**Hydrology and
Earth System
Sciences**

Inclusion of potential vorticity uncertainties into a hydrometeorological forecasting chain: application to a medium size basin of Mediterranean Spain

A. Amengual, R. Romero, M. Vich, and S. Alonso

Grup de Meteorologia, Departament de Física, Universitat de les Illes Balears, Palma de Mallorca, Spain

Abstract. The improvement of the short- and mid-range numerical runoff forecasts over the flood-prone Spanish Mediterranean area is a challenging issue. This work analyses four intense precipitation events which produced floods of different magnitude over the Llobregat river basin, a medium size catchment located in Catalonia, north-eastern Spain. One of them was a devasting flash flood – known as the "Montserrat" event – which produced 5 fatalities and material losses estimated at about 65 million euros. The characterization of the Llobregat basin's hydrological response to these floods is first assessed by using rain-gauge data and the Hydrologic Engineering Center's Hydrological Modeling System (HEC-HMS) runoff model. In second place, the non-hydrostatic fifth-generation Pennsylvania State University/NCAR mesoscale model (MM5) is nested within the ECMWF large-scale forecast fields in a set of 54 h period simulations to provide quantitative precipitation forecasts (QPFs) for each hydrometeorological episode. The hydrological model is forced with these QPFs to evaluate the reliability of the resulting discharge forecasts, while an ensemble prediction system (EPS) based on perturbed atmospheric initial and boundary conditions has been designed to test the value of a probabilistic strategy versus the previous deterministic approach. Specifically, a Potential Vorticity (PV) Inversion technique has been used to perturb the MM5 model initial and boundary states (i.e. ECMWF forecast fields). For that purpose, a PV error climatology has been previously derived in order to introduce realistic PV perturbations in the EPS. Results show the benefits of using a probabilistic approach in those cases where the deterministic QPF presents significant deficiencies over the Llobregat river basin in terms of the rainfall amounts, timing and lo-

calization. These deficiences in precipitation fields have a major impact on flood forecasts. Our ensemble strategy has been found useful to reduce the biases at different hydrometric sections along the watershed. Therefore, in an operational context, the devised methodology could be useful to expand the lead times associated with the prediction of similar future floods, helping to alleviate their possible hazardous consequences.

1 Introduction

The Spanish Mediterranean area is affected every year by flood events of diverse spatial and temporal scales and, mainly in autumn, most of these can be considered as flash floods. The region contains complex mountainous systems along the coast which makes this area especially prone to such hazardous episodes (Fig. 1). These orographic systems act as natural barriers to the warm and moist Mediterranean air which, together with the early intrusion of Atlantic cold fronts, can induce the generation of quasi-stationary mesoscale convective systems (MCSs) with intense rainfall rates (Romero et al., 2000). Coastal steep streams and high urbanization rates mean short hydrological response times and potentially high personal and material damages. In addition, many small and medium steep streams are ephemeral in this semiarid environment and the increased flows over short periods – with associated high flow velocities and large volumes of sediments – can cause unexpected and extensive flood damage. Thus, it becomes an issue of the maximum regional interest to develop flood forecasting tools which can contribute to implement reliable warning systems before these kinds of events.
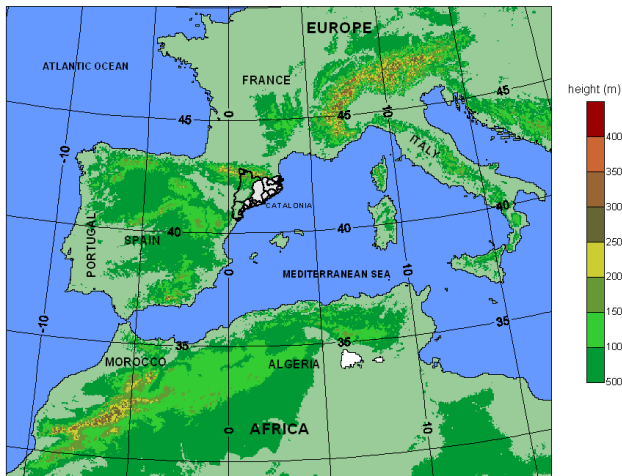
Fig. 1. Geographical locations and main mountain systems of the Western Mediterranean region. Major topographic features are shown according to the colour scale at 500 m intervals starting at 500 m. The thick continuous line shows Catalonia and its internal basins (highlighted with shaded grey).
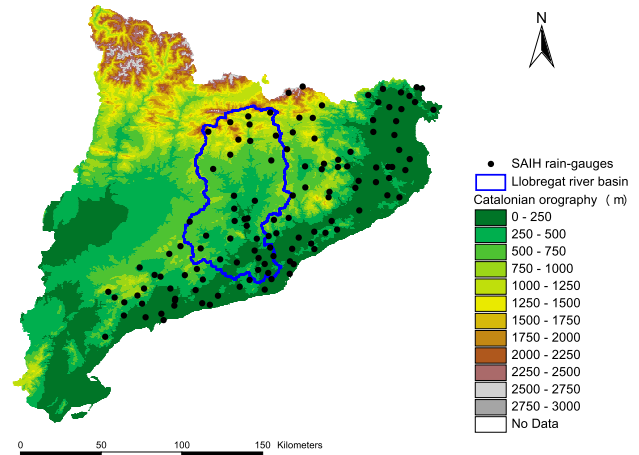

Fig. 2. Distribution of the rain-gauges from the Automatic Hydrological Information System (SAIH) in the internal basins of Catalonia (IBC; shown in Fig. 1). It includes a total of 126 automatic rainfall stations distributed over an area of 16 000 km^2. Llobregat basin is highlighted.

The short-time scales of flash flood events imply that the traditional warning systems based on hydrological models driven by rainfall observations do not provide the timely predictions required to implement the precautionary civil protection measures (Siccardi, 1996). Hence, to further extend the lead times associated to discharge predictions, the forecasting procedures are increasingly dependent upon high-resolution numerical weather prediction (NWP) models (with lead times up to 12–48 h; e.g. Ranzi et al., 2000; Anderson et al., 2002; Ferraris et al., 2002; Ducrocq et al., 2003; Benoit et al., 2003; Diomede et al., 2006). One of the most important challenges for NWP models nowadays is to improve the quantitative precipitation forecasts (QPFs) for hydrological purposes. For example, Younis et al. (2008) have shown that high-resolution operational weather forecasting in combination with a rainfall-runoff model could be useful to determine flash floods more than 24 h in advance.

The Llobregat river basin is a medium size catchment located in Catalonia, north-eastern Spain (Figs. 1 and 2), and it is periodically affected by intense precipitation events resulting in floods which sometimes produce hazardous effects. As a first objective of this work, we are interested in testing the feasibility of HEC-HMS discharge predictions driven by MM5 mesoscale model forecasts over the Llobregat river basin within a real-time hydrometeorological chain forecasting framework. Four notable hydrometeorological episodes occurred over the 1996–2004 period have been selected and the MM5 numerical weather model has been nested within large-scale meteorological forecasts provided by the European Center for Medium Range Weather Forecasts (ECMWF). One of the selected cases –known as the "Montserrat" flash flood event, on 9–10 June 2000 – is a good

example of the catastrophic effects related to such episodes. Many internal basins of Catalonia (IBC, Fig. 1) were overwhelmed as a result of the heavy precipitation produced by a quasi-stationary mesoscale convective system. The subsequent flood produced serious human and material damages over the Llobregat river basin (Llasat et al., 2003). In a previous work, Amengual et al. (2007) studied the "Montserrat" episode by using HEC-HMS runoff simulations driven by high-resolution precipitation fields provided by the MM5 mesoscale model, which was nested within NCEP analyses. This procedure was found useful to simulate the flash flood 24 h in advance with appropriate space and temporal scales.

The flood forecasting process is affected by several sources of uncertainty that should be taken into account for the design of realistic operational systems. These uncertainties arise from the hydrological and meteorological model formulations, their initial and boundary conditions, and in the gap which is still present between the scales resolved by each of the two systems. In addition, quantitative precipitation forecasting of intense episodes is a very difficult task because many factors combine in determining such events and most of the precipitation is of convective nature. NWP models have problems in developing and organizing the convection over the correct locations and times owing to the small-scale nature of many responsible atmospheric features and to their imperfect representation within these models (Kain and Fritsch, 1992; Stensrud and Fritsch, 1994a, b). As widely exposed in Stensrud and Yussouf (2007), the considerable difficulties surrounding the production of accurate QPFs together with the large social impacts of precipitation has led to a call for a focus on probabilistic QPFs (PQPFs) as an important strategy to provide user guidance. The use of an

ensemble of mesoscale model forecasts has become a valuable tool to further span the space of possible results. This is motivated by the fact that forecasts are sensitive to the uncertainties found in the initial and boundary conditions, model formulations, or both. Forecasting ensemble strategies can provide better short-range prediction guidance as long as they take into account these aspects. Several methods are being used to create perturbed initial conditions, such as breeding of growing modes (Toth and Kalnay, 1993), singular vectors (Buizza and Palmer, 1995; Molteni et al., 2006) and Monte Carlo approaches (Mullen and Baumhefner, 1988; Houtekamer and Derome, 1995: Du et al., 1997). The use of multi-physics ensemble strategies in order to take into account the model imperfections are widely described, for example, in Stensrud et al. (2000) and Jones et al. (2007).

Nowadays, many efforts are being devoted to explore the potential of using Ensemble Prediction Systems (EPS) for flood forecasting in order to extend the warning lead times for large river basins and to develop suitable probabilistic hydrological predictions techniques (de Roo et al., 2003; Gouweleeuw et al., 2005; Schaake et al., 2006, 2007). As an example, the European Flood Alert System (EFAS) is a transnational European effort to provide deterministic and probabilistic flood forecasting information to several regional hydrological services (Thielen et al., 2009; Bartholmes et al., 2009). Recent studies have shown that flood forecasting systems based on ensembles of quantitative precipitation forecasts or high-resolution rainfall simulations can also be of great value for medium and small size river basins (Roulin, 2007; Amengual et al., 2008; Diomede et al., 2008).

The second objective of this work is to study the impact on the hydrometeorological modelling chain of the uncertainties found in the initial and boundary conditions of the mesoscale model forecasts. For this purpose, an EPS has been generated by means of spatial shifts and variations of intensity of the upper-level precursor synoptic scale trough – a PV anomaly – for each case under study. Essentially, we try to encompass the uncertainties associated with an imprecise representation of the upper-level precursor troughs on the synoptic scale. With this method, it is possible to assess the effects on the driven runoff forecasts of the spatial and temporal uncertainties found in the QPFs. A previous PV error climatology has been derived in order to introduce realistic (i.e. with an appropriate range) PV perturbations in the EPS. Furthermore, the short-range prediction guidance value of a probabilistic hydrometeorological chain framework against a deterministic one in the Llobregat river basin can be assessed with this ensemble method.

The rest of the paper is structured as follows: Sect. 2 consists of a brief description of the study area and the rain and flow gauge networks; Sect. 3 describes the hydrometeorological episodes; Sect. 4 presents the hydrological tools used for the basin characterization and to set-up the runoff simulations; Sect. 5 contains the meteorological tools applied to forecast the intense rainfall events and to design the en-
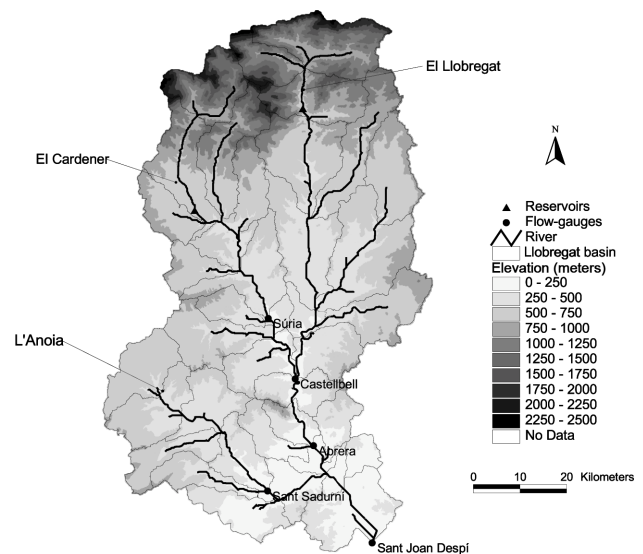

Fig. 3. Digital terrain model of the Llobregat river basin with a cell size of 50 m. It displays the basin segmentation, main tributaries, stream-gauges (circles) and reservoirs (triangles).

semble prediction system; Sect. 6 presents and discusses the results; and Sect. 7 provides an assessment of the presented methodology which includes future directions for the later development of the system.

2 The study area

2.1 Overview of the Llobregat river basin

Among the internal hydrographic catchments of Catalonia, the Llobregat river basin is the most important regarding to its size, river length, mean flow and population living inside (Figs. 2 and 3). This catchment is formed by the Llobregat river and its main tributaries, the Anoia and the Cardener. Llobregat basin extends from the Pyrenees – with heights over 3000 m, through the pre-Pyrenees – with a height transition from 750 m to 200 m in the pre-coastal range. The last section of the river flows through the Mediterranean orographic systems which consist of two mountainous aligments almost parallel to the coast line: the pre-coastal and coastal ranges. The basin has a total drainage area of 5040 km^2 and a maximum length close to 170 km (for a more detailed description of the watershed see Amengual et al., 2007).

The Llobregat river basin contains a broad range of climatological regimes expressed on quite diverse annual rainfalls depending on the altitude. These rainfalls can range from values exceeding 1000 mm in the Pyrenees (over 1000 m), 700 mm over the pre-Pyrenees (with elevations between 600 and 1000 m) and hardly reaching 500 mm for altitudes below 500 m. The rainfall regime is typical of the western Mediterranean climate, with several heavy-rainfall episodes per year

occurring mainly in autumn and occasionally in the other seasons. These intense daily precipitation events can represent a large fraction of the annual amounts.

2.2　The rain and stream gauge networks

Validation of QPFs is not always straightforward, in particular for exceptional events. If rain-gauge networks are not dense enough, these are not able to resolve the small-scale features of the highly variable precipitation fields driving floods. Furthermore, a point comparison among the observed and simulated rainfall fields is not always appropriate for hydrological purposes that emphasize an integrated value over the watershed surface (Chancibault et al., 2006). In this study, the spatial and temporal distributions of the forecasted rainfall fields are evaluated against the observed rainfall patterns at the subcatchment scale by using a set of continuous and categorical verification indices using the sub-basins as rainfall accumulation units.

Available raw precipitation consists of 5-min rainfall data recorded at 126 stations inside the Internal Basins of Catalonia (IBC) and distributed over an area of 16 000 km² (Fig. 2). These stations belong to the Automatic Hydrological Information System (SAIH) network of the Catalan Agency of Water (ACA). Out of the 126 stations, about 40–50 lie inside the Llobregat basin or near its boundaries.

The flow forecasts resulting from the one-way coupling between the meteorological and hydrological models will be compared against the rain-gauge driven runoff simulations, thus employing the hydrological model as an advanced validation tool. This approach has been found especially suitable for the evaluation of high-resolution simulated precipitation fields (Benoit et al., 2000; Jasper and Kaufmann, 2003; Chancibault et al., 2006; Amengual et al., 2008).

Runoff in the Llobregat river basin is available at five flow gauges (Fig. 3) also integrated in the SAIH network, located in: (i) Súria town, on the Cardener river, with a dranaige area of 940 km² and elevation from 250 m at gauge level to 2350 m in the Pyrenees; (ii) Sant Sadurní d'Anoia city, on the Anoia river, with a drainage area of 736 km² and elevations from 125 m at gauge level to 850 m at headwater; and (iii) Castellbell (3340 km²), (iv) Abrera (3587 km²) and (v) Sant Joan Despí (4915 km²) towns along the Llobregat river. Unfortunatelly, only a partial amount of the stream-flow data is available for this study owing to the malfunction of some gauges for some of the analysed hydrometeorological episodes.

3　Description of the hydrometeorological episodes

We have selected four intense precipitation episodes which resulted in floods of remarkable magnitude from 1996, when the SAIH network was deployed, until 2004. These case studies are a sample of different intense rainfall episodes pro-

duced by prototype atmospheric dynamical situations. The first two cases were linked with sustained and stratiform-like rainfalls over long periods of time, which resulted in remarkable discharges at the Llobregat basin outlet. The last two cases produced important and sudden rising flows owing to their convective nature, bearing heavy precipitations in short temporal scales. Next, the causative atmospheric patterns as well as the characteristics of the rainfall distributions and subsequent floods are briefly summarized.

3.1　Case 1: 16–17 November 1996

This episode involved the intrusion of an Atlantic surface cold front and a positively tilted upper-level trough that contributed to the generation of a mesoscale cyclone over the western Mediterranean Sea. The cyclone was located off the Catalonian coast and advected warm and moist air toward Catalonia (Fig. 4a and b). The whole episode led to maximum precipitation values above 100 mm inside the Llobregat catchment from 14 November 00:00 UTC to 18 November 00:00 UTC, though the most of the amounts were concentrated on the last two days, with maxima close to 80 mm inside the basin. This was an example of sustained stratiform-like precipitation rates over a 48 h period distributed between the pre-coastal and the Pyrenees ranges (Fig. 5a).

The maximum flow discharge recorded at Súria was close to 133 m³ s⁻¹ at 19:40 UTC on 17 November 1996. In the Anoia affluent, only a moderate flow increase was observed with a maximum of about 67 m³ s⁻¹ at 10:45 UTC. In Castellbell town, a remarkable peak discharge up to 972 m³ s⁻¹ at 23:10 UTC was produced owing to the important rainfall amounts registered in the north-western and central areas of the basin (figures not shown). In Abrera, data was not available and, in Sant Joan Despí, a maximum peak discharge of 1250 m³ s⁻¹ was recorded on 18 November 1996 at 03:15 UTC (Fig. 6a).

3.2　Case 2: 17–18 December 1997

This event was characterized by a deep depression centered to the north-west of Spain and a general south-westerly flux at all levels (Fig. 4c). In Catalonia, there was warm and moist advection from the south-southeast that favoured important precipitations on the south-faced ranges of the Pyrenees owing to the important orographic forcing (Fig. 4d). Precipitation amounts reached 232 mm from 15 to 19 December 1997 00:00 UTC, albeit maximum quantities were observed during the last 48 h period. From 17 to 19 December 00:00 UTC, the Llobregat river basin collected a maximum amount up to 202 mm and a 24 h accumulation above 120 mm. Maximum rainfall affected the northern part of the watershed (Fig. 5b). Again, this flood event was linked to sustained stratiform-like precipitation.

The maximum peak flow recorded at Súria town was 179 m³ s⁻¹ on 18 December 1997 at 18:15 UTC (figure not
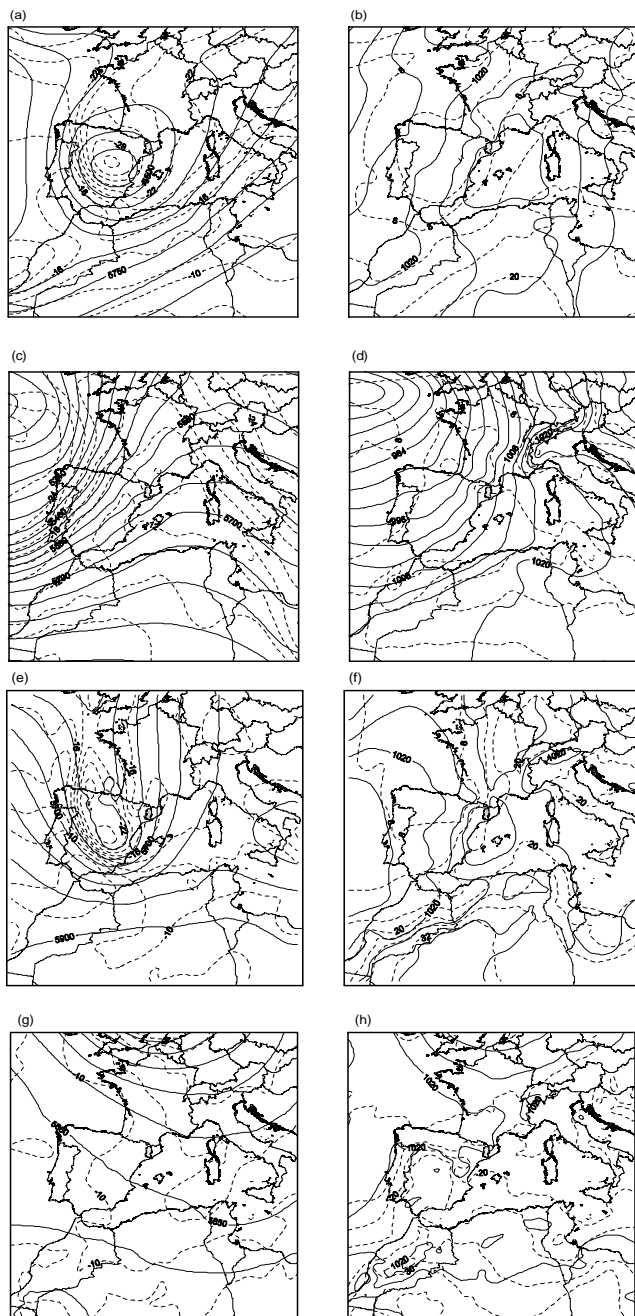
Fig. 4. ECMWF forecast maps. Geopotential height (continuous line, in gpm) and temperature (dashed line, in °C) at 500 hPa for: **(a)** 17 November 1996 at 00:00 UTC, **(c)** 18 December 1997 at 00:00 UTC, **(e)** 10 June 2000 at 00:00 UTC and **(g)** 30 August 2004 at 00:00 UTC. Sea level pressure (continuous line, in hPa) and temperature at 925 hPa (dashed line, in °C) for: **(b)** 17 November 1996 at 00:00 UTC, **(d)** 18 December 1997 at 00:00 UTC, **(f)** 10 June 2000 at 00:00 UTC and **(h)** 30 August 2004 at 00:00 UTC. The Llobregat river basin is highlighted in shaded grey.
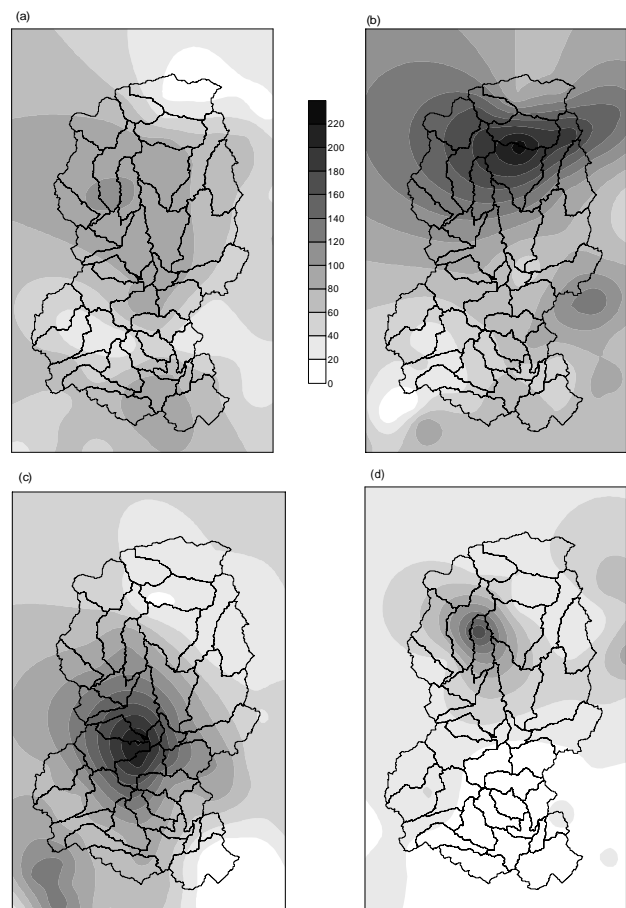

Fig. 5. Observed accumulated precipitation in the Llobregat basin (in mm according to the scale) for: **(a)** 16–17 November 1996, **(b)** 17–18 December 1997, **(c)** 9–10 June 2000 and **(d)** 29–30 August 2004 episodes.

shown). At Castellbell stream-gauge, the peak discharge was $503 \, \text{m}^3 \, \text{s}^{-1}$ at 20:00 UTC (Fig. 6b). At Anoia, Abrera and Sant Joan Despí flow-gauges, no data were available on this case.

3.3 Case 3: 9–10 June 2000

The synoptic situation for this case was very similar to our first case study. The approach of an Atlantic low-level cold front and a negatively tilted upper-level trough contributed to the generation of a mesoscale cyclone over the Mediterranean Sea between the Catalonian coast and the Balearic Islands (Fig. 4e and f). The subsequent advection of warm and humid air from the Mediterranean together with the Atlantic front and the complex orography of the region resulted in the generation of a quasi-stationary convective system (Martín et al., 2007). Hence, the most hazardous heavy rainfall lasted about six hours on 10 June 2000, from 00:00 to 06:00 UTC. The extraordinary rise of the Catalonia internal river basin flow regimes produced serious damages. Some of the most
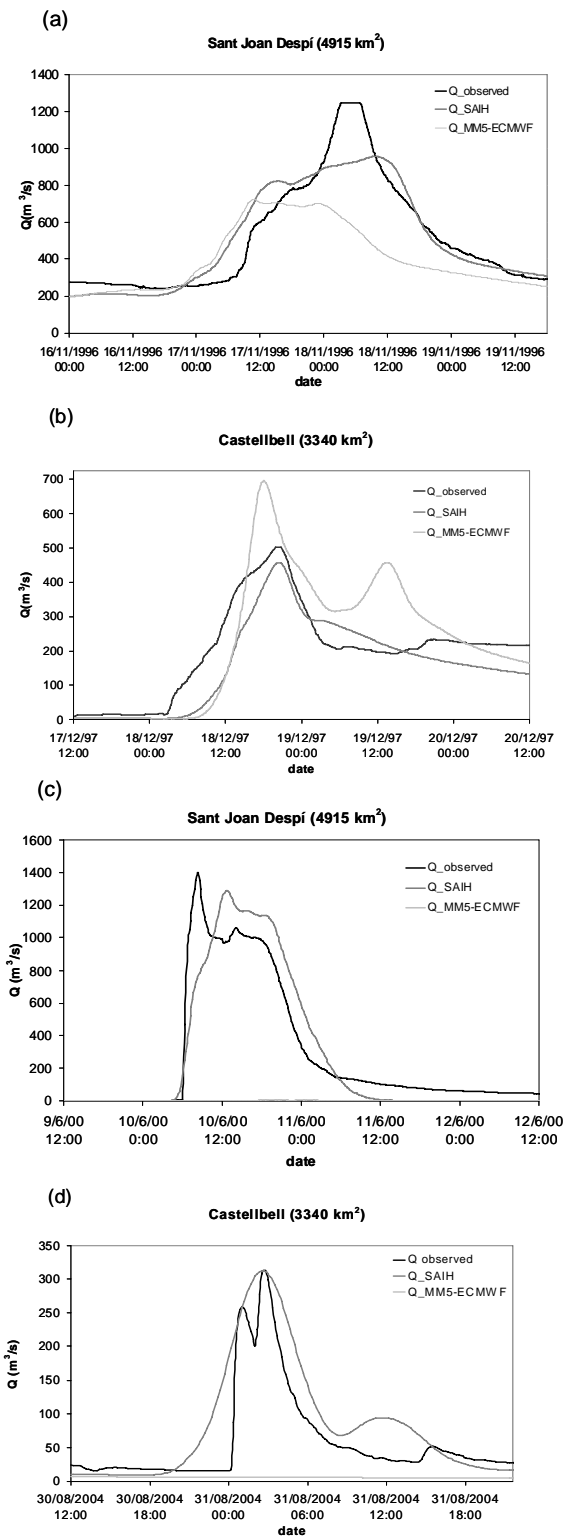
Fig. 6. Observed, SAIH rain-gauge driven, and MM5 control simulation driven runoff discharge for: **(a)** 16–17 November 1996 episode, **(b)** 17–18 December 1997 episode, **(c)** 9–10 June 2000 episode and **(d)** 29–30 August 2004 episode. The hydrographs are shown at the stream-gauges indicated in Table 1.

notable disasters consisted of the partial destruction of the infrastructure of a monastery in Montserrat's mountain (720 m) where about 500 people had to be evacuated. The episode caused five fatalities and material losses estimated at about 65 million euros (Llasat et al., 2003).

The most remarkable hydrometeorological feature of this case – known as the "Montserrat" flash flood event – was the high intensity of the convective rainfall: the accumulated hourly quantities were above 100 mm with a 6 h maximum up to 224 mm, mainly affecting the pre-coastal ranges (Fig. 5c). The maximum discharge observed at Súria was $260 \, \mathrm{m^3 \, s^{-1}}$ at 10:25 UTC on 10 June 2000. In the Anoia affluent, the peak discharge was $270 \, \mathrm{m^3 \, s^{-1}}$ at 04:45 UTC. In Castellbell town, several peak discharges were observed and the maximum of these reached $1000 \, \mathrm{m^3 \, s^{-1}}$ at 06:00 UTC. In Abrera, the peak discharge was close to $1200 \, \mathrm{m^3 \, s^{-1}}$ at 10:50 UTC (figures not shown). Finally, in Sant Joan Despí city, the maximum peak flow was up to $1400 \, \mathrm{m^3 \, s^{-1}}$ at 08:15 UTC (Fig. 6c).

3.4 Case 4: 29–30 August 2004

Figure 4g and h depicts a synoptic situation not clearly defined and with a weak dynamical forcing. Relatively cold air was present at mid-upper tropospheric levels over Catalonia. At low levels, a relative thermal low located in the southwestern part of Iberian Peninsula and its associated meso-low over the Ebro valley resulted in an easterly flow over Catalonia. Then, the warm and humid Mediterranean air together with the cold air aloft was enough to trigger a very localized and intense convective storm. This synoptic situation is typical of the warm season stormy conditions in Mediterranean Spain. This episode was characterized by accumulated rainfalls over 180 mm upon the Llobregat basin. Again, high intensities of mostly convective precipitation were found, with hourly amounts close to 51 mm and 6 h maximum quantities up to 115 mm from 14:00 to 20:00 UTC on 30 August (Fig. 5d).

With regard to the observed runoff, the maximum peak flow was close to $150 \, \mathrm{m^3 \, s^{-1}}$ on 30 August 2004 at 23:25 UTC for Súria. In Castellbell town, the peak flow was $314 \, \mathrm{m^3 \, s^{-1}}$ on 31 August 2004 at 02:40 UTC (Fig. 6d). Owing to the high spatial localization of this episode, no remarkable flow was observed at the Anoia gauge (the peak discharge was only about $15 \, \mathrm{m^3 \, s^{-1}}$). Unfortunately, in Abrera and Sant Joan Despí, no data were recorded.

4 Hydrological tools

4.1 Rainfall-runoff model and input data

This study is carried out using the physically-based HEC-HMS rainfall-runoff model (USACE-HEC, 1998). The model has been implemented in a semi-distributed and event-based configuration. HEC-HMS utilizes a graphical interface to build the semi-distributed watershed model and to set up

rainfall and control variables for the simulations. Figure 3 depicts the digital terrain model for the Llobregat river basin – with a cell resolution of 50 m – together with the main watercourses and its tributaries, the considered division in subbasins and the location of the available river gauges. After the sensitivity analysis presented in Amengual et al. (2007), the modelled basin was segmented into 39 sub-basins with an average size of 126 km^2 and a total extension of 4915 km^2 upstream from Sant Joan Despí, where the last flow-gauge is installed.

The hydrological model is forced by using a single hyetograph for each sub-basin. Rainfall spatial distributions are first generated from hourly accumulated values recorded at SAIH rain-gauges by applying the kriging interpolation method with a horizontal grid resolution of 1000 m. Then, the hourly rainfall series are calculated for each sub-basin as the area-averaged of the gridded rainfall within each subcatchment. The same methodology is used to assimilate QPFs in HEC-HMS, except that atmospheric model grid point values are used instead of SAIH observations. The kriging analysis method has been used by applying a linear model for the variogram fit (variogram slope = 1, anisotropy ratio = 1, anisotropy angle = 0, no nugget effect). This minimal error variance method is recommended for irregular observational networks and has been commonly used to compute rainfall fields from rain-gauges (Krajewski, 1987; Bhagarva and Danard, 1994; Seo, 1998).

4.2 Theoretical background and basin calibration

The rainfall-runoff model calculates runoff volumes by subtracting from rainfall the water volume that is lost through interception, infiltration, storage, evaporation and transpiration. The loss rate is calculated using the Soil Conservation Service Curve Number (SCS-CN, US Department of Agriculture 1986). This method assumes the storm runoff volumes to be proportional to the rainfall volumes exceeding an initial abstraction threshold, through the ratio of the accumulated infiltration to a storage capacity. With this assumption and the continuity principle the cumulative volume of stormflow becomes nonlinearly related to the excess rainfall volume which is a function of cumulative rainfall, soil cover, land use and antecedent moisture (Chow et al., 1988; Bacchi et al., 2002). The SCS-CN model has been tested on several experimental areas and river basins worldwide and, in Catalonia has been adopted by ACA for their technical studies (ACA, 2001, 2003). A synthetic unit hydrograph (UH) provided by SCS is used to convert rainfall excess into direct runoff on a watershed. The SCS-UH relates the peak discharge with the time to the UH peak through the subbasin area and a conversion constant. Baseflow is calculated by means of an exponential recession method to explain the drainage from natural storage in the watershed (Linsley at al., 1982). The flood hydrograph is routed using the Muskingum method (Chow et al., 1998; USACE-HEC, 2000).

The Llobregat basin contains two reservoirs located in the upstream areas of the Cardener affluent and the Llobregat river (Fig. 3). Therefore, these reservoirs have been modeled using the technical characteristics – storage capacity, maximum outflow, maximum elevation and initial level – available from the aforementioned technical reports issued by ACA. The detention ponds have been modelled by introducing a reservoir element that follows the elevation-storage-outflow relationship series which depends on the characteristics of the dam, the outlet and the spillway, besides the initial elevation of the water level.

The calibration of the rainfall-runoff model is carried out using six events: the four episodes under study and two additional cases available for the period comprised between 1996 and 2004. Table 1 shows the selected flood events, the hydrological model simulation periods and the observed maximum precipitation and flow inside the Llobregat river basin. The malfunction of the flow-gauge network at Abrera and Sant Joan Despí for some of these episodes (Table 1) has limited to some extent the calibration of the lower part of the catchment. In addition, the low available number of high magnitude flood cases is a problem for the calibration process. To mitigate this problem, it has been applied a cross-calibration methodology, i.e. for each of our case studies, the rest of available hydrometeorological episodes are used to calibrate the model. Thus, in effect, four different calibrations of the model are performed, using, in each case, fully independent flood events.

Calibration of the infiltration parameters for each independent episode combines a manual procedure, where the SCS curve numbers are derived from field measurements and normal antecedent moisture conditions (ACA, 2001), and an automatic procedure. The automatic procedure uses as objective function the peak-weighted root mean square error and applies the univariate-gradient search algorithm method (USACE-HEC, 2000). In addition, the flood wave celerity for the main streams is considered as a calibration index – by means of the K parameter – owing to the nature of these kinds of episodes characterized by high flow velocities. Baseflow recession parameters have also been calibrated.

The rainfall-runoff model has been run for Case 1 from 14 to 20 November 1996 00:00 UTC. The simulation for Case 2 has been extended from 15 to 22 December 1997 00:00 UTC. The 'Montserrat' event has been modelled with a 96 h simulation, from 9 to 13 June 2000 at 00:00 UTC. And finally, Case 4 has been run from 29 August to 2 September 2004 00:00 UTC. These simulation periods completely encompass the flood events and the subsequent hydrograph tails. All the experiments have been run with a 10 min time-step, first forced by the SAIH rain-gauge data, and then by the MM5 deterministic and probabilistic QPFs (Sect. 6). Note that owing to the large duration of the rainy periods for the first two cases, the rainfall-runoff model has been run with the antecedent 48 h rain-gauge data before to force it with the QPFs in order to optimize the initial conditions of the basin.

Table 1. Summary of the episodes used for the calibration of the hydrologic model. Note that the observed flow at the basin outlet in Sant Joan Despí is not available for the second, fifth and sixth cases.

| Flood event	Simulation period	Maximum obs. rainfall (mm)	Maximum obs. flow (m³ s⁻¹)
16–17 Nov 1996	14/11/1996–20/11/1996 00:00 UTC (144 h)	102.1	1250.0 (Despí)
17–18 Dec 1997	15/12/1997–22/12/1997 00:00 UTC (168 h)	232.0	502.7 (Castellbell)
9–10 Jun 2000	09/06/2000–13/06/2000 00:00 UTC (96 h)	224.0	1400.0 (Despí)
19–20 Oct 2001	17/10/2001–24/10/2001 00:00 UTC (168 h)	84.5	254.4 (Despí)
4–5 Dec 2003	03/12/2003–07/12/2003 00:00 UTC (96 h)	63.6	436.9 (Castellbell)
29–30 Aug 2004	29/08/2004–02/09/2004 00:00 UTC (96 h)	178.3	313.5 (Castellbell)
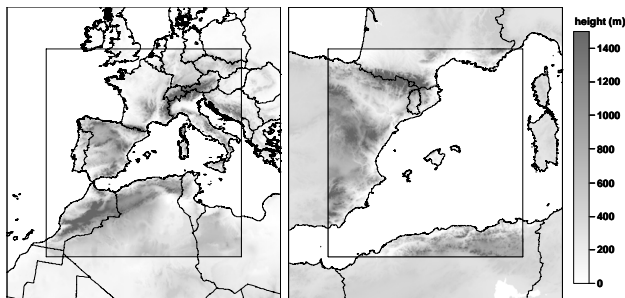
Fig. 7. Configuration of the two computational domains used for the MM5 numerical simulations (inner squares, with horizontal resolutions of 22.5 and 7.5 km for left and right, respectively). The Llobregat river basin is highlighted as thick line.

5   Meteorological tools

The non-hydrostatic MM5 numerical model is used to perform the meteorological simulations. It is a high resolution short-range weather forecast model developed by the Pennsylvania State University (PSU) and the National Center for Atmospheric Research (NCAR) (Dudhia, 1993; Grell et al., 1995). The model domains are configured as in the real-time operational version used at the University of the Balearic Islands (UIB; see http://mm5forecasts.uib.es). Simulations are designed using 30 vertical σ-levels and two spatial domains with 121×121 grid points (Fig. 7). Their respective horizontal resolutions are 22.5 and 7.5 km. In particular, the second domain spans the entire Catalonian territory and contiguous land and oceanic areas, and is used to supply the high-resolution rainfall fields to drive the hydrologic simulations. The interaction between the domains follows a two way nesting strategy (Zhang et al., 1986).

To initialize the model and to provide the time-dependent boundary conditions, the European Center for Medium Range Weather Forecasts (ECMWF) large-scale meteorological grid forecasts are used, with a spatial resolution of 0.3^0 and an update frequency of 6 h. The deterministic MM5 experiments (also labelled as 'control' simulations) for the four hydrometeorological episodes follow the same physics op-

tions as in the UIB operational runs. To parameterise moist convective effects, the Kain-Fritsch parametrization scheme (Kain, 2004) is used in both domains. Explicit microphysics are represented in both domains with prediction equations for cloud and rain water fields, cloud ice and snow allowing for slow melting of snow, supercooled water, graupel and ice number concentration (Reisner et al., 1998). The planetary boundary layer physics is formulated using a modified version of the Hong and Pan scheme (Hong and Pan, 1996). Surface temperature over land is calculated using a force-restore slab model (Blackadar, 1979; Zhang and Anthes, 1982) and over sea it remains constant during the simulations. Finally, long- and short-wave radiative processes are formulated taking into account the cloud cover (Benjamin, 1985).

With the purpose of generating the ensemble of perturbed runs, the invertibility principle of Ertel potential vorticity (PV; Hoskins et al., 1985) is applied. In particular, we are interested in studying the sensitivity of the hydrometeorological events to uncertainties in the precise representation of the upper-level precursor troughs on the synoptic scale for each hydrometeorological episode, being aware that small scale aspects of the circulation are propitious to analysis or forecast errors. The piecewise PV inversion scheme described in Davis and Emanuel (1991) is then used as a clean approach to manipulate the upper-level synoptic troughs in the model initial and boundary conditions. What it is necessary is a simple identification of the anomalous PV signatures of the troughs and then the balanced mass and wind fields associated with that PV element can be used to alter the meteorological fields in a physically consistent way (effectively, a change in the structure or position of the troughs). This method has already shown its value for assessing the predictability of extreme floods in the western Mediterranean area (e.g. Romero, 2001; Homar et al., 2002; Romero et al., 2006).

In order to introduce realistic perturbations in the EPS, a previous PV error climatology (PVEC) has been derived. This climatology allows to introduce the perturbations on the ECMWF forecast PV fields in the appropriate error range. The PVEC is calculated using a large collection of MEDEX cyclones (MEDEX, MEDiterranean EXperiment on cyclones that produce high impact weather in the Mediterranean;
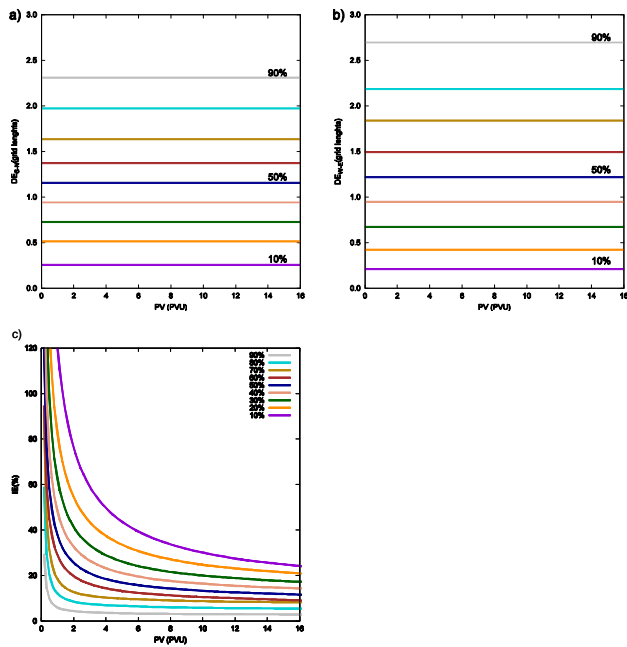
Fig. 8. Percentile levels at 300 hPa of: **(a)** south-north displacement PV errors, **(b)** west-east displacement PV errors and **(c)** PV intensity errors.

further information at http://medex.inm.uib.es). In particular, 19 cyclonic episodes – which comprise 56 days covered every 6 h by ECMWF grid analyses and forecasts – are used to provide the displacement and intensity error climatologies of the PV fields in the study region at the standard pressure levels (i.e. 100, 200, 300, 400, 500, 700, 850, 925 and 1000 hPa).

The displacement error (DE) corresponds to the minimum displacement of the ECMWF 24 h forecasts PV field necessary to obtain the maximum local correlation with the ECMWF analysis PV field (referred to as "optimal displacement"). Displacements of the forecast PV field local matrix – which corresponds to a domain of 21×21 grid points or 450×450 km – in all directions up to 10 grid points (i.e. 225 km) are tested. The minimum displacement which shows a local maximum correlation between the displaced forecast and analysis fields is assigned to the matrix central grid point as the corresponding displacement error. This method is repeated for each domain grid point along with its associated local matrix. Figure 8a and b displays the statistics of displacement errors at 300 hPa as percentile levels. These figures only show the results for the south-north (and west-east) directions owing to the symmetry found between the S-N (W-E) and N-S (E-W) displacement errors. It is also worth to note the independency of DE on the PV values. DE percentile levels have been adjusted to a constant function of the PV values.

The intensity error (IE) corresponds to the local difference between the optimally displaced ECMWF 24 h forecast PV field and the ECMWF analysis field. It is computed for each domain grid point as the difference in the 21×21 grid points PV average between the aforementioned optimally displaced forecast and analysis fields and it is expressed as an error percentage relative to the analysis PV average. Figure 8c shows the percentile levels of intensity errors at 300 hPa. It is found a symmetry between positive and negative values (not shown), and therefore, absolute values are used to formulate the intensity error. For each isobaric surface, IE percentile levels have been adjusted to a potential function of the PV values. It has been imposed a superior limit to the intensity error corresponding to $IE_{max} = 200\%$, since the relative error tends to infinity for PV tending to zero.

Finally, the PV error climatology (DE and IE percentiles) is used to implement the EPS by randomly perturbing the PV field in the ECMWF forecasts used to nest the MM5 simulations. The perturbations are applied along the areas with the most intense PV values and gradients, usually found within and around the upper-level precursor troughs.

This ensemble of simulations for each hydrometeorological event is a first approximation to the problem of incorporating the spatial and temporal uncertainties of the rainfall forecasts into a medium size catchment as the Llobregat basin. The MM5 simulations cover a 54 h period for each of the flood episodes under study, starting at 00:00 UTC 16 November 1996 for Case 1; 00:00 UTC 17 December 1997 for Case 2; 00:00 UTC 9 June 2000 for Case 3; and 00:00 UTC 29 August 2004 for Case 4.

6 Results and discussion

6.1 Runoff simulations driven by SAIH rain-gauge data

SAIH rain-gauge derived rainfalls of the four hydrometeorological events (recall Fig. 5) are used to drive the calibrated HEC-HMS model for each episode in a single evaluation runoff simulation according to the methodology described in Sect. 4. The skill of the resulting runoff simulations is expressed in terms of the Nash-Sutcliffe efficiency criterion (NSE; Nash and Sutcliffe, 1970), a "goodness-of-fit" measure widely used in hydrological model validation (Dolciné et al., 2001; Jasper et al., 2003). The performance of the runoff simulations is also checked by means of the relative error of total volume at flow-gauge sites, expressed as percentage (%EV). %EV>0 and %EV<0 would indicate an over and underestimation of the volume by the model, respectively. In addition, the relative error in percentage of the peak discharge has also been calculated (%EP). Hereafter, the results are only commented at the stream-gauges indicated in Table 1 in order to summarize the discussion.

It is worth to note that the rainfall-runoff model reasonably simulates the hydrological response of the basin in terms of

Table 2. NSE efficiency criterion, percentage of error in volume (%EV) and peak flow (%EP) for the calibration episodes at the stream-gauges indicated in Table 1.

Flood event	NSE	%EV	%EP
16–17 Nov 1996	0.87	1.3	−21.7
17–18 Dec 1997	0.84	−21.3	−0.9
9–10 Jun 2000	0.95	−3.4	−12.2
19–20 Oct 2001	0.55	10.3	4.2
4–5 Dec 2003	0.77	−25.1	−3.4
29–30 Aug 2004	0.43	−17.6	−29.4

the observed peak flows, runoff volumes and peak flows for most of the six hydrometeorological events used for the calibration (Table 2). Table 3 shows the statistical skill scores of the single evaluation rain-gauge driven runoff simulations when compared against the observed flows for the study cases (Fig. 6). In particular, NSE widely exceeds 0.8 for the 16–17 November 1996 episode with only a very slight underestimation of the runoff volume. However, SAIH driven runoff simulation does not capture satisfactorily the observed peak discharge and its underestimation is remarkable (Table 3, Fig. 6a). With regard to the 17–18 December 1997 experiment, the peak discharge is slightly underestimated and the time to peak is well simulated, but it can also be observed a noticiable error in %EV. Although NSE score is penalized due to this last feature, it still overpasses 0.7 (Table 3 and Fig. 6b). With regard to the "Montserrat" event, rain-gauge driven runoff simulation yields a suitable performance: NSE score is above 0.8 together with small relative errors in the volume runoff and the peak flow. These results point out a goodness-of-fit for the main peak discharge, even though a significant delay occurs in its timing (Fig. 6c). Finally, the 29–30 August 2004 episode depicts the worst statistical indices in terms of NSE and %EV among the whole set of selected floods. NSE is slightly below 0.6 and a clear overestimation in the simulated flow volume can be observed (above 40%). This wide overestimation in the rain-gauge driven runoff volume clearly is against the NSE score, since the maximum peak discharge and its timing are well estimated. However, the evaluation run only simulates an envelop of the two observed peaks (Table 3 and Fig. 6d).

6.2 Runoff simulations driven by MM5 deterministic QPFs

Prior to analysing the runoff simulations, a set of continuous verification indices have been used to evaluate the skill of the MM5 deterministic runs over the Llobregat river basin (Fig. 9). The spatial and temporal distributions of the forecasted rainfall volumes have been compared against the observed precipitation volumes by using the correlation coefficient (r), root-mean-square error (RMSE) and mean abso-
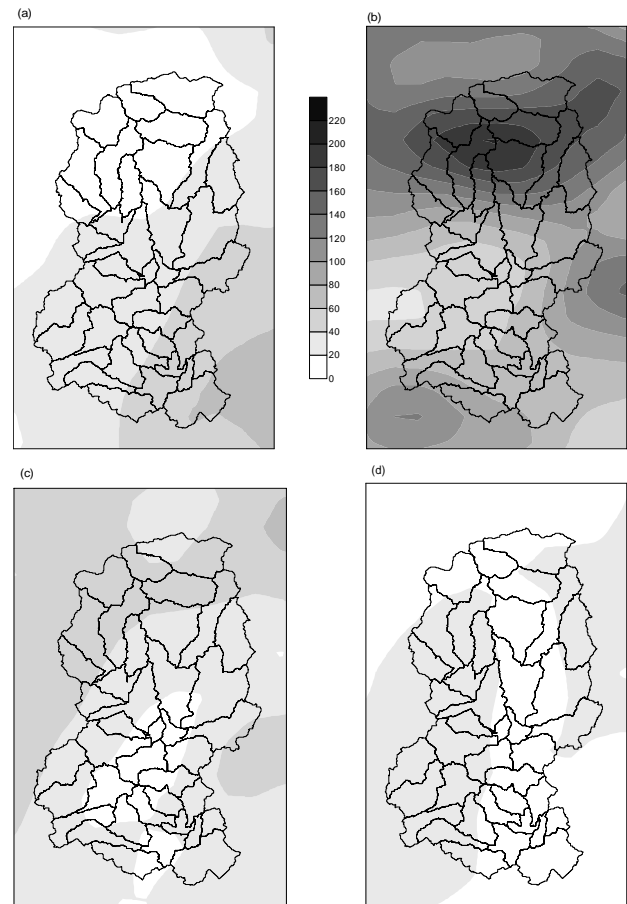

Fig. 9. Spatial distribution of the accumulated precipitation for the MM5 control forecast simulations in the Llobregat basin (in mm according to the scale) for: **(a)** 16–17 November 1996, **(b)** 17–18 December 1997, **(c)** 9–10 June 2000 and **(d)** 29–30 August 2004 episodes.

lute error (MAE) statistical scores. To evaluate the spatial distributions, the 39 sub-basins of the Llobregat catchment are used as areal accumulation units for the precipitation volumes. Thus, the statistical scores provide information on the general performance of the mesoscale model to forecast the cumulative rainfall volumes at the sub-basin scale. The temporal distributions are computed using hourly rainfall volumes over the whole basin. These hourly discretizations were found optimum to model the times of concentration of the basin when it is affected by intense rainfalls (Amengual et al., 2007).

In terms of the spatial and temporal distributions and the maximum amounts of precipitation, it appears that only the 17–18 December 1997 experiment shows a good similarity with the observed pattern over the Llobregat river basin, although the forecasted maximum cover a wider area (compare Figs. 5b and 9b). Therefore, the spatial and temporal correlation coefficients increase owing to this pattern agreement,

Table 3. NSE efficiency criterion, percentage of error in volume (%EV) and peak flow (%EP) for the SAIH rain-gauge driven and the MM5 control driven runoff simulations. Statistical indices are shown for the flow-gauges indicated in Table 1.

Flood event	NSE SAIH	%EV SAIH	%EP SAIH	NSE MM5 control	%EV MM5 control	%EP MM5 control
16–17 Nov 1996	0.86	−1.1	−23.6	0.34	−21.6	−42.2
17–18 Dec 1997	0.72	−23.2	−9.1	0.43	6.7	38.2
9–10 Jun 2000	0.84	1.1	−7.8	−0.05	−99.7	−99.7
29–30 Aug 2004	0.57	41.2	−0.4	−0.58	−90.4	−97.8

Table 4. Correlation coeficient, RMSE (in Hm^3) and MAE (in Hm^3) of the spatial and temporal rainfall volume distributions yielded by the set of MM5 control simulations.

Flood event	r spatial	RMSE spatial	MAE spatial	r temporal	RMSE temporal	MAE temporal
16–17 Nov 1996	0.46	3.4	2.5	−0.09	4.4	3.5
17–18 Dec 1997	0.84	7.1	4.5	0.67	8.4	5.6
9–10 Jun 2000	0.28	8.7	6.8	0.11	24.2	15.8
29–30 Aug 2004	0.37	1.9	1.3	−0.17	5.4	3.3

but RMSE and MAE skill scores penalize the wider over-forecasting of the maximum amounts (Table 4). MM5 control forecast for the 16–17 November 1996 episode shows a considerable shift of the predicted maximum amounts southeastwards of the watershed (Table 4, Fig. 9a): the maxima are located off the Catalonian coast, and the low part of the basin collects up to 80 mm. The timing of the forecasted precipitation shows remarkable deficiencies as well. The remaining deterministic forecasts display a poor skill in reproducing the observed accumulated precipitations and their spatial and temporal distributions over the catchment. MM5 control simulations are clearly deficient and maximum bulks of precipitation are located quite far away from the Llobregat basin (Table 4, Fig. 9c and d). The maximum quantities are located in the north-western part of the Pyrenees range for the 9–10 June 2000 episode, whereas for the 29–30 August 2004 episode, these fall over the Mediterranean Sea between the Catalonian coastline and the Balearic Islands (figures not shown).

The aforementioned deficiencies are propagated into the MM5 driven runoff forecasts. These runs do not capture in a suitable way the observed peak flows. The severe underestimation in the QPFs for the June 2000 and August 2004 episodes results in a complete miss of the floods by the runoff simulations (Table 3, Fig. 6c and d). For the 16–17 November 1996 episode, it can be appreciated a remarkable underforecasting of the observed runoff volume and the peak flow (Table 3). The simulated maximum peak discharge is also affected by the biases found in the rain-gauge driven runoff evaluation run, but still it reaches a considerable maximum of 722.8 m^3 s^{-1}. This fact is mainly due to have driven the hydrologic model using the previous 48 h observed rainy period in order to intend reproducing an operational hydrometeorological chain framework (Fig. 6b). The 17–18 December 1997 MM5 control driven runoff simulation is the most suitable among all the study cases. However, the excessive rainfall accumulations inside the Llobregat river basin result in an overestimation of the peak flow – with a value close to 700 m^3 s^{-1}, even though the volume presents a small error (Table 3, Fig. 6b). We have also simulated this event including the previous 48 h observed precipitation.

It appears that the skill of the hydrometeorological predictions is highly dependent on the flood event as well as on the initial and boundary conditions of the databases. For example, in Amengual et al. (2007), the MM5 mesoscale model was initialized using NCEP analyses – rather than ECMWF forecasts – to study the "Montserrat" flash-flood. It was found that the control simulation reasonably reproduced the spatial and temporal distributions of the observed rainfall, the total precipitated water and the runoff volume over the Llobregat river basin.

It is worth to note that Anderson et al. (2002) pointed out that runoff predictions for use in emergency management directives may not need to match exactly the peak discharges or the timing, but must reach suitable thresholds so as to cause the appropriate directives to be enacted. Ferraris et al. (2002) argued similar comments within an operational civil context for the Tanaro river basin, northwestern Italy. Within a theoretical civil protection framework for the Llobregat river basin, a warning should be issued as long as the forecasted peak discharge exceeds a pre-determined threshold in the analyzed hydrometric sections. Therefore, the MM5 deterministic runs for the November 1996 and December 1997 events could be considered suitable at this aim. However,
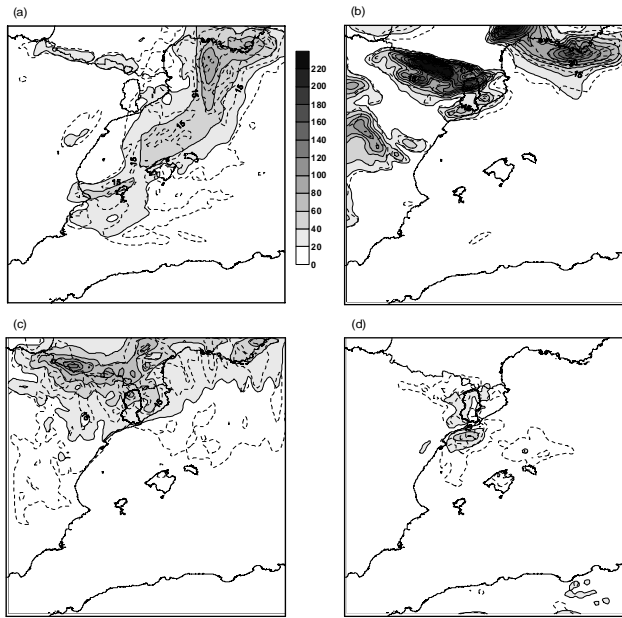
Fig. 10. Ensemble mean (shaded contours, in mm) and ensemble standard deviation (dashed line, in mm at 7.5 mm intervals) for the accumulated precipitation over the second MM5 computational domain for: **(a)** 16–17 November 1996, **(b)** 17–18 December 1997, **(c)** 9–10 June 2000 and **(d)** 29–30 August 2004 hydrometeorological episodes. The Llobregat river basin is highlighted as thick line.

large errors in runoff forecasts have been found for the June 2000 and August 2004 episodes which would not allow the triggering of any warning. The difficultties encountered by NWP models to forecast localized convectively-driven rainfall systems become evident. The arduous reproduction of the precise location and timing of these kinds of events poses great challenges to overcome, even more if affect small and medium size basins. Next, the value of our ensemble forecasting system to convey the external-scale uncertainties encountered in the hydrometeorological chain will be analysed.

6.3 Runoff simulations driven by MM5 probabilistic QPFs ensemble

Following the PV inversion method described in Sect. 5, twenty additional MM5-perturbed mesoscale runs are produced and used to drive new runoff simulations of each hydrometeorological episode. This population, together with the previous MM5-control experiment, becomes a useful expression of a probabilistic quantitative precipitation forecast in order to investigate the effects on the hydrometeorological chain of the uncertainties present in the atmospheric initial and boundary conditions.

For the discussion of the results, we have considered all the members of the MM5 ensemble to be equally-like and the subsequent driven runoff peak flows are represented as cumulative distribution functions (CDFs) plotted in a Gumbel

chart (Ferraris et al., 2002; Taramasso et al., 2005; Fig. 11). Together with the peak flows, we include a vertical black line which represents the observed maximum discharge for each hydrometeorological event at a particular hydrometric section. Although no hydrometeorological forecasting chain is currently implemented for civil-protection purposes in the Llobregat river basin, we have considered suitable for the present study the introduction of two hypothetical warning discharge thresholds. Then, the additional vertical light and dark grey lines in Fig. 11 denote the maximum peak discharges for return periods of 5 and 10 years (labelled as $Q_{p(T=5\,\text{yrs})}$ and $Q_{p(T=10\,\text{yrs})}$, respectively). These peak flows have been extracted from the aforementioned technical report issued by ACA (ACA, 2001). The first and second warning discharge thresholds are 168.5 m^3 s^{-1} and 359.9 m^{3}s^{-1} at Castellbell flow-gauge, and 266.3 m^3 s^{-1} and 555.4 m^3 s^{-1} at Sant Joan Despí hydrometric section.

A set of categorical verification indices has been used in order to test the quality of the PQPFs at sub-basin scale. Specifically, the probability of detection (POD), the false alarm rate (F), the frequency-bias score (BIAS) and the relative-operating-characteristics (ROC) score are computed for diverse precipitation and runoff volume thresholds (Jolliffe and Stephenson, 2003; Wilks, 2006). These skill scores are calculated using all the hydrometeorological episodes in order to increase the statistical significance of the results as follows:

– the rainfall and driven runoff volumes of the MM5 control runs are compared against the observed rainfall and SAIH rain-gauge driven runoff volumes, respectively;

– the rainfall and driven runoff volumes of the ensemble means are employed;

– the rainfall and driven runoff volumes by the individual members of the ensembles are used for the comparison with the observed rainfall and rain-gauge driven runoff volumes.

All these rainfall and runoff volumes are accumulated at hourly time-steps and using the sub-basins as accumulation units. For the driven discharge volumes, the scores are calculated using only the hourly data produced at each sub-basin that is non-zero in at least one of the two compared series in order to prevent an artificial improvement of the ROC values. It is worth to remark that the ensemble mean, statistically, should provide a better forecast that any individual ensemble member, because errors in the individual forecasts tend to cancel when averaged (Epstein, 1969; Leith, 1974). Therefore, statistical scores for the ensemble means of rainfall and runoff volumes have also been computed in order to highlight the possible benefits of a simple ensemble average in comparison with the control experiments. Next, a brief discussion case by case is presented.
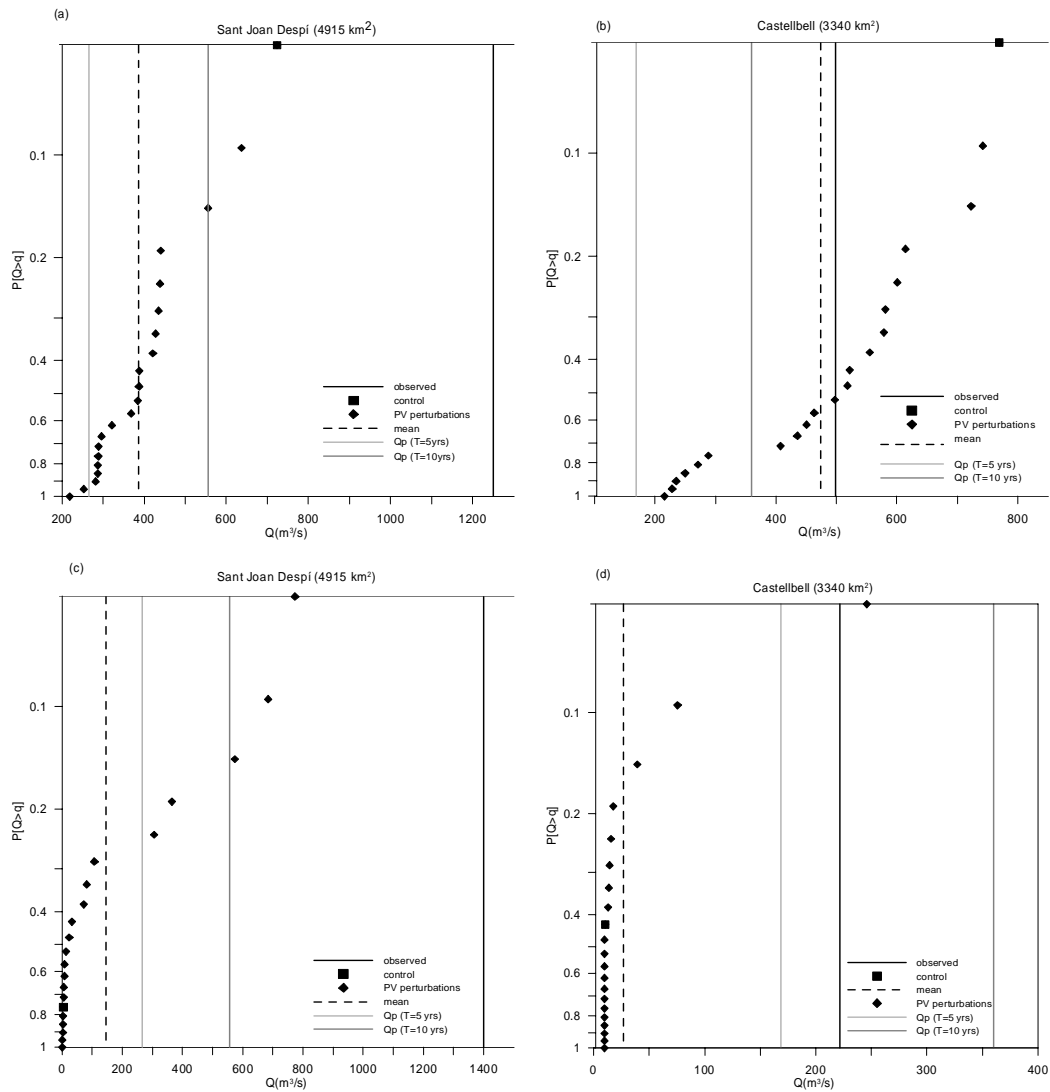
Fig. 11. Peak discharge exceedance probability plotted on a Gumbel chart for: **(a)** 16–17 November 1996, **(b)** 17–18 December 1997, **(c)** 9–10 June 2000 and **(d)** 29–30 August 2004 hydrometeorological episodes. The cumulative distribution functions of peak discharge are shown at the stream-gauges indicated in Table 1. The vertical black line indicates the observed maximum peak flow, the vertical dashed black line indicates the maximum peak discharge from the ensemble mean. Light and dark grey vertical lines show the maximum peak flows corresponding to return periods of 5 and 10 years, respectively.

6.3.1 The 16–17 November 1996 episode

Figure 10a depicts the ensemble mean and standard deviation of the accumulated forecasted rainfall over the MM5 inner domain. The rainfall pattern is clearly distributed over a maritime strip along the Spanish Mediterranean coastline and the maximum quantities are found over the Gulf of Lyons and southern France. Relatively small amounts are collected upon the Llobregat river basin. Figure 11a depicts that the best member of the ensemble is the control experiment, even though the simulated peak discharge presents an important underforecasting. Then, none of the members of the ensemble is able to reproduce the observed peak discharge at the

basin outlet. It appears that the introduction of realistic PV perturbations does not achieve to alter substantially the rainfall control pattern when the biases are so important and the maximum rainfall amounts are located so far away from the Llobregat river basin.

The peak discharge exceedance probabilities for $Q_{p(T=5\,\mathrm{yrs})}$ and $Q_{p(T=10\,\mathrm{yrs})}$ are close to 0.9 and 0.15, respectively. Albeit the first level of the theoretical warning is widely exceeded, this fact can be attributed to a couple of reasons: first, the antecedent baseflow is rather close to $Q_{p(T=5\,\mathrm{yrs})}$ for the driven runoff simulations; and second, previous 48 h observed rainfall is used in the driven runoff ensemble. These results point out the poor forecasting

skill of this event and the few benefits obtained from the hydrometeorological model ensemble strategy. It is worth to note that the Spanish Agency of Meteorology (AEMET) issues a warning when the probability of occurrence of an extreme event exceeds 0.2. Following this criterion, no second level warning would have been released by a hypothetical real-time forecasting hydrometeorological chain.

6.3.2 The 17–18 December 1997 episode

With regard to the 17–18 December 1997 episode, the MM5 ensemble accumulates substantial precipitations for the whole of the Pyrenees range, including large areas over the Llobregat river basin (Fig. 10b). In addition, coastal areas of the Gulf of Lyons are affected by remarkable precipitations as well. It seems clear that the south-westerly flow together with the orographic enhancement of the Pyrenees determine the location of the intense rainfalls and the mesoscale model ensemble correctly captures this physical mechanism. A high degree of spatial consistency occurs among the members of the ensemble as reflected in the standard deviation field of Fig. 10b. Figure 11b shows that some members of the ensemble are very close to the observed peak flow, thus correcting the overforecasting of the control driven runoff peak flow. It is worth noting that the probabilities of exceedance of $Q_{p(T=5\,\mathrm{yrs})}$ and $Q_{p(T=10\,\mathrm{yrs})}$ are 1.0 and greater than 0.7, respectively. Furthermore, the exceedance probability for the observed maximum discharge is above 0.5. Then, the probabilistic PV perturbed ensemble proves to be suitable for civil protection procedures for this case.

6.3.3 The 9–10 June 2000 episode

Recall that the control run missed completely this extreme case, since the greatest precipitation amounts were mainly located over the north-western part of the Pyrenees range, far away from the Llobregat river basin. Therefore, the control driven runoff simulation was not able to produce any remarkable signal. Figure 10c depicts an important spread of the rainfall cumulative values over north-eastern Spain within the mesoscale model ensemble. It seems that the "Montserrat" hydrometeorological event is highly sensitive to the PV perturbations. This episode was characterized by an important orographic enhancement mechanism in the organization, development and duration of the quasi-stationary mesoscale convective system, but the atmospheric dynamical forcing played an essential role as well. Dynamical factors are linked to the upper-level synoptic trough and midto-upper-tropospheric flow, which are emphasized by the PV perturbation approach. As consequence, the biases found in the control run are partially corrected, and in spite that none of the members of the ensemble is able to forecast the observed peak flow for this exceptional event, some of them yield important peak discharges (Fig. 11c). The probabilities

Table 5. ROC scores and the 95% percentile confidence intervals for the control simulation, ensemble mean and ensemble members of all the hydrometeorological experiments, for hourly rainfall and runoff volumes.

	control	mean	ensemble
rainfall volumes	0.62 (0.56–0.68)	0.65 (0.61–0.69)	0.58 (0.54–0.62)
runoff volumes	0.79 (0.73–0.84)	0.79 (0.73–0.84)	0.87 (0.83–0.89)

of exceeding the first and second warning discharge thresholds are close to 0.3 and 0.15, respectively. At least, a first theoretical warning could have been issued, even though the magnitude of the "Montserrat" episode would have remained completely undetected.

6.3.4 The 29–30 August 2004 episode

This summer episode was driven by physical processes of local convective origin. It was found for the MM5 deterministic run a misplacement and underforecasting of the rainfall amounts. Figure 10d depicts that the ensemble mean from rainfall is mainly located off the Catalonian coastline, south-eastwards from the Llobregat river basin, but precipitation amounts are also collected within the catchment. However, the NWP ensemble underforecasts the observed rainfall amounts. This fact can be noticed on the cumulative distribution functions of peak discharges at Castellbell (Fig. 11d): only one member of the ensemble is sufficiently accurate to reproduce the observed discharge. Thus, the probabilities of exceeding both the observed maximum peak and the predetermined thresholds – $Q_{p(T=5\,\mathrm{yrs})}$ and $Q_{p(T=10\,\mathrm{yrs})}$ – lie below 0.1. Obviously, the triggering of any emergency management directives would have been suppressed.

It appears that the hydrometeorological events with an enhanced synoptic-scale dynamical component are more sensitive to the PV perturbations, whereas the events with a more important orographic involvement in their development become more insensitive. It is worth to remember that the technique for generating the EPS is based on randomly perturbing the upper-level precursor trough in a realistic magnitude. Therefore, if the biases in the deterministic QPFs are important, the great precipitation errors will be hardly corrected by the technique.

6.3.5 POD, F, BIAS and ROC skill indices

Table 5 and Figs. 12a–c and 13a–c depict the results for the categorical verification scores and their 95% confidence intervals. That is, a bootstrap test (Diaconis and Efron, 1983) with 1000 repetions has been applied for each experiment in order to better establish the uncertainties around the statistical indices. Regarding the rainfall accumulations, the ensemble mean proves to be the best for the POD at low and medium thresholds, although the increased skill for the POD
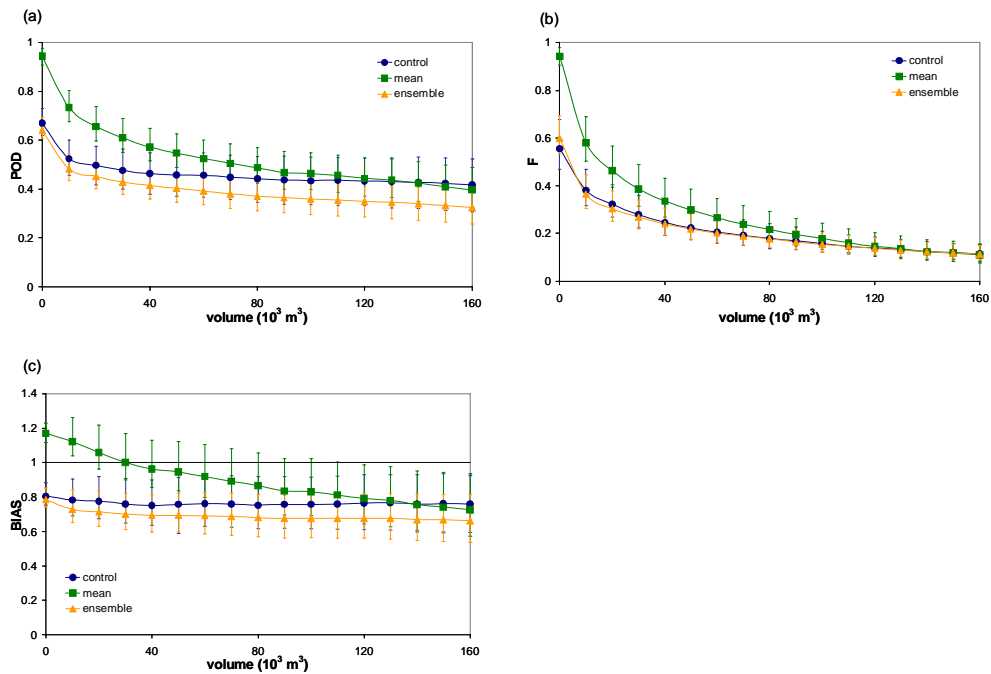
Fig. 12. POD, F and BIAS skill scores and the 95% percentile confidence intervals for different precipitation volume thresholds obtained by the ensemble of MM5 simulations for all the hydrometeorological experiments.
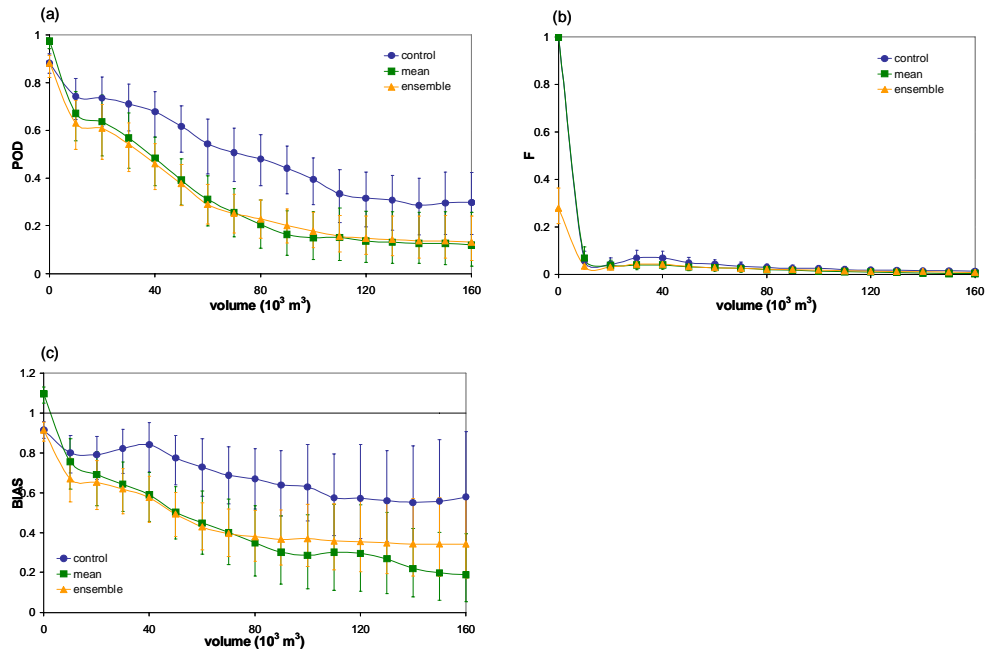

Fig. 13. POD, F and BIAS skill scores and the 95% percentile confidence intervals for different runoff volume thresholds obtained by the ensemble of MM5 driven runoff discharge simulations for all the hydrometeorological experiments.

also induces a rise of the F score. A tendency in underforecasting the rainfall amounts at mid- and high thresholds can be observed attending to the BIAS index as well. The ensemble mean presents the highest ROC score as a consequence of the smoothing of the individual forecasts errors when averaging the rainfall fields (Table 5). ROC score for the control runs exceeds the ROC score obtained from the ensembles, owing to a higher probability of detection of the deterministic

runs at all the volume thresholds together with similar false alarm rates for both sets of experiments. A systematic under-forecasting of the precipitation amounts at all thresholds are found for the deterministic and the probabilistic experiments (Fig. 12c).

With respect to the driven runoff forecasts, Table 5 shows the highest performance in terms of ROC score for the full ensemble of all the hydrometeorological experiments owing to the smallest false alarm scores at small volume values. However, in terms of the probability of detection, the control simulations appear to be superior. It is worth to note an overall underforecasting of the full ensemble of runoff experiments at all the thresholds, although the control simulations show a better behaviour for this feature (Fig. 13a–c). This general underprediction of the runoff volumes clearly benefits the false alarm rates which become smaller, and therefore, leading to higher ROC scores. In addition, no improvement of the ROC scores is found when comparing the ensemble means against the deterministic simulations.

7   Conclusions and further remarks

The feasibility of HEC-HMS runoff forecasts driven by QPFs provided by the MM5 mesoscale model has been assessed over the Llobregat medium size river basin. This issue is fundamental to help extend the forecast lead times beyond the concentration time of the watershed. Any gain in these times could be used to trigger warning and emergency procedures before extreme floods in order to alleviate their possible hazardous consequences. The analysis has focused on four intense rainfall episodes which resulted in flood events of different spatial and temporal scales. These kinds of intense precipitation events – often highly localized and convectively driven – present short recurrence periods in the Spanish Mediterranean area and deserve special attenttion.

Deterministic quantitative precipitation forecasts for the four cases have been obtained by forcing the NWP model with the ECMWF large-scale forecasts, thus mimicking an operational hydrometeorological context. In addition, an ensemble of MM5 experiments has been introduced by means of realistic spatial shifts and changes in intensity of the upper-level synoptic-scale trough for each case study. The performance of the control QPFs has been evaluated from a comparison of the forecasted and observed rainfall volumes in space and time over the sub-basins by means of a set of continuous verification scores. The probabilistic QPFs have been evaluated following the same procedure, but using a set of categorical verification indices. This methodology allows to assess the performance of the forecasted rainfall patterns at the catchment scale. The one-way coupling between the meteorological and hydrological models has been regarded as a validation tool as well.

Most of the deterministic quantitative precipitation forecasts have shown significant deficiencies over the Llobregat river basin in terms of the rainfall amounts, their timing and localization. Obviously, these deficiencies have a major impact on the one-way coupling between the meteorological and hydrological models. Our ensemble strategy has been found useful to reduce these biases at the analyzed hydrometric sections – Castellbell and Sant Joan Despí –, for most of the hydrometeorological episodes. However, the probabilistic approach would have indicated surpassion of pre-defined theoretical warning runoff thresholds for only the 17–18 December 1997 case. It is worth to note that the procedure presented in this work has been applied to a limited number of cases – although this population is representative of different dynamical atmospheric situations bearing important precipitation amounts – owing to the restricted number of important floods available in the databases. It must be expected that a future application and evaluation of this procedure for a broader climatology could better establish its confidence levels. Then, we could apply the presented methodology onto a larger sample of events in order to use the results from an operational perspective.

The value of our PV perturbed ensemble to encompass in a realistic way the uncertainties intrinsic to the QPFs has been proven. It has also been possible to highlight the sensitivity – or insensitivity – to the PV perturbations of different dynamical and orographic factors. The external-scale meteorological uncertainties have been reflected through moderate spatial and temporal variations of the forecasted rainfall patterns as well as in significant changes of the precipitation amounts. Obviously, if the control simulations are characterized by large biases, our ensemble strategy presents limitations when trying to address these forecast errors.

Finally, it is important to note that the hydrological model's parameters related to the initial conditions (the infiltration mechanism), the dynamical formulation (the channel routing) and the baseflow modelization have been calibrated by using six events which include our four study cases. These events are characterized by important discharges at different hydrometric sections of the Llobregat river basin. However, the lack of flow data at some gauges for some of these events and the limited number of floods of remarkable magnitude available in the ACA database have posed difficultities in the calibration and evaluation of the rainfall-runoff model. We have applied a cross-calibration and evaluation procedure in order to optimize the skill of the hydrological model before such episodes. Again, it would be desirable to get a wider climatology of remarkable flood events affecting the basin. The future increase of the number of recorded cases in the SAIH database together with a larger number of stream-gauges operating inside the Llobregat river basin will permit an improvement of the forecast and warning schemes. All these issues appear to be of the maximum interest in order to develop the most suitable hydrometerological forecasting chain system for the Llobregat river basin.

References

Agència Catalana de l'Aigua: Delimiting of flood plain for the draft of the INUNCAT. Internal Basins of Catalonia. Hydrological calculations and determination of the flood wave discharge, Vol. 2, Departament de Medi Ambient, Generalitat de Catalunya, 88 pp., 2001.

Agència Catalana de l'Aigua: Technical recommendations for flood studies at local area, Departament de Medi Ambient, Generalitat de Catalunya, 106 pp., 2003.

Amengual A., Romero R., Gómez M., Martín A., and Alonso S.:A hydrometeorological modeling study of a flash flood event over Catalonia, Spain, J. Hydrometeor., 8, 282–303, 2007.

Amengual A., Romero R., and Alonso S.: Hydrometeorological ensemble simulations of flood events over a small basin of Majorca Island, Spain, Q. J. R. Meteorol. Soc., 134, 1221–1242, 2008.

Anderson, M. L., Chen, Z. Q., Kavvas, M. L., and Feldman, A.: Coupling HEC-HMS with atmospheric models for prediction of watershed runoff, J. Hydrol. Eng., 7, 312–318, 2002.

Bacchi, B., Buzzi, A., Grossi, G., and Ranzi R.: Flood forecasting in a midsize catchment in the southern Alps: Recent experiences on the use of couple meteorological and hydrological model, Proceedings of the third EGS Plinius conference, Baja Sardinia, Consiglio Nazionale delle Ricerche, Roma, Italy, 201–208, 2002.

Bartholmes, J. C., Thielen, J., Ramos, M. H., and Gentilini, S.: The european flood alert system EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, Hydrol. Earth Syst. Sci., 13, 141–153, 2009,
http://www.hydrol-earth-syst-sci.net/13/141/2009/.

Benjamin, S. G. and Seaman, N. L.: A simple scheme for improved objective analysis in curved flow, Mon. Wea. Rev., 113, 1184–1198, 1985.

Benoit, R., Pellerin, P., Kouwen, N., Ritchie, H., Donaldson, N., Joe, P., and Soulis, E. D.: Toward the use of coupled atmospheric and hydrologic models at regional scale, Mon. Wea. Rev., 128, 1681–1706, 2000.

Benoit, R., Kouwen, N., Yu, W., Chamberland, S., and Pellerin, P.: Hydrometeorological aspects of the Real-Time Ultrafinescale Forecast Support during the Special Observing Period of the MAP, Hydrol. Earth Syst. Sci., 7, 877–889, 2003,
http://www.hydrol-earth-syst-sci.net/7/877/2003/.

Bhagarva, M. and Danard, M.: Application of optimum interpolation to the analysis of precipitation in complex terrain, J. Appl. Meteor., 33, 508–518, 1994.

Blackadar, A. K.: High resolution models of the planetary boundary layer, Adv. Environ. Sci. Eng., 1, 50–85, 1979.

Buizza, R. and Palmer, T. N.: The singular-vector structure of the atmospheric general circulation, J. Atmos. Sci. Eng., 52, 1434–1456, 1995.

Chancibault, K., Anquetin, S., Ducrocq, V., and Saulnier, G.-M.: Hydrological evaluation of high-resolution precipitation forecasts of the Gard flash flood event (8–9 September 2002), Q. J. R. Meteorol. Soc., 132, 1091–1117, 2006.

Chow, V. T., Maidment, D. R., and Mays L. W.: Applied hydrology, McGraw-Hill International Editions, Civil engineering series, 572 pp., 1988

Davis, C. A. and Emanuel, K. A.: Potential vorticity diagnosis of ciclogenesis, Mon. Wea. Rev., 119, 1929–1953, 1991.

De Roo, A., Gouweleeuw, B., Thielen, J., Bates, P., Hollingsworth, A., et al.: Development of a European Flood Forecasting System, Int. J. River Basin Management, 1(1), 49–59, 2003.

Diaconis, P. and Efron, B.: Computer-intensive methods in statistics, Sci. Am., 248, 116–130, 1983.

Diomede, T., Davolio, S., Marsigli, C., Miglietta, M. M., Moscatello, A., and Papetti, P.: Discharge ensemble predictions based on multi-model precipitations forecasts, Geophy. Res. Abstr., 8, EGU06-A-06931, 2006.

Diomede, T., Nerozzi, F., Paccagnella, T., and Todini, E.: The use of meteorological analogues to account for LAM QPF uncertainty, Hydrol. Earth Syst. Sci., 12, 141–157, 2008,
http://www.hydrol-earth-syst-sci.net/12/141/2008/.

Dolciné, L., Andrieu, H., Sempere-Torres, D., and Creutin D.: Flash flood forecasting with coupled precipitation model in mountainous mediterranean basin, J. Hydrol. Eng., 6, 1–9, 2001.

Du, J., Mullen, S. L., and Sanders, F.: Short-range ensemble forecasting of quantitative precipitation, Mon. Wea. Rev., 125, 2427–2459, 1997.

Ducrocq, V., Lebeaupin, C., Thouvenin, T., and Anquetin, S.: Mesoscale modelling of a cooding storm. Application to the extreme flood of Gard, 5th EGS Plinius conference, Ajaccio, Corsica, France, 2003.

Dudhia, J.: A nonhydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front, Mon. Wea. Rev., 121, 1493–1513, 1993.

Epstein, E. S.: Stochastic dynamic prediction, Tellus, 21, 739–759, 1969.

Ferraris, L., Rudari R., and Siccardi, F.: The uncertainty in the prediction of flash floods in the Northern Mediterranean environment, J. Hydrometeor., 3, 714–727, 2002.

Gouweleeuw, B. T., Thielen, J., Franchello, G., De Roo, A. P. J., and Buizza, R.: Flood forecasting using medium-range probabilistic weather prediction, Hydrol. Earth Syst. Sci., 9, 365–380, 2005,
http://www.hydrol-earth-syst-sci.net/9/365/2005/.

Homar, V., Romero, R., Ramis, C., and Alonso, S.: Numerical study of the October 2000 torrential precipitation event over eastern Spain: analysis of the synoptic-scale stationarity, Ann. Geophys., 20, 2047–2066, 2002,
http://www.ann-geophys.net/20/2047/2002/.

Hong, S.-Y. and Pan, H.-L.: Nonlocal boundary layer vertical diffusion in a medium-range forecast model, Mon. Wea. Rev., 124, 2322–2339, 1996.

Hoskins, B. J., McIntyre, M. E., and Robertson, A. W.: On the use and significance of isoentropic potential vorticity maps, Q. J. Roy. Meteor. Soc., 111, 877–946, 1985.

Houtekamer, P. L. and Derome, J.: Methods for ensemble prediction, Mon. Wea. Rev., 123, 2181–2196, 1995.

Jasper, K. and Kaufmann, P.: Coupled runoff simulations as validation tools for atmospheric models at the regional scale, Q. J. Roy. Meteor. Soc., 129, 673–692, 2003.

Jolliffe, I. T. and Stephenson D. B.: Forecast verification. A practitioner's guide in atmospheric science, John Wiley and Sons Ltd., Chichester, West Sussex, England, 240 pp., 2003.

Jones, M. S., Colle, B. A., and Tongue, J. S.: Evaluation of a mesoscale short-range ensemble forecasts system over the northeast United States, Wea. Forecast., 22, 36–55, 2007.

Kain, J. S.: The Kain-Fritsch convective parameterization: an update, J. Appl. Meteor., 43, 170–181, 2004.

Krajewski, W. F.: Cokriging radar-rainfall and rain gauge data, J. Geophys. Res., 92(D8), 9571–9580, 1987.

Leith, C. E.: Theoretical skill of Monte Carlo forecasts, Mon. Wea. Rev., 102, 409–418, 1974.

Linsley, R. K., Kohler, M. A., and Paulhus J. L. H.: Hydrology for engineers, 3rd edition, McGraw-Hill International Editions, New York, 508 pp., 1982.

Llasat, M. C., Rigo, T., and Barriendos, M.: The 'Montserrat-2000' flash flood event: a comparasion with the floods that have ocurred in the northeast iberian peninsula since the 14th century, Int. J. Climat., 23, 453–469, 2003.

Martín, A., Romero, R., Homar, V., de Luque, A., Alonso, S., Rigo, T., and Llasat, M. C.: Sensitivities of a flash flood event over Catalonia: A numerical analysis, Mon. Wea. Rev., 135, 651–669, 2007.

Molteni, F., Buizza, R., Palmer, T. N. and Petroliagis, T.: The ECMWF ensemble prediction system: Methodology and validation, Q. J. Roy. Meteor. Soc., 122, 73–119, 1996.

Mullen, S. L. and Baumhefner, D. P.: Sensitivity to numerical simulations of explosive oceanic cyclogenesis to changes in physical parameterizations, Mon. Wea. Rev., 116, 2289–2329, 1988.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part I: A discussion of principles, J. Hydrol., 10(3), 282–290, 1970.

Ranzi, R., Bacchi, B., Grossi, G., Buzzi, A., Malguzzi, P., Ratto, C., and Corazza, M.: Previsioni di piena mediante un modello idrologico e un modello meteorologico ad area limitata: alcune esperienze applicative durante l'esperimento MAP-SOP, Atti del XXVII Convegno di Idraulica e Costruzioni Idrauliche, Genova, 12–15 settembre 2000, Vol. II, 385–393, 2000.

Reisner, J., Rasmussen, R. J., and Bruintjes R. T.: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model, Q. J. Roy. Meteor. Soc., 124B, 1071–1107, 1998.

Romero, R., Doswell III, C. A., and Ramis, C.: Mesoscale numerical study of two cases of long-lived quasistationary convective systems over eastern Spain, Mon. Wea. Rev., 128, 3731–3751, 2000.

Romero, R.: Sensitivity of a heavy rain producing Western Mediterranean cyclone to embedded potential vorticity anomalies, Q. J. Roy. Meteor. Soc., 127, 2559–2597, 2001.

Romero, R., Martín, A., Homar, V., Alonso, S., and Ramis, C.: Predictability of prototype flash flood events in the Western Mediterranean under uncertainties of the precursor upper-level disturbance, Adv. Geosci., 7, 55–63, 2006,
http://www.adv-geosci.net/7/55/2006/.

Roulin, E.: Skill and relative economic value of medium-range hydrological ensemble predictions, Hydrol. Earth Syst. Sci., 11, 725–737, 2007,
http://www.hydrol-earth-syst-sci.net/11/725/2007/.

Schaake, J., Franz, K., Bradley, A., and Buizza, R.: The Hydrologic Ensemble Prediction EXperiment (HEPEX), Hydrol. Earth Syst. Sci. Discuss., 3, 3321–3332, 2006,
http://www.hydrol-earth-syst-sci-discuss.net/3/3321/2006/.

Schaake, J., Hamill, T. M., Buizza, R., and Clark, M.: HEPEX, the Hydrological Ensemble Prediction Experiment, B. Am. Meteorol. Soc., 88, 1541–1547, 2007.

Seo, D. J.: Real-time estimation of rainfall fields using radar rainfall and rain gage data, J. Hydrol., 208, 37–52, 1998.

Siccardi, F.: Rainstorm hazards and related disasters in the western Mediterranean region, Rem. Sens. Rev., 14, 5–21, 1996.

Stensrud, D. J. and Fritsch, J. M.: Mesoscale convective systems in weakly forced large-scale environments. Part II: Generation of a mesoscale initial condition, Mon. Wea. Rev., 122, 2068–2083, 1994a.

Stensrud, D. J. and Fritsch, J. M.: Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational forecasting, Mon. Wea. Rev., 122, 2084–2104, 1994b.

Stensrud, D. J., Bao, J.-W., and Warner, T. T.: Using initial and model physics perturbations in short-range ensemble simulations of mesoscale convective events, Mon. Wea. Rev., 128, 2077–2107, 2000.

Stensrud, D. J. and Yussouf, N.: Reliable probabilistic quantitative precipitation forecasts from a short-range ensemble forecasting system, Wea. Forecast., 22, 3–17, 2007.

Taramasso, A. C., Gabellani, S., and Parodi, A.: An operational flash-flood forecasting chain applied to the test cases of the EU project HYDROPTIMET, Nat. Hazards Earth Syst. Sci., 5, 703–710, 2005,
http://www.nat-hazards-earth-syst-sci.net/5/703/2005/.

Thielen, J., Bartholmes, J., Ramos, M.-H., and de Roo, A.: The European Flood Alert System –Part 1: Concept and development, Hydrol. Earth Syst. Sci., 13, 125–140, 2009,
http://www.hydrol-earth-syst-sci.net/13/125/2009/.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: The generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993.

US Army Corps of Engineers Hydrologic Engineering Center: HEC-HMS Hydrologic Modeling system user's manual, USACE-HEC, Davis, California, 188 pp., 1998

US Army Corps of Engineers Hydrologic Engineering Center: Hydrologic Modeling system HEC-HMS. Technical reference manual, USACE-HEC, Davis, California, 157 pp., 2000.

US Department of Agriculture: Urban hydrology for small watersheds. Technical release 55 of the Nature Resources Conservation Service, US Department of Agriculture, Washington DC, USA, 1986.

Hydrol. Earth Syst. Sci., 13, 793–811, 2009

www.hydrol-earth-syst-sci.net/13/793/2009/

Wilks, D.S.: Statistical methods in the atmospheric sciences, 2nd edition, Elsiever Academic Press Publications, International Geophysics Sciences, 627 pp., 2006.

Younis, J., Anquetin, S., and Thielen, J.: The benefit of high-resolution operational weather forecasts for flash flood warning, Hydrol. Earth Syst. Sci., 12, 1039–1051, 2008, http://www.hydrol-earth-syst-sci.net/12/1039/2008/.

Zhang, D. L. and Anthes, R. A: A high resolution model of the planetary boundary layer. Sensitivity tests and comparisons with SESAME-79 data, J. App. Meteor., 21, 1594–1609, 1982.

Zhang, D. L. and Fritsch, J. M.: Numerical simulation of the meso-β scale structure and evolution of the 1977 Johnstown flood. Part I: Model description and verification, J. App. Meteor., 43, 1913–1943, 1986.

www.hydrol-earth-syst-sci.net/13/793/2009/

Hydrol. Earth Syst. Sci., 13, 793–811, 2009