Nat. Hazards Earth Syst. Sci., 12, 2631–2645, 2012
www.nat-hazards-earth-syst-sci.net/12/2631/2012/
doi:10.5194/nhess-12-2631-2012
© Author(s) 2012. CC Attribution 3.0 License.



Natural Hazards and Earth System Sciences

Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for Mediterranean heavy precipitating events

B. Vié[1], G. Molinié[2], O. Nuissier[1], B. Vincendon[1], V. Ducrocq[1], F. Bouttier[1], and E. Richard[3]

[1]Groupe d'étude de l'atmosphère météorologique (GAME) – URA1357, Météo-France/CNRS, Toulouse, France
[2]LTHE, Laboratoire d'études des transferts en hydrologie et environnement – (CNRS, UJF, IRD, INPG), Grenoble, France
[3]Laboratoire d'Aérologie, CNRS et Université de Toulouse, Toulouse, France

Correspondence to: B. Vié (benoit.vie@meteo.fr)

Received: 28 February 2012 – Revised: 26 June 2012 – Accepted: 6 July 2012 – Published: 21 August 2012

Abstract. An assessment of the performance of different convection-permitting ensemble prediction systems (EPSs) is performed, with a focus on Heavy Precipitating Events (HPEs). The convective-scale EPS configuration includes perturbations of lateral boundary conditions (LBCs) by using a global ensemble to provide LBCs, initial conditions (ICs) through an ensemble data assimilation technique and perturbations of microphysical parameterisations to account for part of model errors. A probabilistic evaluation is conducted over an 18-day period. A clear improvement is found when uncertainties on LBCs and ICs are considered together, but the chosen microphysical perturbations have no significant impact on probabilistic scores.

Innovative evaluation processes for three HPE case studies are implemented. First, maxima diagrams provide a multi-scale analysis of intense rainfall. Second, an hydrological evaluation is performed through the computation of discharge forecasts using hourly ensemble precipitation forecasts as an input. All ensembles behave similarly, but differences are found highlighting the impact of microphysical perturbations on HPEs forecasts, especially for cases involving complex small-scale processes.

1 Introduction

The north-western Mediterranean basin is frequently hit by Heavy Precipitating Events (HPEs), mainly during the fall. The mesoscale convective systems associated with these HPEs typically produce over 200 mm in 6–24 h. Such intense rainfall events occurring over the small and steep Mediterranean hydrological catchments can trigger catastrophic flash-floods, threatening both people and property. The most dramatic of these events have been extensively studied (see Buzzi et al., 1998; Nuissier et al., 2008; Ducrocq et al., 2008, among others). It was shown that their predictability is strongly affected by the complex interactions of different small scale processes, such as convective instability or microphysical processes.

Thus, although Convection-Permitting Models (CPMs) simulate realistic precipitating systems, forecasting those systems precisely remains a great challenge, and still the hydrological runoff forecasts for such small catchments are very sensitive to both the rainfall maximum intensity and location. It is, therefore, essential to evaluate the uncertainty of convective-scale forecasts.

Ensemble prediction is now a well-known tool for quantifying uncertainties of weather forecasts. While global, medium-range Ensemble Prediction Systems (EPSs) are operational since the 1990s to assess the predictability of large-scale atmospheric flows, the design of convection-permitting EPSs adapted to the evaluation of the predictability of local, high-impact weather is still at an early stage. The different sensitivity to Initial Conditions (ICs), the faster growth of convective-scale perturbations due to the more nonlinear physical parameterisations, the need to account for the uncertainty of the Lateral Boundary Conditions (LBCs), as well as the much higher computing time required by CPMs make it difficult to adapt the methods used to generate global EPSs.

Thus, the design of a convection-permitting EPS remains an open question.

Recent studies on convection-permitting EPSs have explored various techniques to generate ensembles. Hohenegger et al. (2008) compared different downscaling procedures to generate both ICs and LBCs for their convection-permitting EPSs and found that the added value of the CPM forecasts varied with the synoptic-scale conditions. For their model set-up, the impact of uncertainty on LBCs became predominant after 12 h. Hohenegger and Schär (2007a) focused on the impact of differences in the initial conditions and compared a shifted initialisation technique to perturbations in the initial temperature field. They showed that all methods had a similar impact and identified the same region of lower predictability. Gebhardt et al. (2011) designed three ensembles to sample the uncertainties on LBCs and model physics, separately at first, and then together. They found that physics perturbations are dominant over the first 4 to 6 h of simulation. The LBCs have generally more impact on the convection-permitting ensemble spread after 6 h, although in some cases physics perturbations have a strong impact on longer forecast ranges as well. They also concluded that sampling both sources of uncertainty together increased the ensemble quality. Fresnay et al. (2012) studied the sensitivity of a CPM to the auto-conversion, accretion and evaporation tendencies, for a Mediterranean HPE. Perturbations of these microphysical tendencies mostly impacted rainfall intensity, but sometimes displaced the precipitating systems as well. Clark et al. (2009, 2010), using data from the 2007 NOAA Hazardous Weather Testbed Spring Experiment, discussed the factors influencing the growth of ensemble spread and highlighted the value of convection-permitting ensemble forecasts compared to regional EPSs, despite the lower number of members.

Vié et al. (2011) assessed the relative impact on CPM forecasts of uncertainties associated with convective-scale ICs and synoptic-scale LBCs. Comparing distinct ensembles over both a 1-month period and case studies of HPEs, they showed that the impact of these two sources of uncertainty were different. Initial perturbations mostly impact short forecast ranges, while uncertainty coming from the LBCs rapidly becomes predominant (after 12 h in their experiment set-up) and accounts for most of the ensemble spread. However, even if initial perturbations have generally little impact beyond 12 h, they remain important for some of the HPE case-studies. Both ensembles had satisfying probabilistic scores, but suffered from a strong lack of spread, especially for low-level parameters, known to strongly influence Mediterranean HPEs.

The present work assesses the benefit of accounting for these two sources of uncertainty in a single convection-permitting EPS, both in terms of ensemble spread and probabilistic scores. In an effort to sample model errors, a convection-permitting ensemble including perturbations of the microphysics scheme is also evaluated.

Table 1. Characteristics of the different ensemble experiments.

Experiment name	LBCs	Data assimilation	Microphysical perturbations
E1	PEARP EPS	Unperturbed obs.	No
E2	ARPEGE deterministic fc.	Perturbed obs.	No
E3	PEARP EPS	Perturbed obs.	No
E4	PEARP EPS	Perturbed obs.	Yes

Designing and comparing convection-permitting EPSs for rare, high-impact events also raises the question of probabilistic forecast verification and evaluation. In addition to the common probabilistic evaluation using ensemble scores such as rank histograms and Relative Operating Characteristics (ROC) curves, an innovative evaluation is performed over three case studies of Mediterranean HPEs. The computation of maxima diagrams provides a multi-scale assessment of simulated and observed rainfall, as in Ceresetti et al. (2012), in relation with the structure of the precipitating systems. Precipitation forecasts are also used as input for hydrological ensemble runoff forecasts, to assess the value of our ensembles for flash-flood forecasting.

This paper is structured as follows. In Sect. 2, the forecasting system is detailed. Section 3 presents the results of the probabilistic evaluation of the ensemble forecasts over an 18-day period. The evaluation of the ensembles for the three case studies is performed in Sect. 4, using both maxima diagrams and hydrological discharge forecasts. Conclusions are drawn in the last section.

2 Methodology

2.1 The convection-permitting model AROME

The operational CPM Application of Research to Operations at MEsoscale (AROME) from Météo-France was used in this study. The AROME forecasting system is extensively described in Seity et al. (2010). It is based on adiabatic, non-hydrostatic equations from the limited-area ALADIN model. An horizontal grid-spacing of 2.5 km and 41 vertical levels are used. AROME uses physical parameterisations from the research model Meso-NH (Lafore et al., 1998), including a bulk, one-moment microphysics scheme following Caniaux et al. (1994) which represents six water species. An eddy diffusivity Kain-Fritsch (EDKF) scheme (Pergaud et al., 2009) is used for shallow convection parameterisation, and the turbulent scheme follows Cuxart et al. (2000).

AROME has its own 3D-VAR data assimilation scheme, with background and observation statistics adapted to its fine resolution (Yan et al., 2009; Boniface et al., 2009). Ground-based observations, as well as satellite data and doppler radial winds from the weather radar network are assimilated.

Table 2. Multiplying factors applied to the auto-conversion, accretion and evaporation tendencies, for each member of the E4 ensemble.

Ens. member	0	1	2	3	4	5	6	7	8	9	10
Autoconversion	1	0.75	0.75	0.75	0.75	1.25	1.25	1.25	1.25	1	1
Accretion	1	0.75	0.75	1.25	1.25	0.75	0.75	1.25	1.25	1	1
Evaporation	1	0.75	1.25	0.75	1.25	0.75	1.25	0.75	1.25	0.5	1.5

Table 3. Brier Skill Score computed for different precipitation thresholds for the E1, E2 and E3 ensembles. The reference is the deterministic AROME forecast. Bold numbers for each threshold show the best performing ensemble.

threshold (mm)	0.5	2	5	10	20	30	40	50	75	100
E1	0.33	0.29	0.27	**0.26**	**0.33**	0.35	0.36	0.27	0.29	0.26
E2	0.29	0.28	0.27	0.22	0.23	0.26	0.26	0.23	0.31	0.31
E3	**0.35**	**0.31**	**0.29**	**0.26**	**0.33**	**0.38**	**0.38**	**0.29**	**0.34**	**0.33**

2.2 The ensemble forecasts

Clark et al. (2011) showed that most of the probabilistic quantitative precipitation forecast skill for a convection permitting ensemble is obtained with around 10 members for forecast lead times up to 30 h. All the ensembles described below are composed of 11 parallel AROME data assimilation cycles, with a 3-hourly data analysis frequency. One 24-h forecast is issued each day at 12:00 UTC. Two ensembles sample the uncertainty separately on LBCs and convective-scale ICs. Another ensemble combines the representation of both uncertainty sources, the last one also accounts for part of model uncertainty as well as through the perturbation of microphysical tendencies. Table 1 summarises the characteristics of these four ensembles.

The E1 ensemble (named AROME-PEARP2 in Vié et al., 2011) samples the uncertainty coming from the imperfect LBCs by driving the 11 AROME data assimilation cycles and forecasts with the members of the Météo-France global, short-range EPS, called Prévision d'Ensemble ARPEGE (PEARP). The PEARP members are first downscaled using the regional model ALADIN to prevent a large gap in resolution. The E2 ensemble (named AROME-PERTOBS in Vié et al., 2011) assesses the impact of uncertainty on convective-scale ICs through an ensemble data assimilation technique, as in Berre et al. (2006) and Houtekamer et al. (1996). The analysis error is sampled by the cycled assimilation of randomly perturbed observations (every 3 h), creating different ICs for each of the 11 E2 forecasts. Each member has LBCs provided by the operational, deterministic ALADIN forecast.

The E3 ensemble combines the two methods to generate an ensemble accounting for both uncertainty on the LBCs and the ICs at convective scale. Each of the 11 members of E3 uses randomly perturbed observations in its data assimilation cycle, as in E2, and uses LBCs provided by one member of the PEARP ensemble as in E1. Figure 1 describes the numerical set-up for the E3 ensemble. The 24-h forecasts issued daily at 12:00 UTC, and the preceding assimilation cycles, are drawn in black. Running assimilation cycles are shown in light grey.

Based on the E3 ensemble, E4 in addition accounts for model errors in its representation of forecast uncertainty. Perturbations of the warm rain microphysical parameterisation, as in Fresnay et al. (2012), are introduced during the data assimilation and forecasts. Auto-conversion of cloud droplets into raindrops, accretion of cloud droplets by raindrops and rain evaporation processes are perturbed by applying a multiplying factor to each tendency. This factor is constant in both space and time, that is, for one given member of the ensemble, the multiplying factor is the same at each grid point and throughout the whole data assimilation and forecast period. The perturbations selected for each member ranged from 0.5 to 1.5 (Table 2).

An ensemble, E5, has been specifically designed to study the impact of microphysical parameterisation perturbations alone, using a single set of initial and lateral boundary conditions for all members.

3 Statistical evaluation

The ensembles are evaluated over the same period used in Vié et al. (2011), which corresponds to 18 consecutive days, from 15 October 2008 to 1 November 2008 inclusive. This period includes days with different atmospheric conditions. Precipitation and lightning observations (not shown) indicate that convective activity occurred on 20–24 October 2008 and from 30 October to 1 November 2008. The statistical evaluation of our ensembles was carried out for 24-h accumulated precipitation, surface and low-level parameters. Definition of the scores used in this study can be found in Vié et al. (2011).

Rank histograms are a measure of ensemble spread. An ensemble with an adequate spread would produce a flat histogram, a U-shaped histogram highlights a lack of spread in the ensemble forecast. Figure 3 shows rank histograms

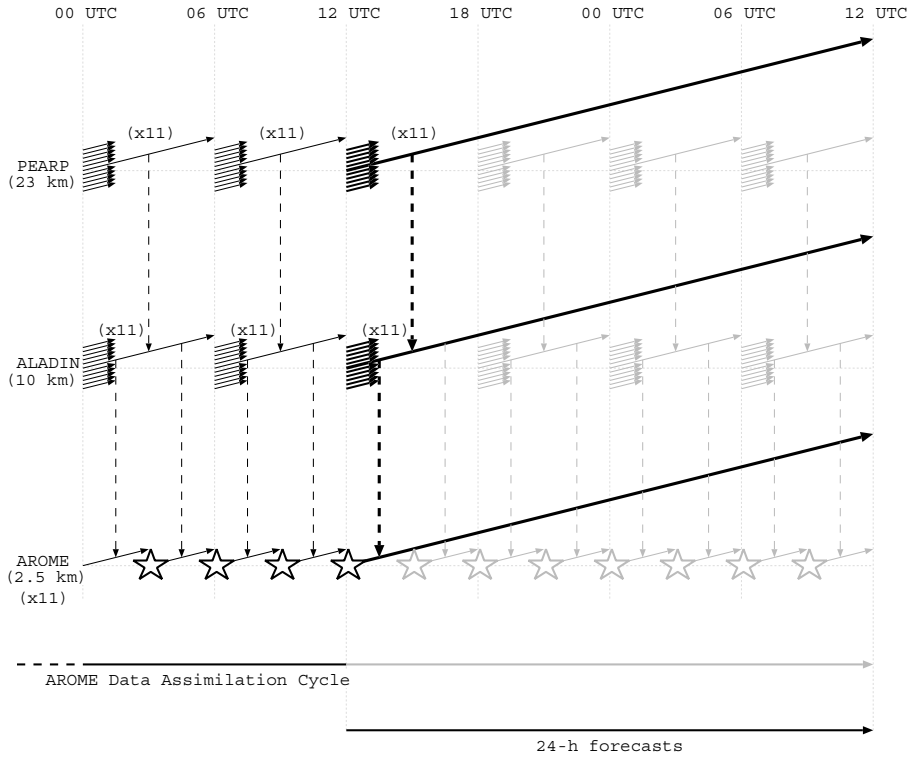


Fig. 1. The E3 ensemble experiment. Each nested model run is schematised along one row. Each plain (thick) arrow indicates the behavior of the eleven ensemble members (x11) short (long) runs. Their coupling along time is figured by the vertical dashed arrows. Each PEARP member provides LBCs to one ALADIN downscaling forecast, itself providing LBCs to one AROME data assimilation cycle with data analysis using randomly perturbed observations (stars) every 3 h. Each day, 24-h AROME forecasts are run at 12:00 UTC. The continuing assimilation cycles after 12:00 UTC are shown in light grey.

for the three ensembles E1, E2 and E3, computed for wind speed at 925 hPa, against the operational AROME analysis, for forecast ranges of 3, 6, 12 and 24 h. Although the E3 ensemble is still under-dispersive at all forecast ranges, the combined use of the ensemble data assimilation technique and different coupling conditions for each member clearly yields a better spread than each method used separately. It is especially interesting that the improvement is larger for intermediate forecast ranges (6 to 12 h), when the impact of initial spread is already largely reduced and the use of different LBCs begins to produce significant spread.

Brier Skill Scores (BSS) computed for different thresholds of 24-h accumulated precipitation, using the operational deterministic AROME forecast as reference, are shown in Table 3. A perfect forecast has BSS = 1. Positive BSSs for all three ensembles and every threshold highlight the added value of a probabilistic forecast compared to a single deterministic forecast. The E3 ensemble has better scores than any of the other two ensembles for each precipitation threshold (equal to E1 for 10 and 20 mm), which again shows the benefit of sampling both sources of uncertainty simultaneously.

Relative Operating Characteristics (ROC) and reliability diagrams (computed for precipitation intervals of $0 < 0.5 < 2 < 5 < 10 < 20 < 50 < +\infty$, and wind intervals of $0 < 1 < 2 < 4 < 6 < 8 < 10 < +\infty$) are shown in Fig. 4 for 24-h accumulated precipitation and 10-m wind speed, computed against ground-based observations (hourly rainfall amounts provided by automated raingauges, 10-m wind speeds measured by automated land-surface stations). ROC curves for E1 and E3 are very close, both for wind speed and precipitation. Both ensembles, thus, have a similar resolution. However, reliability diagrams, especially for precipitation (Fig. 4c), show that the E3 ensemble has a better reliability.

Figure 4 also shows that the addition of microphysical tendencies perturbations in the E4 ensemble has very little impact on the probabilistic scores. This is confirmed by rank histograms shown in Fig. 5 among other scores and parameters computed for this ensemble (not shown). The microphysical perturbations applied in the E4 ensemble focus only on warm microphysical processes. Thus, one can expect an impact of these perturbations for precipitating days, much less for days with no rain. For this reason, no significant effect was found on probabilistic scores applied to the whole period.

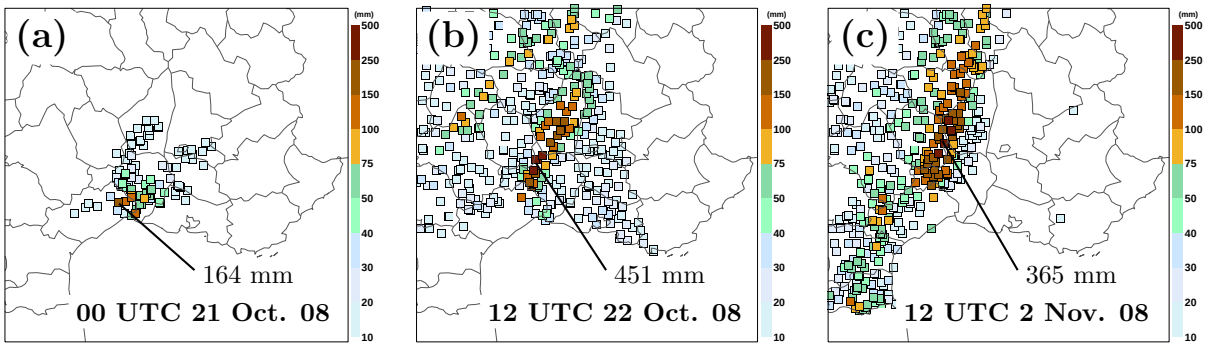


Fig. 2. Observed 24-h accumulated precipitation (mm) at **(a)** 00:00 UTC 21 October 2008, **(b)** 12:00 UTC 22 October 2008 and **(c)** 12:00 UTC 2 November 2008.

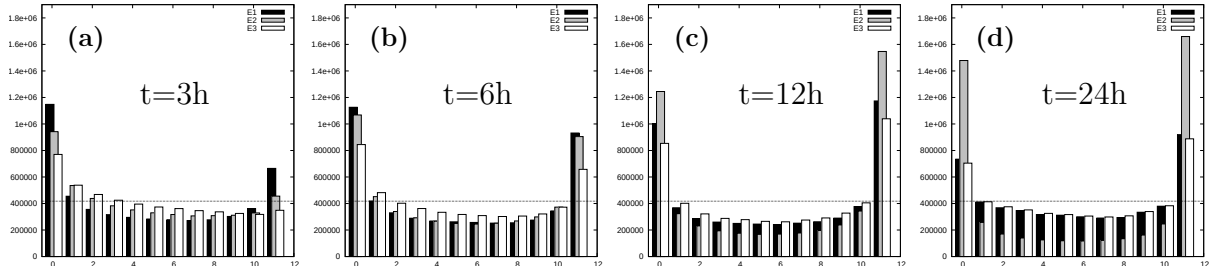


Fig. 3. Rank histograms for the 925-hPa wind speed from 15 October 2008 to 1 November 2008 at a forecast range of **(a)** 3, **(b)** 6, **(c)** 12 and **(d)** 24 h. E1 (black), E2 (gray) and E3 (white).

4 Heavy Precipitating Events case studies

During this 18-day period, three HPEs occurred on 20, 21–22 October and 1–2 November 2008, as evidenced by the 24-h accumulated precipitation observations for these events (Fig. 2).

On 20 October 2008 (Case 1), a quasi-stationary MCS formed over the plain upstream of the Massif Central foothills, between 13:00 UTC and 17:00 UTC. Synoptic scale conditions showed only a weak baroclinic activity and the MCS was driven mainly by mesoscale mechanisms involving interaction with a low-level cold pool. Due to the convective activity beginning around 06:00 UTC, 24-h ensemble forecasts were issued at 00:00 UTC for this case, to simulate the whole event.

On 21–22 October 2008 (Case 2), localised convective cells formed between 10:00 UTC an 16:00 UTC over the Massif Central, ahead of a cold front moving south-eastward. Then, convective activity decreased with the cold front approaching south-eastern France, to evolve into an organised convective line from 21:00 UTC. The convective system merged with the cold front and intensified until 06:00 UTC, then the system decayed and moved south-eastward over the sea. This system produced 450 mm of precipitation in 24 h locally.

On 1–2 November 2008 (Case 3), convection formed in connection with a large trough over western France. A very strong low-level jet, bringing moist, unstable air, was lifted by the Massif Central. An upper-level low, at the west of the region, induced a strong divergent flow over south-eastern France. Rainfall amounts up to 365 mm in 24 h were observed.

4.1 Ensemble precipitation forecasts

Figure 6a–c shows the ensemble average and spread of 24-h accumulated precipitation forecasts for Case 3, for the E1, E3 and E4 ensembles, respectively. There are only minor differences between the three ensembles. Since this case is largely driven by the synoptic-scale forcing, it was expected that members using the same LBCs would behave similarly, and perturbations of initial conditions or physics parameterisations have little impact. For some members, the difference between E1 and E3 is slightly greater than between E3 and E4. For instance, for member 4 (Fig. 6d–f), E3 and E4 have similar maximum precipitation accumulation, a little higher than in E1, and both produced a secondary precipitation line south-east of the main precipitation area.

Case 2 involves both mesoscale processes and interactions with synoptic-scale conditions. Figure 7 shows ensemble averages of 24-h accumulated precipitation (a–c) and individual forecasts from members 8 and 9 (d–i). There are more significant differences between ensembles for this case, showing the greater impact of perturbations on initial

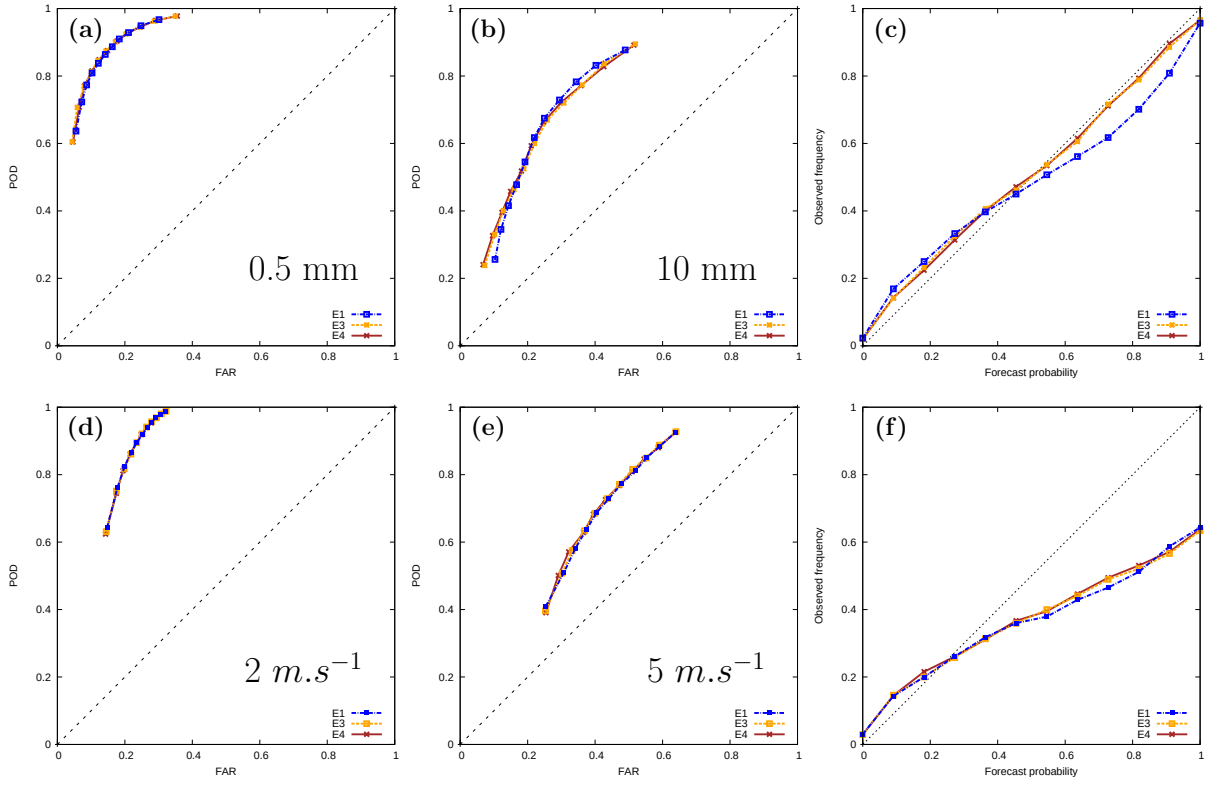


Fig. 4. ROC curves for precipitation thresholds of **(a)** 0.5 and **(b)** 10 mm and for 10-m wind speed thresholds of **(d)** 2 and **(e)** 5 m s^{-1}. Reliability diagrams for **(c)**precipitation and **(f)** 10-m wind speed.

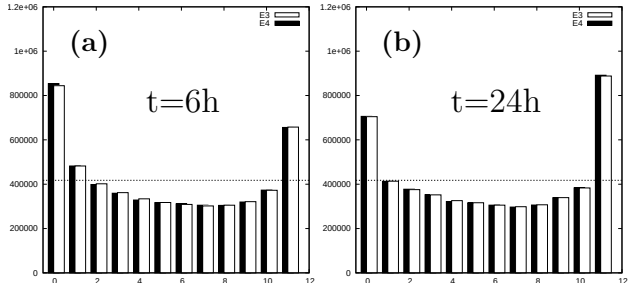


Fig. 5. Same as Fig. 3 for E3 (white) and E4 (black), at forecast ranges of **(a)** 6 and **(b)** 24 h.

conditions than for Case 3. E1 produces higher rainfall totals, as shown by the ensemble mean, as well as forecasts from members 8 and 9. For all ensembles, the ensemble spread reaches similar values and is colocated with maximum rainfall amounts, but E3 and E4 have a slightly more extended 20 mm spread region. Differences between the E3 and E4 ensembles remain small, and are especially much smaller than between E1 and E3 (see for instance members 8 and 9, Fig. 7d–i). Perturbations of the microphysical parameterisations again bring little improvement to these probabilistic forecasts.

The quasi-stationary MCS on 20 October 2008 (Case 1) is mainly driven by mesoscale processes. It is, therefore, expected that the impact of initial conditions and microphysical perturbations will be larger than for the previous two cases. Ensemble averages of 24-h accumulated precipitation (Fig. 8a–c) show higher rainfall amounts for E1. For all ensembles, the ensemble spread highlights that the uncertainty is larger on the south-eastern side of the precipitating system. Members from different ensembles show greater differences than in the previous cases, highlighting the more important impact of perturbations of ICs and microphysical parameterisations. However, the impact of microphysical perturbations still seems smaller, for instance in members 5 and 8 (respectively, in Fig. 8d–f and g–i).

Overall, combining different LBCs and perturbations to the observations in the mesoscale data assimilation produces significant differences. As Vié et al. (2011) found, the impact of the addition of initial conditions perturbations in the E3 ensemble depends on the atmospheric conditions, and is larger for days with a weaker synoptic-scale circulation. The perturbation of microphysical parameterisations has a much smaller impact, noticeable only for Case 1 when the system is mainly driven by a rain evaporation induced cold-pool. To further assess the impact of microphysical perturbations, an E5 ensemble forecast is issued for Case 1. Unique initial and

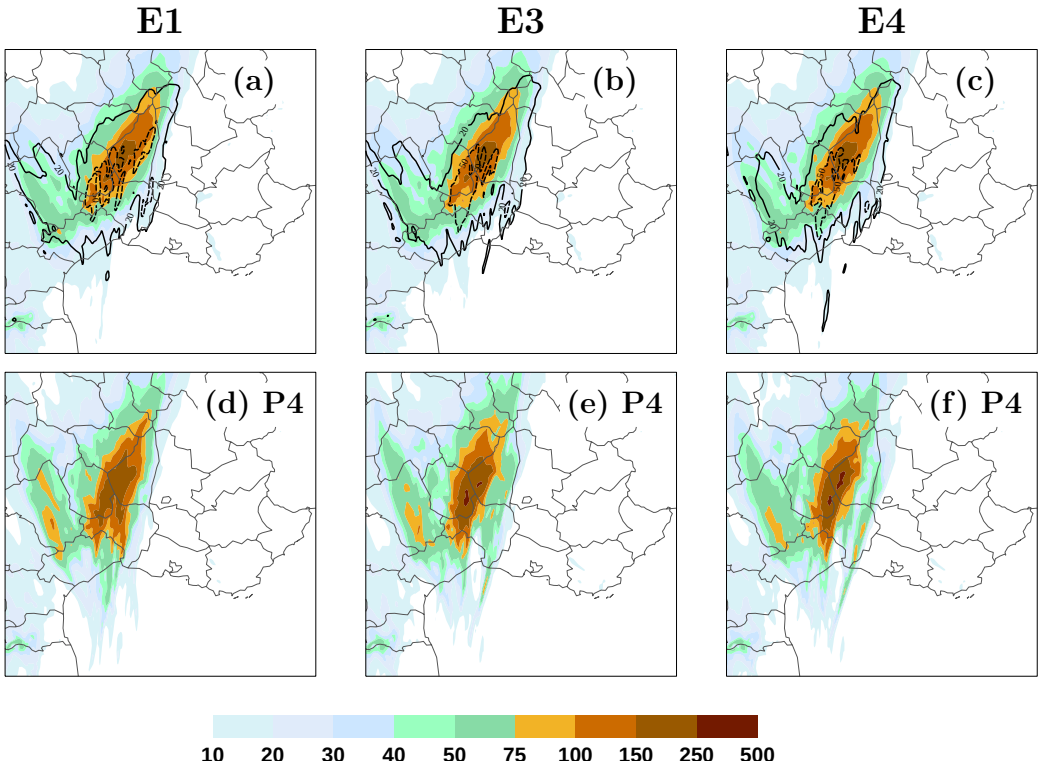


Fig. 6. 24-h accumulated precipitation forecasts (mm) for Case 3, at 12:00 UTC 2 November 2008. **(a–c)** Ensemble mean (colour shading) and standard deviation (solid contour: 20 mm, dashed contour: 50 mm), and **(d–f)** precipitation forecast by member 4, from the E1 (left), E3 (centre) and E4 (right) ensembles.

lateral boundary conditions for all members come from the member 8 of the E3 ensemble. The 24-h accumulated precipitation produced by the E3 member 8 is shown in Fig. 8h. These initial and lateral boundary conditions were chosen because the issued forecast simulated an intense system southward over the plains.

This case involved interactions between the moist, unstable low-level jet and a low-level cold pool, pushing the convective system southward. The microphysical perturbations are expected to have the most impact on this kind of simulation where mesoscale processes are important. In this case, one could expect the rain microphysical perturbations to impact the evaporative cooling and, therefore, affect both rainfall intensity and position. Figure 9 shows 24-h accumulated precipitation forecasts from the 11 E5 ensemble members, as well as the ensemble mean and standard deviation in panel (l). It shows moderate differences on the intensity and spatial structure of rainfall, as well as on the position of the precipitating systems. The ensemble spread is clearly located on the south-eastern side of the precipitating system, showing that the microphysical perturbations have more impact on the triggering of convective cells.

Figure 10 shows, for all ensembles, the virtual potential temperature ensemble spread, as well as the 292 K contour of the ensemble average of the virtual potential temperature, at 15:00 UTC 20 October 2008. Microphysical perturbations in the E5 ensemble indeed have an impact at the edge of the low-level cold air region, where the new convective cells are triggered. However, even in this case where mesoscale processes play an important role, the impact of microphysical perturbations is weaker than that of lateral boundary conditions and even more initial conditions. In this case, the cold air pool which helps the triggering of convection over the plains is present in initial conditions (at 00:00 UTC 20 October 2008, not shown).

4.2 Scale-dependent analysis of heavy rainfall

Ramos et al. (2005) introduced severity diagrams, a scale-dependent analysis of extreme rainfall events based on their return period. In order to avoid extrapolating rainfall intensities by means of probability density functions of extreme values, we perform here an assessment of maximum rainfall intensities, through maximum intensity diagrams similarly to Ceresetti et al. (2012). Therefore, rainfall intensities of the studied events are accumulated by means of moving average operations over all the possible combinations of the discrete temporal and spatial scales.

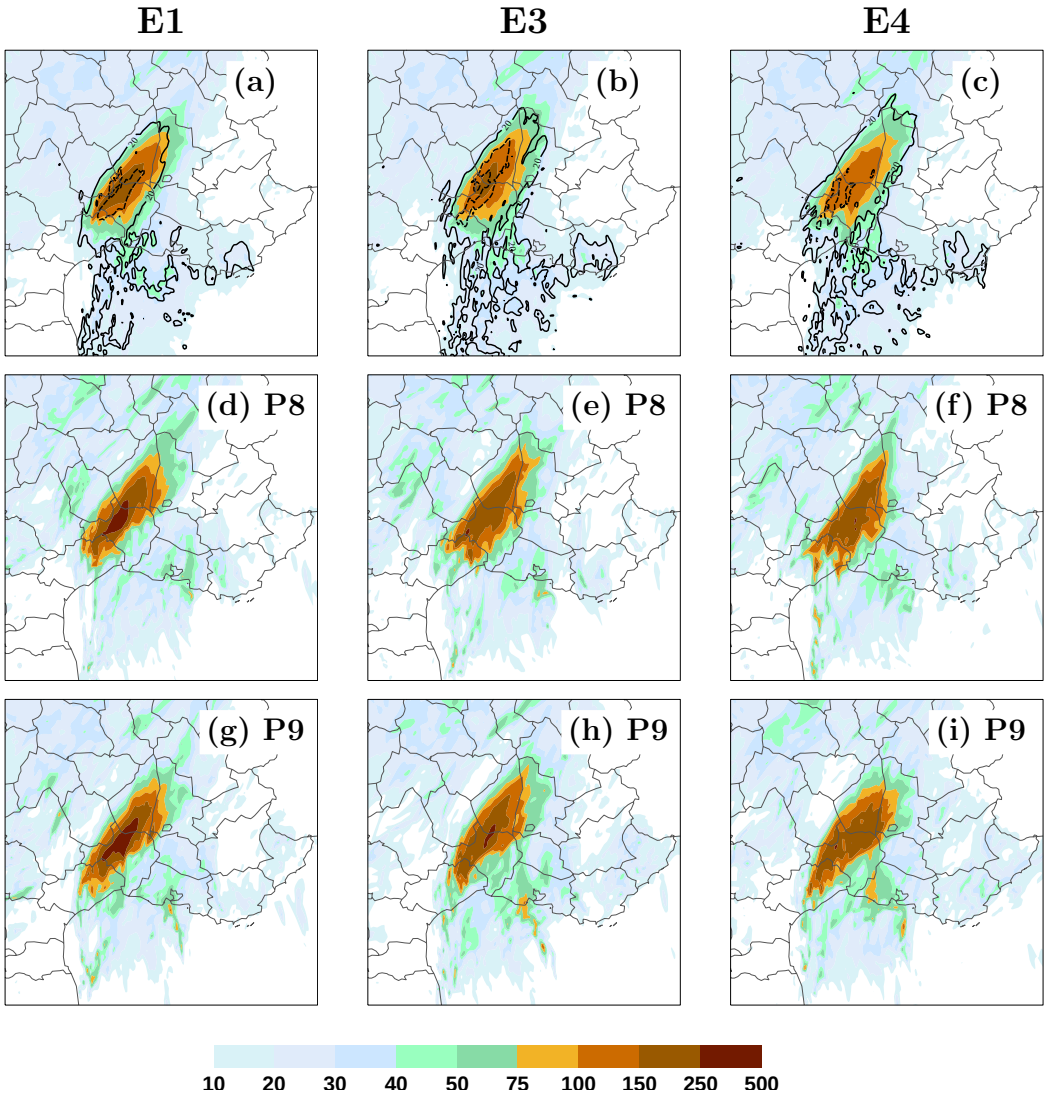


Fig. 7. Same as Fig. 6, for Case 2 at 12:00 UTC 22 October 2008, for members 8 **(d–f)** and 9 **(g–i)**.

The shape of maxima diagrams gives information about the characteristics of the observed or simulated precipitating system.

Figure 11a–c shows the maximum observed rainfall intensities during the three HPEs for a range of spatial and temporal scales. Case 1 yields the lowest intensities, with a maximum around 55 mm h^{-1} for small spatial and temporal accumulation scales, while Cases 2 and 3 reach over 70 mm h^{-1}.

The maximum diagrams also allow a scale-dependent comparison and give information about the characteristics of the observed or simulated precipitating system. For Case 1, rainfall intensities remain high for accumulation over large areas for very short integration duration (one to two hours), but decrease very rapidly with a growing accumulation time. For the other two cases, rainfall intensities for accumulation durations longer than 3 h remain more important, and the diagrams display a more symmetric decrease in intensity with increasing spatial or temporal scales. Berne et al. (2009) have depicted the rain-cell structures highlighting the advection role. The symmetric figures for Cases 2 and 3 show the stationarity of the convective systems, while Case 1 shows some advection.

Figure 12 shows, for each ensemble and each case, the normalised difference between the average forecast maximum intensity and the observed maximum intensity (defined as $\frac{\text{mean(max[sim])-max[obs]}}{\text{max[obs]}}$). These diagrams show no differences between the E1, E3 and E4 ensembles. The structure and behaviour of precipitating cells and systems are, therefore, not significantly changed by the different perturbations.

For Case 1, Fig. 12a–c shows that the maximum intensities simulated for scales between 30 and 80 km, and accumulation durations of 1 to 3 h, are underestimated by around

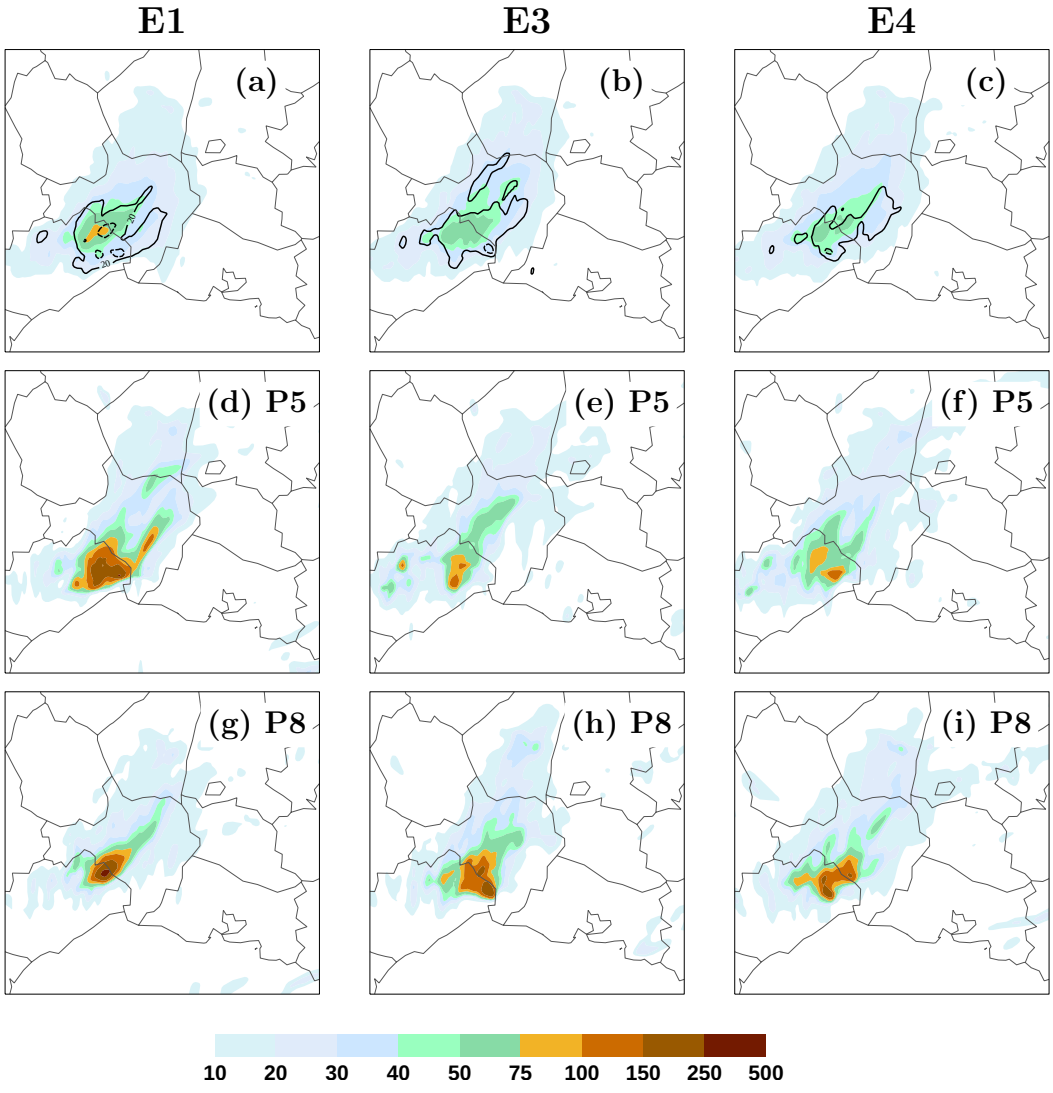


Fig. 8. Same as Fig. 6, for Case 1 at 00:00 UTC 21 October 2008, for members 5 **(d–f)** and 8 **(g–i)**.

50 %. For smaller spatial scales, the simulated maxima are close to the observed ones. This indicates that the convective cells simulated for this case have a correct intensity, but they did not organise into a convective system as large as was observed. Simulations for Case 1 also exhibit an overestimation of maximum intensities for larger spatial and temporal scales (50 km, 18 h). This may be caused by disorganised convective cells producing precipitation in a larger region than was observed, and the simulations producing too much light rain. This is consistent with time series of average precipitation, which show that the peak of precipitation is underestimated, but precipitation is overestimated before and after the peak (see Vié et al., 2011, their Fig. 13b).

The ensembles perform better for Case 2 (Fig. 12d–f), with only an underestimation of about 30 % for spatial scales under 30 km and an overestimation for spatial scales around 50 km. Ensemble forecasts for Case 3 (Fig. 12g–i) produce a more symmetric maxima diagram, with an underestimation of maximum intensities for small spatial and temporal scales and a strong overestimation of intensities for large scales. This shows that the intensity of convective cells is only slightly underestimated, but the simulated precipitating system has a wider extension than the observed one. In both cases, the simulated precipitating region is more extended than in the observations and covers most of the Cévennes. This may indicate that the orographic forcing is too strong in the numerical model.

4.3 Hydrological ensemble discharge forecasts

An evaluation of the AROME ensembles performance for flash-flood forecasting was performed through the computation of hydrological ensemble discharge forecasts using the

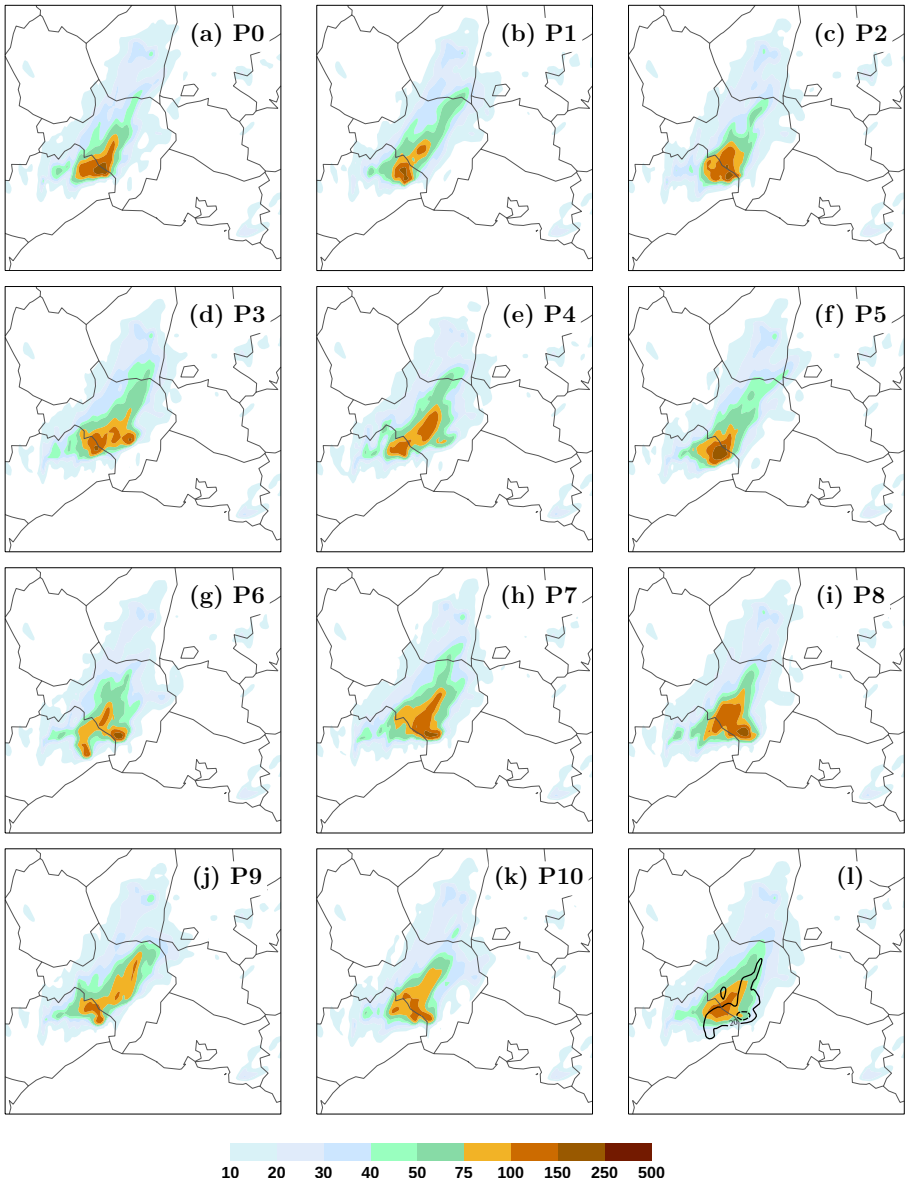


Fig. 9. 24-h accumulated precipitation forecasts (mm) for Case 1, at 00:00 UTC 21 October 2008, from each member of the E5 ensemble. Panel (l) shows the ensemble mean (colour shading) and 20 mm (solid) and 50 mm (dashed) standard deviation contours.

ISBA-TOPMODEL coupled system. This system is a full coupling between the land surface model ISBA, which manages soil water budget on soil columns, and TOPMODEL which takes care of the lateral redistribution of soil moisture based on the topography. This coupled system is run in forecasting mode, using hourly precipitation data (as well as other surface parameters, such as humidity and temperature) from the AROME ensemble members to produce runoff forecasts. Initial conditions are prepared by a 48-h run before the event, started from a larger scale soil analysis (for soil moisture and temperature) and driven by the observed precipitation data during these 48 h. For more details on this forecasting chain, refer to Vincendon et al. (2011) (their Sect. 2.1).

Discharge forecasts were performed for three catchments of the Cévennes-Vivarais region, at Vallon Pont d'Arc for the Ardèche river, at Bagnols-sur-Cèze for the Cèze river and at Boucoiran for the Gardons river (Fig. 1 of Vincendon et al., 2011). Since the precipitating system on 20 October 2008 (Case 1) produced rainfall over the plains south of these three catchments, no hydrological forecasts are performed for this case. Ensemble discharge forecasts were computed for Cases 2 and 3, on 21–22 October and 1–2 November 2008, for the E1, E3 and E4 ensembles.

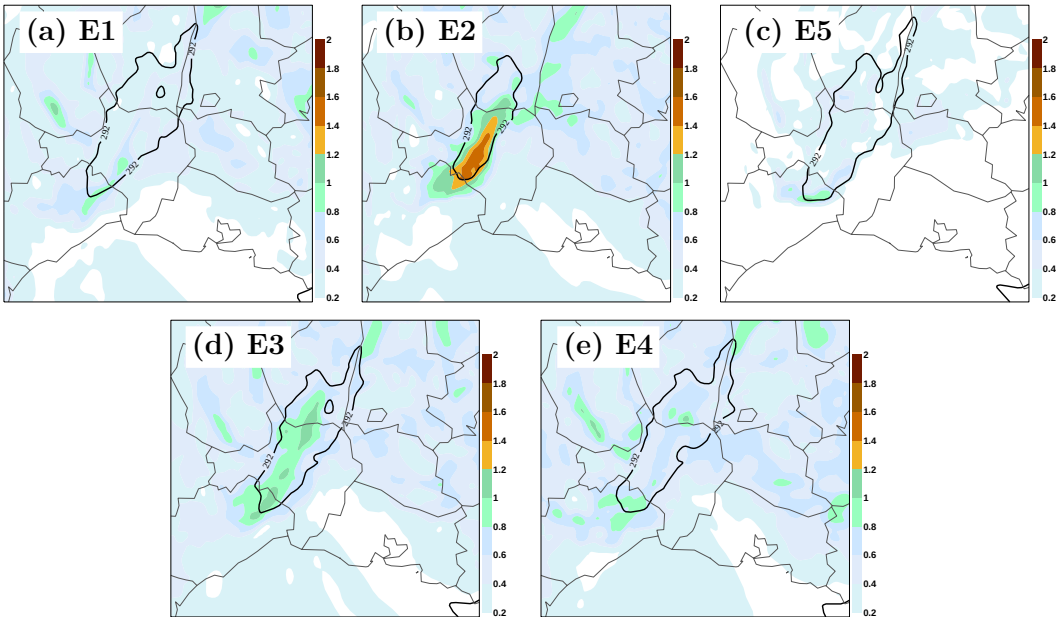


Fig. 10. Virtual potential temperature, ensemble spread (colour shading) and 292 K contour of the ensemble mean (black), at 15:00 UTC 20 October 2008, for **(a)** E1, **(b)** E2, **(c)** E5, **(d)** E3 and **(e)** E4.

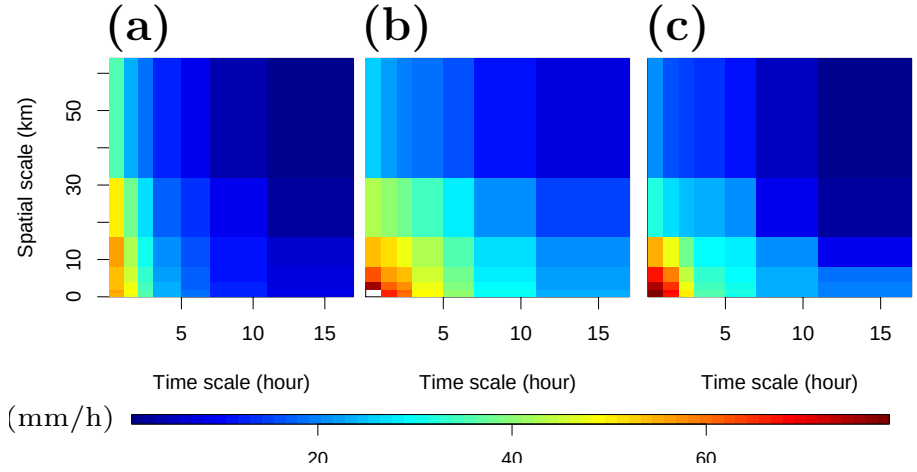


Fig. 11. Maxima diagrams: maximum observed rainfall intensities (mm h^{-1}) as a function of accumulation duration and area, for **(a)** Case 1, **(b)** Case 2 and **(c)** Case 3.

Discharge simulations for Case 3 do not show significant differences between the three ensembles. Forecasts for the Gardons river for each ensemble are shown in Fig. 13. As detailed in the previous section, most of the uncertainty on this case emerges from the synoptic-scale conditions. LBCs have a dominant impact on the AROME forecasts, so that the three ensembles produced similar precipitation forecasts, and the hydrological forecasts behave similarly as well.

Figure 14 shows discharge forecasts for the E1, E3 and E4 ensembles for the Ardèche, Cèze and Gardons rivers for Case 2. As previously stated, the E1 ensemble produced higher rainfall totals than E3 and E4, which explains that hydrological forecasts driven by the E1 ensemble members simulated stronger discharges. This is especially clear for the Cèze (Fig. 14d–f) and Gardons (Fig. 14g–i) rivers. Differences exist between E3 and E4, although they are again weaker than differences between E1 and E3. Discharge forecasts from both ensembles are close for the Ardèche river (the northernmost catchment). For the Cèze river, and even more for the Gardons (the southernmost catchment), the E3 ensemble members produced a shifted discharge peak either too early or too late at 12:00 UTC and 20:00 UTC

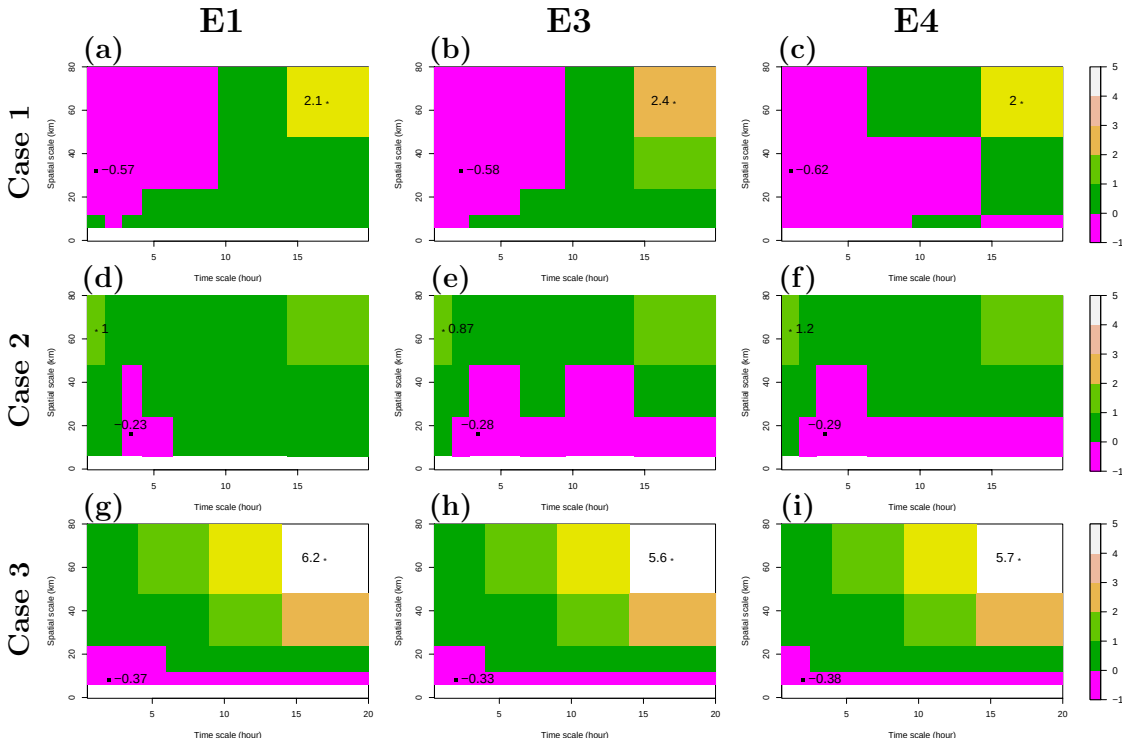


Fig. 12. Normalised anomalies of average forecast maximum intensity relative to the observed maximum intensity, for **(a–c)** Case 1, **(d–f)** Case 2 and **(g–i)** Case 3, and for the **(a,d,g)** E1, **(b,e,h)** E3 and **(c,f,i)** E4 ensembles. Extreme values are plotted on each diagram.

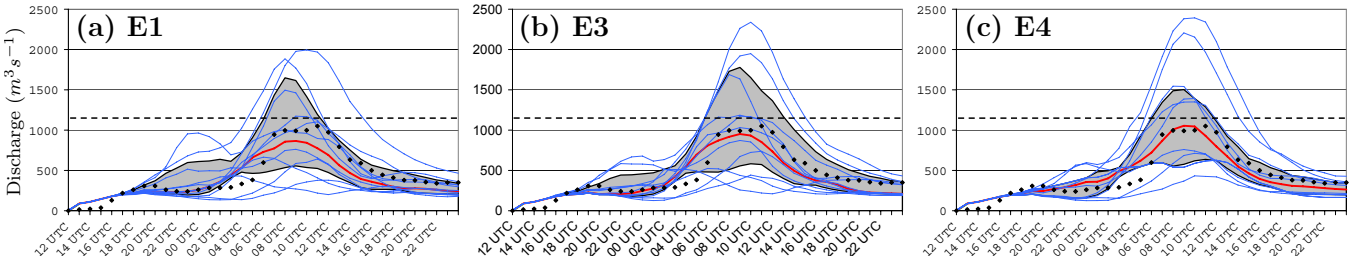


Fig. 13. Observed and forecast hydrological discharge for the Gardons river for Case 3, from 12:00 UTC 1 November 2008 to 00:00 UTC 3 November 2008, for the **(a)** E1, **(b)** E3 and **(c)** E4 ensembles. Hourly observed discharge is plotted as black diamonds, blue lines stand for each ensemble member, the ensemble median is shown in red. The grey shading represents the interquartile range. Dashed lines represent the discharge for a reference flood on each watershed.

22 October 2008 at Bagnols-sur-Cèze (Fig. 14e) and at 02:00 UTC and 12:00 UTC 22 October 2008 at Boucoiran (Fig. 14h). The E4 ensemble produces a single peak, around 16:00 UTC 22 October 2008 at Bagnols-sur-Cèze (Fig. 14f), and 10:00 UTC 22 October 2008 at Boucoiran (Fig. 14i). These differences affect more the southernmost catchment, which is where the convective system is regenerated, that is, where new convective cells are created. This could be explained by the role of the microphysical parameterisation during the triggering of convection, or their impact on mesoscale mechanisms, such as the development of a cold pool, as discussed in Sect. 4.1.

5 Conclusion

Previous research on convection permitting ensemble forecasts by Vié et al. (2011) assessed the impact of uncertainties separately on initial conditions and lateral boundary conditions. The present study focused on the combination of both uncertainty sources in a single ensemble, and on the assessment of part of model errors through the addition of

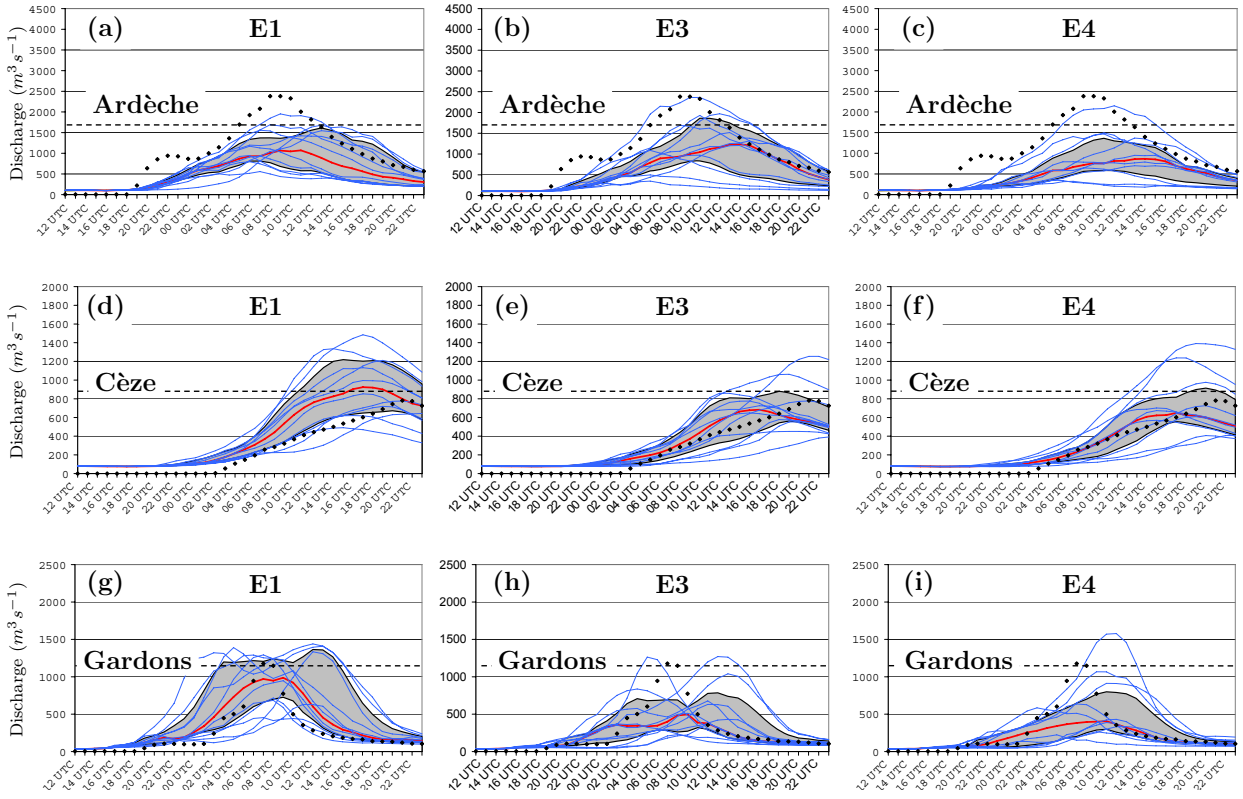


Fig. 14. Same as Fig. 13, for **(a–c)** the Ardèche, **(d–f)** the Céze and **(g–i)** the Gardons rivers, for Case 2, from 12:00 UTC 21 October 2008 to 00:00 UTC 23 October 2008, for the **(a,d,g)** E1, **(b,e,h)** E3 and **(c,f,i)** E4 ensembles.

microphysical perturbations. Following Fresnay et al. (2012), and focusing on HPEs, perturbations were applied to the auto-conversion, accretion and evaporation tendencies.

The ensembles were evaluated using probabilistic scores on an 18-day period, including three HPE case studies. For the three case studies, an innovative evaluation process was set up. Rainfall forecasts were evaluated in a scale-independent framework using maxima diagrams Ceresetti et al. (2012). Furthermore, an hydrological evaluation was conducted through ensemble discharge forecasts for typical Mediterranean coastal watersheds, as in Vincendon et al. (2011).

The statistical evaluation showed the benefit of accounting for both uncertainties on initial and lateral boundary conditions in the E3 ensemble. Probabilistic scores are either as good as, or better than any of the E1 and E2 ensembles taken separately. Moreover, a significant gain in ensemble spread has been found, especially for intermediate forecast ranges, when the E1 and E2 ensembles performed the worst. The statistical evaluation of the E4 ensemble showed no improvement over the E3 ensemble. An impact of the perturbation of the microphysical parameterisation was found for the HPE case studies. This impact is more important when the convective system involves complex mesoscale processes, such as a low-level cold pool on 20 October 2008, than for systems having stronger interactions with the synoptic-scale circulation (on 21–22 October and 1–2 November 2008). However, even for the HPE on 20 October 2008, these perturbations had a weaker impact than perturbations of either initial or lateral boundary conditions.

The hydro-meteorological evaluation of our ensembles on the three case studies confirms that the E1, E3 and E4 ensembles behave quite similarly. Maxima diagrams show that the perturbations used in the ensembles have no noticeable impact on the characteristics of the simulated precipitating systems. For Case 1, they confirm that the peak of precipitation is underestimated, and indicate that simulated convective cells do not organise into a well-formed convective system. For Cases 2 and 3, the forecast convective cells are slightly weaker than observed, but the precipitating system has a wider extension over the Cévennes mountains.

The hydrological discharge forecasts for HPEs on 21–22 October and 1–2 November 2008 also show small differences between ensembles. Differences are more important on 21–22 October 2008, especially for the southernmost watershed. This shows that, despite having little noticeable impact on the 24-h accumulated precipitation forecasts, the microphysical perturbations can affect the convective system initiation and development, at scales that are relevant for flash-flood forecasts. The impact of microphysical perturbations

seems located at the southern end of the precipitating system, where the convection is continuously regenerated by the triggering of new cells.

Overall, the three chosen microphysical perturbations, focusing on precipitation, have no impact on probabilistic scores. They may affect ensemble forecasts of heavy precipitating events, although not as much as initial or lateral boundary conditions. Another approach to sample the model error in ensemble forecasts is detailed in Bouttier et al. (2012). They use the ECMWF stochastic perturbation of physics tendencies (SPPT) scheme, adapted to convective-scale forecasts. They found an improvement in probabilistic scores as well as on case studies, even on low-level fields despite the lack of surface perturbations. To specifically enhance the spread at lower levels, it is planned to investigate perturbations of surface fields and surface-atmosphere fluxes. Further research also focuses on different parameterisations, such as the turbulence one. These scientific issues will be addressed in forthcoming studies.

Acknowledgements. This work was carried out in the framework of the MEDUP project (grant ANR-07-315 VULN-06-001), funded by the Vulnerabilité Milieux et Climat (VMC) programme of ANR. The authors thank the two reviewers for their comments and suggestions.

Edited by: P. Drobinski
Reviewed by: R. Rotunno and another anonymous referee



The publication of this article is financed by CNRS-INSU.

References

Berne, A., Delrieu, G., and Boudevillain, B.: Variability of the spatial structure of intense Mediterranean precipitation, Adv. Water Resour., 32, 1031–1042, 2009.

Berre, L., Stefanescu, S. E., and Belo Pereira, M.: The representation of the analysis effect in three error simulation techniques, Tellus A, 58, 196–209, 2006.

Boniface, K., Ducrocq, V., Jaubert, G., Yan, X., Brousseau, P., Masson, F., Champollion, C., Chéry, J., and Doerflinger, E.: Impact of high-resolution data assimilation of GPS zenith delay on Mediterranean heavy rainfall forecasting, Ann. Geophys., 27, 2739–2753, doi:10.5194/angeo-27-2739-2009, 2009.

Bouttier, F., Vié, B., Nuissier, O., and Raynaud, L.: Impact of stochastic physics in a convection-permitting ensemble, Mon. Wea. Rev., accepted, doi:10.1175/MWR-D-12-00031.1, 2012.

Buzzi, A., Tartaglione, N., and Malguzzi, P.: Numerical simulations of the 1994 Piedmont flood: Role of orography and moist processes, Mon. Wea. Rev., 126, 2369–2383, 1998.

Caniaux, G., Redelsperger, J., and Lafore, J. P.: A Numerical Study of the Stratiform Region of a Fast-Moving Squall Line. 1. General Description and Water and Heat Budgets, J. Atmos. Sci., 51, 2046–2074, 1994.

Ceresetti, D., Anquetin, S., Molinié, G., Leblois, E., and Creutin, J.-D.: Multiscale Evaluation of Extreme Rainfall Event Predictions Using Severity Diagrams, Wea. Forecasting, 27, 174–188, 2012.

Clark, A. J., Gallus Jr., W. A., Xue, M., and Kong, F.: A Comparison of Precipitation Forecast Skill between Small Convection-Allowing and Large Convection-Parameterizing Ensembles, Wea. Forecasting, 24, 1121–1140, 2009.

Clark, A. J., Gallus Jr., W. A., Xue, M., and Kong, F.: Growth of spread in convection-allowing and convection-parameterizing ensembles, Wea. Forecasting, 25, 594–612, 2010.

Clark, A. J., Kain, J. S., Stensrud, D. J., Xue, M., Kong, F., Coniglio, M. C., Thomas, K. W., Wang, Y., Brewster, K., Gao, J., Wang, X., Weiss, S. J., and Du, J.: Probabilistic Precipitation Forecast Skill as a Function of Ensemble Size and Spatial Scale in a Convection-Allowing Ensemble, Mon. Wea. Rev., 139, 1410–1418, 2011.

Cuxart, J., Bougeault, P., and Redelsperger, J.: A turbulence scheme allowing for mesoscale and large-eddy simulations, Quart. J. Roy. Meteor. Soc., 126, 1–30, 2000.

Ducrocq, V., Nuissier, O., Ricard, D., Lebeaupin, C., and Thouvenin, T.: A numerical study of three catastrophic precipitating events over southern France. II: Mesoscale triggering and stationarity factors, Quart. J. Roy. Meteor. Soc., 134, 131–145, 2008.

Fresnay, S., Hally, A., Garnaud, C., Richard, E., and Lambert, D.: Heavy precipitation events in the Mediterranean: Sensitivity to cloud physics parameterisation uncertainties, Nat. Hazards Earth Syst. Sci., in press, 2012.

Gebhardt, C., Theis, S. E., Paulat, M., and Ben Bouallègue, Z.: Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries, Atmos. Res., 100, 168–177, 2011.

Hohenegger, C. and Schär, C.: Predictability and error growth dynamics in cloud-resolving models, J. Atmos. Sci., 64, 4467–4478, 2007a.

Hohenegger, C., Walser, A., Langhans, W., and Schär, C.: Cloud-resolving ensemble simulations of the August 2005 Alpine flood, Quart. J. Roy. Meteor. Soc., 134, 889–904, 2008.

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

Lafore, J., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Hereil, P., Mascart, P., Masson, V., Pinty, J., Redelsperger, J., Richard, E., and de Arellano, J.: The Meso-NH atmospheric simulation system. Part I: Adiabatic formulation and control simulations, Ann. Geophys., 16, 90–109, 1998, http://www.ann-geophys.net/16/90/1998/.

Nuissier, O., Ducrocq, V., Ricard, D., Lebeaupin, C., and Anquetin, S.: A numerical study of three catastrophic precipitating events over southern France. I: Numerical framework and synoptic ingredients, Quart. J. Roy. Meteor. Soc., 134, 111–130, 2008.

Pergaud, J., Masson, V., and Malardel, S.: A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction, Bound.-Layer Meteor., 132, 83–106, 2009.

Ramos, M. H., Creutin, J.-D., and Leblois, E.: Visualization of storm severity, J. Hydrol., 315, 295–307, 2005.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, Mon. Wea. Rev., 139, 976–991, 2010.

Vincendon, B., Ducrocq, V., Nuissier, O., and Vié, B.: Flash-flood ensemble forecasts based on convective-scale NWP, Nat. Hazards Earth Syst. Sci., 11, 1529–1544, 2011, http://www.nat-hazards-earth-syst-sci.net/11/1529/2011/.

Vié, B., Nuissier, O., and Ducrocq, V.: Cloud-resolving ensemble simulations of Mediterranean heavy precipitating events: uncertainty on initial conditions and lateral boundary conditions, Mon. Wea. Rev., 139, 403–423, 2011.

Yan, X., Ducrocq, V., Jaubert, G., Brousseau, P., Poli, P., Champollion, C., Flamant, C., and Boniface, K.: The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: a case-study from COPS IOP 9, Quart. J. Roy. Meteor. Soc., 135, 1788–1800, 2009.