Spatially distributed flood forecasting in flash flood prone areas: Application to road network supervision in Southern France

Jean Philippe Naulin, Olivier Payrastre, Eric Gaume

Spatially distributed flood forecasting in flash flood prone areas: Application to road network supervision in Southern France.

J.-P. Naulin[a], O. Payrastre[a], E. Gaume[a,*]

[a] *LUNAM Université, IFSTTAR, GERS, F - 44344 Bouguenais cedex, France*

Abstract

Accurate flood forecasts are critical to an efficient flood event management strategy. Until now, hydro-meteorological forecasts have mainly been used to establish early-warnings in France (meteorological and flood vigilance maps) or over the world (flash-flood guidances). These forecasts are typically limited either to the main streams covered by the flood forecasting services or to specific watersheds with specific assets like check dams, which in most cases are well gauged river sections, thus leaving aside large parts of the territory. This paper presents a distributed hydro-meteorological forecasting approach, which makes use of the high spatial and temporal resolution rainfall estimates that are now available, to provide information at ungauged sites. The proposed system intended to detect road inundation risks had initially been developed and tested in areas of limited size. This paper presents the extension of such a system to an entire region (i.e. the Gard region in Southern France), including over 2,000 crossing points between rivers and roads and its validation with respect to a large data set of actual reported road

*Corresponding author. Tel.:+33(0)2 40 84 58 84, Fax: +33(0)2 40 84 59 98
Email address: eric.gaume@ifsttar.fr ()

inundations observed during recent flash flood events. These initial validation results appear to be most promising. The eventual proposed tool would provide the necessary information for flood event management services to identify the areas at risk and adopt appropriate safety and rescue measures: i.e. pre-positioning of rescue equipment, interruption of the traffic on the exposed roads and determination of safe access or evacuation routes. Moreover, beyond the specific application to the supervision of a road network, the research undertaken herein also provides also results for the performance of hydro-meteorological forecasts on ungauged headwaters.
Keywords: Flash flood, forecast, distributed model, event management, ungauged watershed

1. Introduction

Europe's Mediterranean region is frequently affected by severe flash floods, which can be characterized by their very fast dynamics. According to Gaume et al. (2009), these so called "flash floods" probably represent the most destructive natural hazard in the Mediterranean zone. The catastrophic flooding that occurred in the Var region in southern France in June 2010 once again illustrates the destructive effects of such floods. Furthermore, the lack of anticipation of these events combined with their very rapid evolution limit the efficiency of event management and rescue operations to a considerable extent. If good preparation is able to help authorities cope with such events, the availability of short-term forecasts could offer a major improvement, which is essential to providing the necessary information for a correct diagnosis of the situation as well as the organization of rescue operations.

Several systems are currently used to produce operational flood warnings in Europe. One recently developed approach consists of introducing Ensemble meteorological predictions as input to hydrological and/or hydraulic models in order to generate river discharge forecasts (Cloke and Pappenberger, 2009). This approach lies at the core not only of the European Flood Alert System (EFAS) developed by the European Joint Research Center (Thielen et al., 2009), but also of the advanced Hydrological Prediction Service proposed by the U.S. NOAA (McEnery et al., 2005) or of the Vigicrue system run by the SCHAPI in France (Tanguy et al., 2005). These methods however mostly apply to gauged sections of relatively large rivers with forecasting lead times of at least 12 to 24 hours and are not suited to flash flood conditions (Borga et al., 2009). Flash flood forecasting must cope with the complex spatio-temporal patterns of the generating rainfall events which significantly limits the accuracy and relevance of meteorological forecasts. Recent improvements in the EFAS approach (Alfiery et al., 2012) have shown that relevant flash flood warnings may be expected for medium-sized watersheds (i.e. greater than 500 km^2). Since damage often occurs in headwater catchments a few km^2 in size (Ruin et al., 2008), it is also necessary to provide information in these areas that are often ungauged. This implies implementating rainfall-runoff models based on distributed rainfall observations, with shorter lead times and limited calibration possibilities.

A variety of methods have been developed to produce short-term flash flood forecasts or alerts on small streams (1 to 500 km^2). The most common of these are based on rainfall thresholds associated with risk levels, such as Flash Flood Guidance (Carpenter et al., 1999; Georgakakos, 2006) or

Bayesian decision approach (Martina et al., 2006). Other methods, based on distributed rainfall-runoff simulations and discharge thresholds (Reed et al., 2007; Younis et al., 2008; Javelle et al., 2010) have been developed yet based on limited tests and validation at ungauged sites. Moreover, the selected alert thresholds are tipically uniform over a given territory, or else determined based on knowledge of the local consequences of flooding. In an ideal case, flood forecasting methods should be combined with an exposure rating method so as to provide valuable outputs at a large spatial scale and not just at a small number of strategic points.

More recently, Versini et al. (2010a,b) proposed a prototype of a road inundation warning system (RIWS) adapted to flash floods. This system also relies on distributed rainfall-runoff simulations, though several reasons have led to focusing on road inundations as a specific consequence of flash floods. First, roads are particularly vulnerable: they generally are the assets affected first and nearly half of all flash flood victims are car passengers trapped in their vehicles by the rapid rise of water (Drobot et al., 2007; Ruin et al., 2007). Second, the road network is sufficiently dense throughout the territory and moreover inundations are often reported. These reports provide indirect information on flood location and intensity. Third and last, warnings of possible road inundations may be of strategic value for emergency services in identifying the areas at risk and selecting safe accesses or evacuation routes. The RIWS is intended to provide a real time estimation of the road submersion risk at each intersection between roads and the stream network. These intersections mainly correspond to bridges, dams, culverts (for channelized streams) or fords. The computed risk level results from the com-

parison of forecasted discharges using a rainfall-runoff model and discharge thresholds established according to the estimated susceptibility to flooding of each single intersection.

The initial version of the RIWS has been developed and tested on a territory with limited areal (240 km^2, see Figure 1) under relatively favorable conditions: homogeneous hydrological behavior, and the availability of extensive information on road susceptibility to flooding.

The work presented herein is aimed at both, extending the RIWS to the entire Gard region (5000 km^2), which corresponds to an actual pre-operational test and further developing and testing this tool using a larger data set.

A spatial extension of the initial RIWS presents two main difficulties. The first one is the implementation of a rainfall-runoff model at high spatial resolution, mainly on ungauged catchments of a limited area (less than 2 km^2 on average), and on a territory that features a wide array of hydrological and hydraulic conditions: variability of soil and land use characteristics (in some cases with karstic conditions), a stream network including small upstream watercourses and downstream rivers with broad floodplains. The model should also take into account the spatial heterogeneity of rainfall inputs, with computation times being compatible with a real-time application. The second difficulty lies in the definition of the susceptibility of roads to flooding in areas where little or no data on past reported inundations are available, in comparison with the areas initially selected (Versini et al., 2010a).

The paper's next section will present the case study and the developments introduced in the initial RIWS methodology in an effort to answer the

aforementioned questions. Several tests of the expanded system will then be provided, based on a dataset that includes 10 recent flooding events, which lead to observations of significant disruptions to the road network. These initial results will then be discussed as a preliminary to proposing further system improvements.

2. Presentation of the Gard region and available datasets

2.1. The Gard region

The Gard region is an administrative jurisdiction located in southeastern France. Its boundaries are presented in Figure 1: the northwest part of this area includes the Cévennes Mountains, a medium elevation chain culminating at Mont Lozère (1,699 m); on the other hand, the southeast part corresponds to plateaus and plains extending to the Mediterranean coast, with elevations varying between 0 and 200 meters. This region displays typical Mediterranean climate, as revealed by heavy rainfall events during the autumn season. The 10-year daily rainfall accumulation ranges from 100 mm near the Mediterranean coast to more than 270 mm in the Cévennes foothills (CNRS-INPG, 1997).

This territory is the French region most frequently affected by flash floods (Gaume et al., 2009). Over the last 25 years three exceptional rainfall events have caused considerable damage and losses. In 1988 the city of Nîmes was completely devastated by a flood associated with rainfall accumulations in excess of 400 mm within 6 hours (Desbordes et al., 1989), causing 9 deaths. The extraordinary flood event of the 8^{th} and 9^{th} of September 2002 affected the entire Gard region, with rainfall accumulations that locally exceeded 600

mm in 12 hours, causing 24 deaths in addition to economic losses estimated at €1.2 billion (Delrieu et al., 2004). Lastly, a major event was also observed on September 2005 6^{th}-8^{th} in the southern part of the region, with local rainfall accumulations of about 300 mm on the 6^{th} and 250 mm on the 8^{th}, responsible for significant inundations in Nîmes, but more predominantly in the downstream plains (Maréchal et al., 2007).

This region is divided into six main watersheds covering $4/5^{th}$ of the Gard territory: the Gardon (1858 km^2), the Cèze (1355 km^2), the Vidourle (798 km^2), the Vistre (600 km^2), the Hérault (817 km^2), and the Dourbie (468 km^2) (see Figure 1). Except for the Vistre, these rivers all originate in the Cévennes Mountains, which correspond to primary geological formations. Over their downstream part, the Gardon, the Vidourle and the Cèze rivers all cross a plateau area, characterized by sedimentary bedrock mostly composed mainly of marls and locally of karstified limestone. The extreme southern part of the Gard region corresponding to the Vistre watershed and extreme downstream part of the Vidourle River, is a flat and low elevated plain, covered by thin and diversified Quaternary alluvial formations.

2.2. Road network and road inundation inventories

The road network considered herein includes all road sections except for local connections at the district scale. Since 2007 this road network has been managed in large part by the Gard Departmental Council whose oversight includes maintenance operations, traffic management and the construction of new road sections. This involvment extends to field surveillance during floods and, if necessary, closing the inundated roads to the traffic, and information dissemination on road traffic conditions to other authorities and to the media.

The total number of intersections between roads and rivers depends on the river network definition. This number lies close to 2,200 when the Carthage [1] hydrographic database is used with an average upstream watershed area of less than 2 km^2 (Fig. 2). The RIWS will be required to provide an estimated discharge value every 15-min at each of these intersections.

One important specificity of the region considered in this analysis is the availability of a relatively thorough information (when compared to other regions) regarding road network susceptibility to flooding. This information relies on two distinct sources.

First, a systematic inventory of critical road sections, known from field experience and human memory, was conducted after the major 2002 flood. This initial compilation, called the PICH, basically covers a 40-year period and identifies 160 road sections displaying a potential risk of submersion. This PICH inventory also includes an estimation of the flood return period of each identified vulnerable road section. This compilation however only covers the northwest part of the Gard region, and does not provide event-specific information. The PICH has been used to adjust the susceptibility rating method in the RIWS. Note that only 71% of the points reported in the PICH actually correspond to intersections identified in Figure 2: the remaining 29% may correspond to local road sections not considered herein, or else to inundations that cannot be associated with a river intersection: low points and or lateral overflow of a stream. Consequently, the RIWS will

[1] The BD CARTHAGE® database, produced by the French National Geographic Institut, the MEDD (French Ministry of Ecology and Sustainable Development) and the various Water agencies, constitutes the hydrographic reference in France.

not be able to detect all types of localized inundations or disruptions but will capture the majority of them.

Second, since 2002, inventories of the road disruptions associated with single flood events have gradually been developed. These inventoried road closures are mainly caused by road submersions but might also sometimes be induced by snow, landslides or rock falls, or even road accidents. The most exposed roads may also be closed preventively. This information collection system at the event scale has been progressively implemented and has just recently reached maturity: these inventories have only become systematic since 2007. Ten event reports are ultimately available over the period 2002-2008; they have been presented in Table 1. Nonetheless, this information is heterogeneous and probably far from being exhaustive. The first reports did not indicate the timing or exact locations and causes of the road closures; reports on more recent events tend to be more accurate in space and time. These event reports of road disruptions have still been used to evaluate the RIWS performance.

According to both the PICH and the event-specific reports, 528 of the 2,172 intersections represented in Figure 2 have been inundated at least once over the last 40 years.

2.3. Rainfall and stream flow data

The Gard region is covered by a relatively dense network of rain gauges (about 1 gauge every 150 km^2) and two weather radars providing high spatial and temporal resolution quantitative precipitation estimates (QPE). The available operational datasets have been subjected to a thorough quality control and can be considered of above-average accuracy (Boudevillain et al.,

2011). The results presented in this paper are based on an interpolation of rain gauge measurements through simple kriging with a climatological isotropic variogram (Lebel et al., 1987). Radar quantitative precipitation estimates would probably be more appropriate in this context of flash floods. Due however to the evolution in radar data post-processing and to the radar disruptions for certain specific events, radar data of homogeneous quality were only available for 5 out of the 10 reported flooding events selected for this study (see Table 1). A comparison of the RIWS results with various types of rainfall inputs, based on these 5 events (Naulin et al., 2011), showed that the detailed location and levels of risk alarms triggered are significantly modified by the type of inputs (radar-based QPEs or kriged rain gauge data), but that the overall detection performance based on skill scores were not highly altered. The results obtained with the kriged rainfall data were selected for this paper in order to maintain an extensive evaluation dataset with relatively homogeneous rainfall inputs. The generated rainfall fields display a 1-kilometer spatial resolution and a 1-hour temporal resolution: they have been divided into four equivalent 15-minute time steps to match the needs of the rainfall runoff-model.

The Gard region counts 23 stream gauges located forthe most part on the major streams. The stream gauge measurements have been used herein to verify the performances of the rainfall-runoff model prior to its application in the RIWS. Nevertheless, these measurements yield little information regarding the accuracy of the rainfall-runoff simulations on the upstream tributaries.

2.4. Selection of the flood event dataset

The 13 events with a maximum measured point rainfall accumulation exceeding 50 mm/day over the period 2007-2008 were selected in order to test the RIWS. Four additional events from the period 2002-2006 with available event-specific reports of road disruptions were added to the dataset. The final event dataset is presented in Table 1: it includes a wide variety of hydro-meteorological configurations: from exceptional and generalized events, to more localized and/or lower magnitude events. Seven of the selected events did not induce any reported disruptions, while the other events illustrate the features and frequency of floods inducing road inundations across this region.

3. Description of the RIWS

3.1. General principle

Versini et al. (2010a,b) proposed to consider the risk of submersion as the combined consequences of the magnitude of the forecasted discharge in a given river section and the susceptibility to flooding of the corresponding road-river intersection. This susceptibility is determined using the estimated flood return period for the road; it can be compared to the theoretical return period of the forecasted discharge. The RIWS therefore includes three separate components, namely: i)a rainfall-runoff model producing the forecasted discharges at each intersection, ii) a road-river intersection susceptibility rating method, and iii) a method for estimating the discharge return periods at each intersection (i.e. the SHYREG regional flood frequency method Lavabre et al. (2003)).

Since the flood return period depends to a great extent on the local configuration of each road-river intersection, it might not be determined accurately without specific observations which are generally unavailable. Versini et al. (2010a) therefore proposed to simplify the description of susceptibility and sort intersections into just four susceptibility classes (from high to very low susceptibility), with each class assigned its own distribution of flood return periods. For a given discharge return period, the probability of inundation (i.e. the flood risk level) thus differs for each susceptibility class.

Lastly, three inundation risk levels, corresponding to an increased probability of flooding, may be generated by the RIWS at each computation time step and intersection, on the basis of the contingency table 2.

The first application presented by Versini et al. (2010a,b) is based on the Cinecar rainfall-runoff model (Gaume et al., 2004), which has been applied using homogeneous production parameters given the limited extent of the considered areas (240 km^2). The susceptibility rating method, i.e. the process of sorting intersections into the four susceptibility classes, was based on three morphological parameters: local slope of the river bed, upstream watershed area, and elevation. Versini et al. (2010a) observed that the more vulnerable road sections identified in the PICH inventory were in fact concentrated on smaller slopes and lower elevation areas with large upstream watersheds (Fig. 3). The four susceptibility classes were defined using combinations of thresholds for these three parameters, with the high susceptibility class only containing PICH points and the very low susceptibility class containing no PICH points (see Table 4).

In the application presented herein, the general principles of the RIWS

have been preserved. The extension of this system to the entire Gard territory, which involves a wider variety of local configurations, has however required introducing some changes into both the rainfall-runoff model and susceptibility rating method. The next two sections will present in detail the corresponding methodological evolution. The return periods in Table 2 have remained unchanged. The selected discharge return period estimation method is based on a stochastic simulation of long rainfall-runoff series, yet is not directly correlated with the rainfall-runoff model used in the RIWS. This situation may cause biases; therefore, the RIWS was tested on the 2007-2008 continuous period (including 13 significant rainfall events but only 6 inundating events) to evaluate the possible system under or over-reaction (Section 4.2).

3.2. Modifications of the rainfall-runoff model

The CINECAR model has once again been selected here to compute the discharges every 15-min at the 2,172 intersections identified in the Gard region. This model is distributed and based on a representation of the watershed as a ramified series of stream reaches with a rectangular cross-section, to which both left and right hand hillslopes with schematic rectangular shapes are connected (Fig. 4).

This model only depicts the rapid runoff and is suited for simulating the rising limb and peak phases of significant flood events. The Soil Conservation Service - Curve Number (SCS-CN) model has been used to compute runoff rates and the corresponding effective rainfall on the hillslopes at each computation time step. The effective rainfall is then propagated on both the hillslopes and downstream river network using the kinematic wave model,

which has been resolved by the method of characteristics according to an approach proposed by Borah et al. (1980). The model is flexible and its space and time scales can easily be adapted to meet user needs. Moreover, the model is capable of accounting for the spatial variability of rainfall. For this particular application, the hyetographs corresponding to each individual hillslope have been extracted from interpolated rainfall fields.

This model, initially developed to support the hydrological analysis of data collected after extreme flash flood events, has provided satisfactory results for reconstructing the floods that occurred in September 2002 in the Gard (Delrieu et al., 2005; Gaume et al., 2004). Due to model simplicity, its computation time is limited: less than 60 sec. CPU for a 12-hours event over the entire Gard region. Moreover, CINECAR has a limited number of calibration parameters. The widths of river reaches constitute the main parameters controlling the transfer function. The roughness coefficients are considered to be constant and their values are set at 0.05 for channels and 0.1 for hillslopes. Channel widths can be determined through field surveys or an analysis of aerial photographs. For the sake of simplicity, fixed average channel width values $W_{I,T}$, related to the the Strahler order I of the reach, and the return period T of the simulated discharge, have been introduced according to a simple relation:

$$W_{I,T} = W_{1,T} \cdot I^2 \tag{1}$$

with $W_{1,T}$ being the channel width for reaches of Strahler order 1 (see Table 3 for the calibrated values in the Gard).

Some modifications of the transfer function have been introduced in order

to extend the model to the region, namely: a model simulating the effects of the five major check dams, the diffusive wave model and Hayami method (Hayami, 1951; Moussa, 1996) for river reaches with slopes of less than 1% (to simulate the attenuation of the flood waves in the downstream part of the river network).

The Curve Number (CN) value is the second important calibration parameter and controls the temporal evolution of runoff rates and, as a consequence, the flood magnitude. In the initial version, the CN value was assumed to be homogeneous over the limited application areas; its variation in space has been taken into consideration for the RIWS extension to the entire Gard region. This step is based on the USDA proposed method (USDA, 1986) to determine the CN values as a function of land use, soil types (Corine Land Cover data base and European Soil Database) and during the 5 days preceding rainfall accumulation. As opposed to the conclusions of other recent works (Huang et al., 2007; Soulis et al., 2009), this method has led to satisfactory results, without requiring further adjustments, for the major recent flood events (Table 1) at the 23 gauged sites of the Gard region.

Most of the target sites (intersections) are located on small ungauged upstream watersheds. The hydrological model which has only undergone slight adjustments, has in fact been mostly validated with respect to measured data. No real systematic calibration has been performed. The objective here was to obtain a model that could produce homogenous results at the regional level while avoiding site specific models requiring an intensive calibration phase. As such, the ultimatemodel performance is satisfactory for the major events selected to test the RIWS (Fig. 5): with an average Nash

criterion computed for each single event equal to 0.49.

These results can be compared with the validation outcome of a conceptual rainfall runoff model calibrated on each specific stream gauge observed series. The GR4 model, chosen for such a comparison, is a lumped model developed by Edijatno and Michel (1989) and advanced by Perrin et al. (2003). This hourly rainfall-runoff model requires four parameters to be automatically calibrated with respect to measured data.

A comparive test between Cinecar and GR4 has been conducted for the Anduze stream gauge on the Gardon River: i.e. the largest and highest quality discharge series in the region. A 2–year series was used to calibrate the GR4 model. The obtained Nash score derived over a 6–year validation period reached 0.59, which lies in the average range of values reported in the literature for applications of continuous rainfall-runoff models (Perrin, 2002). The corresponding "event–specific" Nash criteria are then compared with the Cinecar criteria for the largest validation floods in Figure 6.

Results indicate that while the Cinecar model is less efficient for minor events, the two models present a similar performance for the larger events (i.e. unit peak discharge exceeding 2 $m^3/s/km^2$) which do constitute the focus of the present study and more generallyl of flood forecasting models for this region.

Since the RIWS is based on thresholds and not directly on simulated discharge values, its performance will mainly be controlled by the correct forecast of the magnitude or, more precisely, the correct position of the simulated discharge when compared to the discharge quantiles. Figure 7 shows that the Cinecar model predicts peak discharges with typically the same

range of estimated return periods as the observed peak discharges. Except for some few obvious overestimations, the differences between simulated and observed discharges do not exceed one return period class. In considering the additionnal filtering introduced when computing risk levels (Table 2), it is anticipated that these differences will have little impact on the RIWS performance.

3.3. Modifications to the susceptibility rating method

Extending the RIWS to the entire Gard region also required adapting the susceptibility rating method to account for the large spatial variability of observed inundations and geomorphic characteristics throughout the region.

The distribution of PICH points into the four susceptibility classes, as presented in Table 4, reveals that the efficiency of a sorting method based on morphological parameters decreases when moving from the initial area considered by Versini to the entire region covered by the PICH inventory (approx. half of the entire Gard region), with the proportion of PICH points in susceptible classes markedly decreasing.

A detailed analysis shows that the geographical distribution of PICH points, as presented in Figure 2, is very heterogeneous at the PICH territorial scale. The initial sorting method is not able to capture this variability since it generates susceptibility classes with a relatively homogeneous geographical density. In order to improve the results of the susceptibility rating results, the sorting method has therefore been applied separately to three sub-areas displaying significantly different PICH point densities:

- the mountainous area, where the density of PICH points is very low (i.e. less than 5 % of the intersections). The ntersections located in

this area are only assigned to the low and very low susceptibility classes depending on local slope, altitude and upstream catchment area;

- the plateau and plain area with an intermediate PICH point density and where the intersection points may be assigned to any of the three lowest susceptibility classes (very low, low and average);

- the identified flood plains in the plateau and plain area, with a very high PICH point density (exceeding 30 %). In these areas, all intersections are assigned to the high and very high susceptibility classes.

The results presented in Table 4 indicate that this application to three homogeneous sub-areas improves the sorting results to a fair extent, which in turn leads to decreasing the proportion of intersections considered as vulnerable (as reflected by an increase in the very low susceptibility class). A higher proportion of intersections is also assigned to the high and average susceptibility classes with an equivalent proportion of PICH points in these classes. To validate this evolution in the susceptibility rating method, a comparison of RIWS performances will be presented in Section 5.1 below.

Furthermore, the distributions of flooding frequencies for intersections placed in each susceptibility class may be estimated on the basis of the PICH inventory. The dispersion in these distributions, combined with the uniform risk thresholds presented in Table 4, will logically lead the RIWS to generate a significant number of false alarms. Ideal detection rates and False Alarm Ratios of the system (FAR, see definition in Section 4.2) linked to the dispersion of flooding frequencies in each susceptibility class can be estimated in an *a priori* manner for each risk level, i.e.:

- Risk level 3: Inundation highly probable (over 40% of the alarms correspond to actual inundations, hence FAR<60%, i.e. a discharge return period threshold exceeding the reported inundation return period for about 40% of all intersections listed in the class);

- Risk level 2: Inundation possible with a significant probability (over 20% of alarms corresponding to actuel inundations, FAR<80%);

- Risk level 1: Inundation possible with low probability (over 10% of alarms corresponding to actual inundations, FAR<90%);

These values will serve as a benchmark for theRIWS performance evaluation. The relatively high theoretical false alarm ratios are due to the difficulty involved in evaluating the susceptibility of intersections to flooding on the basis of available data; these figures provide an idea of the possible performances of the RIWS performance if the rainfall-runoff forecasts were nearly perfect.

4. Evaluation of the Road Inundation Warning System (RIWS)

For starters, it is worth noting that over the period 2007-2008, during which the reports on flood events are certainly comprehensive, the RIWS only produces warnings (risk levels 1 to 3) for the 6 events with actually reported road inundations. This finding confirms that the RIWS is neither under-reacting nor over-reacting and is moreover capable of distinguishing the most significant flood events. Given these initial results, it seems that the rainfall-runoff model estimates discharges with the correct order of magnitude and that the risk thresholds were appropriately selected and estimated.

Given this initial positive outcome, the results presented below are mainly focused on a detailed performance analysis of RIWS during each inundating event. These results comprise 10 events if the 2002-2006 period is included (Table 1). The events examined have very distinct characteristics, i.e. variable intensities and total rainfall amounts, more or less concentrated in space, thus allowing for an in-depth analysis of the RIWS performance.

4.1. Evaluation criteria

The comparison between computed risk levels and reported floodings will be based on standard criteria used to evaluate meteorological forecasts, computed based on the contingency Table 5 and then calculated separately for each risk level.

From this table, four scores will be calculated as follows (Schaefer, 1990; Wilks, 2006):

- Probability of detection (POD) :

$$POD = \frac{H}{H + M} \tag{2}$$

- False Alarm Ratio (FAR):

$$FAR = \frac{Fa}{H + Fa} \tag{3}$$

- Probability of False Detection (POFD) which relates the number of false alarms to the total number of non-flooded points and offers an idea of the legibility of warnings:

$$POFD = \frac{Fa}{Cn + Fa} \tag{4}$$

- The Peirce Skill Score (PSS) has been used as a global score.

$$PSS = POD - POFD \tag{5}$$

A perfect forecasting system should feature both a POD and a PSS equal to 1 and a FAR and a POFD of 0. Yet, the limitations of the road inundation observations must be considered when applying these skill scores to evaluate the RIWS. First, a comparison between simulated stream discharges and reported road disruptions reveals possible inconsistencies along with the lack of exhaustiveness in the reports. Road sections are often reported as closed for relatively moderate floods and not for more intense ones over the considered period, according to the rainfall-runoff simulations. If we were to consider that the rainfall-runoff model, despite its lack of accuracy, is at least able to correctly rank the discharges between events most of the time, this confirms the limitations of available inundation inventories. Such inconsistencies are present in 41% of the reported road disruptions; they have two main potential origins:

- The road closure reports are not likely to be comprehensive.
- The reasons behind road closures are sometimes unknown, hence some road sections present in the inventory may not have been submerged but instead affected by other phenomena (e.g. landslide, snow, accidents).

Furthermore, the inundation location may have imprecisely logged. Reported closures can sometimes be attributed to several intersections; in this situation, if at least one possible corresponding intersection has been targeted by a warning, then all forecasts would be considered as correct (either

461 H or Cn). Otherwise, if no warning had been issued, only one missed alarm
462 would be counted. Such a case is nevertheless rare: this simplification has
463 limited impact on the evaluation criteria.

464 It is also important to note that in the case of events presenting just a few
465 inundated roads, the POD scores are very sensitive to the number of actual
466 detections. For instance, if only 10 roads are reported to be flooded, each
467 detected point increases the POD score by 10%.

468 *4.2. RIWS performance*

469 Examples of POD and FAR obtained by the system are shown in Figure 8.
470 The false alarm ratios are very high which is clearly correlated with the
471 difficulties involved in defining the susceptibility to flooding of intersections.
472 Section 3.3 pointed out that given the limitations of this susceptibility rating,
473 FAR ratios of about 90% for risk level 1 are to be expected. The FAR ratios
474 observed here lie in this same range (around 90%). Improvements to the
475 rainfall-runoff model can therefore only exert a small impact on the FAR
476 scores.

477 The POD scores depend on the flood event under consideration. A de-
478 tailed analysis shows that the POD is correlated with the total number of
479 reported inundations, which serves as an indicator of the flood magnitude.

480 The POD are very high for the extreme floods of 2002 and 2005, in
481 exceeding or nearing 90%. In such cases, the susceptibility rating plays a less
482 important role since the return period of the computed as well as observed
483 discharges at the affected intersections generally exceeds 10 or even 50 years
484 (Fig. 8). For the less intense events, the discharge values are lower and both
485 the risk level and corresponding warnings are much more dependent on the

susceptibility rating of the intersections, which is the weakest part of the RIWS, thus explaining lower and more variable POD values.

For risk level 2, Figure 8 also confirms the correlation between the number of reported road closures (and hence the local magnitudes of floods and the POD) and the September 2007 exception, which affected a particularly susceptible area according to the susceptibility rating method: i.e. the southern part of the region at low elevation over an area with expansive floodplains. Once again, FAR values basically match the forecasts of section 3.3 (on the order of 80 %).

Some POD scores nevertheless appear to be on the low side. On December 13^{th} 2008, the relatively high number of road closures appears to be inconsistent with the magnitude of floods according to the simulated as well as to the observed discharges. These closures may have been due to other factors than inundation, in particular snow accumulation. This phenomenon has been mentioned in the logs of the road management services with no further details. On December 29^{th} 2008, the limited POD value is related to the underestimation of the discharges as a result of an incorrect evaluation of the initial wetness state of the watershed. A discharge assimilation procedure is under study in order to resolve the initialization problems of this rainfall-runoff model.

On the basis of these results, it is now possible to conclude that the RIWS, in its present form, is able to correctly assess not only the magnitude of the floods (with the exception of some particular events) but also the consequences of floods in terms of the number of inundated roads and risk levels. On the other hand, RIWS fails to accurately identify the actually affected

road/river intersections. Figures 9 and 10 illustrate this last point and the information content in RIWS output. On the wholel, the affected areas of the Gard region have been correctly depicted for these two events. The risk levels are clearly contrasted between these two floods, thus indicating localized and significant consequences for the September 2007 flood and a larger area affected during October-November 2008, yet with a lower density of high risk warnings (basically limited to the most susceptible areas: i.e. the valleys of the main streams). These forecasts match field observations: localized yet significant inundations in 2007 and sparsely scattered local problems in 2008.

It has also been verified that despite the short forecasting lead time (no longer than 15 min. on the headwater catchments given that the system is based on rainfall measurements), the first warnings have for the most part been delivered well ahead of the first field report. Such a system, if operationally implemented, could, for all the aforementioned reasons, be of great utility for the real-time evaluation of situations as well as for organizing the rescue response and managing the road network. It is currently tested by the Gard Departmental Road Management Office and the initial feedback has been positive overall.

5. Discussion of RIWS components

5.1. Utility of the susceptibility model

Given the limitations of available road inundation inventories, the susceptibility assessment is probably the weakest RIWS component. In order to evaluate its real utility to the system and feasible improvements, two tests have been conducted:

- The system was run with discharge thresholds of identical return periods for all intersections: without introducing any susceptibility ratings. Risk level 3 has been associated with a 2-year return period, the level 2 with a 10-year return period and level 1 with a 50-year return period.

- Susceptibility has been defined on the basis of the inundations actually observed during the 10 test events: This is the pseudo perfect case. The intersections inundated once have been assigned to the low susceptibility class, those inundated twice to the moderate class and those inundated more than twice to the high class. The intersections not reported as inundated have been assigned to the very low susceptibility class.

This last procedure does not claim to establish a valid susceptibility layer, but rather has been implemented in order to evaluate the system with a susceptibility assessment close to perfect and moreover to illustrate potential improvements that could be achieved if reported inundations were to be assimilated.

System performance was compared using the two susceptibility rating method configurations presented in Section 3.3. The results are shown in Figure 11. Though far from perfect, the susceptibility rating method clearly improves system performance, as the improved susceptibility rating method leads to higher scores. A detailed analysis also indicates that the scores of the system without susceptibility rating appear to be extremely variable with a high POD, but also POFD and FAR scores for high intensity flood events and scores close to zero for less intense events. The susceptibility rating method introduces stability and nuance into the RIWS response, thereby

producing warnings for every significant event. The magnitude and location of the floods determine both the warning density and the level of risks (see Fig. 9 and 10) yet have less influence on the detection scores.

This "pseudo perfect" configuration furthermore indicates that an improvement in susceptibility rating, through data assimilation for instance, would clearly improve the overall performance, mainly by reducing the proportion of false detections. The room for improvement is indeed significant, given that the PICH inventory used to calibrate of the susceptibility rating method is likely to be far from perfect. As an example, Table 6 lists the number of road sections included in recent event-specific reports of road disruptions that had not been identified in the PICH inventory: 228 road sections are applicable, with 13 of them having been closed more than twice. This finding exposes the probable lack of comprehensiveness in the PICH inventory.

5.2. Comparison of the rainfall-runoff model configurations

The rainfall-runoff model developed for the Gard region has evolved from its first configuration with homogeneous CN parameters to a distributed CN assessment, according to the USDA methodology. The transfer function has also been improved by implementating the diffusive wave model on the downstream reaches of the river network. These changes have led to improving the Nash criteria computed on the gauged sections (Fig. 5). However the effect of these improvements on the RIWS performance, while not totally undetectable, still appears to be very limited as illustrated in Figure 12.

This confirms that only orders of magnitudes of the discharges are necessary to deliver relevant alarms in the context of the RIWS. Overall system

performance is therefore relatively insensitive, even with significant improvements, to the rainfall-runoff model. One positive conclusion drawn is that useful forecasts can be derived from output of hydro-meteorological models affected by significant uncertainties. A more disappointing conclusion however is that the observed local consequences of flood events in ungauged watersheds, like road inundations, do not provide useful indirect measurements of discharges when evaluating rainfall-runoff models.

6. Conclusion

A prototype of a road inundation warning system (RIWS), initially tested on a limited spatial scale, has been extended to an entire French region and has yielded promising results. The resulting system is able to provide warnings at 2,172 intersections between roads and rivers every 15 minutes. Tests conducted in batch mode on 10 recent flood events with observed road inundations have revealed the system's capacity to forecast the location and magnitude of floods and their consequences with relatively good accuracy, yet it remains unable to locate the affected road sections with precision. The main limitation of this system pertains to poor knowledge on the actuall susceptibility of individual road sections to flooding. A data assimilation method, based on the comparison of field reports and forecasted risk levels, is currently being tested. Early results are encouraging and in time this method could lead to system improvements. For the interested readers, a demonstrator (`prediflood.lthe.fr/carto.php`) has been developed to provide the simulated warnings and field observations for 5 of the tested events.

Beyond the specific application to the supervision of a road network, this

research has also produced results relative to hydro-meteorological forecasts and monitoring at ungauged sites. It appears that rainfall-runoff models may offer valuable information for flood event management over small space and time scales even on ungauged watersheds. With a limited calibration and validation effort, rainfall-runoff models could at least provide relatively accurate flood magnitude forecasts that would certainly be sufficient for many event management decisions as has been illustrated herein. To some extent, the philosophy behind the approach presented in this paper has similarities with the philosophy of the EFAS system (Younis et al., 2008; Thielen et al., 2009; Alfieri et al., 2012) though at different time and space scales and with a focus on flash floods. Such approaches contrast with the traditional forecasting methods, which are focused on gauged locations and are definitely forging new avenues in the field of hydrological forecasting.

Acknowledgements

This work benefited from the support of the French National Research Agency, contract number ANR-08-RISKNAT-006-01/PREDIFLOOD. The authors thank Marco Borga and the two anonymous reviewers for their very helpful comments on earlier versions of this paper.

References

Alfieri, L., Thielen, J., Pappenberger, F., 2012. Ensemble hydro-meteorological simulation for flash flood early detection in southern switzerland. Journal of hydrology 424-425, 143153.

631 Borah, D., Prasad, S., Alonso, C., 1980. Kinematic wave routing incorporat-
632 ing shok fitting. Water resources research 3, 529–541.

633 Borga, M., Creutin, J.D., Gaume, E., Martina, M., Todini, E., Thielen, J.,
634 2009. Flood Rik Management : Research and Practice. Taylor and Francis
635 Group. chapter Flash flood risk management : Advances in hydrological
636 forecasting and warning. pp. 1305–1314.

637 Boudevillain, B., Delrieu, G., Galabertier, B., Bonnifait, L., Bouilloud, L.,
638 Kirstetter, P.E., Mosini, M.L., 2011. The Cévennes-Vivarais Mediter-
639 ranean Hydrometeorological Observatory database. Water Resources Re-
640 search 47, 6.

641 Carpenter, T., Sperfslage, J., Georgakakos, K., Sweeney, T., Fread, D., 1999.
642 National threshold runoff estimation utilizing gis in support of operational
643 flash flood warning systems. Journal of Hydrology 224, 21–44.

644 Cloke, H., Pappenberger, F., 2009. Ensemble flood forecasting : a review.
645 Journal of hydrology 375, 613–626.

646 CNRS-INPG, 1997. Atlas expérimental des risques de pluies intenses,
647 cévennes-vivarais. Laboratoire détude des Transferts en Hydrologie et En-
648 vironnement, Grenoble, France. 21p.

649 Delrieu, G., Kirstetter, P.E., Nicol, J., Neppel, L., 2004. L'événement plu-
650 vieux des 8-9 septembre 2002 dans le Gard : Estimation des précipitations
651 par radars et pluviomètres. La Houille Blanche 6, 93–98.

652 Delrieu, G., Nicol, J., Yates, E., Kirstetter, P.E., Creutin, J.D., Anquetin,
653 S., Obled, C., Saulnier, G.M., Ducrocq, V., Gaume, E., Payrastre, O.,

Andrieu, H., Ayral, P.A., Bouvier, C., Neppel, L., Livet, M., Lang, M., du Chtelet, J.P., Walpersdorf, A., Wobrock, W., 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 hydrometeorological observatory. Journal of Hydrometeorology 6, 34–52.

Desbordes, M., Durepaire, P., Gilly, J., Masson, J., Maurin, Y., 1989. 3 octobre 1988, inondations sur nmes et sa région, manifestation, causes et conséquences. ed. C. Lacour.

Drobot, S., Benight, C., Gruntfest, E., 2007. Risk factors for driving into flooded roads. Environmental Hazards 7, 227–234.

Edijatno, Michel, C., 1989. Un modèle pluie-débit trois paramètres. La Houille Blanche 2, 113–121.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičova, L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velascom, D., Viglione, A., 2009. A compilation of data on european flash floods. Journal of Hydrology 367, 70–78.

Gaume, E., Livet, M., Desbordes, M., Villeneuve, J.P., 2004. Hydrological analysis of the river aude, france, flash flood on 12 and 13 november 1999. Journal of Hydrology 286, 135–154.

Georgakakos, K.P., 2006. Analytical results for operationnal flash flood guidance. Journal of Hydrology 317, 81–103.

677 Hayami, S., 1951. On the propagation of flood waves. Technical Report 1.
678 Disaster prevention research institute. Kyoto university, Japan.

679 Huang, M., Gallichand, J., Dong, C., Wang, Z., Shao, M., 2007. Use of soil
680 moisture data and curve number method for es-timating runoff in the loess
681 plateau of china. Hydr. Process. 21, 1471–1481.

682 Javelle, P., Fouchier, C., Arnaud, P., Lavabre, J., 2010. Flash flood warning
683 at ungauged locations using radar rainfall and antecedent soil moisture
684 estimations. Journal of Hydrology 394, 267–274.

685 Lavabre, J., Fouchier, C., Folton, N., Gregoris, Y., 2003. SHYREG: une
686 méthode pour l'estimation régionale des débits de crues. Application aux
687 régions méditerranéennes franaises. Ingénierie-EAT numéro spécial Risque
688 naturel et aménagement du territoire, 97–111.

689 Lebel, T., Bastin, G., Obled, C., Creutin, J., 1987. On the accuracy of
690 areal rainfall estimation : A case study. Water Ressources Research 23,
691 2123–2134.

692 Maréchal, J., Ladouche, B., Drfliger, N., 2007. Karst flash flooding in a
693 mediterranean karst, the example of fontaine de nmes. Engineering Geol-
694 ogy 99, 138–146.

695 Martina, M., Todini, E., Libralon, A., 2006. A bayesian decision approach
696 to rainfall thresholds based flood warning. Hydrol. Earth Syst. Sci. 10,
697 413–426.

698 McEnery, J., Ingram, J., Duan, Q., Adams, T., Anderson, L., 2005. NOAAs

699 advanced hydrologic prediction service; building pathways for better sci-
700 ence in water forecasting. American Meteorological Society BAMS, 375–
701 385.

702 Moussa, R., 1996. Analytical hayami solution for the diffusive wave flood
703 routing problem with lateral inflow. Hydrological processes 10, 1209–1227.

704 Naulin, J., Gaume, E., Payrastre, O., 2011. Distributed flood forecasting
705 for the management of the road network in the gard region (france), in:
706 Proceedings of a symposium held in Exeter, UK, April 2011) (IAHS Publ,
707 2011).

708 Perrin, C., 2002. Vers une amélioration d'un modèle global pluie-débit au
709 travers d'une approche globale comparative. La Houille Blanche 6, 84–91.

710 Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious
711 model for streamflow simulation. Journal of Hydrology 279, 275–289.

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

715 Ruin, I., Creutin, J.D., Anquetin, S., Lutoff, C., 2008. Human exposure to
716 flash floods - relation between flood parameters and human vulnerability
717 during a storm of september 2002 in southern france. Journal of Hydrology
718 361, 199–213.

719 Ruin, I., Gaillard, J.C., Lutoff, C., 2007. How to get there? assessing mo-

720 torist's flash flood risk perception on daily itineraries. Environmental Haz-
721 ards 7, 235–244.

722 Schaefer, J.T., 1990. The critical success index as an indicator of warning
723 skill. Weather and forecasting 5, 570–575.

724 Soulis, K.X., Valiantzas, J.D., Dercas, N., Londra, P.A., 2009. Investigation
725 of the direct runoff generation mechanism for the analysis of the scs-cn
726 method applicability to a partial area experimental watershed. Hydrology
727 and Earth System Sciences 13, 605–615.

728 Tanguy, J., Carriere, J., Trionnaire, Y., Schoen, R., 2005. Reorganisation
729 of flood forecasting in france - réorganisation de l'annonce des crues en
730 france. La Houille Blanche 2, 44–48.

731 Thielen, J., Bartholmes, J., Ramos, M.H., de Roo, A., 2009. The European
732 Flood Alert System - Part 1: Concept and development. Hydrol. Earth
733 Syst. Sci. 13, 125–140.

734 USDA, 1986. Urban hydrology for small watersheds. Technical Report 55.
735 United States Department of Agriculture, Natural Resources Conservation
736 Service.

737 Versini, P., Gaume, E., Andrieu, H., 2010a. Application of a distributed
738 hydrological model to the design of a road inundation warning system for
739 flash flood prone areas. Nat. Hazards Earth Syst. Sci. 10, 793–803.

740 Versini, P., Gaume, E., Andrieu, H., 2010b. Assessment of the vulnerability
741 of roads to flooding based on geographical information - test in a flash

flood prone area (the gard region, france). Nat. Hazards Earth Syst. Sci. 10, 793–803.

Wilks, D., 2006. Statistical Methods in the Atmospheric Sciences, Second Edition. International Geophysics series.

Younis, J., Anquetin, S., Thielen, J., 2008. The benefit of high-resolution operational weather forecasts for flash flood warning. Hydrol. Earth Syst. Sci. 12, 1039–1051.

List of Figures

List of Tables

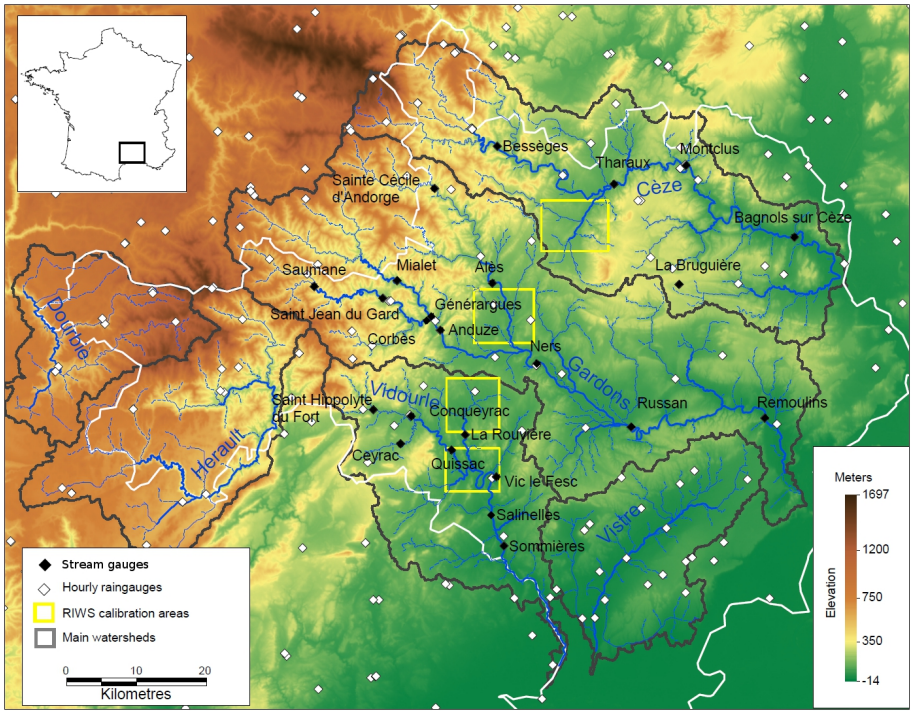

Figure 1: Gard department (shown in white), main streams and watersheds (black), initial test areas of the RIWS (white windows).

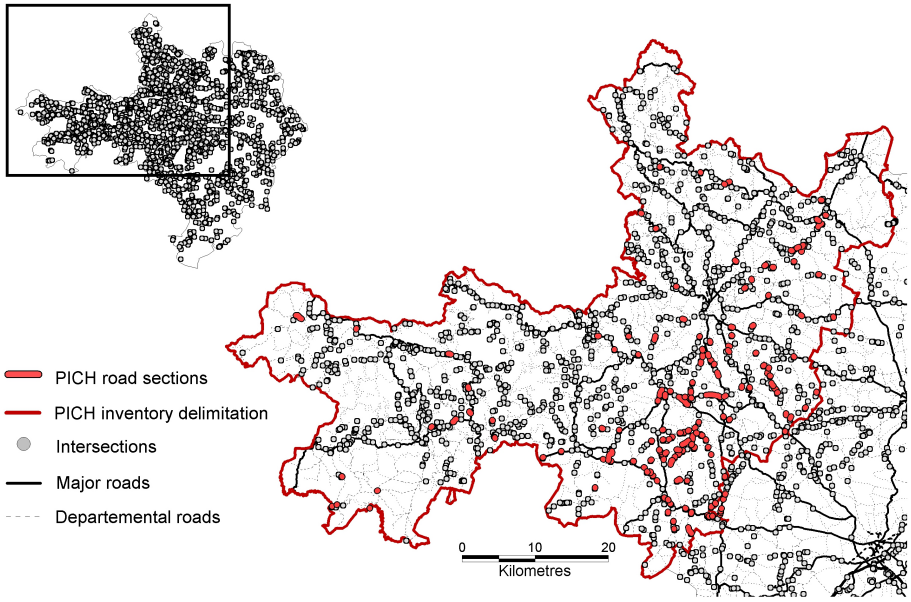

Figure 2: Presentation of the PICH inventory and identification of intersections between road and river networks in the Gard region.

Table 1: The set of test events for the RIWS.

Dates of events	Event duration (days)	max rainfall (mm)	% area > 200 mm	Nb. flooded roads
08-10/09/2002	3	679	74	367
01-04/12/2003	4	380	61	112
02-05/11/2004	4	139	0	11
04-12/09/2005	6	403	76	198
29-30/09/2007	2	204	0	26
19-23/11/2007	5	293	13	14
16-19/09/2007	4	62	0	0
03-11/10/2007	8	100	0	0
06-12/12/2007	7	70	0	0
01-07/09/2008	7	100	0	0
10-15/09/2008	6	60	0	0
07-10/10/2008	4	110	0	0
19-22/10/2008	4	497	15	18
01-02/11/2008	2	514	17	19
10-13/11/2008	4	50	0	0
13-15/12/2008	3	114	0	32
29-31/12/2008	3	218	0.5	54

Table 2: Contingency table for the determination of inundation risk levels: 0 (no risk), 1, 2 and 3.

	Discharge return period (years)				
Susceptibility	0.5	1	10	50	
High	0	2	3	3	3
Medium	0	1	2	3	3
Low	0	0	1	2	3
Very low	0	0	0	0	1

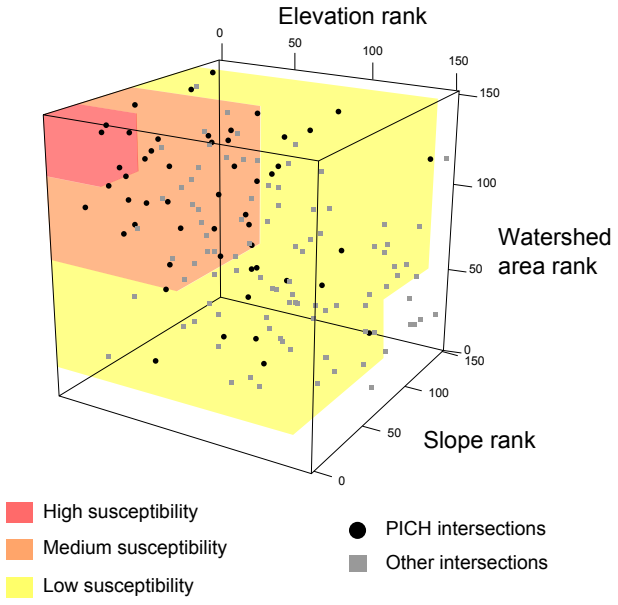

Figure 3: Illustration of the susceptibility sorting method applied by Versini on the 150 intersections of the initial testing areas.

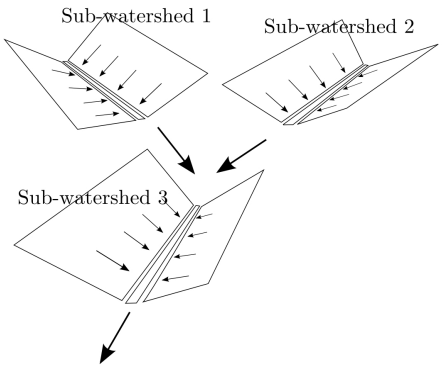

Figure 4: Schematic representation of a watershed in the CINECAR model.

Table 3: Average widths of river reaches according of their Strahler order I and return period T of the simulated stream discharge.

Return period of	Width of the stream (m)				
the discharge Q	I=1	I=2	I=3	I=4	I=5
$Q < Q_2$	4	16	36	64	100
$Q_2 < Q < Q_{10}$	8	32	72	128	200
$Q > Q_{10}$	12	48	108	192	300

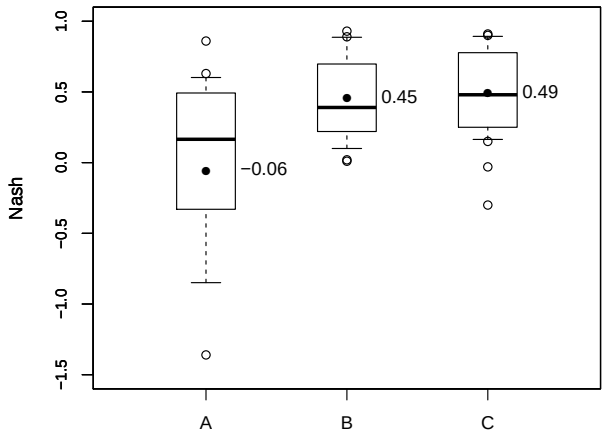

Figure 5: Distribution of the "event-specific" Nash criteria obtained at the 23 stream gauges for the validation events: A) initial version of the CINECAR model with homogeneous CN, B) CN estimated according to the USDA method, and C) introduction of the diffusive wave.

Table 4: Efficiency of the various sorting methods as applied over the entire PICH inventory territory: A) reference results obtained with the initial method on the area considered by Versini, B) initial method applied to the entire PICH territory, and C) application of the method with a sub-area distinction in the PICH territory.

	Susceptibility	Susceptibility class			
	method	High	Average	Low	Very low
Distribution of	A	0%	17%	58%	25%
intersections between	B	4.5%	9.9%	67.4%	18.3%
the susceptibility classes	C	6.2%	15.6%	53.3%	24.2%
Proportion of PICH	A	100%	60%	37%	0%
points in each	B	37.7%	25.4%	6.2%	0%
susceptibility class	C	40.1%	23.5%	4%	0%

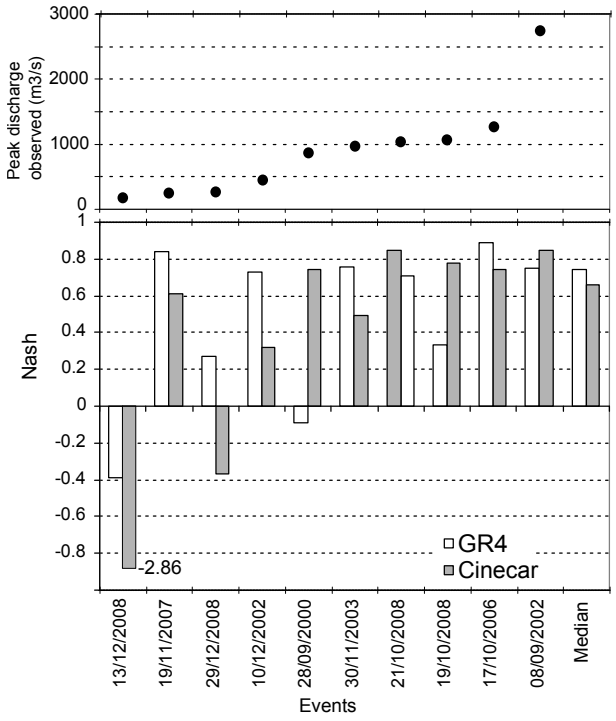

Figure 6: Comparison of Cinecar and GR4 model performance at the gauged Anduze station. The magnitude of the peak discharge is indicated in the top figure.

Table 5: Contingency table for the evaluation of the warning system performance: H (Hits), M (Misses), Fa (False Alarms), and Cn (Correct negatives).

		Reported flooding	
		Yes	No
Risk level	Yes	H	Fa
reached	No	M	Cn

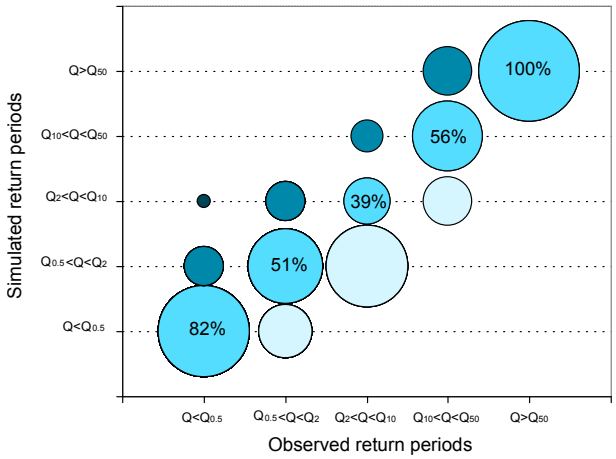

Figure 7: Comparison between the return period ranges of the observed and simulated peak discharges. Q_X stands for the discharge quantile of return period X.

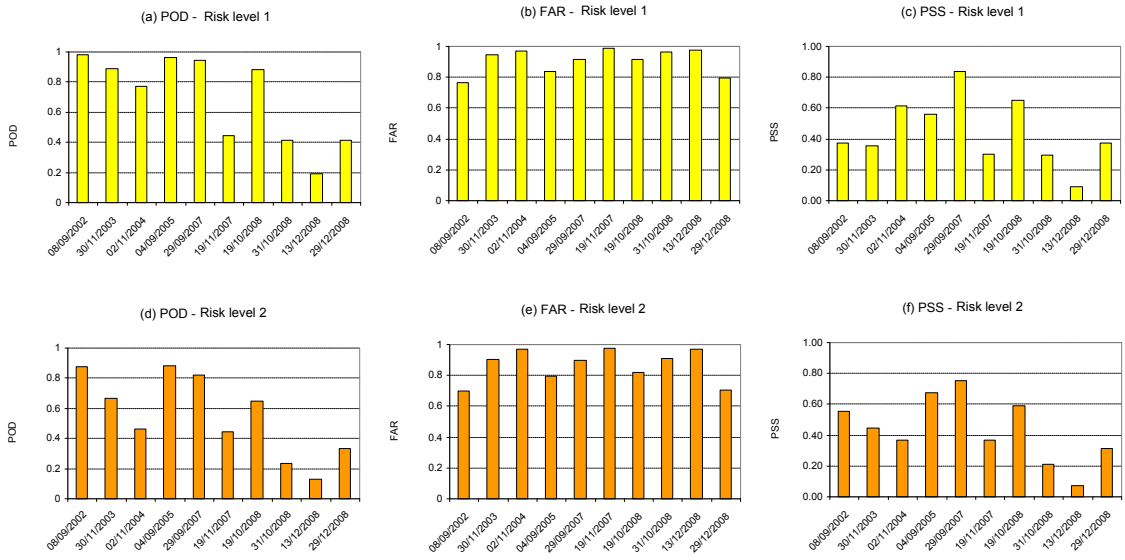

Figure 8: POD, FAR and PSS values obtained for risk levels 1 and 2

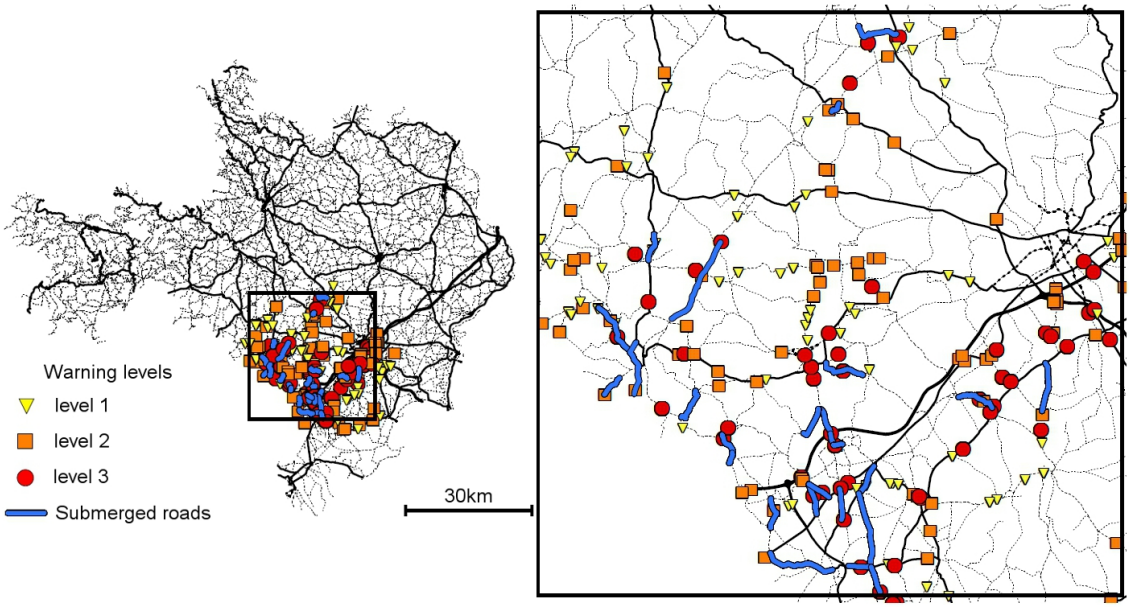

Figure 9: September 2007 flood. Map of the maximum forecasted risk levels and reported inundated roads.

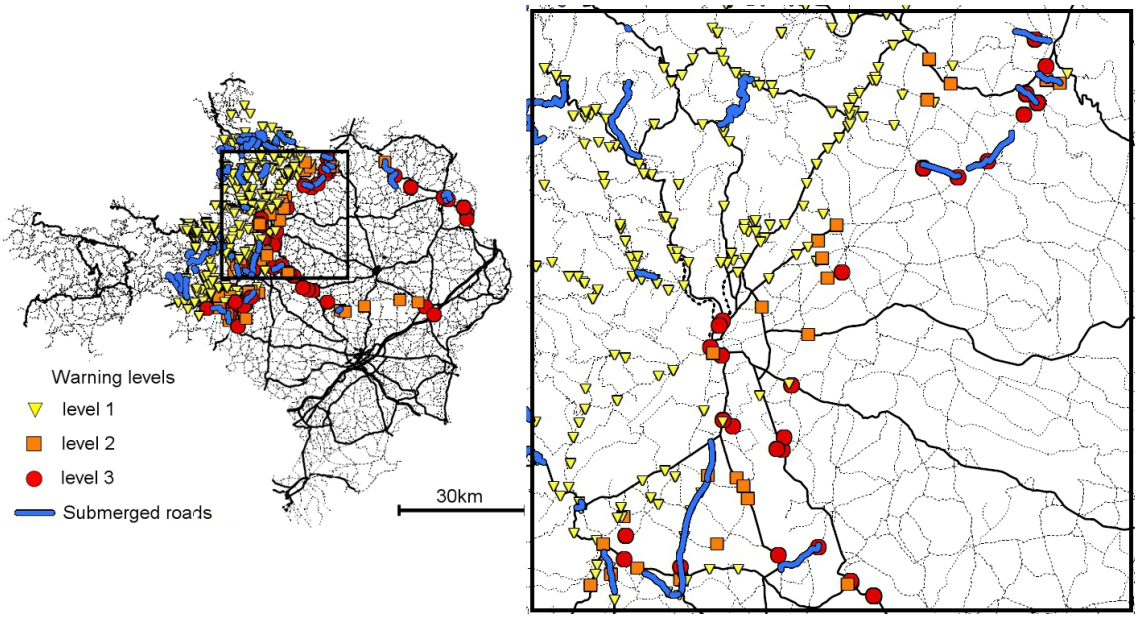

Figure 10: October 2008 flood. Map of the maximum forecasted risk levels and reported inundated roads.

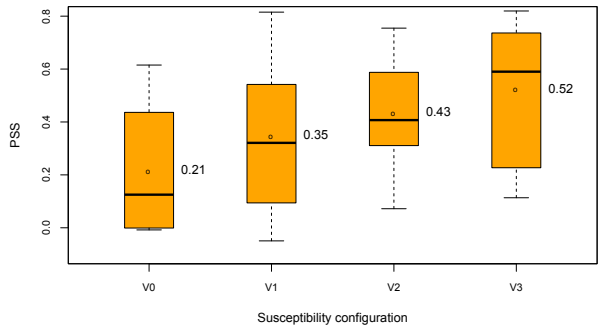

Figure 11: Box plot of PSS obtained for the risk level 2 with various susceptibility model configurations: V0) without susceptibility rating, homogeneous thresholds, V1) initial susceptibility rating method, V2) modified susceptibility rating method, V3) pseudo-perfect case, susceptibility determined based on the 10 tested events. The box plots synthesize scores obtained for the 10 tested events. The average value is indicated by the point.

Table 6: Proportion of PICH points among intersections identified in the recent event-specific reports of road disruptions (conducted since 2002), in areas where the PICH inventory is available.

Number of inundations	Non PICH	PICH
1	189	32
2	26	25
3	6	13
4	6	5
5	1	4
6	0	1
7	0	1
Total	228	81

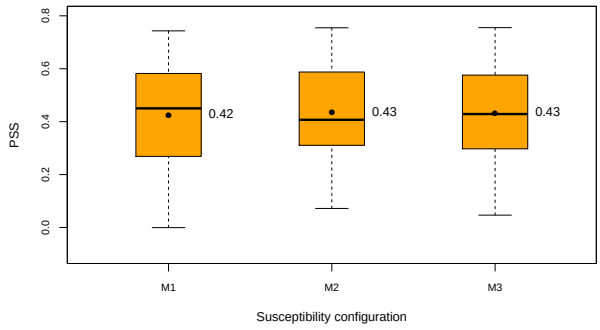

Figure 12: Comparison of the PSS distribution obtained with various rainfall-runoff model configurations for risk level 2: A) initial version of the CINECAR model with homogeneous CN, B) CN estimated according to the USDA method, and C) introduction of the diffusive wave.