Multi-scale evaluation of extreme rainfall-event predictions using severity diagrams

D. Ceresetti, S. Anquetin, G. Molinié, E. Leblois, J.D. Creutin

Multiscale Evaluation of Extreme Rainfall Event Predictions Using Severity Diagrams

DAVIDE CERESETTI, SANDRINE ANQUETIN, AND GILLES MOLINIÉ

Laboratoire d'études des Transferts en Hydrologie et Environnement, and CNRS, UJF, IRD, INPG, Grenoble, France

ETIENNE LEBLOIS

CEMAGREF, Lyon, France

JEAN-DOMINIQUE CREUTIN

Laboratoire d'études des Transferts en Hydrologie et Environnement, and CNRS, UJF, IRD, INPG, Grenoble, France

(Manuscript received 14 December 2010, in final form 28 June 2011)

ABSTRACT

Observations and simulations of rainfall events are usually compared by analyzing (i) the total rainfall depth produced by the event and (ii) the location of the rainfall maximum. A different approach is proposed here that compares the mesoscale simulated rainfall fields with the ground rainfall observations within the multiscale framework of maximum intensity diagrams and severity diagrams. While the first simply displays the maximum rainfall intensity of an event at a number of scales, the second gives the frequency of occurrence of the maximum rainfall intensities as a function of the spatial and temporal aggregation scales, highlighting the space–time scales of the event severity. For use in a region featuring complex relief, severity diagrams have been generalized to incorporate the regional behavior of heavy rainfall events. To assess simulation outputs from a meteorological mesoscale model, three major storms that have occurred in the last decade over a mountainous Mediterranean region of southern France are analyzed. The severity diagrams detect the critical space–time scales of the rainfall events for comparison with those predicted by the simulation. This validation approach is adapted to evaluate the ability of the mesoscale model to predict various types of storms with different regional climatologies.

1. Introduction

In recent years, Mediterranean storms have caused serious damage to life and property at many locations in southern Europe. These events have hit coastal as well as mountainous areas and have involved various spatial and temporal scales. Recent research has shown that fatalities occur both in small and large basins presenting considerably different space and time dynamics (Ramos et al. 2005; Ruin et al. 2008). The variability of heavy rainfall events changes with the scale of analysis: two events that occur at two different scales cannot be compared only on the basis of average rainfall depth or rainfall intensity (Bousquet et al. 2006; Yates et al. 2007).

These considerations highlight the need for an objective evaluation of the impacts of storms at a variety of spatial and temporal scales that better takes into account the storm structure, its critical scales, and its hydrological impacts.

Similar considerations apply to the evaluation of meteorological models. The performance of mesoscale models is usually assessed by comparing simulated rainfall fields at a spatial resolution of the order of 1 km^2 to ground measurements from rain gauges (with a collecting area of 1000 cm^2). Different types of scores have been proposed to compare model outputs and observations. In the literature (e.g., Mason 1989; Ducrocq et al. 2002; Venugopal et al. 2005), equitable threat scores derived from contingency tables such as the probability of detection (POD), false alarm rate (FAR), or bias (FBIAS) mainly qualify the capability of the model to simulate rainfall depths and storm locations. Most of these indicators focus on the

Corresponding author address: Sandrine Anquetin, CNRS/LTHE, Université Grenoble 1, BP 53, 38041 Grenoble, CEDEX 9, France.
E-mail: sandrine.anquetin@ujf-grenoble.fr

DOI: 10.1175/WAF-D-11-00003.1

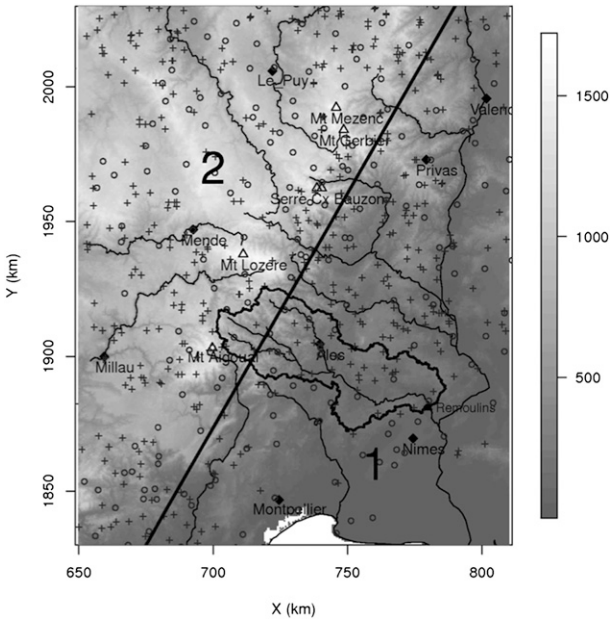

FIG. 1. Main features of the analysis region: elevation above sea level, main rivers (thin solid lines), main mountain peaks (triangles), and main cities (diamonds). Ground measurement network: daily network (circles) and hourly network (crosses). The solid rough circle in boldface indicates the boundaries of the Gard basin, with Remoulins at the outlet. The diagonal separates the mountainous (1) and flat (2) subregions. The rectangle identifies the area where the analysis was carried out.

most intense values. They do not provide any diagnostic on the capability of reproducing the whole storm structure at different scales (Zepeda-Arce et al. 2000; Yates et al. 2007).

In this paper we propose a new technique for multiscale comparisons of rainfall events based on the work of Ramos et al. (2005). The aim of ''severity diagrams'' is to represent the magnitude of a storm over a range of spatial and temporal scales within a normalized framework adapted to interevent comparisons. Considering ''severity'' in probabilistic terms, severity diagrams represent the return period of a storm for different durations and surface areas.

The use of severity diagrams requires knowledge of the extreme-rainfall distribution in the region. A preliminary comparison between observations and simulations can be achieved by analyzing ''maximum intensity diagrams,'' an intermediate indicator representing the maximum rainfall intensity for each spatial and temporal scale. Maximum intensity diagrams compare different events, but are scale dependent. Severity diagrams, thanks to the transformation of the maximum rainfall intensity into a return period, allow both interscale comparisons of single storms and comparisons between storms.

In the area studied by Ramos et al. (2005), a dense high-resolution rain-gauge network and climatic homogeneity simplified spatial rainfall frequency estimation. In the present paper, we need a more refined approach since the study area, the Cévennes-Vivarais region (southern France), features heterogeneous extreme rainfall behavior. Close to the Mediterranean Sea, this region with its complex topography is particularly prone to heavy rainfall events and flash floods (Delrieu et al. 2005; Nuissier et al. 2008; Ducrocq et al. 2008).

The construction of the above transformation is here based on the coupled application of areal reduction factors (ARFs) (NERC 1975; Rodriguez-Iturbe and Mejía 1974; Bacchi and Ranzi 1996; Sivapalan and Blöschl 1998; Asquith and Famiglietti 2000) and intensity–duration–frequency (IDF) curves (Burlando and Rosso 1996; Koutsoyiannis et al. 1998).

The paper is structured as follows. The region and the dataset used for this study are presented in section 2. The methodology is developed in section 3. Section 4 introduces the three cases. In section 5, the severity diagrams are plotted for three storms using the observed and simulated rain fields. For a particular rainfall event, the severity diagram is analyzed to explore the hydrological effects of the storm. Section 6 presents the advantages and limits of severity diagrams for the evaluation of meteorological models. A summary of our work and conclusions follow in section 7.

2. Data description

The studied region is located in southern France, bounded by the Massif Central to the west and north, the Mediterranean Sea to the south, and the Rhône River to the east. The area covers a window of about 160×200 km^2, containing a coastal zone, a large plain, a mountainous region, and a high plateau (Fig. 1).

The rainfall in the region has been widely studied (Lebel and Laborde 1988; Bois et al. 1997; Molinié et al. 2012, hereafter MCABC). Two kinds of rainfall data are used in this study: (i) the observed rain-gauge data provided by the Cévennes-Vivarais Hydro-Meteorological Observatory (OHMCV) and (ii) the rainfall fields simulated by Météo-France using the Mesoscale Nonhydrostatic (Méso-NH) model.

The rain-gauge rainfall dataset includes hourly and daily rainfall intensity series. The daily rainfall database (Fig. 1) comes from 225 stations providing more than 30 yr of continuous records. The hourly rainfall database includes 150 continuous rain-gauge records over the 1993–2008 period. Figure 2 reports the average density of the hourly and daily rain-gauge networks as a function of elevation.

The simulated rainfall data are available from Météo-France for specific severe rainy events thanks to the Forecast and Projection in Climate Scenario of

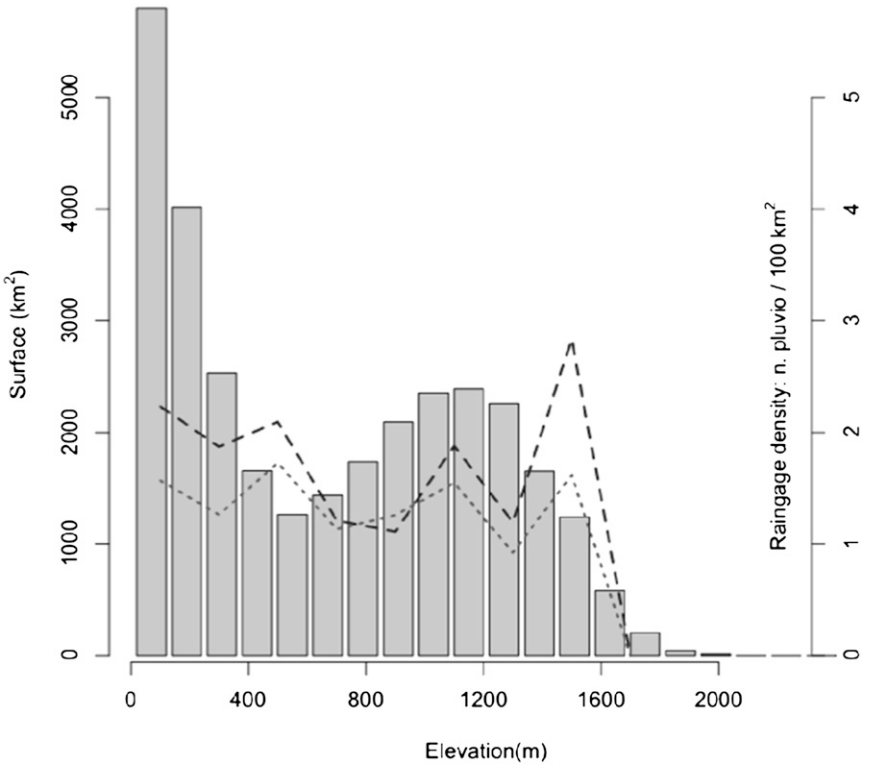

FIG. 2. Rain-gauge distribution in the Cévennes-Vivarais region. The histogram represents the surface area (left axis) associated with each elevation band. The daily (dashed line) and hourly (dotted line) rain-gauge densities are plotted as a function of elevation (the values can be read along the right axis).

Mediterranean Intense Events: Uncertainties and Propagation on Environment (MedUP) research program. They are the product of the Méso-NH cloud-resolving model (Lafore et al. 1998). Méso-NH is run on two-way nested grids at 9.5- and 2.4-km resolutions, respectively. The coarser Méso-NH domain is driven by 6-hourly Aire Limitée Adaptation Dynamique Développement International (ALADIN) analyses, linearly interpolated in time, to provide lateral conditions. The finer-scale Méso-NH domain is centered over the northwestern Mediterranean, where the studied rainy events were initiated, and includes the study region. The rainfall fields are provided at a space–time resolution of 2.4 km–1 h. Three simulated events are studied and will be fully described in section 4. The model configuration, successfully tested for simulations of intense Mediterranean rainfall events (Ducrocq et al. 2002), is the same for the three events. Below, we will provide the main characteristics of the simulations (Table 2). For more details, refer to Lebeaupin et al. (2006) for event 1, Yan et al. (2009) for event 2, and Nuissier et al. (2008) for event 3. The initial conditions are provided either by the large-scale operational analysis of the French Action de Recherche Petite Echelle Grande Echelle (ARPEGE) NWP model or by the finescale initialization procedure, three-dimensional variational data assimilation (3DVAR) ALADIN, as described in Ducrocq et al. (2000).

3. Methodology

The use of severity diagrams requires three steps, as schematized in Fig. 3, involving a study of the extreme spatial rainfall climatology. The first step (section 3a) consists of building the spatial rainfall database by interpolating the rain-gauge observations. From this database, we derive ARFs, expressing the rainfall intensity decrease with increasing surface area (section 3b). As severity diagrams require the frequency of occurrence of spatial rainfall intensities for any accumulation duration and surface area, a continuous ARF model must be coupled with IDF relationships. The IDF model adopted in this study is described in section 3c.

a. Spatial rainfall database

The spatial rainfall database was built from the hourly rainfall intensity database. The point rainfall observations were accumulated over durations of 2, 4, 8, 12, and 24 h.

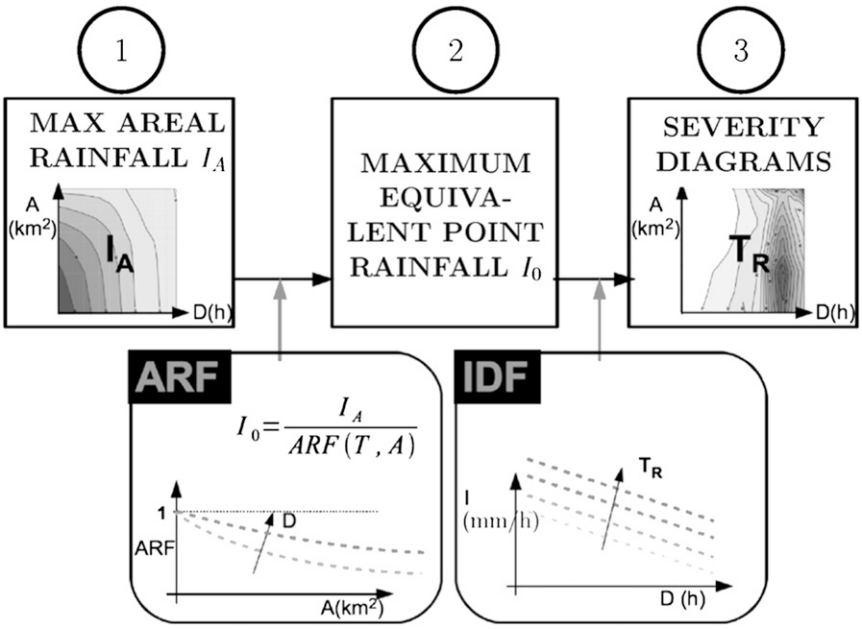

FIG. 3. Sketch illustrating the severity diagram computation. Step 1 consists of computing the maximum rainfall intensities for a range of space and accumulation duration scales. Using the ARF relationship, the maximum areal intensities are converted into point intensities (step 2). The arrow indicates how the ARF curve behaves as the accumulation duration D increases. IDF relationships are used to determine the return period of intensities for any space and accumulation duration pairs yielding severity diagrams (step 3). In the IDF diagram, the arrow shows how the rainfall intensity increases with the return period T_R.

For each accumulation duration, the spatial rainfall database was built in three steps:

(i) *Definition of rainy events*—The number of working rain gauges was determined along with the average rainfall depth measured by the network. When these two indicators were lower than a fixed threshold, the field was rejected because the event was probably weak and local.

(ii) *Determination of the spatial structure of the rainfall field*—For every field retained, the variogram of the positive rainfall values was computed and a spherical model was inferred. If the variogram underwent large variations for specific distance ranges, then a ''climatological variogram'' was forced with the field (Lebel and Laborde 1988). Such a pattern of behavior was usually due to an insufficient number of points (typically <30) being used to determine a reliable empirical average of the sample variograms.

(iii) *Interpolation*—Using the variogram model defined above, the kriging interpolation was computed over a regular grid. A grid spacing of 2×2 km^2 was chosen, consistent with the resolution of the meteorological model output.

b. Areal reduction factors

The maximum rainfall intensity of a storm for a given accumulation duration decreases with increasing surface area. This property can be used to derive a probable rainfall intensity level for a given surface area A when a single observation is available, by applying areal reduction factors.

The ARF curves are then computed as the ratio between areal rainfall and point rainfall for a given return period, T_R:

$$\mathrm{ARF}(A, D, T_R) = \frac{I_A(A, D, T_R)}{I_0(0, D, T_R)}, \tag{1}$$

where I_A is the areal rainfall over the area A, for duration D, and I_0 is the point rainfall for the same duration D. By definition, the fixed-area ARF curves continuously decrease with increasing surface area, and their maximum is 1 (corresponding to the storm center $A = 0$, where the areal and point rainfall maxima are equal).

The empirical ARF curves are derived from the spatial rainfall database. Omolayo (1993) showed that statistically significant ARF curves can be derived within a fixed-area

framework, while ARFs based on a storm-centered approach are significantly underestimated.

In practice, storm-centered ARFs evaluate the rainfall intensity decay of selected events over concentric windows of increasing size, while fixed-area ARF curves rely on the maxima over moving windows of increasing size. In the former approach, concomitance between the maxima observations is not required. The loss of physical significance of fixed-area ARF curves is compensated by the gain in statistical significance.

According to empirical evidence (Bacchi and Ranzi 1996), the ARF decreases with the frequency of occurrence of the event. The ARF of heavy rainfall events (i.e., low frequencies) decreases with increasing surface area (A) faster than the ARFs of mixed heavy and regular rainfall events (i.e., high frequencies). Weak rainfall events can extend over large regions, leading to flatter ARF curves, while more intense deep convective events, for instance, are more localized, leading to steeper ARF curves. To consider independent and identically distributed (iid) samples, each ARF is computed on the basis of the highest 32 rainfall events over the 16-yr period for each surface area and therefore corresponds to a return period of 0.5 yr.

Due to the limited sample set size, the ARF dependence on T_R cannot be assessed for large return periods in this study. We therefore assume that ARF curves are independent of the return period T_R, as was usually done in previous studies (e.g., NERC 1975):

$$\mathrm{ARF}(A, D) = \frac{I_A(A, D)}{I_0(0, D)}. \tag{2}$$

Up to this point, the ARF curves have been computed empirically for a discrete number of surface areas and accumulation durations. Since we are interested in the relationship between the point and spatial rainfall for any surface area and accumulation duration, we adopt the ARF model proposed by De Michele et al. (2001).

De Michele et al. (2001) worked on a space–time self-similar model of annual maxima. The model introduced the use of a "dynamic scaling" coefficient, expressing the relationship between spatial and temporal scales in the definition of the rainfall intensity. The concept of dynamic scaling, originally introduced by Venugopal et al. (1999) in agreement with the Taylor hypothesis of "frozen turbulence" (Taylor 1938), is physically consistent at small space–time scales (sizes of up to 100 km^2 and accumulation durations of less than 1 h). The Taylor hypothesis implies that the temporal variation at fixed locations can be reinterpreted as a spatial variation (Deidda 2000). At larger scales, dynamic scaling is not related to the physics of the phenomenon, conserving only its statistical significance.

TABLE 1. Scale-invariant ARF model parameters (from De Michele et al. 2001) for region 1 (flat land) and region 2 (mountainous terrain).

Region	ω	a	b	υ
1: Flat area	0.00632	0.55	0.34	0.84
2: Mountainous region	0.00234	0.52	0.14	0.64

The ARF formulation proposed by De Michele et al. (2001) is

$$\mathrm{ARF}(D, A) = \left[1 + \omega\left(\frac{A^a}{D^b}\right)\right]^{-\upsilon/b}, \tag{3}$$

where υ is the scaling exponent of point rainfall with time, ω is a homogenization parameter, and a and b express the power-law decay of ARF curves with the surface area and duration, respectively. The dynamic scaling factor z is related to this expression by the relation $z = a/b$.

Equation (3) gives values that agree substantially with the empirical results presented by NERC (1975). When fitting this model, it is important to take into account the possible undersampling due to the rain-gauge-network density. Considering the average rain-gauge density of $1/50\ \mathrm{km}^{-2}$, fitting will not take into account surface areas less than 50 km^2 and durations less than 2 h.

The regional heterogeneity of extreme rainfall behavior is the main factor limiting the interpretation of spatial rainfall occurrences for large areas. Orography, in particular, forces the anisotropy and increases the temporal persistence of rainfall fields (Prudhomme and Reed 1999; Haberlandt 2007; Berne et al. 2009; Godart et al. 2009). To accurately compute ARF curves, we split our study region into two domains that are assumed to be quasi-homogeneous in terms of extreme rainfall behavior (Fig. 1), based on the results of previous studies (Ceresetti et al. 2010). We therefore consider a flat subregion (region 1), located in the southeast part of the study region, extending up to the foothills of the Cévennes Mountains, and a mountainous subregion (region 2), composed of the mountain ridge and the Massif Central highlands located in the northwest part of the study region. The anisotropy of ARF curves has not been taken into account due to the limited sample set size. The results of the fitting are shown in Table 1 and are plotted in Fig. 4.

For both subregions, Fig. 4 shows, as expected, a regular decrease in areal reduction factors as the area of the involved surface increases, and a corresponding increase in ARFs with duration (corresponding to a similar parameter a for the two regions). The accumulation duration of 1 h was not used for the fitting due to the undersampling issues described above. Except for the 1-h accumulation

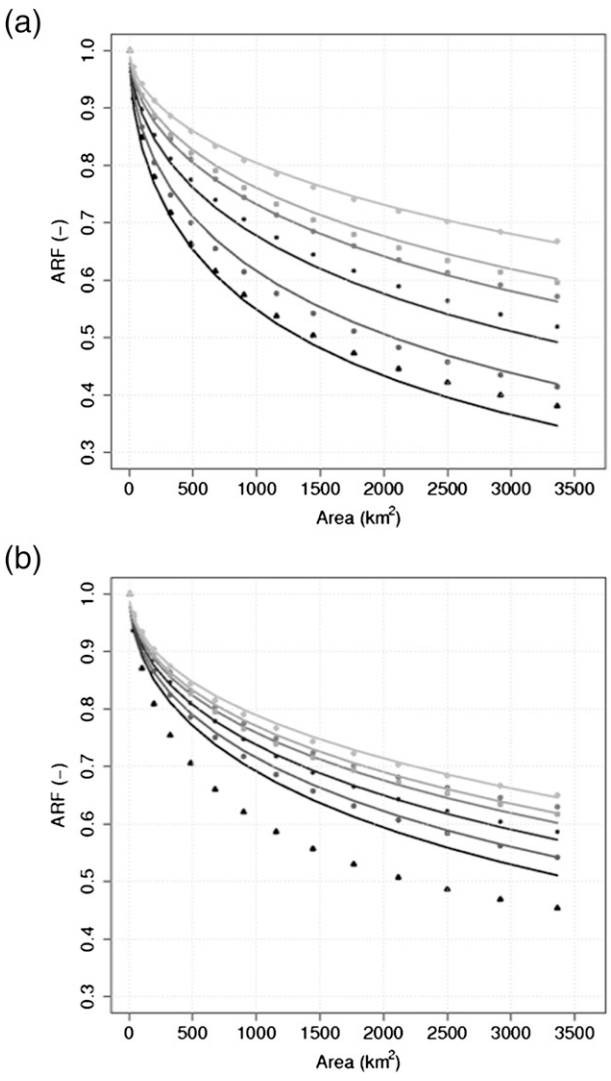

FIG. 4. Modeled ARF curves and empirical ARF values (dots) for different durations as a function of the area: (a) region 1, flat land; and (b) region 2, mountainous terrain. In both diagrams, the duration of 1 h is not used for ARF model fitting, but the empirical as well as modeled curves are plotted. The different durations are indicated from top to bottom by diamonds, 24 h; squares, 12 h; circles, 8 h; small circles, 4 h; circles, 2 h; and triangles, 1 h. Durations lower than 2 h and surface areas lower than 50 km^2 have not been taken into account for fitting due to undersampling of volumes.

duration, the model fits the experimental data well. The main difference between the two subregions is the effect of the accumulation duration on ARF curves, which is significantly lower in the mountainous subregion (the model shows the differences in the values of the *b* parameter in the two regions). This phenomenon could be physically explained by the persistence of mesoscale convective systems over mountainous regions (Sénési et al. 1996; Ducrocq et al. 2003; MCABC).

c. Intensity–duration–frequency curves

To construct severity diagrams, a continuous IDF model is required to estimate the frequency of occurrence of given rainfall observations when knowing their intensity and accumulation duration. We used the daily rainfall database to build a regional IDF model. Since we analyze three uncommonly heavy events, the challenge is in estimating the corresponding long-return periods for such subdaily events. Given that the longest daily series features a 50-yr dataset, the uncertainty in the estimation of return periods higher than 100 yr is too great to provide a reliable value for the return period.

Based on work carried out on the scale invariance of IDF and depth–duration–frequency curves (Burlando and Rosso 1996; Bendjoudi et al. 1997; Menabde et al. 1999; Borga et al. 2005), we were able to implement a scale-invariant model for IDF curves. The model assumes that, for the region of interest, the maxima are distributed according to a generalized extreme value distribution (GEV). Ceresetti et al. (2010, manuscript submitted to *Water Resour. Res.*, hereafter CMC) have shown that, at least in the 4–100-h range, the maximum rainfall intensities are scale invariant and they therefore propose a regional GEV simple-scaling model.

Combining the ARF dynamic scaling model fitting of section 3b and the simple-scaling IDF model of CMC, we derive the intensity–duration–area (IDA) curves for the two subregions of the domain in agreement with De Michele et al. (2001). Figure 5 shows the IDA obtained on the basis of the 32 heaviest observations from 16 yr of data, which therefore correspond to a return period on the order of 0.5 yr. It is easy to see that the rainfall intensity decreases with the area and with the accumulation duration because of the smoothing introduced by spatial and temporal integrations. The IDA model satisfactorily reproduces the empirical behavior of extreme values in space and time.

Based on the study presented in this section, we now have a continuous IDA model that is able to provide an estimation of the probability of occurrence of any spatial rainfall observation within the study region. We now turn to the description of the events analyzed in this study.

4. Description of the events

Three storms are studied: event 1 occurred on 3 December 2003, event 2 on 6 September 2005, and event 3 on 8–9 September 2002. These events differ in terms of the structure, extension, and location of the rainfall intensity maxima. They therefore represent an assorted selection of the extreme meteorological situations observed in the region. For the three events, the observed and

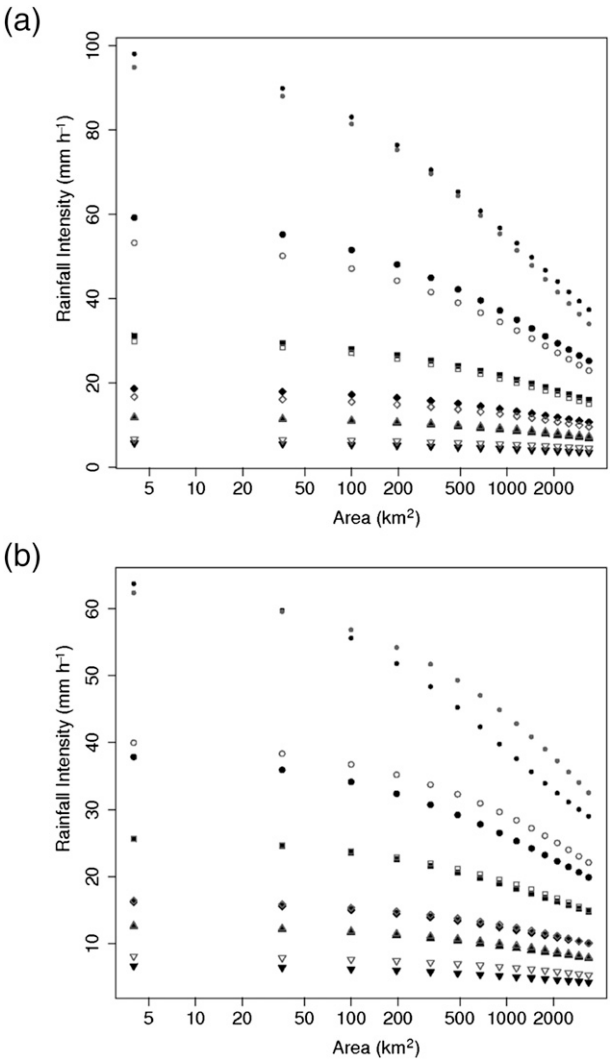

FIG. 5. The IDA model obtained by combining IDF and ARF models: comparison between observed and modeled events based on De Michele et al. (2001). From top to bottom, the plotted accumulation duration is 1, 2, 4, 8, and 24 h. Filled symbols represent empirical data and open symbols represent model results. The two graphs are for (a) region 1 (flat land) and (b) region 2 (mountainous terrain).

simulated rainfall amounts are analyzed. The rainfall output of the model is provided on a regular grid in geographic coordinates (size of about 2.4×2.4 km^2). It is regularized in a Cartesian grid (2×2 km^2) by nearest-neighbor interpolation. In the following, the main synoptic features of each event are described.

As explained by Lebeaupin et al. (2006), event 1 lasted a total of 4 days, starting 30 November 2003 and ending 4 December 2003. On 1 December, an upper-level low pressure area centered over Spain favored an intense southerly flow over southern France. A cold surface front was established from northern to southeastern France. The frontal perturbation formed a mesoscale convective system (MCS) that remained until 3 December, resulting in the 3-day accumulated precipitation reaching about 300 mm over the Rhône valley. The embedded convection inside the frontal perturbation was most active on 3 December. This study focuses on this specific day (from 0000 to 2400 UTC; see Table 2).

Event 2 is between 5 and 9 September 2005 (Yan et al. 2009), where several precipitating systems affected the southeast of France leading to an accumulated rainfall depth of greater than 300 mm in most of the region of interest. During the night between 5 and 6 September, heavy precipitation fell over the west of the Gard basin, reaching over 300 mm in the city of Nîmes (Fig. 6). Despite this heavy rainfall, the runoff process was limited by the initial dryness of the soil. Weaker precipitation was observed on 7 September, followed by a precipitation event coming from the Mediterranean Sea on the morning of 8 September, affecting the Gard basin. The intensity of the rainfall event increased during the afternoon, producing a total rainfall amount of 220 mm near Nîmes. The high soil moisture level in this second phase of the event caused significantly higher runoff.

As described in Delrieu et al. (2005) and Nuissier et al. (2008), the first convective cells of event 3 appeared over the Mediterranean Sea around 0400 UTC on 8 September 2002. Four hours later, the convection formed an MCS just south of the Gard basin and moved northward. The convective system remained over the same region for approximately 24 h (from 1200 UTC 8 September to 1200 UTC 9 September). This is the temporal window analyzed in this study. During this period, a high-level cloud shield displayed a V shape with the tip of the V facing the upper-level southerly flow (Nuissier et al. 2008). Beneath the cloud shield, the convective precipitation mainly affected the Gard region, while stratiform precipitation extended farther to the north. Over this period, the

TABLE 2. Characteristics of the numerical simulations.

	Event 1 Lebeaupin et al. (2006)	Event 2 Yan et al. (2009)	Event 3 Nuissier et al. (2008)
Initialization	ARPEGE	3DVAR ALADIN	3DVAR ALADIN
Initial time	0000 UTC 3 Dec 2003	0000 UTC 6 Sep 2005	1200 UTC 8 Sep 2002
Duration (h)	24	18	24

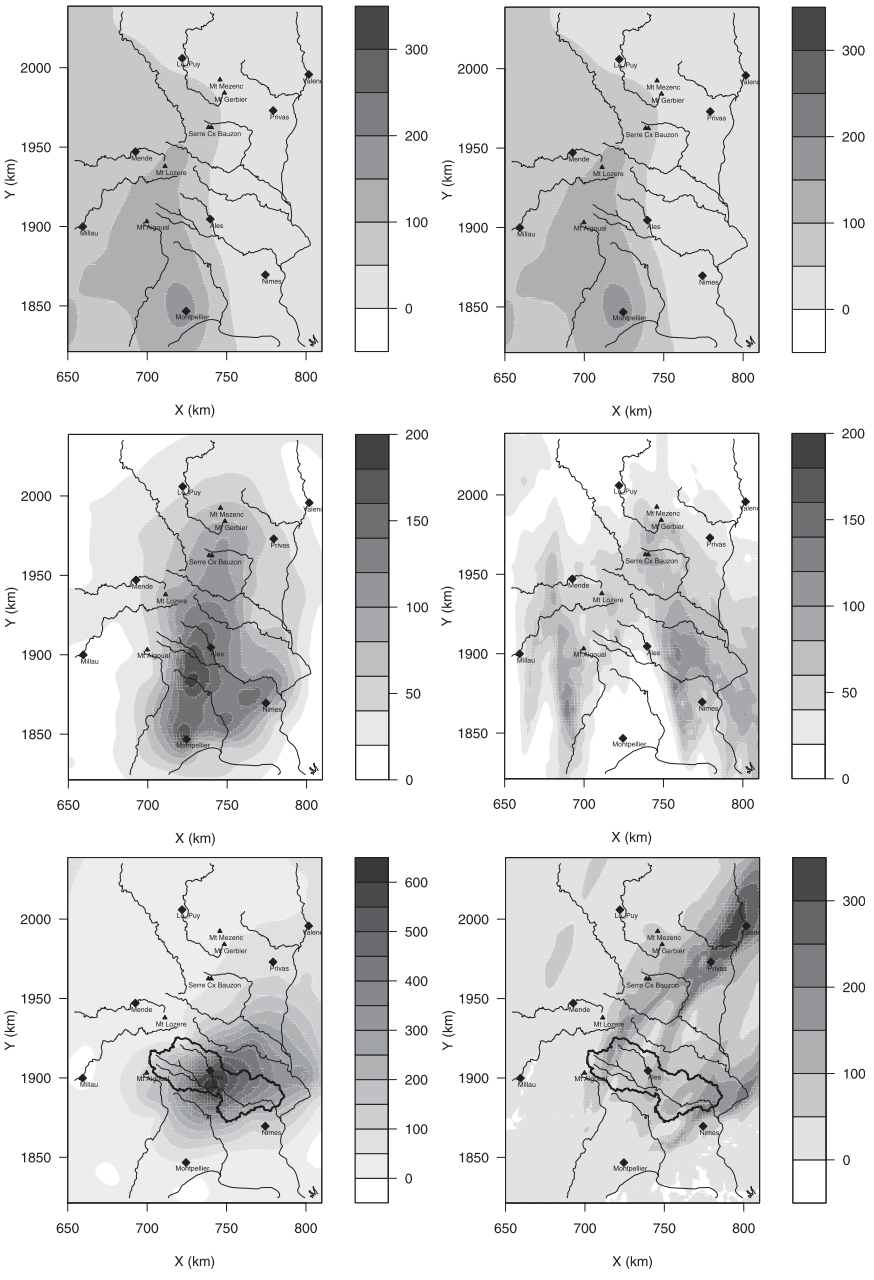

FIG. 6. Comparison between (left) observed and (right) simulated accumulated rainfall depths for (top to bottom) the three events considered. (bottom) Note the different scales used for the observed and simulated fields of event 3.

MCS was oriented southwest–northeast, in line with the prevailing upper-level tropospheric flow. Late in the night of 8 September, the convective system took on a north–south orientation and began moving in the northwestward direction. The precipitating system decayed late in the morning of 9 September.

Event 3 was exceptional from many points of view. The intensity of the event was extreme: the maximum precipitation was around 600 mm in 24 h near Alès (Fig. 6). The area affected by heavy precipitation (at least 200 mm) was considerable, extending over more than 3000 km^2, covering the whole Gard basin as well as the Massif Central foothills in the Ardèche region. River discharges were exceptional, especially for the Gard and Vidourle Rivers, where peaks higher than twice the 10-yr discharges were recorded (Delrieu et al. 2005). The event

TABLE 3. Contingency tables and definitions of the thresholded statistics for the three events.

	Obs ≥ threshold	Obs < threshold
Forecast ≥ threshold	***A***	***b***
Forecast < threshold	***C***	***d***

$$\mathrm{POD} = \frac{a}{a+c};\ \mathrm{FAR} = \frac{b}{a+b};\ \mathrm{ACC} = \frac{b+d}{a+b+c+d};\ \mathrm{FBIAS} = \frac{a+b}{a+b+c+d};$$

SCORE	Event 1			Event 2			Event 3		
Quantile	70%	80%	90%	70%	80%	90%	70%	80%	90%
POD	0.79	0.78	0.84	0.58	0.54	0.21	0.93	0.80	0.29
FAR	0.32	0.43	0.62	0.54	0.60	0.64	0.01	0.02	0.14
ACC	0.82	0.84	0.85	0.67	0.74	0.88	0.92	0.79	0.31
FBIAS	1.17	1.36	2.19	1.28	1.36	0.58	0.94	0.82	0.34

caused damage estimated at 1.2 billion U.S. dollars and led to 25 fatalities.

We will now compute the maximum intensity and severity diagrams of the three events. We use event 3 to test the ability of the approach to assess the hydrological impacts of a storm, isolating the area where the damage was the most important (i.e., the Gard River basin; see Fig. 6).

In Fig. 6, the observed and simulated rainfall depths are plotted for the three events, revealing a number of general features:

- The study domain is large enough to include the entire event.
- The simulated rainfall depths are underestimated for events 2 and 3, whereas the model overestimates the rainfall located in the mountainous area during event 1.
- The locations of the rainfall maxima are relatively well estimated by the model for event 1 and less so for events 2 and 3.
- The spatial extent of the rainfall pattern is captured by the model. However, the model does not correctly reproduce the intensities and location patterns of the rainfall fields (Fig. 6).

Table 3 shows the main thresholded scores (probability of detection, POD; false alarm ratio, FAR: forecast accuracy, ACC) of the simulations for the three events. These scores come from contingency tables (Yates 1984). The threshold levels have been chosen in agreement with Yates et al. (2007) as the 70th, 80th, and 90th percentiles. According to these statistics, event 1 is modeled with good accuracy, even though a nonnegligible false alarm ratio is found for the 90th percentile. Event 2 is poorly reproduced, with low PODs and elevated FARs for the three thresholds. For event 3, the mislocated maximum rainfall depth clearly causes very low scores for the 90th quantile. Considering that both the POD and ACC scores are very poor, it seems that the extreme rainfall has been incorrectly estimated for this event. We emphasize the difficulty of using these scores within a multiscale framework, in particular for hydrologic purposes.

5. Multiscale evaluation of the simulated events

We will now demonstrate the importance of a complete multiscale evaluation of the simulated fields. For this, we first describe the maximum intensity diagrams. They allow a preliminary multiscale comparison between events based on rainfall intensities.

Remember that the observed field is obtained by interpolating ground measurements made at rain-gauge sites. As a consequence, the maximum intensity and severity diagrams of the observed fields are also affected, over certain ranges, by undersampling errors of the measurement network. The spatial undersampling affects areas in the range 0–50 km^2, for which the rain-gauge density is inadequate. For durations less than 4 h, rainfall intensity data are affected by uncertainties due to the resolution of the time series (1 h).

a. Maximum intensity diagrams

The maximum intensity diagrams report the maximum rainfall intensity recorded during the event for each accumulation duration and surface area. These diagrams, based on rain-gauge observations, are given in Figs. 7a, 8a, and 9a for the three events, respectively. The corresponding diagrams based on the simulations are plotted in Figs. 7b, 8b, and 9b. The diagrams in Figs. 7 and 8 indicate that, for events 1 and 2, the observed and simulated maximum intensities are in good agreement for accumulation durations longer than 4 h. On the other hand, the maximum intensities of event 3 (Fig. 9a) seem to be largely underestimated by the simulation at any scale.

Small spatial and temporal scales reveal interesting features for events 1 and 3. The diagrams in Figs. 7a, 9a, and 9b show deviations from the monotonic decrease of maxima with space and time scales. The deviation observed for event 1 for durations in the range of 4–6 h and surface areas

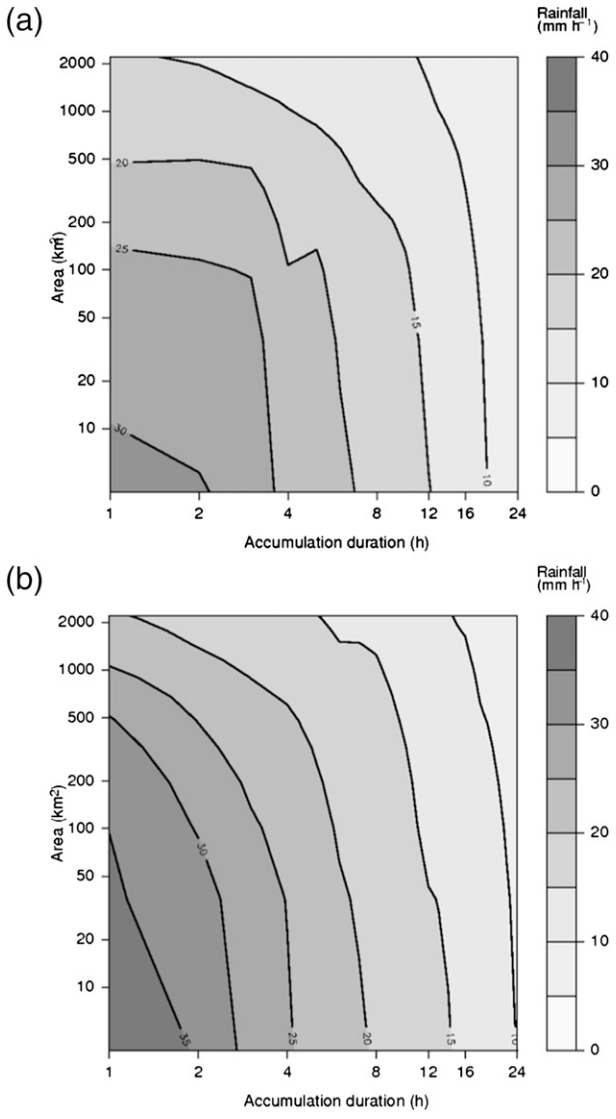

FIG. 7. Maximum intensity diagram for event 1, from 0000 UTC 3 Dec to 0000 UTC 4 Dec 2003 for (a) observed hourly rainfall and (b) simulated rain fields.

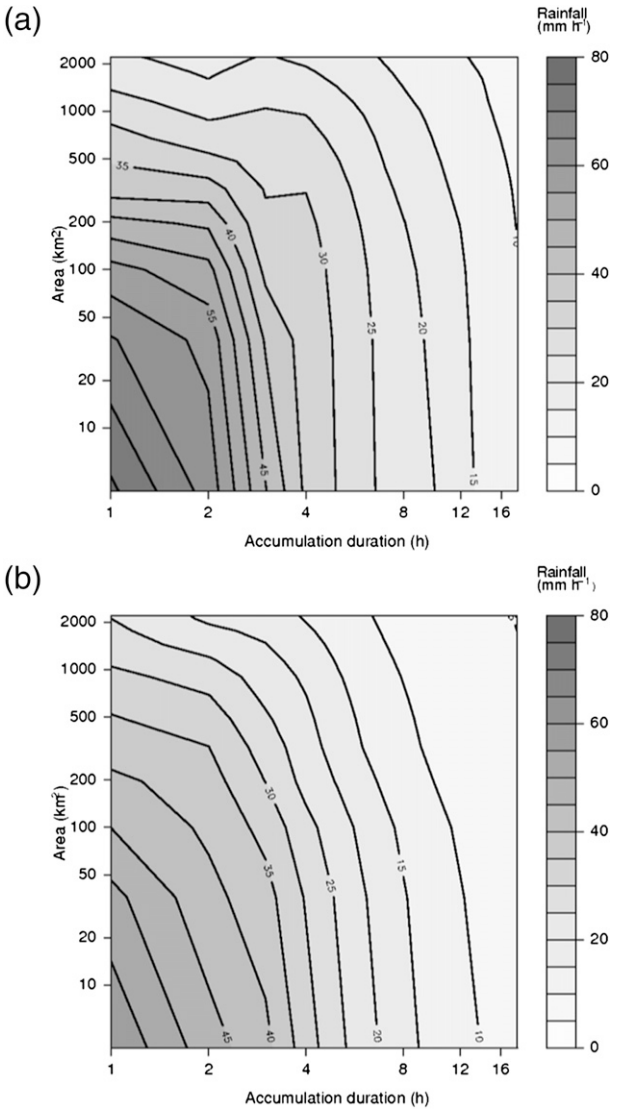

FIG. 8. Maximum intensity diagram for event 2, from 0000 to 1800 UTC 6 Sep 2005 for (a) observed hourly rainfall and (b) simulated rain fields.

between 20 and 200 km^2 is not present in the diagram of the simulated field. For event 3, the diagrams of simulated (Fig. 9b) and observed rain fields (Fig. 9a) both show small-scale deviations, but they appear at different scales.

b. Storm severity: A forecast qualification approach

The severity diagrams allow multiscale comparisons in terms of return period. Figure 10a displays the severity diagram of observations for event 1. The observed severity presents a maximum higher than 300 yr involving large temporal and spatial scales (time scales ranging from 8 to 16 h and for surface areas up to 400 km^2). For small surface areas and durations, the event did not provide significant severities. The simulated fields provided by Méso-NH yield a severity diagram (Fig. 10b) similar to the observed diagram (Fig. 10a). The maximum severity is of the same order of magnitude, as well as the spatial extent of the event. On the other hand, the critical time scale has been overestimated (14–18 h compared to 10–12 h for the observations). In general, the main features of the event seem to be well reproduced by the simulation in scale and return period.

For event 2, the severity diagrams from the observations (Fig. 11a) and from the simulation (Fig. 11b) are very different in terms of magnitude. Figure 11a shows that the absolute severity maximum is recorded at scales lower than 4 h and 50 km^2. A similar pattern is found for the simulation (Fig. 11b). In addition,

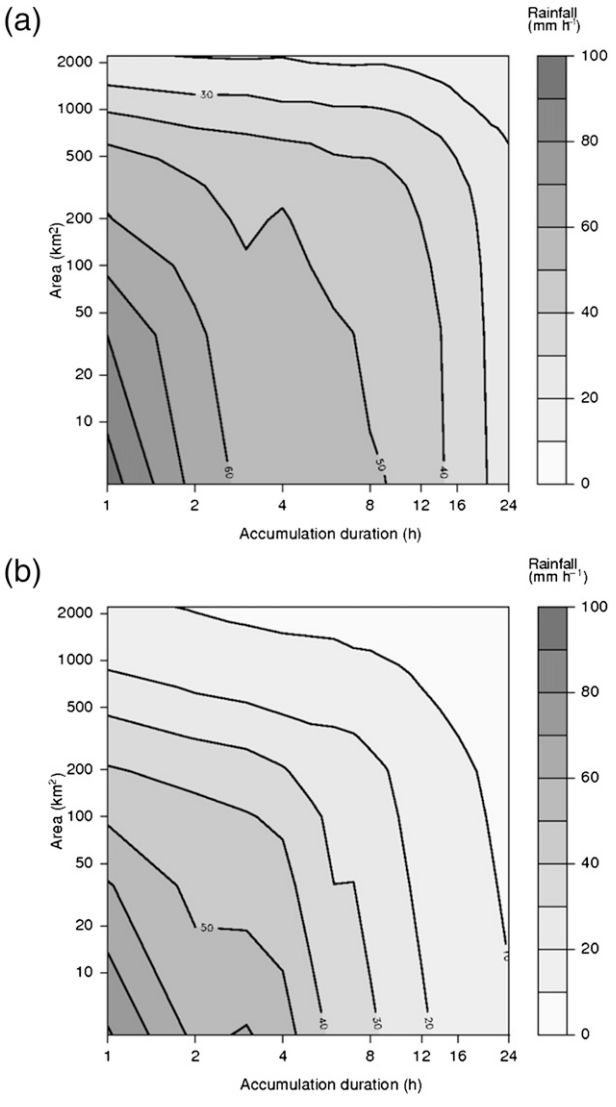

FIG. 9. Maximum intensity diagram for event 3 (over the Gard basin), from 08.09.2002 12 h UTC to 09.09.2002 12 h UTC for (a) observed hourly rainfall and (b) simulated rain fields.

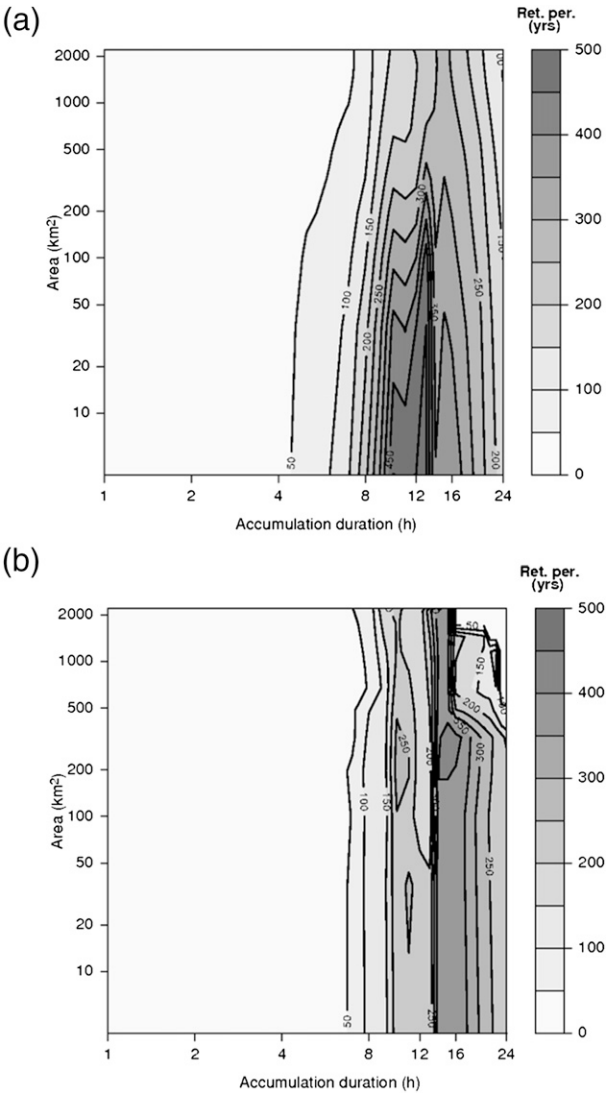

FIG. 10. Severity diagram for event 1, from 0000 UTC 3 Dec to 0000 UTC 4 Dec 2003 for (a) observed hourly rainfall and (b) simulated rain fields.

a secondary maximum is observed for the duration of 8–12 h and for surface areas lower than 500 km^2, corresponding to a rainfall amount higher than 200 mm. This severity peak is not reproduced by the simulated data.

These two cases clearly demonstrate the importance of drawing severity diagrams when evaluating atmospheric models. As opposed to maximum intensity diagrams, severity diagrams highlight the time and space structure of the storm, simplifying a multiscale comparison between observations and simulation outputs.

Event 3 is by far the largest storm in our records. Consequently, the return period associated with the event cannot be correctly assessed due to the uncertainties involved in the extrapolation of the extreme behavior for large return periods. Therefore, the severity of this storm is not analyzed over the whole studied region but in the most stricken area, the Gard basin, bearing in mind the hydrological consequences of this extreme rainfall.

c. *Hydrological aspects of storm severity: 8–9 September 2002*

During the second half of the twentieth century, studies on extreme flood events showed that river flow data could not provide reliable estimations for large quantiles (Guillot and Duband 1967; Guillot 1993). Because of the limits of flow measurements, it is preferable to study the occurrence of the rainfall process. Even if the rainfall–runoff relation modifies the impacts of storms on a basin,

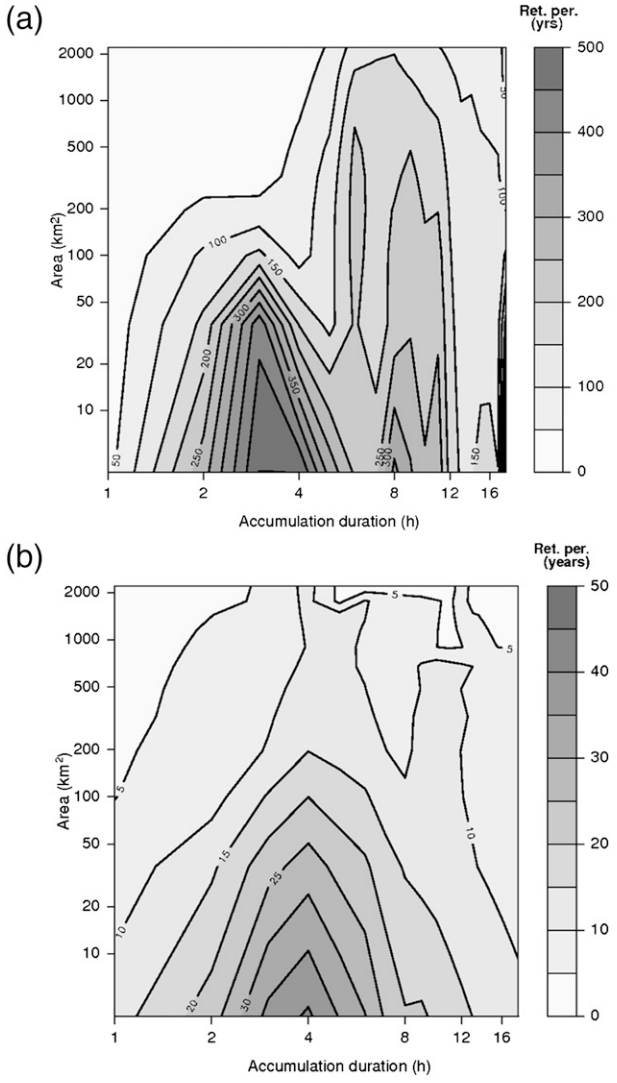

FIG. 11. Severity diagram for event 2, from 0000 to 1800 UTC 6 Sep 2005 for (a) observed hourly rainfall and (b) simulated rain fields.

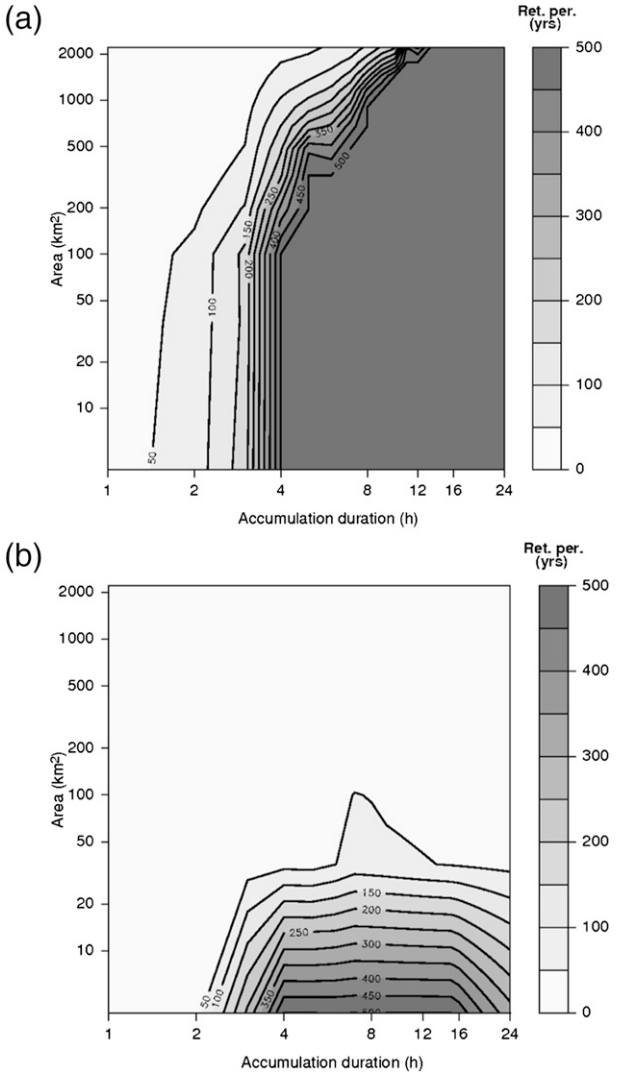

FIG. 12. Severity diagram for event 3 (within the Gard basin), from 1200 UTC 8 Sep to 1200 UTC 9 Sep 2002 for (a) observed hourly rainfall and (b) simulated rain fields. Due to the large uncertainties in return period estimation, the absolute severity values for return periods > 500 yr are not reported.

for rainfall events characterized by return periods higher than 10 yr, the transfer function between runoff and rainfall can reasonably be considered equal to 1. This supports the use of severity diagrams for the assessment of the impacts of a storm over a basin. From a practical point of view, we consider only the rainfall field over the basin of interest.

In the present case (event 3), the severity is estimated with respect to the Gard basin at the Remoulins outlet (indicated in Fig. 1), with a maximum surface area of 2200 km^2. The severity analysis will be conducted on spatial scales lower than the basin surface area. To avoid the analysis of spurious return periods, the severity diagram has therefore been limited to maximum severities of 500 yr. The severity diagram related to the observations is reported in Fig. 12a. Despite the large uncertainties in the evaluation of the return period, the severity shows a sharp increase with the accumulation duration, already reaching severities larger than 500 yr for the 4-h duration. The critical scales of the event are reached for an accumulation duration of 16–24 h and a surface area of 500 km^2. This means that a small subregion within the basin (fortunately not corresponding to a catchment basin) received an extremely heavy rainfall.

The severity diagram for the simulation is reported in Fig. 12b. It is clear that the simulation provides much lower severities. The critical space–time scales of the event have not been properly reproduced by the simulation.

In Fig. 12b, the severities are negligible for surface areas > 50 km^2. This result is likely due to the wrong location of the simulated rainfall maxima (as can be seen in the panels concerning event 3 in Fig. 6), leading to poor performance from a hydrological point of view.

6. Discussion

The multiscale analysis of maximum rainfall intensities and return periods (severity) reveals that severity diagrams provide more sensitive diagnostics than maximum intensity diagrams:

- The simulated rainfall fields of events 1 and 2 are of good quality in terms of maximum intensities over a large range of scales (Figs. 7 and 8).
- The severity diagrams of these two events (Figs. 10 and 11) show significant differences between simulated and observed rainfall fields.

These differences can have a dual origin. The first lies in the nonlinear transformation of rainfall intensity into frequency of occurrence. The return period is obtained as $1/(1-P)$, where P is the cumulative density function of the extreme value relation, obtained by a double exponentiation (in the simplest case) of the intensity. [See Kotz and Nadarajah (2000) for further information.] Therefore, a small difference in maximum intensities is amplified when transformed into a return period.

The second origin of the differences lies in the fact that the severity is highly dependent on the location. Taking for example daily rainfall, CMC show that the 100-yr level varies from 100 mm (over the Massif Central plateau) to 400 mm (over the mountain ridge and the southeastern foothills). Therefore, the same storm occurring at two different locations can give very different severities. Severity diagrams incorporate this effect and their analysis can be of interest when the simulated rainfall fields are used as input to hydrological models and when the storm covers the entire region.

This analysis also reveals that the two diagnostics (severity and maxima diagrams) are complementary. For events 1 and 3, the maximum intensity diagrams highlight the deviation from a monotonic decrease of maxima with increasing space and time scales. Some of these deviations are not readable in the corresponding severity diagrams. On the other hand, severity diagrams can be used to delineate the critical scales of an event providing additional information for comparing the actual and simulated events. For event 1, for instance, the critical time scale is slightly overestimated by the model (16 instead of 12 h). The spatial extent of the region affected by return periods of the order of 200 yr has been correctly modeled, since in both panels of Fig. 10 it is possible to detect severities higher than 200 yr, along a vertical cut, from 0 to 2000 km^2. On the other hand, the time and space scales of return periods highest than 400 yr are poorly reproduced. For event 2, even though the observed and simulated maximum intensities are similar (Fig. 7), the severity has an order of magnitude of difference, which means that the simulation essentially failed to identify the correct location of rainfall cells.

Summarizing the results obtained by comparing the severity diagrams for the three cases analyzed, we can conclude that the Méso-NH meteorological model reproduces many of the features of the events. However, regarding the largest integrated rainfall values, the model (i) underestimates in the plain region, (ii) overestimates over the foothills due to orographic effects, (iii) poorly reproduces some of the small-scale features of the events, and (iv) mislocates the maxima, making it impossible to use the simulations in hydrology.

Throughout our discussion, we have not expressly interpreted the diagrams for time scales lower than 2 h and spatial scales lower than 50 km^2. The reason is that, at these scales, the observation network is subject to undersampling problems leading to uncertainties in the measurement of point rainfall and, to a greater extent, spatial rainfall. The same problems affect the simulated fields, since the time resolution of the rainfall fields is 1 h and their spatial resolution is 6.25 km^2. Note that these undersampling problems affect not only the event measurements but also the estimation of the IDF and ARF curves that are required to define the climatology of extreme rainfall events.

7. Conclusions

The main purpose of the present work was to show the utility of (i) a multiscale assessment of the largest values in simulated rainfall fields and (ii) an assessment of the severity (i.e., return period) of rainfall events at multiple scales. This work proposes an extension of severity diagrams (Ramos et al. 2005) to larger surface areas and accumulation durations. We also introduce maxima intensity diagrams, a preliminary diagnostic showing the maximum rainfall intensity of the event at each space–time aggregation scale. The use of severity diagrams within a regional context offers the possibility of illustrating the effects that an incorrect positioning or an over–underestimation of the maxima rainfall depth has on the severity of an event. It also detects the spatial–temporal scales at which the model has limited performance and provides support for evaluating whether the space–time patterns of the largest values in given storms have been correctly identified.

From these analyses, it turns out that the severity diagram has three main merits: (i) it is useful for an objective

comparison of the violence of a given storm over different space–time scales, detecting its critical scales; (ii) it provides a useful complementary indicator for comparison between severe events, as shown by Ramos et al. (2005); and (iii) it offers an innovative way to evaluate the ability of the mesoscale model to reproduce the space–time structure of the most intense parts of rainfall events.

Despite the above many benefits, the use of severity diagrams presents certain limitations: (i) the spatial heterogeneity of the extreme rainfall climatology in the region may make it impossible to assign a unitary and reliable severity value to spatially extended rainfall observations, (ii) the maximum occurrence frequency that can be assigned to an observation strongly depends on the available rainfall database and should not exceed the observation period, and (iii) with severity diagrams, the space–time coordinates are lost in favor of a multiscale description.

Acknowledgments. J. Labalette is deeply acknowledged for his valuable contribution to this study. This work has been performed within the framework of the French ANR-MEDUP project on the Forecast and projection in climate scenario of Mediterranean intense events: Uncertainties and Propagation on environment. The climatological data were collected and provided by the Cévennes-Vivarais Hydro-Meteorological Observatory (OHMCV). We are indebted to Dr. O. Nuissier for providing us with the results of the Méso-NH simulations carried out at Météo-France.

REFERENCES

Asquith, W., and J. Famiglietti, 2000: Precipitation areal-reduction factor estimation using an annual-maxima centered approach. *J. Hydrol.,* **230,** 55–69.

Bacchi, B., and R. Ranzi, 1996: On the derivation of the areal reduction factor of storms. *Atmos. Res.,* **42,** 123–135.

Bendjoudi, H., P. Hubert, and D. Schertzer, 1997: Interpretation multifractale des courbes intensité-durée-fréquence des précipitations. *C. R. Acad. Sci. Paris,* **325,** 323–326.

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

Bois, P., C. Obled, M.-F. De Saintignon, and H. Mailloux, 1997: Atlas expérimental des risques de pluies intenses: Cévennes-Vivarais. LTHE-LAMA Tech. Rep., Pôle Grenoblois d'étude et de Recherche pour las Prévention des Risques Naturels, 19 pp.

Borga, M., C. Vezzani, and G. Dalla Fontana, 2005: Regional rainfall depth–duration–frequency equations for an Alpine region. *Nat. Hazards,* **36,** 221–235.

Bousquet, O., C. Lin, and I. Zawadski, 2006: Analysis of scale dependence of quantitative precipitation forecast verification: A case study over the Mackenzie River basin. *Quart. J. Roy. Meteor. Soc.,* **132,** 2107–2125, doi:10.1256/qj.05.154.

Burlando, P., and R. Rosso, 1996: Scaling and multiscaling models of depth–duration–frequency curves for storm precipitation. *J. Hydrol.,* **187,** 45–64.

Ceresetti, D., G. Molinié, and J.-D. Creutin, 2010: Scaling properties of heavy rainfall at short duration: A regional analysis. *Water Resour. Res.,* **46,** W09531, doi:10.1029/2009WR008603.

Deidda, R., 2000: Rainfall downscaling in a space–time multifractal framework. *Water Resour. Res.,* **36,** 1779–1794.

Delrieu, G., and Coauthors, 2005: The catastrophic flash-flood event of 8–9 September 2002 in the Gard region, France: A first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory. *J. Hydrol.,* **6,** 34–52.

De Michele, C., N. Kottegoda, and R. Rosso, 2001: The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resour. Res.,* **37,** 3247–3252.

Ducrocq, V., J.-P. Lafore, J.-L. Redelsperger, and F. Orain, 2000: Initialization of a fine-scale model for convective-system prediction: A case study. *Quart. J. Roy. Meteor. Soc.,* **126,** 3041–3065.

——, D. Ricard, J. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Wea. Forecasting,* **17,** 1236–1256.

——, G. Aullo, and P. Santurette, 2003: Précipitations intenses et les inondations des 12 et 13 novembre 1999 sur le sud de la France. *Meteorologie,* **42,** 18–27.

——, O. Nuissier, D. Ricard, C. Lebeaupin, and T. Thouvenin, 2008: 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.

Godart, A., S. Anquetin, and E. Leblois, 2009: Rainfall regimes associated with banded convection in the Cévennes-Vivarais area. *Meteor. Atmos. Phys.,* **103,** 25–34.

Guillot, P., 1993: The arguments of the GRADEX method: A logical support to assess extreme floods. *Extreme Hydrological Events: Precipitation, Floods and Droughts,* Z. W. Kundzewicz et al., Eds., Vol. 1, *Proceedings of the Yokohama Symposium,* AHS Publ. 213, 287–298.

——, and D. Duband, 1967: La méthode du GRADEX pour le calcul de la probabilité des crues à partir des pluies. *Journées de la SHF,* Société Hydrologique de France, Paris, Société Hydrologique de France, Vol. Rapport 7, question 1.

Haberlandt, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J. Hydrol.,* **332,** 144–157.

Kotz, S., and S. Nadarajah, 2000: *Extreme Value Distributions: Theory and Applications.* Imperial College Press, 185 pp.

Koutsoyiannis, D., D. Kozonis, and A. Manetas, 1998: A mathematical framework for studying rainfall intensity–duration–frequency relationships. *J. Hydrol.,* **206,** 118–135.

Lafore, J. P., and Coauthors, 1998: The Méso-NH atmospheric simulation system. Part I: Adiabatic formulation and control simulations. *Ann. Geophys.,* **16,** 90–109.

Lebeaupin, C., V. Ducrocq, and H. Giordani, 2006: Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts. *J. Geophys. Res.,* **111,** D12110, doi:10.1029/2005JD006541.

Lebel, T., and J. P. Laborde, 1988: A geostatistical approach for areal rainfall statistics assessment. *Stochastic Hydrol. Hydraul.,* **2,** 245–261, doi:10.1007/BF01544039.

Mason, I., 1989: Dependence of the critical success index on sample climate and threshold probability. *Aust. Meteor. Mag.,* **37,** 75–81.

Menabde, M., A. Seed, and G. Pegram, 1999: A simple scaling model for extreme rainfall. *Water Resour. Res.,* **35,** 335–339.

Molinié, G., D. Ceresetti, S. Anquetin, J.-D. Creutin, and B. Boudevillain, 2012: Rainfall regime of a mountainous

Mediterranean region: Statistical analysis at short time steps. *J. Appl. Meteor. Climatol.,* in press.

NERC, 1975: Meteorological Studies. Vol. 2. Flood Studies Report, Natural Environmental Research Council, 10 pp.

Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, 2008: 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.

Omolayo, A., 1993: On the transportation of areal reduction factors for rainfall frequency estimation. *J. Hydrol.,* **145,** 191–205.

Prudhomme, C., and D. Reed, 1999: Mapping extreme rainfall in a mountainous region using geostatistical techniques: A case study in Scotland. *Int. J. Climatol.,* **19,** 1337–1356.

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

Rodriguez-Iturbe, I., and J. M. Mejía, 1974: On the transformation of point rainfall to areal rainfall. *Water Resour. Res.,* **10,** 729–735.

Ruin, I., J.-D. Creutin, S. Anquetin, and C. Lutoff, 2008: Human exposure to flash-floods relation between flood parameters and human vulnerability during a storm of September 2002 in southern France. *J. Hydrol.,* **361,** 199–213.

Sénési, S., P. Bougeault, J. Chèze, P. Cosentino, and R. Thepenier, 1996: The Vaison-la-Romaine flash flood: Mesoscale analysis and predictability issues. *Wea. Forecasting,* **11,** 417–442.

Sivapalan, M., and G. Blöschl, 1998: Transformation of point rainfall to areal rainfall: Intensity–duration–frequency curves. *J. Hydrol.,* **204,** 150–167.

Taylor, G., 1938: The spectrum of turbulence. *Proc. Roy. Soc. London,* **164A,** 476–490, doi:10.1098/rspa.1938.0032.

Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, 1999: Evidence of dynamic scaling in space–time rainfall. *J. Geophys. Res.,* **104** (D24), 31 599–31 610.

——, S. Basu, and E. Foufoula-Georgiou, 2005: A new metric for comparing precipitation patterns with an application to ensemble forecasts. *J. Geophys. Res.,* **110,** D08111, doi:10.1029/2004JD005395.

Yan, X., V. Ducrocq, P. Poli, M. Hakam, G. Jaubert, and A. Walpersdorf, 2009: Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall. *J. Geophys. Res.,* **114,** D03104, doi:10.1029/2008JD011036.

Yates, E., J.-D. Creutin, S. Anquetin, and J. Rivoirard, 2007: A scale-dependent quality index of areal rainfall prediction. *J. Hydrometeor.,* **8,** 160–170.

Yates, F., 1984: Tests of significance for 2×2 contingency tables. *J. Roy. Stat. Soc.,* **147A,** 426–463.

Zepeda-Arce, J., E. Foufoula-Georgiou, and K. Droegemeier, 2000: Space–time rainfall organization and its role in validating quantitative precipitation forecasts. *J. Geophys. Res.,* **105,** 10 129–10 146.