Hydrol. Earth Syst. Sci., 12, 1039–1051, 2008
www.hydrol-earth-syst-sci.net/12/1039/2008/

Hydrology and Earth System Sciences

The benefit of high-resolution operational weather forecasts for flash flood warning

J. Younis[1], S. Anquetin[2], and J. Thielen[1]

[1]Institute for Environment and Sustainability, European Commission DG Joint Research Centre, Ispra, Italy
[2]Laboratoire d'Etude des Transferts en Hydrologie et Environnement, Grenoble Université (CNRS, UJF, IRD, INPG), France

Received: 3 January 2008 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 February 2008
Revised: 20 June 2008 – Accepted: 22 June 2008 – Published: 30 July 2008

Abstract. In Mediterranean Europe, flash flooding is one of the most devastating hazards in terms of loss of human life and infrastructures. Over the last two decades, flash floods have caused damage costing a billion Euros in France alone. One of the problems of flash floods is that warning times are very short, leaving typically only a few hours for civil protection services to act. This study investigates if operationally available short-range numerical weather forecasts together with a rainfall-runoff model can be used for early indication of the occurrence of flash floods.

One of the challenges in flash flood forecasting is that the watersheds are typically small, and good observational networks of both rainfall and discharge are rare. Therefore, hydrological models are difficult to calibrate and the simulated river discharges cannot always be compared with ground measurements. The lack of observations in most flash flood prone basins, therefore, necessitates the development of a method where the excess of the simulated discharge above a critical threshold can provide the forecaster with an indication of potential flood hazard in the area, with lead times of the order of weather forecasts.

This study is focused on the Cévennes-Vivarais region in the Southeast of the Massif Central in France, a region known for devastating flash floods. This paper describes the main aspects of using numerical weather forecasting for flash flood forecasting, together with a threshold – exceedance. As a case study the severe flash flood event which took place on 8–9 September 2002 has been chosen.

Short-range weather forecasts, from the Lokalmodell of the German national weather service, are used as input for the LISFLOOD model, a hybrid between a conceptual and physically based rainfall-runoff model. Results of the study indicate that high resolution operational weather forecasting combined with a rainfall-runoff model could be useful to determine flash floods more than 24 h in advance.

1 Introduction

Flash floods are rapidly developing floods with devastating effects for the environment and high risk of loss of life. In Mediterranean Europe, flash flooding is classified as one of the most devastating hazards in terms of human life loss and infrastructures (Gruntfest and Handmer, 1999). Over the last two decades, flash floods caused a billion Euros of damage in France alone (Hauet et al., 2003, in French). Apart from the economic impact, flash floods are life threatening: 11 victims were reported in the Nîmes event (1988), 58 during the Vaison la Romaine flash flood (1992) (Sénési et al., 1996), 35 for the Aude storms (1999) (Ducrocq et al., 2003) and 24 for the 2002 Gard event (Delrieu et al., 2005; Hauet et al., 2003) which is presented here. Vulnerability to flash floods will probably increase in the coming decades due to evolving land use and the modification of the pluviometric regime associated with the evolution of the climate (Parry et al., 2007; Palmer and Raisanen, 2002; Rosso and Rulli, 2002).

A flash flood is typically the consequence of a short-duration storm event. The term "flash" refers to the rapid response, with water levels in the drainage network reaching a crest within minutes to a few hours after the onset of the rain event, leaving extremely short time for warning. Flash flood generating storms can accumulate more than 200 mm of rainfall in less than 6 h, over natural watersheds ranging in area from 25 to 2500 km^2 (Creutin and Borga, 2003; Collier, 2007). Over built-up areas of 1 to 100 km^2, flash floods

Correspondence to: J. Thielen
(jutta.thielen@jrc.it)

can be produced by storms of even shorter duration with accumulations of over 50 mm in less than one hour (Creutin and Borga, 2003; Collier, 2007). The rising rate of waters of several m.h^{-1}, and the flow velocity of several m.s^{-1}, make these floods far more dangerous for human lives than large river floods (excluding dam breaks). Furthermore, the danger also comes from the rarity of the phenomenon, which demands a new observation strategy, as well as new forecasting methodology.

The meteorological conditions leading to flash floods are mostly severe convective systems that typically develop under potentially unstable conditions released by very localized trigger mechanisms. Due to their very localized nature, the observation of these events with a gauging network is problematic. Weather radars can provide better spatial rainfall estimations. However, it has been demonstrated that the more intense the rainfall, the less reliable the rainfall estimates from radar become (Austin, 1987; BASC, 2005; Borga et al., 2002; Ciach et al, 2000). Therefore accurate monitoring and prediction of severe storm rainfall intensities continue to be a major challenge. Prediction of these events with numerical meteorological models is even more difficult (Fritsch et al., 1998; Anquetin et al., 2005; Chancibault et al., 2006; Yates et al., 2007) due to the strong interaction of different physical and micro-physical processes across different scales.

One accepted method for predicting flash floods in ungauged river basins is so called "flash flood guidance" (Georgakakos, 2006). Flash flood guidance is a general term referring to the average rain needed over an area and during a specific time in order to initiate rapid flooding in small streams. Depending on the method, in addition the antecedent soil moisture or precipitation from previous days is taken into account.

However, there is no unique and simple theory about the runoff production on watersheds during flash flood events. The main reason is that a variety of processes can be involved which are usually grouped in two types of overland flow: saturation excess (Dunne process) or infiltration excess (Horton runoff). Due to the high heterogeneity and space variability of the watershed characteristics (land use, soil type and depth, sub-soil, local slope, upstream contributing area) and to antecedent moisture conditioning, these processes are likely to be active at the same time in various combinations (Smith and Goodrich, 2005).

Therefore, in order to forecast flash floods reliably, the temporal and spatial resolutions of the meteorological and hydrological model should ideally be linked also at small scales. Recognition of the need to couple meteorological and hydrological processes in interpretative studies and in the development of predictive models for flash floods has been demonstrated (Anquetin et al., 2004).

In this study a regional approach for flash flood warning also in ungauged river basins is being proposed. The concept is based on the principle of discharge threshold exceedances as opposed to a rainfall exceedance (Georgakakos, 2006). A discharge threshold exceedance approach is currently being explored for river flood forecasting (Thielen et al., 2008; Ramos et al., 2007), but has also been recently applied for flash floods, for instance by Reed et al. (2007).

In this, the first feasibility study, a deterministic approach has been used. For future studies it is envisaged to test the approach within a probabilistic framework and use high resolution ensembles from limited area models. Similar studies in this direction have shown positive results (Chen and Yu, 2007).

In Sect. 2 of this paper, the September 2002 case study is described, and the main characteristics of the study area are presented. The hydrological model and the threshold concepts are explained in Sect. 3. In Sect. 4, results are illustrated for (a) the long-term simulations used for the threshold calculation, (b) the forecasting performance for the case study, and (c) a 6-month control forecasting run. Conclusions and the proposed way forward are illustrated in the final Section.

2 The flash flood event of 8–9 September 2002

2.1 The Cévennes – Vivarais region in Southern France

The Cévennes-Vivarais region (Fig. 1) is situated Southeast of the Massif Central, the V-shaped Hercynian mountain range of the central part of France, and covers 85 000 km^2; i.e. one sixth of the country's total area. The aspect is southeasterly, sloping from the Mediterranean shore and the Rhône Valley. Elevation varies from sea level up to 1700 m (Mount Lozère) over roughly 70 km. The area is characterized by relatively small catchments (few hundreds of km^2) with a short response time of less than 12 h. The main Cévennes Rivers are the Vidourle, Ardèche, Cèze, and Gard. They are characterized by a typical Mediterranean hydrological regime with very low level of water in summer and floods occurring mainly during the autumn.

For this study the four watersheds corresponding to the above four rivers are analyzed. The outlets at which comparison of observed with simulated data is performed, are listed in Table 1. Only those meteorological and hydrological stations with long-term records from 1990 to 2002 were chosen. In total, 11 meteorological stations and 15 discharge stations were selected (Fig. 1). Radar information was used from 2002.09.08, 06:00 onward for a period of 36 h.

2.2 Description of the 8–9 September 2002 case study

The 8–9 September 2002 heavy precipitation event was responsible for one of the most important floods ever recorded in the Cévennes–Vivarais region. It caused 24 casualties and economic damage estimated at 1.2 billion euros (Huet et al., 2003).

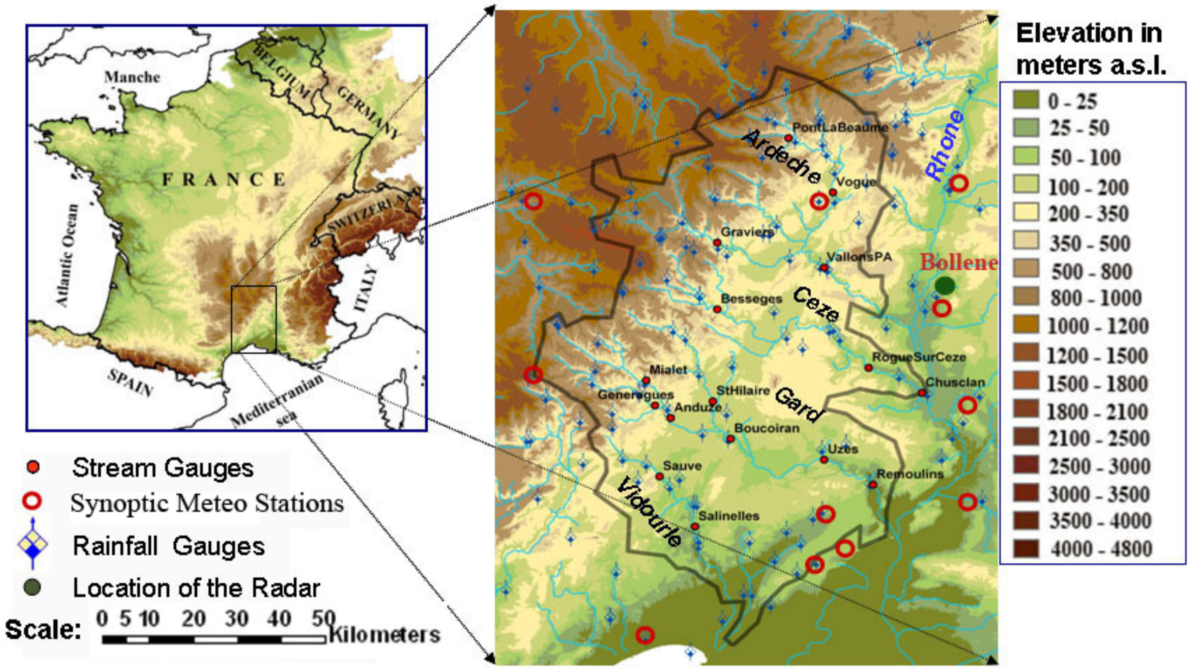

Fig. 1. Map of the topography of France and zoom into the study area of the Cévennes region with the catchments used for this study. Rainfall gauges from high density network (hourly data, blue flags), stream gauging stations (red dots), and synoptic meteorological stations (daily data, red circles). The location of the Bollène radar is also indicated in the map.

The event started early in the morning of 8 September 2002, with first convective cells forming over the Mediterranean Sea. The convection progressed northward to form, inland, a mesoscale convective system over the Gard. The quasi-stationary system remained over the region until the following morning, and then evolved eastward together with a surface front. For more detail on the event, refer to Delrieu et al. (2005). For the whole event, the raingauge network locally recorded 24 h of accumulated rainfall greater than 600 mm (Fig. 2a), which is also estimated by the radar observation (Fig. 2b).

For the Gard and Vidourle river watersheds, the peak discharges were observed to be two times higher than the ten years return period specific discharge (Delrieu et al., 2005; Chancibault et al., 2006). In Fig. 3, the maximum specific discharge recorded or retrieved from a post-event field experiment (Delrieu et al., 2005) illustrates the intensity of the hydrological response of the watershed within the region. Most of the estimated peaks indicate specific discharges of more than 5 $m^3\,s^{-1}\,km^{-2}$, with some over 20 $m^3\,s^{-1}\,km^{-2}$. These are the most important values ever reported for watersheds of similar areas in France (Delrieu et al., 2005). In this region, the 10 years return period discharge for such catchments is about 2 $m^3\,s^{-1}\,km^{-2}$. Figure 3 also indicates the characteristic size of the watershed affected by the flood, for which detailed rainfall fields have to be correctly forecasted.

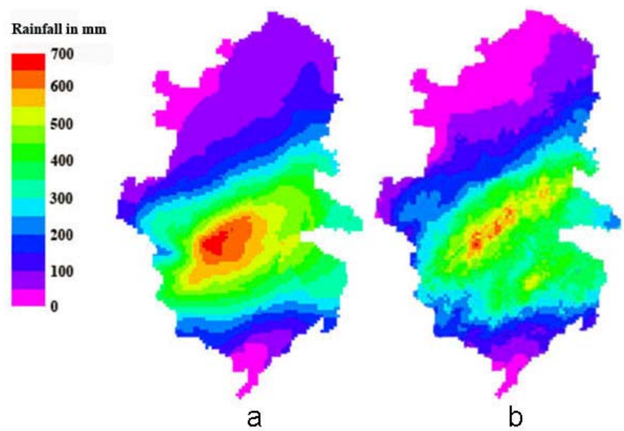

Fig. 2. Accumulated 48 h observed rainfall from 20020908 as observed by gauges **(a)** and by radar **(b)**. The locations of the gauges used for the interpolation are as shown in Fig. 1 and the region also as delineated in Fig. 1.

A more detailed description of the case study can be found in Delrieu et al. (2005), Chancibault et al. (2006) and Nuissier et al. (2008).

Table 1. List of river basins, discharge stations and their upstream areas used for this study. In addition the observed maximum daily discharge ($Q_{\max}$) for the study period 19990-2002 is given, the recorded maximum discharge at this station ($Q_{\max\_recorded}$) and at what date, the 2, 5 and 10 year observed return periods ($RP_{2,5,10\,yrs}$), and the high threshold calculated from the observed daily discharges.

Basin	Station	Upstream area [km^2]	$Q_{\max,(1990-2002)}/Q_{\max\_recorded}$ $RP_{2\,yrs}/RP_{5\,yrs}/RP_{10\,yrs}$ [m^3/s]	$Q_{obsHigh}$ (1990–2002) [m^3/s]
Ardèche	Ard102-Pont de Labeaume-V5004010	264	404/531(8/11/82) 220/320/390	134
	Ard105-Vogue-V5014010	636	830/890(23/10/77) 420/600/720	233
	Ard501-Chambonas-V5045020	476	608/878(10/11/76) 280/450/570	118
Cèze	Ceze102-Besseges-V5424010	230	262/475(23/10/77) 130/200/250	64
	Ceze104-Roque_sur_Ceze-V5474010	1060	1170*/1200(24/10/77) 410/670/840	171*
	Ceze106-Chusclan-V5474020	1180	1250*/1250 (21/10/94) -/-/-	209*
Gard	GAR103-Corbès-V7135010	263	416/510(23/10/77) 150/230/280	94
	GAR104-Généragues-V7124010	251	353/450(23/10/77) 150/220/270	82
	GAR203-St. Hilaire-Brethmas -V7155040	328	302***/302(20/10/94) -/-/-	88***
	GAR301-Boucoiran-V7164010	1087	424***/ (no info) -/-/-	236***
Vidourle	Vid103-Le Vidourle a Sauve-Y3414010	190	417** (9/9/2002) 120/180/220	47**
	Vid105- Le Vidourle a Salinelles-Y3444010	539	978**/(no info)	63**

*Total gap in data records >3 years;
**total gap in data records 3< years <6;
*** total gap in data records >6 years

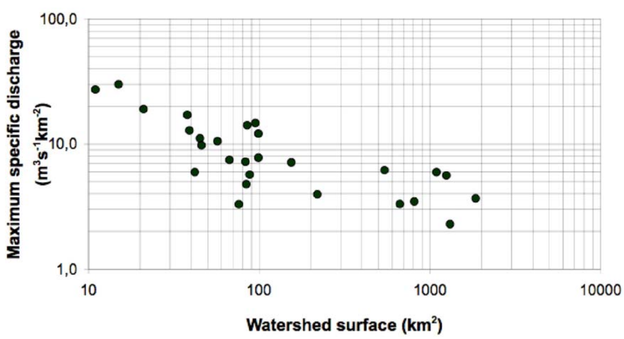

Fig. 3. Observed maximum specific discharges for the flash flood event in September 2002 in the Cévennes-Vivarais region. The plot is based on the data collected in the post-event field experiment as described in Delrieu et al. (2005).

3 Hydrological model, input data and methodologies

3.1 The hydrological LISFLOOD model

The hydrological model used in this study is the LISFLOOD model, a hybrid between a conceptual and physical rainfall-runoff model combined with a routing module in the river channel. LISFLOOD has been specifically designed for large river catchments (de Roo et al., 2001) but has also been applied to smaller watersheds (Everhardus et al., 2002). A brief model description is given below. More details can be found in the LISFLOOD manual (van der Knijff and de Roo, 2006).

For the simulation of fast subsurface flow through macropores (preferential flow), it is assumed that the fraction of the water on the soil surface contributing to preferential flow is a non-linear function of the relative saturation of the top-soil, and that the importance of preferential flow increases as the top-soil gets wetter. For the remaining water that falls on the soil surface, infiltration and surface runoff are simulated using the Xinanjiang approach (Zhao and Liu, 1995; Todini, 1996). The moisture fluxes out of the top- and sub-soil are calculated assuming that the flow is entirely gravity-driven.

The groundwater system is described using two parallel interconnected linear reservoirs, similar to the HBV-96 model (Lindström et al., 1997). The upper zone represents a mix of fast groundwater and subsurface flow, including flow through macropores. The lower zone has a much slower response and generates the base flow. Routing of water through the river channel is simulated with the kinematic wave (Chow et al., 1988). Whenever possible parameters in LISFLOOD are linked to physical properties, e.g. soil or land use properties. For five parameters, however, default values are proposed but need to be calibrated for better model performance (Feyen, 2005; van der Knijff and de Roo, 2006). Analysis of model parameter uncertainty and its impact on discharges simulated by the LISFLOOD model is presented in Feyen et al. (2007).

The JRC stores all maps required to set up LISFLOOD at 5 km and 1 km resolutions. For this study the model has been set up at a regional scale, including all basins, with a 1 km grid, because 5 km would be too coarse. Time steps are adapted to the resolution of the available input variables. Daily time-steps are used for the long term simulations, and hourly time-steps for the detailed case study calculations. Since the aim of the study is to test the approach for ungauged river basins, the available discharge data have been used for comparison and validation only, and not for calibration. For calibration, the default values for the model (listed in van der Knijff and de Roo, 2006), have been applied to all basins and for both long-term and forecasting simulations.

3.2 Input data for the study

For the determination of the hydrological regime over the previous years, synoptic meteorological station data from the data archive of the AgriFish unit[1] at the DG Joint Research Centre have been used. This database holds reliable meteorological data from about 2000 stations across Europe since 1990. In the study area 11 meteorological stations are available for which daily values of temperature and rainfall have been reported and potential evaporation estimated.

For the case study hourly meteorological and hydrological data are available from the databank of the Observatoire Hydrométéorologique Méditerranéen/Cévennes-Vivarais (OHM-CV)[2], which was initiated in 2000 in order to study intense Mediterranean storms leading to devastating flash floods.

Rainfall data estimated from radar were derived from the Bollène 2002 experiment (Chapon, 2006; Boudevillain et al., 2006), designed by DSO/Météo-France and LTHE[3]. This experiment aimed at evaluating the benefit of a radar volume-scanning strategy (8 elevation angles in 5 mn) for radar quantitative precipitation estimation in Cévennes-Vivarais. The three Plan Position Indicators (PPIs) corresponding to the elevation angles needed for the operational products in real time of the Bollène radar (Météo-France; Fig. 1) were complemented by two sets of five PPIs (elevation angle from 0.4° to 18°), alternated every 5 mn. This protocol allowed a good sampling of the atmosphere at a 10 mn sampling interval. The data available at $1\times1\,\mathrm{km}^2$ resolution was the average reflectivity and the mean absolute reflectivity difference averaged over the individual radar polar bins. Innovative algorithms were developed in order to identify and correct various errors sources in quantitative precipitation estimation in mountainous regions (i.e. radar calibration, ground clutter identification, vertical profile of reflectivity versus rainfall). Calibration based on rainfall observations has not been used. For more details on each step of the radar data processing see Boudevillain et al. (2006).

High-resolution operational weather forecasting data are provided to the JRC for research, by the German national weather service (DWD). In 2002 the Lokalmodell of the DWD had a grid spacing of 7 km, a forecasting lead time of 48 h, and the outputs are provided as hourly data. The DWD forecasts are provided every 12 h starting at 00:00 UTC and 12:00 UTC.

For the long-term simulations, discharge data from 15 stream gauging stations were selected from the Banque HYDRO[4], the French national database for discharge. At these stations discharge records are available from 1990 onwards. Few stations have complete records, while others have gaps up to 6 years (Table 1). Only those stations were selected where the influence of hydrological structures such as reservoirs, can be assumed to be small. In addition, for 12 out of these 15 stations, also hourly data are available from the OHM-CV for the 2002 event.

3.3 Methodology of discharge threshold exceedances

In order to issue a flood warning a decision-making element needs to be incorporated: is the discharge going to exceed a critical threshold or not? The determination of the critical thresholds is crucial, in particular when dealing with watersheds where only few or no discharge measurements are available. For this study a model-consistent approach which has been tested previously for early flood warning in large river basins (Ramos et al., 2007; Thielen et al., 2008) is proposed:

1. A long time series simulation based on observed meteorological data is calculated with LISFLOOD. Obviously, the denser the station network, the better rainfalls and subsequent discharge peaks can be captured.

2. For each grid point, the discharges from the long-term time series are evaluated statistically for threshold values, e.g. for return periods or quantiles. Due to the rel-

[1] http://agrifish.jrc.it

[2] http://www.lthe.hmg.inpg.fr/OHM-CV/

[3] Laboratoire d'étude des Transferts en Hydrologie et Environnement

[4] http://www.hydro.eaufrance.fr/

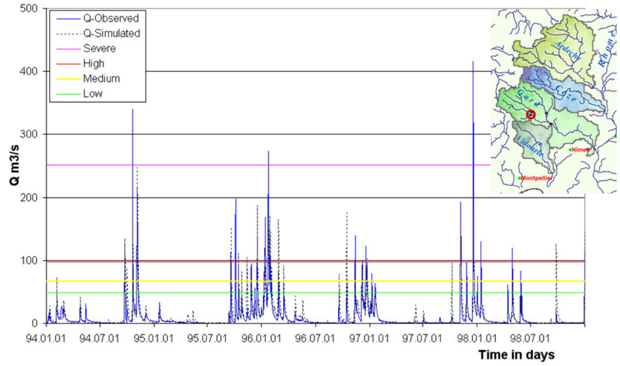

Fig. 4. Observed (Blue) and simulated (Red) Gard discharges at the Corbès (262 km^2) from 1/1/1994 to 31/12/1998.

atively short time series for which reliable meteorological data are available for this study (1990–2002), the quantiles approach was used. Discharges are ranked from highest to lowest and certain cut-off values are chosen as critical thresholds. The highest discharge is defined as the *severe* threshold level. That corresponding to the 99% highest discharge is chosen as the *high* threshold level (comparison with observations has shown that this corresponds frequently with observed 1–2 year return periods), the 98% as *medium* and the 97% as *low*.

3. Comparison with observations is performed through the exceedance of the different corresponding thresholds. For instance for the "high" threshold, it is $Q_{\mathrm{obs}} > Q_{\mathrm{threshold\,High\,obs}}$ when $Q_{\mathrm{sim}} > Q_{\mathrm{thresholds\,High\,sim}}$,

The major advantage of this approach is that any systematic over- or under-prediction of the model is compensated for. If the model tends to over-estimate discharges in a given river reach, for example because of a non-optimized parameterization or lack of processes such irrigation or reservoir operations, this would be reflected in the thresholds as well as in the forecasts. In this way the relative difference of simulated discharge to simulated thresholds, but not the actual values, are evaluated. The same procedure was used for the calculation of observed thresholds. The limitation of the approach is that EFAS may produce reasonable results in terms of threshold exceedance while being seriously offset from the observed hydrographs. Such quantitative discrepancies need to be identified and reduced over time to ensure that the system remains credible. For visualization, the critical thresholds are coded by different colors (Table 2).

For those stations where observed discharge data are available, the same method has been applied to calculate the corresponding critical values $Q_{c\,\mathrm{obs}}$.

One of the main drawbacks for this study lies with the different time and space resolutions that are imposed by the availability of the data. While for the determination of the

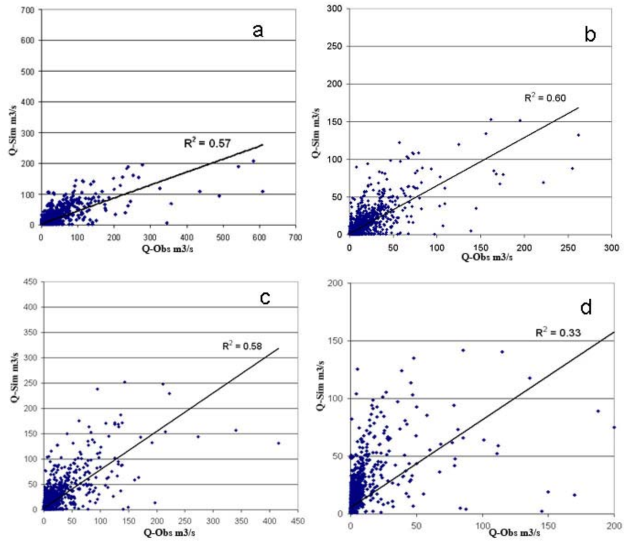

Fig. 5. Scatterplots for the stations in the Ardèche **(a)**, Cèze **(b)**, Gard **(c)** and Vidourle **(d)** with the daily observed discharges in m^3/s on the x-axis and simulated discharges in m^3/s on the y-axis during the 1990 to 2002.

thresholds only daily meteorological and hydrological data are available over a sufficient long time, flash floods themselves develop on shorter time scales. It is therefore likely that the calculated thresholds are low compared to the actual peak discharges occurring during flash floods. Long-term studies with hourly data were not available for this study but are planned in future work. In a similar way, the weather forecasting data neither entirely match the resolution of the climatological network nor the high-resolution network. Too low thresholds can lead to a high number of false alarms that could be reduced if the data used to determine the thresholds were of higher resolution.

4 Results

4.1 Hydrological regime and calculation of thresholds

The hydrological regime of the river basins in the Cévennes-Vivarais region is illustrated for the example of the Gard basin, a typical medium-size catchment of the Cévennes region. Figure 4 illustrates clearly the concentration of peak discharges in autumn and spring. Despite the coarse meteorological station network data used as input, the simulations capture the periods of high flows reasonably well. Although peaks are both over- and under-estimated, the scatter plot in Fig. 5 shows clearly that the model tends to under-estimate the discharges. Particularly severe is the under-estimation of simulated discharge in the example of the Ardèche. In contrast, in the Vidourle, the model tends to over-estimate discharges in particular for low flows. As emphasized before, it is not the absolute discharges that are of interest in this study

Table 2. EFAS thresholds, their color codes and associated hazard class.

EFAS threshold	Color	Description
S (Severe)	■	Very high possibility of flooding, potentially severe flooding expected
H (High)	■	High possibility of flooding, bankful conditions or higher expected
M (Medium)	■	Water levels high but no flooding expected
L (Low)	■	Water levels higher than normal but no flooding expected

but the exceedances of corresponding thresholds. Therefore the performance of the simulations is not assessed in terms of classical efficiency measures, but rather through the number of hits, false alarms and misses.

The number of threshold exceedances, shown in Fig. 6 as cumulative values over the 12 year period for both simulated and observed data series, shows that the simulated threshold exceedances compare well with the observed ones, in particularly for the Gard and Ardèche. While during certain periods differences can occur, with increasing number of years the values approach similar numbers, indicating that the 12 year period is sufficiently for the calculation of the thresholds.

The exceedances of threshold levels for both the simulated and observed discharges have been summarized in contingence tables (Table 3) for each of the four threshold levels. Figure 7 shows, for the same stations as in Fig. 5, the first three components of this contingence table for thresholds high, medium and low. By definition, the "severe" threshold, which is defined as the highest discharge of the time series, cannot be exceeded and is therefore zero in all cases. Positive rejections, the vast majority of the cases, are not plotted to avoid distortion of the graphs.

The splitting of the contingence tables shows mixed results. When analyzing the data according to the lowest alert class, there are generally more hits than false alarms or misses. The numbers of misses and false alarms are about equally high. When looking in more detail at the different threshold classes, the number of hits compared to the number of false alarms tends to decrease from low to high threshold classes.

These results must, of course, be analyzed in view of the uncertainty introduced through the variable quality of the meteorological input data and the events analyzed. In the case of the Vidourle (Fig. 7c) also the large gaps in discharge observations introduce additional uncertainty into the calculation of the observed discharge thresholds. Flash floods are events that are much localized and there is a high probability that the synoptic station network misses the event – hence

Table 3. Definition of contingence table for threshold N.

	$Q_{\text{obs}} \geq Q_{\text{threshold N, obs}}$	$Q_{\text{obs}} < Q_{\text{threshold N, obs}}$
$Q_{\text{sim}} \geq Q_{\text{threshold N, obs}}$	H (Hit)	F (False)
$Q_{\text{sim}} < Q_{\text{threshold N, obs}}$	M (Miss)	PR (Positive Rejection)

a high number of misses. More surprisingly is perhaps the high number of false alarms. This can be partially explained with the temporal resolution of the input data and the corresponding daily time step which can easily introduce a 1 day time shift which in this rigid analysis based on daily values then counts as a miss or false alarm. A systematic time shift could not be identified. If the input data were available at a higher temporal resolution and derived from a higher density network, the results would very likely be better.

It can be concluded from the long-term study that with the given input data resolution the simulated discharges tend to be considerably lower than those observed. The approach is capable of detecting the events whereas the intensities of the events are more difficult to simulate reliably. Therefore quantitative analysis of discharges could not be used for flood warning, while the threshold exceedance approach allows better identification of events.

4.2 Forecasting the 8–9 September 2002 event

In LISFLOOD the output of the daily long-term simulations are used as initial conditions for the hourly flood simulations. The model initial conditions have not been updated with observed data between the 00:00 and 12:00 forecast simulations. Comparison between accumulated rainfalls from radar (Fig. 2a), high resolution rain gauge network (Fig. 2b) and weather forecasts (Fig. 8) show that there are big differences in terms of spatial distribution as well as magnitude.

Taking the high resolution rainfall network data as reference, Fig. 2 shows that the radar produced very similar rainfall in terms of quantity and spatial distribution. There is more information on the spatial variability in the radar data

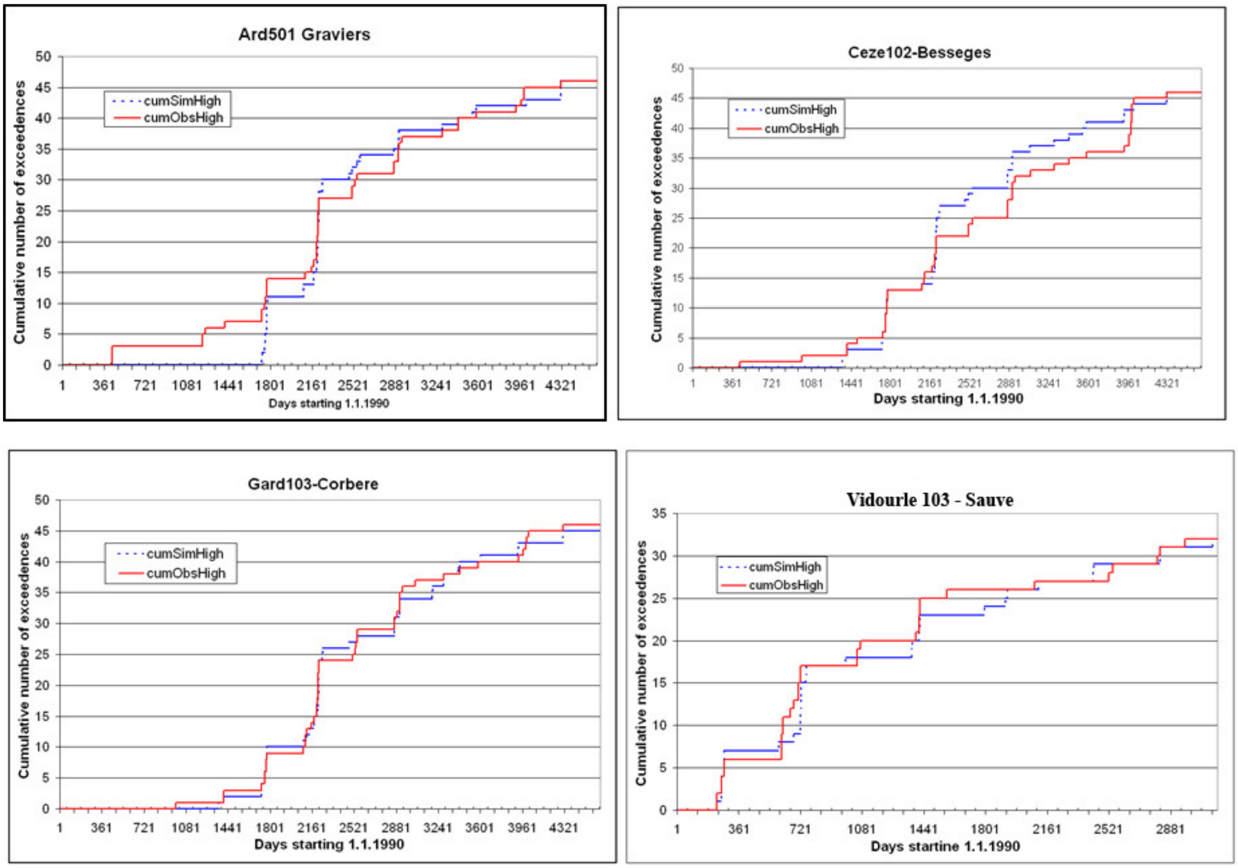

Fig. 6. Cumulative number of threshold exceedances for the observed (solid) and simulated (dashed) time series starting on 1/1/1990 up to 31/12/2001.

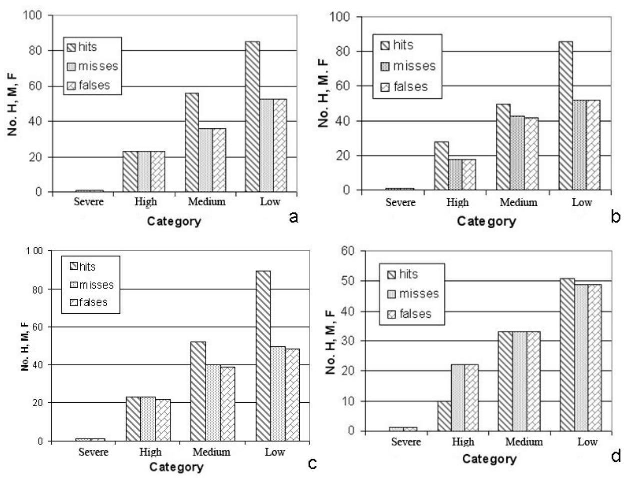

Fig. 7. Contingence tables of hit, false, and missed threshold exceedances for Chambonas in the Ardèche **(a)**, Bessège in the Cèze **(b)**, Corbès in the Gard **(c)** and Sauve in the Vidourle **(d)** for the daily discharge exceedance analysis from 1990–2002.

compared to the interpolated gauge data. In contrast, the daily values of the synoptic gauges do not capture the event in its total spatial distribution. The rainfalls are concentrated in the Southeast of the catchment. Also the rainfall quantity is much too low. As for the weather forecasts, the rainfall patterns are shifted too far north compared to the observation. This is has been observed also for other meteorological models, as documented in Anquetin et al. (2005). Also, the overall volume of precipitation has been under-estimated by the weather forecasting model, which again has been observed previously for other weather forecasting models.

Figure 9a shows hourly hydrographs based on the radar-estimated rainfall and high density rain gauge network. Both simulated discharges are compared against the observed discharge at Générargues in the Gard river. The dynamics of the hydrograph at the onset of the flood is well captured, whereas the recession curve is comparatively too long. The peaks are under-estimated. In Fig. 9a also both the observed and simulated thresholds are shown for a station in the Gard as a representative example how these thresholds relate. In Fig. 9b, the hourly simulated discharges are based on the hourly, 7 km resolution weather forecasts from the German National Weather Service at the same outlet.

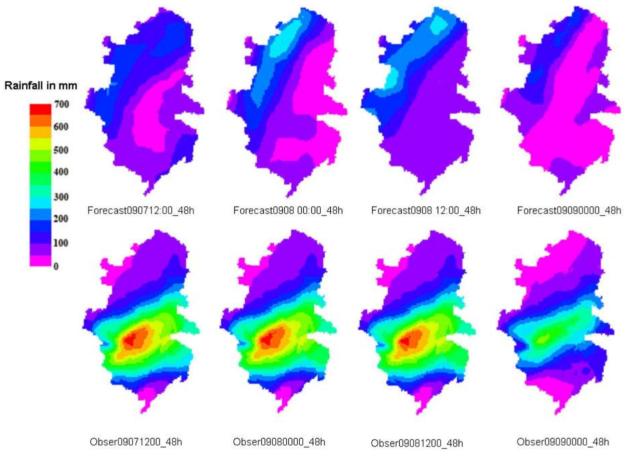

Fig. 8. Accumulated rainfalls from the 20020907 12:00 20020908 00:00, 20020908 12:00 20020909 and 00:00 48 h forecasts, and observed (using 101 rainfall gauging stations (see Fig. 1) for the same period.

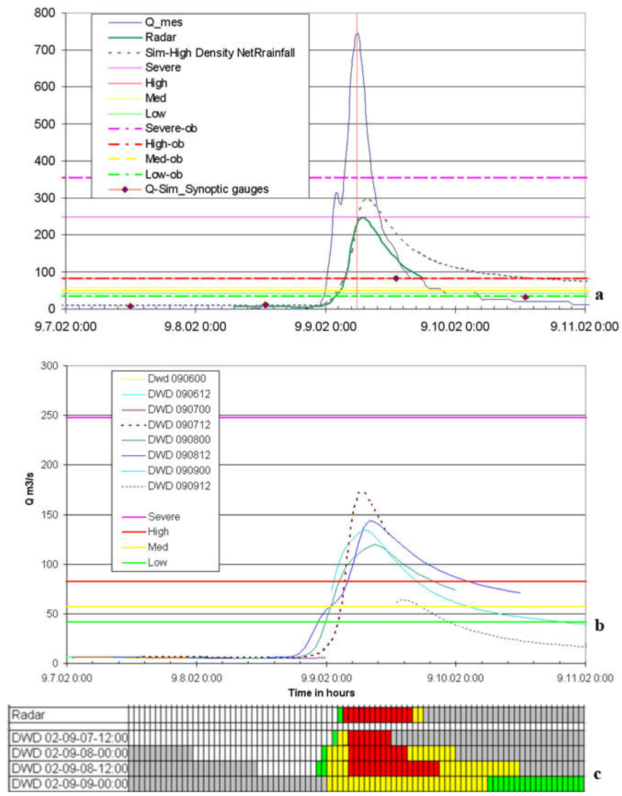

Fig. 9. **(a)** Observed hydrograph at Générargue in the Gard river with the vertical line indicating the time of peak at 06:00 on 9 September, while at 07:00 with Radar(Green) and at 08:00 with the using of high density rainfall gauging network (Grey dashed). **(b)** Hydrographs of forecasted discharges in m^3/s (y-axis) for Générargue (244 km^2) in the Gard river from the 9–12 September 2002 starting at 12:00 in hourly time steps (x-axis). The flood forecasts based on DWD forecasts start on 0906 00:00 until 0909 00:00 in 12 h time intervals. **(c)** Illustrates the visualization of threshold exceedances in hourly time steps for the flood forecasts based on radar data and 48 h DWD Lokalmodell forecasts from 20020907 00:00, 20020907 12:00, 20020908 00:00, 20020908 12:00 and 20020909 00:00. Both observed (solid line) and simulated (dashed lines) exceeded thresholds are color coded as purple (severe), red (high), yellow (medium) and green (low). Grey color indicates no data available.

Figure 9 clearly illustrates the principle of the threshold exceedance. The simulated discharges – even with high resolution radar and gauge data – are three times lower than the observed discharges for which they reach (i.e. measured by radar) or exceed (i.e. measured by gauges) the EFAS severe threshold, indicating serious flooding in the Gard. Although the forecasted rainfalls for the region were shifted too far north compared to the observations and comparatively low rainfalls were forecasted over the Gard basin (Fig. 8), the high threshold is being exceeded with all forecasts from the 7 September 12:00 forecast onwards. Also, the timings of the forecasted peaks correspond well with the observed peak on 9 September at 06:00 o'clock +/– 1–2 h. Thus, although the severity of the event is under-estimated, there is an early warning indication that flooding can be expected within 42 h.

Estimating the run-times of the meteorological and hydrological models as well as data transfer and preparation time of the data to be about 6 h, the lead-time is still of the order of 1 day and more (Fig. 9c). In contrast to the Gard basin, the discharge forecasts for the Ardèche over-estimated the severity of the event (not shown). Clearly, the benefit of these forecasts lies with the potential early warning of the flood event. As the events draw nearer, the flood forecasters would increasingly make use of the observational networks to identify the spatial distribution and the magnitude of the flooding.

Figure 10 illustrates the spatial development of forecasted event. Again the exceeded alert thresholds are visualized with the color coding listed in Table 2. Each panel shows the maximum alert threshold exceeded during the forecasting period in 12 hourly steps. The panel clearly illustrates that the event is first forecasted on 7 September, 12:00, to take place in the upstream areas of all 4 river basins. In the next forecasts the emphasis is mostly in the Gard and Cèze rivers and less in Ardèche and Vidourle. The panel shows how the flooding is forecasted to affect almost the entire basins, well exceeding the severe thresholds over large parts of the river basins. Downstream, towards the outlets, mostly only high thresholds are exceeded.

Compared to the images from radar (Fig. 2) one can see that the spatial distribution of the event was well captured by the forecasts but that there was a tendency to shift the precipitations too far north. As a result, for example the Ardèche was forecast to receive more rainfall compared with observations, while the Gard was forecast to receive less rainfall than observed. This is also reflected in the simulations.

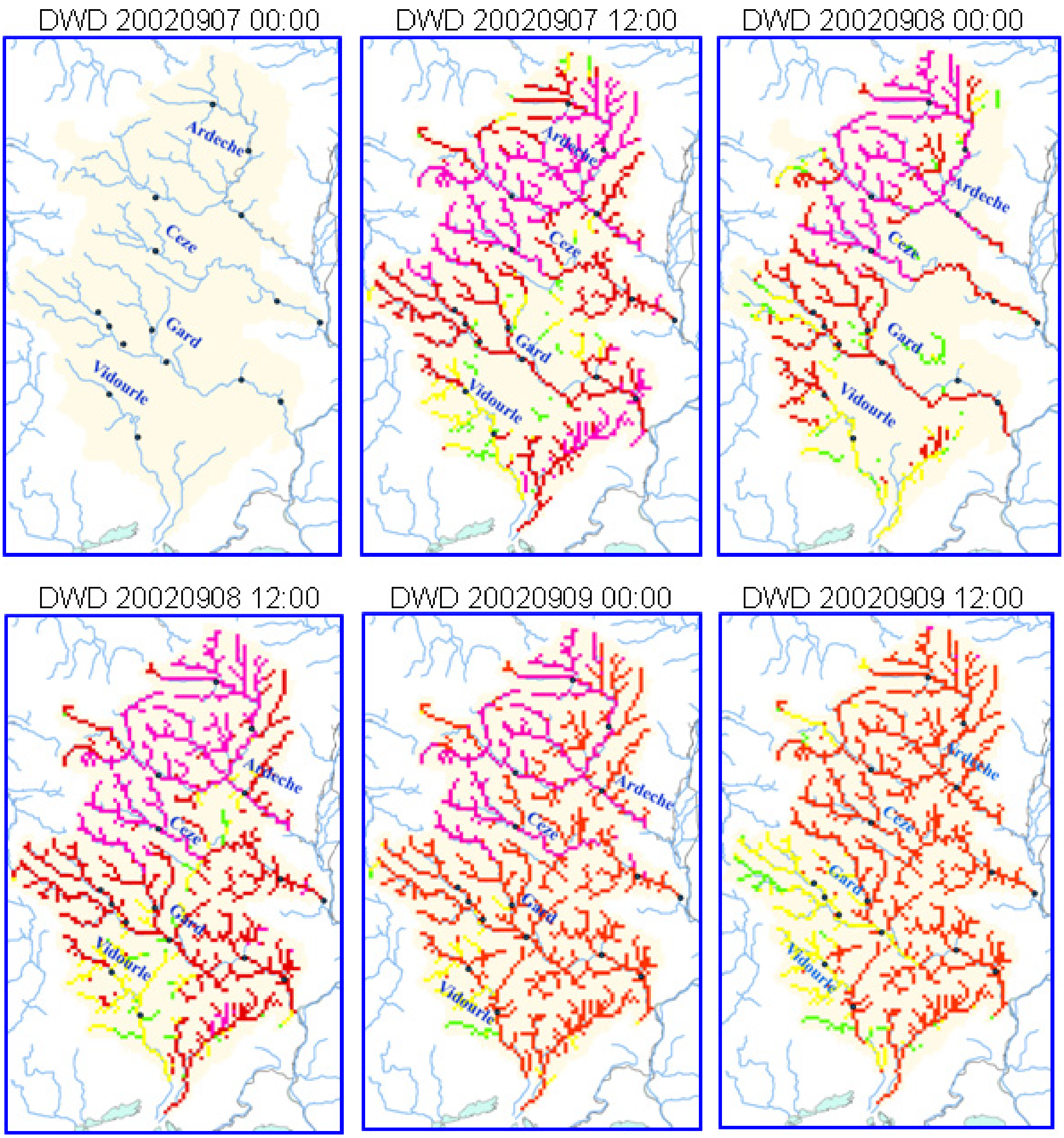

Fig. 10. Summary threshold exceedance maps showing the highest threshold exceeded during the 48 h forecasting time for flood forecasts based on the DWD Lokallmodell weather forecasts from 20020907 00:00 and in 12 hourly steps until 20020909 12:00. The threshold exceedances are color coded with purple (severe), red (high), yellow (medium) and green (low).

In summary one can state that – in this particular case – a system such as the one presented could have provided early warning of the event more than 24 h in advance, with a good estimation of areas affected, timing and magnitude of the event. It appears that the high thresholds derived from the long-term simulations can be used as an indicator for flood events when limited area meteorological model input data are used to drive the simulations.

4.3 Assessments of hits, false alarms and missed events over a 6 months forecasting period

Having demonstrated for this case study that a threshold-exceedance system based on high-resolution operational weather forecasts is capable of detecting flash floods more than 24 h in advance, a 6-months analysis was carried out in order to assess the rate of alarms and missed events. Longer term analyses were beyond the scope of the current feasibility study, but are planned for the future.

For computational and data availability constraints, this study is limited to running all 12:00 DWD forecasts (48 h lead-time) for a six months period from June to December 2002. The forecasts are initialized with the results from a 6 months simulation based on hourly data from the high resolution rain-gauge network. For the 6 months analysis only 9 out of the 15 stream gauging stations were available. For the assessment of the 6 months analysis a flood event is defined as the exceedance of the high alert threshold - both simulated and observed – at least once during a 24 h period. If the high threshold has been exceeded for two consecutive days it is still only counted as one event. If the forecasted hydrographs exceed the simulated high threshold during any time of the observed event, it is counted as a hit.

For this period the OHM-CV has classified 11 significant rainfall events (i.e. 3, 5, and 8–9 September; 9–10, 21, and 30–31 October; 14, 21 and 24 November; 10–11 and 27 December), where a significant event is defined when at least one station has reported more than 50 mm of rainfall during the event. Out of these 11 rainfall events 6 have resulted in a flood event (9 September, 10 October, 22 and 24–27 November, and 11–13 and 29 December) where a flood event is defined as the observed discharge data having exceeding the observed high threshold.

Table 4 summarizes the number of hits, false alarms and misses for the period from 5 June 2002 to 31 December 2002. Positive rejections are not listed.

Over the six-month period there are more false alarms (17) than hits (13) and the number of misses (9) is lowest. From Table 4 it is striking that missed events occurred mostly in the winter months. A closer look at the December events showed that the weather forecasts were shifted too far north, where more precipitation was simulated as snow than was observed. In some cases where both simulated and observed discharges are close to their corresponding thresholds, false alarms or misses can easily occur if one exceeds its threshold by a small amount while the other one stays just below. This may produce misleading results, in particular for stations where the calculation of the observed thresholds has large uncertainties associated with the gaps in data series. This could potentially be overcome by setting a buffer zone around the fixed threshold value. This was not done in this study.

One of the reasons that the number of false alarms is higher than the hits is that the weather forecasts tend to spread the precipitation over larger areas than actually occurs. As a result flooding was simulated in almost all river basins while it only occurred in one or two. The threshold method as presented in this paper does not show whether a threshold is only just exceeded or reached.

In terms of early warning, the missed events are the most serious ones because they do not induce any precautionary measures, while early warning of false alarms can easily be identified through weather observations and radar prior to the event.

Table 4. Number of Hits, False Alarm and Misses out of 178 days of analysis from 5th of June to 31st of December 2002 of flood forecasting based on the 12:00 DWD weather forecasting data.

	Hits	False Alarm	Misses
June	0	0	0
July	0	0	0
Aug	0	0	0
Sep	4	4	0
Oct	2	6	0
Nov	6	6	3
Dec	1	1	6
Summary	13	17	9

It can be concluded from the long-term study that a six-month period for the statistical assessment of hits and false alarms is not sufficiently long. The analysis indicates, however, that the method captures the major flood events, while the forecasted rainfalls are often too widespread resulting in a high number of false alarms. The number of missed events is comparatively low, which is important since, as just stated, missed events are the most undesirable in terms of early flood warning.

The study also highlights again that the usefulness of any flash flood warning system depends very much on spatially and quantitatively good precipitation forecasts.

5 Summary and conclusions

In this paper the feasibility of interfacing short-range numerical weather forecasts with a spatially distributed rainfall-runoff model for early flash flood warning in ungauged river basins has been explored. The methodology is based on flood threshold exceedances, where the thresholds are derived from long-term simulations with an essentially uncalibrated hydrological model. The same model is then used with weather forecast data and the model-consistent thresholds applied for the analysis.

The proposed forecasting strategy addresses a number of shortcomings typically present in flash flood forecasting, namely coarse meteorological station networks and few or no discharge station data.

Results of the study show that by looking at relative differences and model-consistent thresholds, early warning for flash floods can be given with lead-times exceeding 24 h. In the case of the 8–9 September 2002 event, the weather forecasts together with threshold exceedance method enabled the timing and severity of the event to be captured with an absolute lead-time of more than 36 h. In terms of spatio-temporal distribution the event was forecasted too far north, leaving the Gard river basin only with high and not severe threshold values exceeded. Taking into account computing time, processing time and analysis time, the effective lead-time could still have been in the order of 24 h. A six-month analysis confirms that the approach is capable of capturing major events. However, due to the more widely spread forecasted rainfalls, the number of false alarms is relatively high. Therefore the results of such an early warning system should only be used by local flood forecasters as a first indication that a severe event might take place, and should not be distributed to the public.

The number of misses, on the other hand, is comparatively low. This is important since missed events in terms of early warning are more important than false alarms, which can be easily identified in the subsequent hours. In the case of a missed event, however, the benefit of early preparedness measures is lost.

A first attempt of combined analysis of the physical and human responses to this devastating Mediterranean storm (Ruin et al., 2008) shows that most of the casualties were not prepared for (1) the strength of the event and its consequences in terms of water speed, and (2) the time evolution of the storms. In 2002, the warning system, mainly based on the meteorological forecasts, was not designed to give a hydrological signature of the forecasted event. Today, such information could be effectively transmitted through the Schapi[5], a flood forecasting centre that has been recently established following the devastating series of flash floods during the last decade. Together with the Observatoire Hydro-Météorologique Cévennes-Vivarais (OHM-CV) which has now established a high-density measuring network, the performance of such a forecasting system could be greatly improved with better initial conditions, calibrated hydrological model and more realistic thresholds. The radar network could then be used for confirmation of the event and more precise developments. The results show, however, that the principle can also be useful for those areas where no data are available and where the approach could greatly contribute to the preparedness for flash flood events, specifically in terms of awareness, identification of regions at risk, and potential magnitude and timing of the event.

Acknowledgements. This work was carried out in the framework of the FLOODsite project funded by the FP6 Program of the European Commission under the no GOCE-CT-2004-505420. The authors thanked the OHM-CV for providing the radar observations and the discharge data, the Deutscher Wetterdienst for the high-resolution weather forecasting data and the DG JRC AgriFish unit for providing their meteorological data for research purposes.

Edited by: F. Pappenberger

[5]Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations.

References

Anquetin S., Creutin, J. D., Delrieu, G., Ducrocq, V., Gaume, E., and Ruin, L.: Increasing the forecasting lead-time of Weather driven flash floods, Report on http://natural-hazards.jrc.it./downloads/public/2004report_FeasibilityStudyOnFlashfloods.pdf, 2004.

Anquetin, S., Yates, E., Ducrocq, V., Samouillan, S., Chancibault, K., Gozzini, B., Pasi, F., Pasqui, M., Garcia, M., Martorell, M., Romero, R., Silvio, D., Accadia, C., Casaidi, M., Mariani, S., Ficca, G., and Chesa, P.: The 8 and 9 September 2002 flash flood event in France: an intercomparison of operational and research meteorological models, Nat. Hazards Earth System Sci., 5, 741–754, 2005.

Board on Atmospheric Sciences and Climate (BASC): Flash Flood forecasting over complex terrain, the National Academies Press, http://www.nap.edu/catalog.php?record_id=11128, 2005.

Borga, M., Tonelli, F., Moore, R. J., and Andrieu, H.: Long-term assessment of bias adjustment in radar rainfall estimation, Water Resources Research, 38(11), 1226, doi:10.1029/2001WR000555, 2002.

Boudevillain, B., Delrieu, G., Chapon, B., Kirstetter, P.-E., Nicol J., and Andrieu, H.: The Bollène-2002 experiment: innovative algorithms and evaluation of processing strategies for radar QPE in the Cévennes – Vivarais region, Proceeding of the 4th European Conference on Radar in Meteorology and Hydrology, Barcelona, 157–160, 2006.

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

Chapon, B.: Etude des pluies intenses dans la région Cévennes-Vivarais à l'aide du radar météorologique. Régionalisation des traitements radar et analyse granulométrique des pluies au sol, PhD Thesis (in French) University Joseph Fourier, Grenoble, 197 p., 2006.

Chen, S.-T. and Yu, P.-S.: Real-time probabilistic forecasting of flood stages, J. Hydrology, 340(1–2), 63–77, 2007.

Chow, V. T., Maidment, D. R., and Mays, L. M.: Applied Hydrology, McGraw-Hill, Singapore, 1988.

Collier, C.: Flash Flood Forecasting: What are the limits of predictability? Q.J.RM.S., 133, 622A, 3–23, 2007.

Creutin, J. D. and Borga, M.: Radar Hydrology modifies the monitoring of flash flood hazard, Hydrological processes, 17(7), 1453–1456, 2003.

Delrieu G., Ducrocq, V., Gaume, E., Nicol, J., Payrastre, O., Yates, E., Kirstetter, P. E., Andrieu, H., Ayral, P.-A., Bouvier, C., Creutin, J.-D., Livet, M., Anquetin, S., Lang, M., Neppel, L., Obled, C., Parent-du-Châtelet, J., Saulnier, G.-M., Walpersdorf, A., and Wobrock, W.: 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. Hydrometeorology, 6, 34–52, 2005.

De Roo, A., Odijk, M., Schmuck, G., Koster, E., and Lucieer, A.: Assessing the effects of land use changes on floods in the Meuse and Oder catchment, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 26(7), 593–599, 2001.

Ducrocq, V., Aullo, G., and Santurette, P.: Les precipitations intenses des 12 et 13 novembre 1999 sur le sud de la France, La Météorologie, 42, 18–27, 2003.

Everhardus, C., de Roo, A., de Jong, S.: Validation of the LISFLOOD runoff model for a mountainous catchment. Report of the European Commission, Joint Research Centre, EUR 20514 EN, 2002.

Feyen L.: Calibration of the LISFLOOD model for Europe: current status and way forward, European Commission EUR22125 EN, 2005.

Feyen, L., Vrugt, J. A., Ó Nualláinc, B., van der Knijff, J., de Roo, A.: Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model, J, Hydrol., 332(3–4), 276–289, 2007.

Fritsch, J. M., Houze, R. A. J., Adler, R., Bluestein, H., Bosart, L., Brown, J., Carr, F., Davis, C., Johnson, R. H., Junker, N., Kuo, Y. H., Rutledge, S., Smith, J., Toth, Z., Wilson, J. W., Zipser, E., Zrnic, D.: Quantitative precipitation forecasting: report of the 8th prospectus development team, U.S. Weather Research Program, Bull. Amer. Meteorol. Soc., 79, 285–299, 1998.

Georgakakos, K. P.: Analytical results for operational flash flood guidance, J. Hydrol., 317(1–2), 81–103, doi:10.1016/j.jhydrol.2005.05.009, 2006.

Gruntfest E. and Handmer, J.: Coping with flash flood, Ed. NATO Science Series, 322 p, 1999.

Huet, P., Martin, X., Prime, J.-L., Foin, P., Laurain, C., and Cannard, P.: Retour d'expérience des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches-du-Rhône, de l'Ardèche et de la Drôme. Technical report, Ministère de l'Ecologie et du Développement Durable, République Française, 133 pp, 2003.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, 1997.

Nuissier, O. Ducrocq,, V., Ricard, D., Thouvenain, T. and Anquetin, S.: A numerical study of three catastrophic precipitating events over western Mediterranean region (Southern France), Part I: Numerical framework and synoptic ingredients, Q. J. R. Meteorol. Soc., 134(630), 111–130, doi:10.1029/10.1002/qj.200, 2008.

Palmer, T. N. and Räisänen, J.: Quantifying the risk of extreme seasonal precipitation events in a changing climate, Nature, 415, 512–514, 2002.

Parry M., Canziani, O., Palutikof, J., van der Linden, P., and Hanson, C.: Climate Change 2007: Impacts, Adaptation and Vulnerability, Cambridge University Press, 840 p., 2007.

Ramos M.-H., Bartholmes J., and Thielen-del Pozo, J.: Development of decision support products based on ensemble forecasts in the European flood alert system, Atmos. Sci. Lett., 8, 113–119, 2007.

Reed, S., Schaake, J., and Zhang, Z.: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. Journal of Hydrology, 337(3–4), 402–420, 2007.

Rosso R. and Rulli, M. C.: An integrated simulation method for flash flood risk assessment: 2. Effects of changes in land-use under a historical perspective; Hydrology and Earth System Sciences, 6(3), 285–294, 2002.

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

Sénési, S., Bougeault, P., Chèze, J. L., and Cosentino, P.: The Vaison le Romaine Flask flood: Meso-Scale Analysis and Predictability Issues. Wea. Forecasting, 11, 417–442, 1996.

Smith R. E. and Goodrich, D. C.: Rainfall Excess Overland Flow, Encyclopedia of Hydrological Sciences, John Wiley & Sons, Ltd., 1707–1718, 2005.

Thielen, J., Bartholmes, J., Ramos, M.-H., and De Roo, A.: The European Flood Alert System – Part 1: Concept and development, Hydrol. Earth System Sci. Discuss., 5(1), 257–287, 2008.

Todini E.: The ARNO rainfall-runoff model, J. Hydrol., 175, 339–382, 1996.

van der Knijff, J. M., and de Roo, A.: LISFLOOD Distributed Water Balance and Flood Simulation Model. Report of the European Commission, Joint Research Centre, EUR 22166 EN, 2006.

Yates, E., Creutin, J.-D., Anquetin, S., and Rivoirard, J.: A scale dependant quality index of areal rainfall prediction, J. Hydrometeorology, 8(2), 160–170, 2007.

Zhao, R. J. and Liu X. R.: The Xinanjiang model, in: Computer Models of Watershed Hydrology, edited by: Singh, V. P., 215–232, Water Resources Publications, Colorado, USAV, 1995.