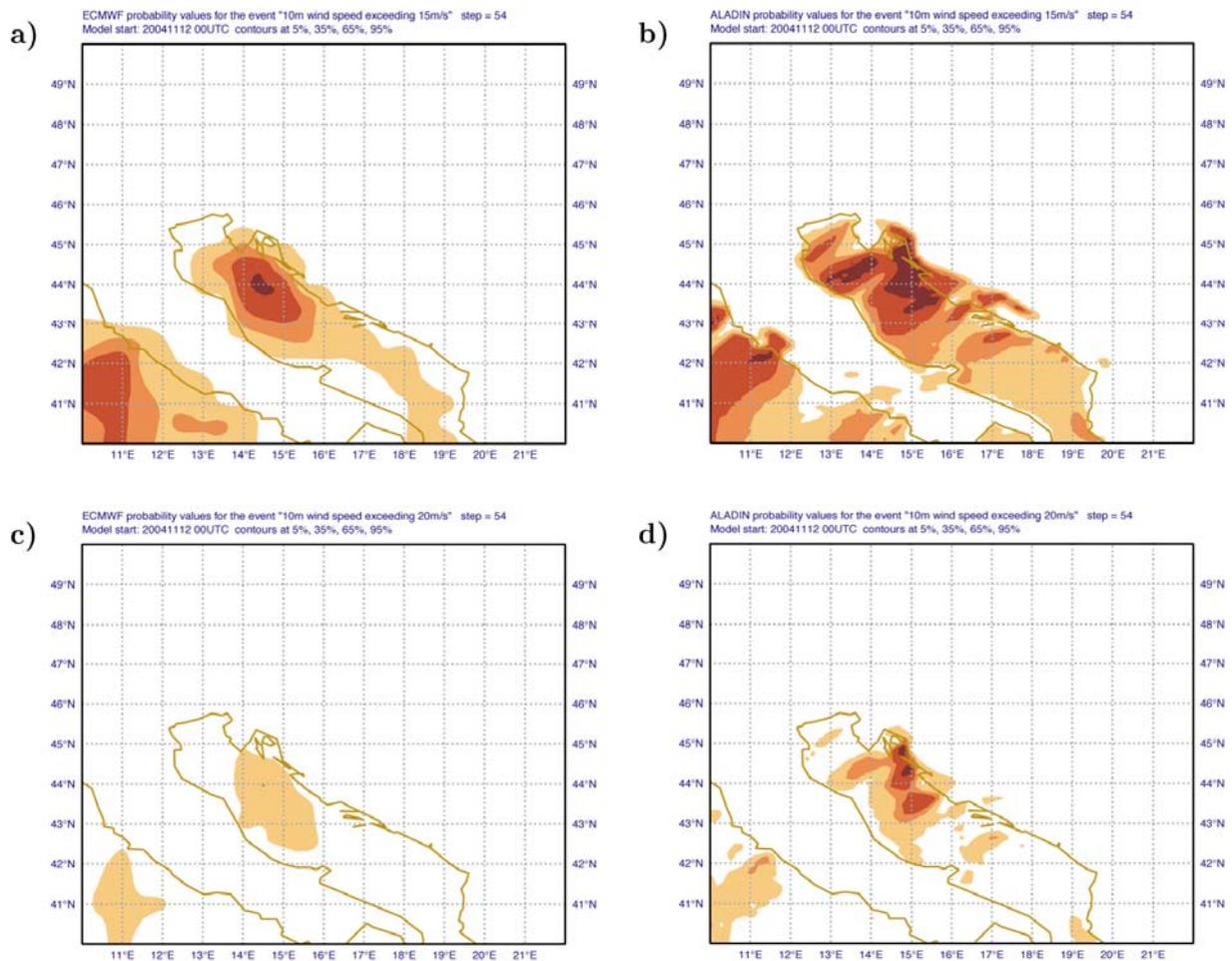


Figure 15 Probabilities for 10 m wind exceeding 15 ms<sup>-1</sup> (top) and exceeding 20 ms<sup>-1</sup> (bottom) in the AUI synoptic case at T+54 for ECEPS (left) and ALEPS (right). Contouring at 5, 35, 65 and 95%.

## 5. Summary and conclusions

Synoptic cases of severe weather in various parts of Croatia are studied by the dynamically downscaling ECMWF global model ensemble forecasts. The four 5-day global forecasts were generated using the ECMWF 51-member EPS (ECEPS). The 3-hourly ECEPS outputs were used to force the Aladin limited area model over the central Europe and the northern Mediterranean domain, thus creating downscaled ensembles (ALEPS) with 51 members each. The ECEPS was run at the TL255 spectral resolution (roughly corresponding to 80 km) with 40 levels in the vertical, and ALEPS was run at the 12.2-km regular grid with 37 vertical levels. The outputs from the both, ECEPS and ALEPS, ensembles were manipulated in the same way in order to determine the impact of dynamical downscaling and possible gains that could be attained for studying synoptic cases of severe weather.

The two measures of modelling error with respect to ECMWF operational analysis (mean absolute difference and mean deviation), defined over the downscaling domain, indicate that the errors in ECEPS and ALEPS upper-air fields are comparable. However, when the  $\omega$  field is excluded from the consideration, the overall picture changes in favour of ALEPS, possibly indicating that dynamical downscaling generally reduces the modelling errors in the free atmosphere. The spatial distribution of model errors indicates that largest

differences between the two models are not necessarily associated with relatively large orographic differences.

When the same clustering algorithm is applied to both sets of ensembles, the resulting difference in the size of clusters appears to be relatively large. Thus, it seems that smaller spatial scales as defined in ALEPS, affect ensemble properties and cause different clustering results, even if the clusters are made of the same members. This is confirmed by examples whereby the error fields, derived from the clusters containing the highest number of common members between ECEPS and ALEPS, look very different. It seems therefore that in the case of dynamical downscaling many common members do not necessarily guarantee similarity between clusters from the two different populations.

When the statistics applied to full ensembles is extended to most populated clusters, it emerges that the clustering reduces the mean variability among cluster members in comparison to the variability of the full ensemble, but will not necessarily reduce the mean cluster error in comparison to mean ensemble error. For 850-hPa wind and both (500 and 700 hPa)  $\omega$  fields, spread of full ensembles and spread of most populated clusters is consistently larger in ALEPS than in ECEPS. Since these fields are generally more susceptible to orographic influences than the other upper-air fields considered, such a result indicates that a better resolved small-scale orography may strongly affect the dispersion among ensemble (cluster) members in ALEPS.

A detailed analysis of various distances among clusters, cluster centroids and representative members reveal that, in approximately one-half of all (parameter/season) combinations considered, dynamical downscaling have caused non-negligible differences between global and regional clusters even when identical clustering algorithm was applied. This implies that the properties identified via representative members in global clusters may not always be extended to regional clusters. In other words, what is found to be representative in global clusters is not necessarily representative in downscaled clusters. Because of such (potential) dissimilarities between global and regional clusters, a careful consideration must be taken when choosing the global representative members for dynamical downscaling, in particular if they serve as proxies for fields that are highly dependent on small-scale orographic features (like for example precipitation discussed in Molteni et al 2001 and Marsigli et al 2001).

Three cases of typical severe weather (or storms) that occasionally occur over the maritime and continental parts of Croatia have been discussed in more details. Overall, the comparison of the results from these synoptic cases yields by no means a unique and unequivocal advantage of the higher resolution model. Whereas in one case the downscaling brings an improvement in the prediction of local rainfall, in the second case almost no improvement is seen, i.e. both global and regional models misplaced the precipitation maximum. In the third synoptic case, the prediction of heavy precipitation by global model was already very good, so that the regional model could add only a little to improve the global model forecast. Our synoptic analysis indicates that a relative success of downscaling could be linked to (at least) the following two factors: a better resolved orography in the regional model, and the spatial/temporal extension (scale) of the synoptic event considered.

In the first synoptic case, a higher orography in the Aladin model acted as an enhanced obstacle to the southwesterly flow over the (perpendicularly oriented) mountains of the eastern Adriatic hinterland. The higher Aladin orography excited a higher precipitation discharge, which was relatively closer to observations than in the global model. In the second case, severe weather was defined over very small spatial scales over

an area of Croatia with no major orographic differences between the global and regional models. Thus, the difference in orography could not play any significant role in the downscaling of this synoptic case. In addition, the event's small spatial and time scales possibly conveyed insufficient pieces of information about severe weather in the global model. The third synoptic case differed from the first and, in particular, from the second case in terms of the spatial extent and intensity of disturbance (winds and precipitation). It covered a much larger portion of the Adriatic Sea and the Balkan Peninsula than the other cases, which is the reason why it has been captured relatively successfully by the global model. However, in this synoptic case the downscaled near-surface winds and local precipitation compare much more favourably with observations than those from the GCM.

Overall, our results indicate that the more detailed and somewhat improved spatial distribution of precipitation in ALEPS could be related to a better representation of high orography in the limited area model. Based on the analysis of the eight different upper-air parameters, no definite conclusion could be made on what parameter constitutes the best proxy for the clustering of precipitation. It has been also demonstrated that, when compared with the precipitation observations, most populated clusters do not always attain correct distribution and amount of local precipitation. Such a result indicates a possible deleterious influence of models' systematic errors on clustering. As a complement to clustering, probabilities in ensemble forecasts have also been discussed. Though somewhat improved, in terms of probabilistic diagnostics the results from limited area model bring no clear-cut benefit over the global model.

Although the error and spread statistics for the full ensembles and most populated clusters indicate what might be an overall impact of dynamical downscaling, when it comes to synoptic case studies it is difficult to generalise some of the results. It has been demonstrated that limited area ensembles in some cases improve the forecast accuracy by bringing more details in the field of interest (precipitation, wind). On the other hand, these details in other cases do not automatically signify an improvement in the quality of ensemble forecasts.

## References

- Anderberg, M.R. (1973): *Cluster analysis for applications*. Academic Press, New York.
- Atger, F. (1999): Tubing: An alternative to clustering for the classification of ensemble forecasts. *Weather and Forecasting*, **14**, 741-757.
- Buizza, R. (1997): Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **125**, 99-119.
- Buizza, R., M. Miller and T.N. Palmer (1999): Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2887-2908.
- Buizza, R. and P. Chessa (2002): Prediction of the US-storm of 24-26 January 2000 with the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.* **130**, 1531-1551.
- Buizza, R. and A. Hollingsworth (2002): Storm prediction over Europe using the ECMWF Ensemble Prediction System. *Met. Appl.*, **9**, 289-305.

- Frogner I-L. and T. Iversen (2002): High-resolution limited-area ensemble predictions based on low-resolution targeted singular vectors. *Q. J. R. Meteorol. Soc.*, **128**, 1321-1341.
- Giard, D. and E. Bazile (2000): Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. *Mon. Wea. Rev.*, **128**, 997-1015.
- Ivančan-Picek, B., D. Glasnović and V. Jurčec (2003): Analysis and Aladin prediction of heavy precipitation event on the Eastern side of the Alps during Map IOP 5. *Meteor. Zeit.*, **12**, 103-112.
- Ivatek-Šahdan, S. and B. Ivančan-Picek (2006): Effects of different initial and boundary conditions in Aladin/HR simulations during MAP IOPs. *Meteor. Zeit.*, **15**, 187-197.
- Ivatek-Šahdan, S. and M. Tudor (2004): Use of high-resolution dynamical adaptation in operational suite and research impact studies. *Meteor. Zeit.*, **13**, 99-108.
- Jung T, E. Klinker and S. Uppala (2004): Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part I: Analyses and deterministic forecasts. *Meteor. App.*, **11**, 343-361.
- Jung T, E. Klinker and S. Uppala (2005): Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part II: Ensemble forecasts. *Meteor. App.*, **12**, 111-122.
- Lorenz, E.N. (1982): Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505-513.
- Marsigli C., A. Montani, F. Nerozzi, T. Paccagnella, S. Tibaldi, F. Molteni and R. Buizza (2001): A strategy for high-resolution ensemble prediction. Part II: limited-area experiments in four Alpine flood events *Q. J. R. Meteorol. Soc.*, **127**, 2095-2115.
- Molteni, F., R. Buizza, T.N. Palmer and T. Petroligis (1996): The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119.
- Molteni, F., R. Buizza, C. Marsigli, A. Montani, F. Nerozzi and T. Paccagnella (2001): A strategy for high-resolution ensemble prediction. Part I: Definition of representative members and global-model experiments. *Q. J. R. Meteorol. Soc.*, **127**, 2069-2094.
- Montani, A., C. Marsigli, F. Nerozzi, T. Paccagnella and R. Buizza (2001): Performance of the ARPA-SMR limited-area ensemble prediction system: two flood cases. *Nonlinear Processes in Geophysics*, **8**, 387-399.
- Montani, A., C. Marsigli, F. Nerozzi, T. Paccagnella, S. Tibaldi and R. Buizza (2003) The Soverato flood in Southern Italy: performance of global and limited-area ensemble forecasts. *Nonlinear Processes in Geophysics*, **10**, 261-274.
- Palmer, T.N., J. Barkmeijer, R. Buizza and T. Petroligis (1997): The ECMWF Ensemble Prediction System. *Meteor. App.*, **4**, 301-304.

Shutts, G.J. (1990): Dynamical aspects of the October 1987 storm: A study of a successful fine-mesh simulation, *Q. J. R. Meteorol. Soc.*, **116**, 1315-1347.

Stanski, H.R., L.J. Wilson and W.R. Burrows (1989): Survey of common verification methods in meteorology. *World Weather Watch Technical Report No. 8 (WMO/TD No. 358)*. World Meteorological Organization, Geneva, Switzerland

Stensrud, D.J., H.E. Brooks, J. Du, S. Tracton and E. Rogers (1999): Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, **127**, 433-446.

Strelec Mahović, N. and D. Drvar (2005): Hailstorm on 04 July 2003 – a case study. *Croatian Meteorological Journal*, **40**, 381-384.

Tudor, M., S. Ivatek-Šahdan (2002): MAP IOP 15 case study. *Croatian Meteorological Journal*, **37**, 1-14.

Tripoli, G.J., C.M. Medaglia, S. Dietrich, A. Mugnai, G. Panegrossi, S. Pinori and E.A. Smith (2005): The 9-10 November 2001 Algerian flood. A numerical study. *Bull. Amer. Meteorol. Soc.*, **86**, 1229-1235.

Van den Hurk BJJM, Viterbo P, Beljaars ACM, Betts AK (2000): Offline validation of the ERA40 surface scheme. *ECMWF Tech Memo 295* (available from: <http://www.ecmwf.int/publications/>)

Wilks, D.S. (1995): *Statistical methods in the atmospheric sciences*. Academic Press, San Diego.