



# **Analysis of the French insurance market exposure to floods: a stochastic model combining river overflow and surface runoff**

D. Moncoulon, D. Labat, J. Ardon, E. Leblois, T. Onfroy, C. Poulard, S. Aji, A. Rémy, A. Quantin

## **► To cite this version:**

D. Moncoulon, D. Labat, J. Ardon, E. Leblois, T. Onfroy, et al.. Analysis of the French insurance market exposure to floods: a stochastic model combining river overflow and surface runoff. Natural Hazards and Earth System Sciences, European Geosciences Union, 2014, 14, p. 2469 - p. 2485. .

**HAL Id: hal-01070525**

**<https://hal.archives-ouvertes.fr/hal-01070525>**

Submitted on 1 Oct 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





# Analysis of the French insurance market exposure to floods: a stochastic model combining river overflow and surface runoff

D. Moncoulon<sup>1,2</sup>, D. Labat<sup>2</sup>, J. Ardon<sup>1</sup>, E. Leblois<sup>3</sup>, T. Onfroy<sup>1</sup>, C. Poulard<sup>3</sup>, S. Aji<sup>1</sup>, A. Rémy<sup>4</sup>, and A. Quantin<sup>1</sup>

<sup>1</sup>CCR, 157 bvd Haussmann, 75008 Paris, France

<sup>2</sup>GET, 14 av. Edouard Belin, 31400 Toulouse, France

<sup>3</sup>IRSTEA, Centre de Lyon, 5 rue de la Doua, CS70077, 69626 Villeurbanne CEDEX, France

<sup>4</sup>Macif, 2 et 4 rue de Pied de Fond, 79037 Niort CEDEX 09, France

*Correspondence to:* D. Moncoulon (dmoncoulon@ccr.fr)

Received: 3 June 2013 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 11 July 2013

Revised: 7 August 2014 – Accepted: 21 August 2014 – Published: 17 September 2014

**Abstract.** The analysis of flood exposure at a national scale for the French insurance market must combine the generation of a probabilistic event set of all possible (but which have not yet occurred) flood situations with hazard and damage modeling. In this study, hazard and damage models are calibrated on a 1995–2010 historical event set, both for hazard results (river flow, flooded areas) and loss estimations. Thus, uncertainties in the deterministic estimation of a single event loss are known before simulating a probabilistic event set. To take into account at least 90 % of the insured flood losses, the probabilistic event set must combine the river overflow (small and large catchments) with the surface runoff, due to heavy rainfall, on the slopes of the watershed. Indeed, internal studies of the CCR (Caisse Centrale de Reassurance) claim database have shown that approximately 45 % of the insured flood losses are located inside the floodplains and 45 % outside. Another 10 % is due to sea surge floods and groundwater rise. In this approach, two independent probabilistic methods are combined to create a single flood loss distribution: a generation of fictive river flows based on the historical records of the river gauge network and a generation of fictive rain fields on small catchments, calibrated on the 1958–2010 Météo-France rain database SAFRAN. All the events in the probabilistic event sets are simulated with the deterministic model. This hazard and damage distribution is used to simulate the flood losses at the national scale for an insurance company (Macif) and to generate flood areas associated with hazard return periods. The flood maps concern river overflow and surface water runoff. Validation of these maps is conducted by comparison with the address located claim data on a small catchment (downstream Argens).

## 1 Introduction

Natural disasters in the world cause significant economic losses, estimated in 2011 at USD 380 billion. Almost a third (USD 113 billion) was covered by insurance (Swiss Re, 2012). The financial losses covered by the insurance market, due to weather-related catastrophes, were estimated in 2011 at USD 60 billion, which represent 50 % of total insurance losses.

Global damages due to climatic disasters have increased dramatically in recent years. In the case of floods, this can be partly explained by an increase of the exposed values in the flood-prone areas (due to growth of population and wealth); an increase of vulnerability due to the growth of industrial dependency to networks (transport, electricity, telephone, etc.); and the cost of protection and a possible influence of global climate change on the frequency of extreme flood events (Bates et al., 2008).

In terms of insurance losses, the following major historical flood events have been estimated by Swiss Re (2014): Hurricane Sandy-related sea surge floods in the US in 2012 (USD 35 billion), Thailand in 2011 (USD 16.2 billion), Germany/Czech Republic in 2013 (USD 4.1 billion) and 2002 (USD 3.1 billion), UK in 2007 (USD 2.9 billion), Switzerland in 2005 (USD 2.6 billion), and Australia in 2011 (USD 2.4 billion). In January 2013, flooding in Brazil, Turkey, South Africa, China, Indonesia and Australia were estimated at several hundred million dollars (Aon Benfield, 2013).

In France, the insurance losses for the major flood events of the last 20 years, estimated by Caisse Centrale de Reassurance (CCR) are the Rhône floods in December 2003 (EUR 950 million), the Xynthia storm sea surge in February 2010 (EUR 770 million), the Gard flash floods in September 2002 (EUR 680 million), and the Argens flash flood in June 2010 (EUR 430 million). The exposure to floods has several causes, and five types of floods have been defined: slow river overflow on large watersheds (e.g., downstream Seine River in 1910); groundwater floods (e.g., Somme River in 2001); flash floods (e.g., Argens overflow in June 2010); surface water runoff floods (e.g., Marseille floods in September 2000) and sea surges (e.g., Xynthia in 2010).

In France, according to law 82-600, 13 July 1982, a natural disaster (Nat Cat) is defined as damage caused by the abnormal intensity of a natural agent when the usual measures to be taken to prevent this damage were not able to prevent its occurrence or could not be taken. All compensations for natural disasters have to satisfy two conditions: a natural disaster must be recognized by an inter-ministerial decree and the property affected must be covered by a “property damage” insurance policy ([www.legifrance.gouv.fr](http://www.legifrance.gouv.fr), 1982). Perils covered by the scheme are not explicitly named in the law, but the most recurrent are flood, ground movements (including subsidence since 1990), avalanches and, since 2000, cyclonic winds in the overseas departments and territories. Earthquakes are considered as a major exposure (in the southeast of France and in the Antilles) but no major event has occurred since 1982. The hazard threshold above which a natural disaster is recognized is not laid out in the official decrees. However, concerning the hydrological flood disasters, legal precedents have converged to a 10 year return period for river flow or a 10 year return period for 1 h to 72 h rain. Thus, in France, the abnormal intensity threshold can be considered relatively low compared, for example, with the application in France of the 2007 European Flood Directive. In this directive, 10 to 30-year return period floods are considered frequent and high probability.

CCR is a reinsurance company owned by the French state. Its main aim is to offer state-guaranteed coverage to insurance companies for extraordinary risks. Among these risks, natural disasters as defined in the Nat Cat scheme represent the major part of the reinsurance premium and losses.

In 2010, the Nat Cat premium for the French insurance market represented EUR 1351 million or 10 % of the global damage insurance premium in France. In the last 20 years, floods and drought together have represented 95 % of the total losses for natural disasters in France, with 58 % (EUR 11.6 billion) for floods and 37 % for drought. The average annual loss for floods is EUR 509 million in this period. The 3 most expensive years for flood damages were 2010 (EUR 1.4 billion), 2003 (EUR 1.2 billion), and 2002 (EUR 1 billion). Those losses were characterized by large interannual variability.

Since 2003, CCR has developed a flood model, ARTEMIS, to estimate insurance losses due to flood just a few days after an event has occurred. This model constitutes the so-called “deterministic” approach. It is used for the communication of insurance loss estimates to the French government and to insurance companies (a simplified description can be found in Moncoulon and Quantin, 2013). It is extremely difficult to predict both when and where hydrological disasters will occur and to estimate their potential impact. Although great progress has been made in short-term meteorological and hydrological forecasting, large uncertainties remain with respect to monthly or yearly predictions. Nevertheless, in 2010, within the framework of the European Flood Directive, a new objective was assigned to CCR by the French Ministry of Finance: to estimate the financial exposure to floods for CCR, for the insurance market and for the French state.

In the same context, flood hazard maps have been generated by the French Ministry of Environment: the highest known floods (flood-prone areas) and the modeled EPRI (Evaluation Préliminaire du Risque Inondation – Preliminary evaluation of the flood risk) following the recommendations of the “Handbook on good practices on flood mapping in Europe” (EXCIMAP, 2007). In parallel, probabilistic flood models with damage estimation are developed by reinsurance brokers and modeling companies. The majority of these models only take into account river overflow, with some of them partially taking into account non-riverine floods. In other European countries, flood hazard mapping methodologies have recently been developed, e.g., in the UK (Bradbrook et al., 2005), in Germany (Falter et al., 2013) or in the Czech Republic (Dráb and Říha, 2010).

To estimate the flood exposure, the authors have built a probabilistic model, homogeneous over the entire country (not including overseas territories), which combines river overflow and surface water runoff. For this purpose, the ARTEMIS flood model is used to simulate the hazard and damages for each single event of a probabilistic event set. The event set originality is to combine a stochastic distribution of river discharges on large catchments with a stochastic distribution of spatialized rain fields on small catchments.

This paper is organized as follows: the first section describes the ARTEMIS flood model; the second section details the build of these two probabilistic flood event sets and their combination to generate a global flood event set. In the third section the results are analyzed and discussed: the insurance loss estimates on a private insurance portfolio and the exceedance probability maps. Finally, the limits of this method are presented.

## 2 Deterministic model ARTEMIS

As explained in the introduction of this paper, the development of the ARTEMIS deterministic model follows a specific

goal: to estimate the Nat Cat insurance losses due to recent flood events. The model must be operational for every flood type (flash flood, fluvial floods, ponding, urban floods and in the coastal floods) everywhere in France.

For more than 10 years, CCR has collected insurance portfolio data in the context of trade relations with its clients. These data represent, for 2013, 328 million risks and 1.8 million claims. The 1995–2010 period is the richest in terms of representativity of risks and claims. Depending on the year, up to 70 % of risks and 50 % of claims for the French insurance market are gathered in the database. All risks and claims are then geolocalized, according to their address, with the following results: 52 % of the risks are located at street number precision (at the exact localization of the address of the risk on the map); 24 % at street center precision (at the centroid of the street); 23 % at the commune level (at the centroid of the commune); and 1 % unlocated.

The flood damage assessment can be performed on different spatial scales (Merz et al., 2010; Messner et al., 2007): macroscale (for example department or region), mesoscale (homogeneous land use or commune level) or microscale (building, infrastructure). The main purpose of the model is to estimate the losses at the flood event scale (macroscale) or, at best, at the commune level (mesoscale).

The estimation of damages at the scale of a single insurance policy (micro scale) would require detailed information on the risk characteristics. According to Merz et al. (2010), information such as business sector, building type, building material, precaution, external response (emergency), and early warning are required at this scale. None of them, except for the business sector, are available in our policy and claim database. Furthermore, this microscale approach must also be combined with a detailed hydrological model.

Merz et al. (2010) has proposed a classification of impact parameters influencing flood damages. These parameters are flood depth, flood duration, flow velocity, contamination, debris, frequency and timing (day/night or holidays). In our model, flood depth and flow velocity are considered as the two main parameters. Others are neglected: the first reason is the difficulty to estimate them with a hydrological model (contamination, debris), the second reason is the lack of data to calibrate damage functions with a statistical approach (flood duration, frequency or timing). Furthermore, this information is never included in the claim insurance database. In addition, as it would be unrealistic to design a detailed topographic and hydrographic database in this large-scale approach, the authors have made the choice to develop a simplified hydrological model.

Using the CCR insurance database, the authors have combined insurance claim data and observed river flood zones to calculate the proportion of claims inside and outside. This study has been conducted on six major events. It appears that the ratio of insurance losses that are located inside a river flood zone varies and depends on the flood type. For slow-onset floods (like the Odet in Quimper in December 2000 or

Meurthe and Moselle 2006), the majority of flood claims are located inside the floodplain (more than 75 %). On the contrary, for flash floods (Gard, September 2002), rapid-onset floods or heavy rainfall storms (Marseille September 2000 or Paris, July 2001) more than half of the claims are generally located outside the river flood areas. A majority of claims seems to result from surface runoff on urbanized areas (Moncoulon, 2014).

In order to estimate the insurance losses due to flood events, ARTEMIS must combine the estimation of damages inside the flooded areas located along a river as well as outside. Given the objectives, the hydrological model has been designed to estimate hazard on both hillslopes and inundated areas along the main streams.

The components of the deterministic flood model are

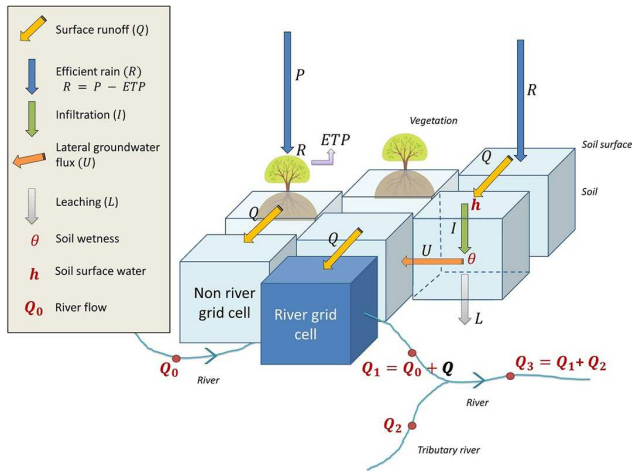
- a hydrological model presenting three components: a rainfall–runoff limited to hillslopes, a propagation model for main streams and an inundation model for floodplains. This model produces a spatial distribution of surface runoff velocity for hillslopes and water level for river banks;
- the vulnerability model, which is mainly represented by the insurance database and aims to describe localization and characterization of the elementary units of the model that are constituted by insurance policies;
- the damage model, which combines the outputs of the hydrological model and the vulnerability data on the basis of statistical distributions.

The two following sections will describe these components.

## 2.1 Flood hazard model

### 2.1.1 General description of the model

The flood model is a rainfall–runoff model distributed on a 50 m grid, coupled with a river discharge model (Fig. 1). The rainfall–runoff model is composed of a production function and a transfer function; the former has been developed using a classical approach, which can be found in several well-known models: GR3H (Edijatno et al., 1999), MERCEDES (Bouvier and Delclaux, 1996), ISBA-TOPMODEL (Bouilloud et al., 2009) and MARINE (Estupina-Borrell, 2009). Each grid cell is schematically described as a succession of vertical reservoirs. ARTEMIS is a double reservoir model: surface and soil water. The soil reservoir is fed at each time step with the infiltration from the surface. It is emptied by hypodermic transfer and groundwater percolation. It also can be emptied by evapotranspiration. The surface reservoir is fed with rainfall and upstream grid inputs. It is emptied by surface runoff and infiltration.



**Figure 1.** Runoff–rainfall and river routing processes for the distributed hazard model used for the deterministic and probabilistic approaches.

The horizontal flows are modeled with two transfer functions (one for surface water and one for hypodermic water). This approach is consistent with the MERCEDES interactive grid approach (Bouvier and Delclaux, 1996). At each time step, the production and transfer functions are computed for every grid cell. The specificity of ARTEMIS takes into account the amount of water entering each grid cell from the upstream cells. This amount of water is thus added to the efficient rain and is available for the production function.

The objective is to calculate a realistic amount of water on the soil surface in order to estimate the maximum hazard value for the insurance policies located on the different modeling grid cells.

The digital elevation model (DEM) is merged with a river database to create river grid cells. The rainfall–runoff model calculates the amount of water entering the river network via surface and hypodermic flows on river grid cells. This approach is similar to the ISBA-MODCOU method (Gomez, 2009; Habets et al., 2010). Then, a very simple 1-D routing model, based on kinematic wave equations with constant velocity, is applied. This choice has been made to simplify the calculation time for extra-large events (some extreme events have a spatial extent which can exceed a quarter of the French territory – for example the Rhône River floods in 2003).

### 2.1.2 Flood event definition

A flood event for CCR is the occurrence of significant flood damages to the insurance market (exceeding an arbitrary threshold of EUR 10 million) due to heavy rainfall or river overflow. The flood event is defined by a geographical location corresponding to a rectangular area characterized by bottom-left and upper-right coordinates. The event duration is determined by two characteristic time limits: 24 h before the first significant rainfall recorded by Météo-France and

24 h after the last significant rain. The 24 h before the significant rains are included to initiate and stabilize state variables (soil surface content and soil reservoir content). The 24 h after the significant rains are included to take partially into account the decrease of river flow and the propagation downstream. Thus, each event has a minimum duration of 48 h. The average duration is 72 h.

An historical event set has been built. It contains all the majors events that occurred in the 1995–2010 period for French metropolitan territory. Each new flood event is added to the event set a few days after its occurrence. Two years after the event, CCR collects a sample of claims that are used for the calibration of the damage model.

### 2.1.3 Rainfall interpolation

The rainfall and evapotranspiration data used in this simulation are provided by Météo-France rain gauges (Table 1). Both values are interpolated on the grid using a Kriging method (Krige, 1951; Arnaud and Lavabre, 2010). For each cell of the grid, the efficient rain  $r$  is calculated as following:

$$r = p - \text{etpr} = p - \text{etp}, \quad (1)$$

where  $p$  is the hourly rainfall (mm) and  $\text{etp}$  is the hourly potential evapotranspiration (mm). The  $\text{etp}$  is considered as a constant value during the day and the daily value will be divided by 24 to estimate the hourly  $\text{etp}$ .

Vegetation interception and depression storage during the rain event are neglected. In the probabilistic approach, this section of the model is bypassed and spatialized rain fields are used as an input for the rainfall–runoff model.

### 2.1.4 Rainfall–runoff

The rainfall–runoff model is a water balance that estimates the water level in the surface reservoir, noted as  $h$  (m) and the soil reservoir content  $\theta$  (mm) at each time step  $t$  for the entire flood event duration. In a simplified approach for large-scale simulations, the surface water balance can be estimated for small time steps and small surface units as follows:

$$\frac{\delta h}{\delta t} = r - i + q_{\text{in}} - q_{\text{out}}(-\text{riv}) \quad \frac{\delta h}{\delta t} = r - i + q_{\text{in}} - q_{\text{out}}(-\text{riv}), \quad (2)$$

where  $h$  is the quantity of surface water (mm),  $t$  is the time,  $r$  is the efficient rain ( $\text{mm h}^{-1}$ ),  $i$  is the infiltration flow ( $\text{mm h}^{-1}$ ),  $q_{\text{in}}$  is the runoff flow of surface water from adjacent cells ( $\text{mm h}^{-1}$ ) and  $q_{\text{out}}$  is the runoff flow of surface water to adjacent cells ( $\text{mm h}^{-1}$ ). The water flow to the river network  $\text{riv}$  is only taken into account for identified river cells.

The soil wetness  $\theta$  is estimated by Eq. (4):

$$\frac{\delta \theta}{\delta t} = i - l + u_{\text{in}} - u_{\text{out}} \quad \frac{\delta \theta}{\delta t} = i - l + u_{\text{in}} u_{\text{out}}, \quad (3)$$

where  $\theta$  (mm) is the soil wetness at time  $t$ ,  $i$  is the infiltration flow ( $\text{mm h}^{-1}$ ),  $l$  is the leaching flow from the soil reservoir to the underground water ( $\text{mm h}^{-1}$ ),  $u_{\text{in}}$  is the groundwater drainage from adjacent cells ( $\text{mm h}^{-1}$ ) and  $u_{\text{out}}$  is the groundwater drainage to adjacent cells ( $\text{mm h}^{-1}$ ).

The water velocity is calculated on each cell using the Manning equation (Manning, 1891). Each square grid ( $50 \times 50 \text{ m}$  grid size DEM) is considered with a constant and homogeneous slope, a constant and homogeneous water level and land cover.

$$v = K \cdot h_t^{\frac{2}{3}} \sqrt{p_t}, \quad (4)$$

where  $K$  is the Manning rugosity coefficient ( $\text{m}^{1/3} \text{ s}^{-1}$ ),  $h_t$  is the quantity of surface water (m) at time  $t$ , and  $p_t$  is the square root of the sum of the square of the eight slope gradients ( $\text{m m}^{-1}$ ).  $K$  is determined based on the land cover. The water velocity is used to compute, for every time step, the proportion of surface water volume leaving a grid cell and reaching the adjacent cells.

Two types of cells are defined on the DEM grid: non-river cells and river cells. A cell is identified as a river cell if a stream segment crosses it. All water amounts entering a river cell by surface and hypodermic flows are added to the river flow model.

The soil reservoir has a volumetric water limit of 200 mm. This maximum soil wetness is constant for the entire grid. Infiltration flow from the soil surface to the soil reservoir is derived at each time step from the Green and Ampt equations (Green and Ampt, 1911):

$$i = \text{Ic} + \frac{K_2 b}{\theta_t}, \quad (5)$$

where  $i$  is the infiltration flow – the volume of water entering a unit of soil surface per unit of time ( $\text{mm h}^{-1}$ ),  $\text{Ic}$  is the asymptotic steady infiltration flow reached when  $t$  becomes large ( $\text{mm h}^{-1}$ ),  $b$  is the infiltration decrease ( $\text{mm}^2 \text{ h}^{-1}$ ) and  $\theta_t$  is the soil wetness expressed in mm.

The infiltration flow is mainly dependent on the amount of water available on the soil surface and on the soil wetness. The infiltration decrease parameter has been modified on each grid cell using a parameter  $K_2$ , with value ranges from 0 to 1 depending on the land cover. This parameter  $K_2$  is used to limit the infiltration in urban areas in order to maximize surface runoff. The results show that the model satisfactorily estimates the preferential flows, which could explain some claims outside the floodplains (see results Sect. 4.2)

At each time step, the following equation is proposed to estimate the groundwater drainage:

$$u = \max(0.1, K_3, \sqrt{\theta_t}), \quad (6)$$

where  $u$  is the groundwater drainage ( $u = 0.1 \text{ mm h}^{-1}$ ) for unsaturated soils and  $K_3$  is a parameter of the model: the horizontal drainage for a saturated soil),  $\theta_t$  is the soil wetness at

time  $t$  (as a ratio of the maximum volumetric water content). The groundwater drainage is then distributed in eight directions depending on the slope. Parameter  $K_3$  is simplified to a unique value for all soil types. The improvement of this parameter is a work in progress.

The leaching to the underground water is a lost volume for the system. This amount of water is determined by the volume of water exceeding the soil reservoir capacity, after groundwater horizontal drainage. This amount of water is divided between exfiltration to soil surface and leaching.

### 2.1.5 Hydrological river routing

A very simple 1-D hydrological river routing method is used to determine river flow, water level and overflow areas in the floodplain (Fig. 1). The river is considered as a succession of river segments which are individually homogeneous in terms of section shape, width and slope.

The transfer is computed in these segments on the basis of the kinematic wave model, which is a simplification of the Barré de Saint-Venant equations. For this large-scale application, the river flow celerity is considered to be constant during the flood event. Its calibration has been implemented for each region on the basis of flood hydrograms.

At each segment, the propagation model is fed by the outputs of the rainfall–runoff model on the corresponding river cell at each time step. This simplified river routing model permits great optimization of the calculation time and allows for the simulation of a dense hydrographic network at a very large scale.

The river discharge at time  $t$  is compared with a threshold defined for each section by the 10-year return period flow or a threshold that characterizes the efficiency of the flood protection structures for that section. The flood protection structure database used for this study has been developed by CCR and is not exhaustive. The flood protection overflow is taken into account when the information exists but the breach risk is not modeled in details.

The volume of water exceeding the threshold for a river segment is propagated on the DEM in eight directions using the Manning equation – same as for rainfall–runoff model (see Sect. 2.1.4). The K-Manning values are estimated from land use (Corine land cover). As mentioned by Bradbrook et al. (2005), the lack of precision in the elevation grid will induce approximations in the expansion of the flood, which will have higher impacts than the uncertainties in river flow estimations.

### 2.1.6 Hazard model parameters

ARTEMIS hazard model has five parameters to adjust: the infiltration function requires three parameters ( $\text{Ic}$ ,  $b$  and  $K_2$ ), the hypodermic transfer function requires one parameter (maximum water speed  $K_3$ ) and the last parameter is modeling time-step duration (e.g., 120s). Kriging parameter  $L$  is

calibrated automatically during rainfall interpolation at each time step. The Manning coefficient  $K$  and parameter  $K_2$  have constant values depending on the land use. These two parameters are not calibrated.

Model calibration is conducive to optimizing river flow prediction and statistical damage law adjustments. First, the hazard model calibration based on river flow estimations is conducted. Second, the statistical damage law adjustments are conducted.

## 2.2 Vulnerability and damages modeling

The damage model is used to estimate the cost of flood events based on information on hazard and vulnerability. Hazard information will indicate flood intensity and vulnerability information will provide flood sensitivity information per insured risk. Without any information on the hazard intensity at the insurance policy scale, the only data available to calibrate the damage function are produced by the hazard model.

As we explained in the general description above, the damage model is composed of statistical regressions built by analyzing correlations between modeled hazard outputs and claim data on a selection of historical events. A given version of the damage functions is thus linked to a given version of the hazard model. Specific damage functions are built for river overflow (characterized by water level inside the flood zone) and for non-riverine floods (characterized by maximum water velocity on each grid cell outside the flood zone).

The classification of elements at risk in the vulnerability model is the following: private household, house tenants, flat holders and flat tenants (for residential risks), professional, industrial and agricultural risks. Professional and industrial risks are separated by their insured values, whereby professional is a mostly commercial risk.

The vulnerability data used in the damage model are

- insured values
- risk location
- floor
- building type (concrete, brick, masonry)
- industrial activity (i.e., line of business).

The hazard model output data used in the damage model are

- water depth for river overflow (m) for risks located in the river flood zone
- water velocity for surface runoff ( $\text{m}^3 \text{s}^{-1}$ ) for other risks.

These values do not represent measurable physical quantities, but the order of magnitude of input data used to calibrate the damage model and to map the exposure.

Two important statistical distributions are calibrated in the damage model:

- the claim frequency, depending on the hazard intensity at the risk address
- the destruction rate, applied to the insured value, in case of claim occurrence.

The first distribution (claim frequency) is calibrated by fitting a logit model to the empirical hazard intensity/occurrence of claim (0 or 1) on the calibration events. The logit model equation is as follows:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right). \quad (7)$$

The logistic regression model is thus:

$$\ln\left(\frac{p}{1-p}\right) = a_0 + a_1x_1 + \dots + a_ix_i, \quad (8)$$

with  $p$  a value between 0 and 1. In this model,  $i = 1$  and  $a_0$  and  $a_1$  are the parameters to be calibrated.  $x_1$  is the hazard intensity.

The second distribution (destruction rates) is a square root function calibrated with the historic claim database. This calibration is studied separately for every risk type (residential, commercial, industrial and agricultural). The input parameter for this distribution is the hazard intensity.

The event set for calibration is selected among the 1995–2010 events. An event becomes a calibration event when both the representativeness of the claim database and the quality of the hazard simulation (based on the river flow model scores) are adequate for modeling. The event set used for the calibration of the model used in the paper is the following: Aude November 1999, Marseille September 2000; Nice December 2000; Brittany December 2000 and January 2001; Gard September 2002 and 2005; Rhône December 2003 and November 2008; Meurthe et Moselle October 2006 and Saint Tropez October 2009.

The cost of flood is calculated for the insurance market or for an insurance company by estimating the claim frequency and the damage ratio for each risk. This calibration step provides reasons to be confident in the simplified approach.

The damage model is a deterministic model except at the end of the whole process: a bootstrap method is applied to the commune level losses (compared to the historical losses) to determine a confidence interval based on the differences between simulations and claim extrapolations. After calibration, the model is used to simulate the whole historical event set (more than 100 events). The error between simulation and claim extrapolation (which is our reference) is then calculated to validate the calibration process. Each time a new version of the model is created, the whole event set is simulated to validate the version.

The error simulation, the so-called second-level uncertainty, is also simulated for each event in the probabilistic (or Monte Carlo) simulations. The comparison between the historical losses and the simulated losses for a selection of flood



**Table 1.** Sources and description of the main input data for the hazard and damage models (deterministic and probabilistic models).

Input data	Source	Description
Digital elevation model (DEM)	IGN BD Alti	DEM with 1 m resolution on a $50 \times 50$ m grid covering the French metropolitan territory
River database	IGN BD Carthage 2012	Selection of 65 000 km of river out of 315 000 km
Measured rainfall data	Météo-France Publiothèque	Rain gauge stations
Evapotranspiration	Météo-France Publiothèque	Penman model applied by Météo-France on rain gauge stations
Modeled rainfall data	Météo-France SAFRAN database	51 years of hourly and daily precipitation on a $8 \times 8$ km grid
Hydrological data	SCHAPI – Ministry of Environment	Historical water level measurements on Banque Hydro river gauges
River protections against flood	CCR	Determination of flood threshold by comparing Nat Cat decrees per commune with historical river flows
Policies and risks locations, insured values and lines of business	CCR insurance database	The database content is estimated as 70 % of market share for 2011 portfolio
Natural disaster claims	CCR insurance database	The database content is estimated as 50 % of market share for 2011 portfolio
Natural disaster recognition per commune	CCR Nat Cat database	Exhaustive database since 1982 for all Nat Cat decrees
Destruction rate curves	CCR	Calibrated on CCR claim database and hazard model outputs

events is shown in Table 6. The distribution of damages for a single event, based on the uncertainties in the simulation of the calibration event set, is shown in Table 7 for an historical event.

### 3 Generation of a probabilistic flood event set

To achieve our goal, which consists in estimating the exposure of the French insurance market to floods, we need to associate a return period to any amount of annual damages. The occurrence exceedance probability (OEP) curve – i.e., the annual probability to exceed, for a single event, a given amount of damages (Lloyd's market association, 2013) – is not consistent with our objectives: most of the insurance and reinsurance contracts are indeed annuals. Analysis of financial exposure for the insurance market must be based on annual damages. The aggregated exceedance probability (AEP) curve – i.e., the annual probability to exceed a total annual loss – has to be calculated.

A flood event is characterized by a single spatial and temporal intensity distribution. It is almost impossible to estimate the global return period of a single event and the same amount of annual damages can be caused by an infinite number of combinations of single events.

To estimate the probability distribution of annual damages, our idea is to use a Monte Carlo approach by simulating a large number of single flood events with the determinis-

tic model and aggregate their losses at the annual timescale. Because our historical event set may not be representative enough, we have to build a fictive event set. Fictive events must be realistic and have to be consistent with historical events statistic behavior.

First, fictive river flows are generated on a selection of flow stations from a river database (Quantin, 2011; Moncoulon and Quantin, 2013; Moncoulon, 2014). This generator is calibrated on historical river discharge records in the Ministry of Environment database (Banque Hydro, 2006). This approach is called the F1 model. The second method uses the fictive rainfall generator SAMPO-TBM (Leblois and Creutin, 2013) calibrated on the Météo-France SAFRAN rainfall database (Durand et al., 1993) to create annual series of hourly rainfall and simulate the flood events associated. This approach is called the F2 model. Both event sets are then combined to create a single library of flood events and create hazard maps and annual damage distribution.

#### 3.1 Description of the stochastic event sets

The methodology chosen in this approach is to combine two independent event sets generated by two independent continuous simulations of hazard values: maximum monthly river flow for F1 and cumulated 72h rainfall for F2. The use of continuous hazard generation allows us to skip the calculation of a single event frequency.

An event set composed of 18 057 flood events has been generated with the F1 model for 1000 years of fictive river flows. These events are located on the entire French territory on the seven major river catchments. The number of events per major catchment is 3125 for Seine-Normandy, 4810 for Rhône-Mediterranean-Corsica, 2786 for Adour-Garonne, 3149 for Loire, 1233 for Brittany, 1540 for Rhine-Meuse and 1414 for Artois Picardy.

Due to the selection of all river flow occurrences exceeding a return period threshold, the number of events is directly dependent on the number of gauge stations in each catchment. With this approach, 18 flood events per year occur on at least 1 river gauge on the French territory. These events have a minimum return period of 10 years (the selected threshold for the study) and can be considered as Nat Cat events. In our approach, the different river flows generated on different flow gauges occur during the same month but not necessarily at the same time. The river flow values are propagated upstream and downstream until another gauge station with a generated value is present on a segment.

On the other hand, an event set composed of 8240 events has been generated by the F2 spatialized rainfall generator for 150 simulated years independently on the 96 CRESTA zones over the French metropolitan territory. In France, a CRESTA zone is a department. The average number of events per year for the F2 distribution is 0.66 per CRESTA. For the entire French territory, 55 F2 model events occur every year (8240 out of 150 years). With this method, independent event sets per CRESTA are created.

Based on these 150 years of fictive rain fields, the copula method gives us 5000 years of correlated 72 h rainfalls in the 96 CRESTA zones. The number of years for spatial and temporal simulations must be higher than in single CRESTA zones to generate all the probabilistic combinations at the entire territory scale.

For the two methods (F1 and F2), the hazard and damages are simulated with the deterministic model for each single event. Each single event in the probabilistic event is associated with a distribution of damages, according to the deterministic model distribution of uncertainties.

All the losses simulated for the events that occur during the same year are summed up to calculate the annual losses. For each year, within the F1 method, the losses of the 12 months and the 7 major catchments are added. For F2, the losses of the 3-day events in the 96 catchments are added to compute the annual losses. We built two independent distributions of annual losses for flood. The F1 only considers river overflow and the F2 is a combination of small catchment river overflow and surface runoff due to heavy rainfall. The originality of this method lies in the combination of both distributions considering independent years (F1 years and F2 years) to build a global flood loss distribution.

This method is applied on the calculation of annual flood losses for the insurance companies or for the whole insurance

market. For this paper, we have chosen to calculate the flood losses for Macif insurance.

### 3.2 F1 model: the generation of fictive river flow

The general principle is to generate years of monthly fictive river flows, in order to detect river overflows for the constitution of a library of multiple single events.

Different types of river flow are available in the Banque Hydro: average daily values or maximum values per month. The maximum values per month are chosen to avoid the underestimation of extreme values due to the average daily data. Statistical distributions are fitted separately for each piece of gauge data. The minimum period of records necessary to be considered is set to 30 years. Therefore, a sample of 802 stations out of 2200 was selected.

The distribution of flow data for a single station is not homogeneous: e.g., river flows in January are strongly different from river flows in August. Furthermore, river flows for 2 successive months are not necessarily independent: a monthly maximum value can occur the last day of month 1 and the first day of month 2. To work on independent and homogeneous data, a set of 9624 variables are defined: 802 stations for 12 months. For each variable, we fit a probability distribution with a maximum-likelihood method among the following: gamma, log-normal, generalized extreme values (GEV) and Weibull. We retain the one that minimizes the BIC (Bayesian information criteria) on the one hand, and that passes goodness-of-fit tests on the other, such as Chi 2 test or the Anderson–Darling test.

The dependence between the 9624 variables is represented by a Gaussian copula (Nelsen, 1999; Quantin, 2011). Sklar's theorem (Sklar, 1959) on copulas allows us to model a multivariate distribution in two separate parts: the individual behavior of each marginal with empirical or fitted distributions and the dependence between marginal with a copula. There are several common families of copulas such as Gaussian or Archimedean. A Gaussian copula has a single parameter: the correlation matrix of the variables. Based on this method, 1000 years of maximum monthly river flows are generated. For each year, 9624 values are generated: 12 maximum monthly river flows for each station.

The intensity of our flood events is not only based on the return period of the river flow but also on the occurrence of many overflows over a large territory. When we generate 1000 years of river flow, the probability to generate a high return period flow on a single station is low. But we will simulate extreme events, in terms of insurance losses, due to the number of rivers integrated in the simulation and the representation of spatial dependency based on the copula.

In our approach, we do not verify the consistency of the hydrological values when we simulate the flood: the river flow, for example, at a station C, downstream of a tributary station B and the upstream station A, will not necessarily verify  $C = A + B$ . This choice has to be made to preserve the

consistency of the statistical values on each river segment. The river flow values generated on a gauge station will be propagated with a constant value on the river segments until another value generated on another gauge is present (Fig. 6).

This library of stochastic river discharges is used to create fictive flood events: an overflow event is created when at least one value of flow for one station is above the 10-year return period value. This 10-year threshold has been chosen to be consistent with the Nat Cat recognition ratio applied in the Nat Cat scheme.

As described in the Banque Hydro, the French metropolitan territory is divided in seven catchment areas: Artois Picardy, Brittany, Adour-Garonne, Loire, Rhine-Meuse, Rhône-Mediterranean-Corsica and Seine-Normandy.

With this method, 1000 years of continuous fictive flows are generated. From these simulated flows, 18 057 events are created. Each event is simulated with the river routing model, which will be forced by the simulated river discharges. The rainfall–runoff model is bypassed in this method. This simplified deterministic model is used to calculate the impacts of each event in terms of hazard and damages. For each event in the event set, the damage uncertainties are estimated based on the error distribution calibrated and validated on the 1995–2010 historical event set. This probabilistic distribution is called F1. The flood events are all dependently generated with the Gaussian copula. Thus, both OEP and AEP curves are built from the flood damage and annual aggregated damages.

### 3.3 F2 model: the generation of stochastic rain fields

#### 3.3.1 Input data

We use an extract of the Météo-France database SAFRAN (reanalysis of atmospheric surface fields): an 8 km resolution over France for the period 1958–2010 for daily and hourly rainfall (Durand et al., 1993). This database was generated by the Météo-France SAFRAN-ISBA-MODCOU analysis system. These data are homogeneous over the period and over the entire territory. In 2011, 184 million pieces of data of daily rainfall and 4416 billion pieces of data of hourly rainfall were recorded in the SAFRAN database and used for the calibration of the method.

A local rainfall climate is considered as a succession in time of several rainfall types, including the dry weather type (i.e., no local rainfall). Homogeneous areas have been defined: for methodological simplification reasons, CRESTA zones (French departments) are used. At every point of a given small catchment, for a given time, only one rainfall type can occur.

Each rainfall type is considered as a random process defined by three parameters: the average ( $\mu$ ) and the standard deviation ( $\sigma$ ) of non-null rainfall and the average spatial rate ( $\tau$ ) of non-null rainfall coverage (i.e., the number of non-

null rainfall SAFRAN grid cells out of the total number of SAFRAN grid cells for the small catchment).

#### 3.3.2 Long-term rain sequences simulation (6 h time steps)

To build rainfall types we use a purpose-oriented typing of observed rainfall data. The sequences of hourly precipitation are summed to 6-hour cumulative values characterized by the average and the standard deviation of non-null rainfall and the average spatial rate of non-null rainfall coverage (see above).

These descriptors are organized by a Kohonen self-organizing-map (Kohonen, 1995) to build several types and allocate individual time steps to one of these types. For this study, 25 types of wet weather and 1 type of dry weather have been defined. One set of parameters is derived for each rainfall type. The qualitative sequence of rainfall types is described by sequencing algorithms. The sequence of rain types in time are analyzed and simulated by using the transition probability matrix from one type to the others.

In the first step, the SAFRAN rain fields are selected in CRESTA zones and analyzed with the Kohonen self-organizing map to create 25 types of wet weather and 1 type of dry weather. Then the transition probability matrix is applied to simulate long-term rain sequences. These rain sequences are characterized, in a CRESTA zone, by a single rain type on the whole surface. The sequences are 6-hour sums of precipitation homogenous for the entire zone. These 6-hour cumulative rainfalls are compared at each time step with a threshold (e.g., the 2-year return period 72-hour rain). When the sum of precipitation exceeds the threshold, a Nat Cat event is identified and created.

#### 3.3.3 Combination of CRESTA zones

Precipitation is generated in CRESTA zones. In order to get precipitation for the entire French territory, we use a Gaussian copula to analyze the dependence between zones. In our study, we want a 72-hour sliding cumulative sum of precipitation for each zone. In our study, the 96 French departments (CRESTA zones) are considered as small catchments. A statistical analysis allows us to assume that the 72-hour cumulative rain fields are temporally independent for each small catchment. We generate daily rainfalls and compute the cumulative sum for every 3 days. To preserve the seasonal pattern, a year of precipitation is simulated and broken down into four parts corresponding to the four quarters of the year. We obtain, for each year, 122 sliding amounts of precipitation that preserve the observed correlation between watersheds.

The simulated damages at the CRESTA scale are aggregated both annually and spatially with the result of the 72-hour copula matrix for the 96 zones. With this method, we build OEP and AEP curves at the whole national scale.

A national-scale event is then defined as the total amount of damages occurring during a 72-hour period over all the CRESTA zones exceeding a threshold.

### 3.3.4 Spatialized rain generator (1 h time steps)

The non-zero precipitation field is derived from a Gaussian field. The indicator field (1 for rain and 0 for dry weather) is obtained by thresholding an independently simulated Gaussian field. The spatio-temporal Gaussian fields are generated by the turning band methods (TBM). This method generates non-conditional stochastic simulations in three dimensions (two in space and one in time) from a large number of one-dimensional simulations called “band” (Matheron, 1973). The homogeneous spatio-temporal simulator is described with more details in Leblois and Creutin (2013). This simulator has been used to prove the importance of rainfall spatial distribution for the design and sizing of flood control structures (Poulard and Leblois, 2008).

As it is calibrated on SAFRAN data, the spatialized rain generator will simulate 6-hour rain fields based on the 6-hour SAFRAN. Then, the 6-hour cumulated rainfall is combined with the indicator field (6 values of no rain (0) or rain (1)) to transform the 6-hour into 1-hour precipitation.

Then, the spatialized hourly rain fields are used as the efficient rain in the rainfall–runoff model. It means that the evapotranspiration (etp) is neglected during the 72-hour flood events. Each event in the library is simulated with the complete deterministic flood model to calculate its impact in terms of hazard and damages (Fig. 6). The initial soil moisture content is dependent on the month of occurrence and is calibrated on the Météo-France soil moisture data. With this method, we build OEP curves at the CRESTA scale for 150 years of simulated rain fields.

## 3.4 Combination of F1 and F2 for exceedance probability mapping

### 3.4.1 Annual losses distribution

The first and second approaches allow us to build two distributions of flood frequency. The probability of occurrence of each event is not calculated but every event is considered as unique in the distribution. These two distributions are not independent: small gauged river overflow may have been simulated with both methods. The F1 distribution presents the highest values for extreme river overflow events (e.g., Seine overflow in Paris, Rhône and Saône in Lyon, Garonne in Toulouse or Bordeaux). Nevertheless, for small river overflow, the F1 approach underestimates the insurance losses: the surface water runoff damages will not be taken into account. For both reasons, F1 and F2 distributions must be combined to build a complete flood distribution.

Our objective, in this study, is to combine both distributions with the following principle: F2 distributions are con-

cerned with mainly small-sized event damage (including surface runoff and small river overflow), while F1 is a large-sized event damage distribution (without surface runoff). F2 will bring the small return period losses, whereas F1 will contribute mostly to the distribution tail.

We propose the following combination method: F1 is the annual loss distribution for the large catchment river overflow approach (1000 years). F2 is the annual loss distribution for the small catchment surface runoff approach (1000 years).  $S$  is a given threshold. The annual probability  $P_{A>S}$  to exceed this threshold is estimated for the F1 distribution by

$$P_{A>S} = \frac{n_{A>S}}{n_A}, \quad (9)$$

where  $n_A$  and  $n_{A>S}$  are respectively the total number of F1 years and the number of F1 years with a loss exceeding the threshold. We define a threshold  $T$ , which is the maximum annual loss of the F2 distribution. We build an event set  $F$  (of  $n_C$  years) combining the entire F2 distribution and the selection of  $n_{A>T}$  most extreme years of F1. If  $n_{A>S}n_{A>T}$  is very small compared to the number  $n_B$  of F2 years, the annual probability to exceed the threshold  $T$  for the combination of F1 and F2 (Eq. 16) will not be significantly different than Eq. (9)

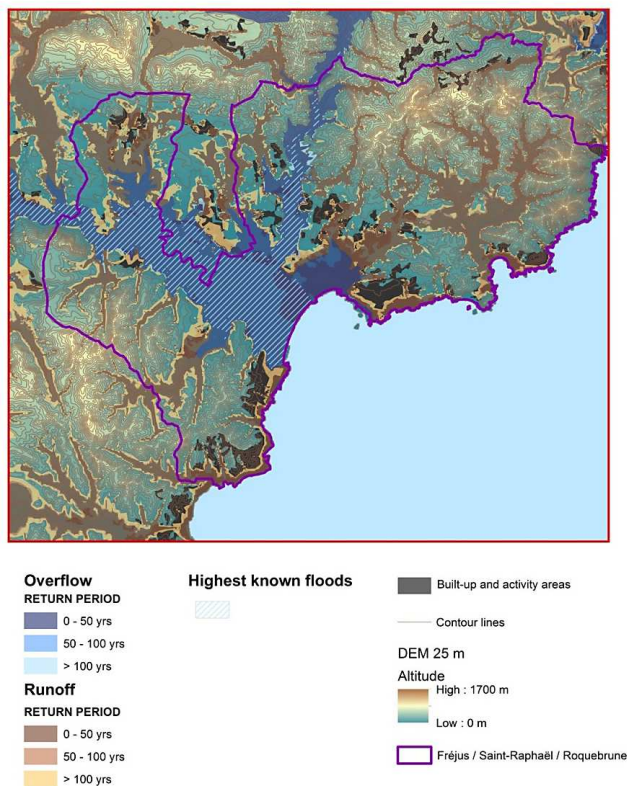
$$P_{C>T} \cong P_{A>T} = \frac{n_{C>T}}{n_B + n_{A>T}}. \quad (10)$$

### 3.4.2 Hazard maps

To build the probabilistic hazard maps, we generate flood maps for each event: river overflow maps for F1 events and both river overflow and surface runoff for F2 events. Then, we use the annual distribution of F1 and F2 events in the event set to generate probabilistic hazard maps. According to the events occurring for each stochastic year in the event set, the hazard maps are overlayed to construct a probabilistic hazard map indicating, at each point of the territory, the return period of a given hazard intensity. These maps correspond to the flood extent map according to the de Moel et al. (2009) classification. The map shown on Fig. 2 is built for the following intensities: non-null water level for river overflow and the water velocity for non-riverine floods. The return period  $R$  indicated on the map is calculated as an exceedance probability:

$$R = \frac{N}{n}, \quad (11)$$

where on each grid cell,  $n$  is the number of years with hazard intensity above the threshold and  $N$  is the total number of years in the simulation.

**CCR Probabilistic Hazard / Highest known floods***Côte d'Azur : Fréjus / Saint-Raphaël / Roquebrune*

**Figure 2.** Probabilistic flood map and official flood-prone areas on the downstream floodplain of the Argens River near Fréjus, Saint-Raphaël and Roquebrune. This exceedance probability map combines the Argens overflow area with the surface runoff at different return periods.

## 4 Results and discussion

### 4.1 Damage simulations for the Macif insurance portfolio

Macif, a French insurance mutual company has joined the CCR flood modeling project by providing high quality data of risks and claims. The Macif portfolio is distributed homogeneously on the French territory. The non-motor Nat Cat premium for Macif was EUR 54.8 million in 2011, which represents a market share of 4.6 % in France. As described in Table 2, the Macif non-motor portfolio is composed of individual risks (93.8 %) and professional risks (6.2 %). The amount of losses per risk category is almost the same for flood as for all perils (floods included), which with floods only represents 39.6 % of the Nat Cat losses over this period.

In this study, we focus on the individual risk category. Regarding these risks, a statistical approach is more robust than on professional or industrial risks for two reasons: the

homogeneity of the insured portfolio and the high number of records in the claim database. The flood claim database analysis shows that the individual risks represent 54.1 % of the total flood losses. The ratio of industrial and commercial damages increases with the event loss amount.

The geocoding quality is described in Table 3. A majority of risks is successfully geocoded at the street number precision (61.8 % of the individual risk premium) or at street center precision (24.7 %). Only 13.5 % of the individual risk portfolio is not precisely located (commune level precision).

The historical flood losses for the period 1996–2011 are estimated by Macif and described in Table 4. The present value of all annual losses was calculated by applying (1) the monetary inflation and (2) the increase in the market share of the Macif portfolio. The average annual loss is EUR 21.4 million, and the 1-in-10-years loss is estimated at EUR 54.7 million. Many flood events occurred during this period: e.g., the Rhône overflow in 2003, the Argens floods in 2010 and 2011, and the Gard in 2002. But in terms of damages, extreme flood events on the most exposed territories have not yet occurred. All historical losses for the Macif include professional risks. These annual losses must thus not be compared with the results of the probabilistic approach.

The results of the simulation of the Macif portfolio for the combination of F1 and F2 are described in Table 5. The probabilistic flood modeling results (Table 5) are higher than the historical flood losses, especially as the model was applied strictly to individual risks.

The probabilistic average annual loss is strongly influenced by the extreme years in the distribution tail. For example, the 1000-year return period event has a simulated loss of EUR 583.3 million for the individual risks in the Macif portfolio, which is very high.

The extreme years are constituted by the occurrence of major floods in several major catchments with high levels of exposed values. Indeed, the F1 model simulates major river overflow in the most important cities: e.g., the Seine in Paris, the Rhône and Saône in Lyon, the Garonne in Toulouse and Bordeaux. The F2 model simulates high losses for territories strongly exposed to heavy rainfall and surface water runoff, especially in the urbanized areas. The “cevenol type” events in the southeast of France are typically concerned with this approach, which is adapted to small catchments and heavy rainfall in short timescales.

This result is a first application of our probabilistic method to estimate the financial exposure for an insurance portfolio. According to these first results, a single analysis of historical events may underestimate the risk, with a short period of records. A long sequence of fictive years has to be produced to take into account the combination of major events in several catchments during the same year.

**Table 2.** Composition of the Macif insurance portfolio and comparison with the global insurance market portfolio, for individual risks, in terms of Nat Cat premium.

Portfolio	Macif		Market		Diff. (%)
Nature of individual risks	Nat Cat premium (EUR)	%	Nat Cat premium (EUR)	%	
Owners of individual houses	21 898 147	42.6	345 859 040	55	−12.4
Tenants of individual houses	3 387 674	6.6	44 738 035	7.2	−0.6
Owners of flats	10 579 481	20.6	64 471 531	10.3	10.3
Tenants of flats	11 153 206	21.7	120 340 637	19.28	2.42
Others	4 362 115	8.5	48 635 606	7.7	0.8
Total individual	51 380 626	100	624 044 849	100	

**Table 3.** Geocoding quality of the Macif individual risk portfolio 2011.

Geocoding results for individual risks	Total premium (EUR)	% of global portfolio
Address	31 737 933	61.8
Street center	12 682 551	24.7
Town center	6 951 406	13.5
Fail	8738	0

**Table 4.** Macif historical flood losses for the period 1996–2011.

Flood losses	Annual flood losses (EUR 2011)
Mean	EUR 21.4 million
10 years return period	EUR 54.7 million
Maximum	EUR 61.5 million

## 4.2 Exceedance probability maps

### 4.2.1 The Argens downstream floodplain

The map on Fig. 2 represents the downstream Argens watershed, near the cities of Roquebrune-sur-Argens, Saint-Raphaël and Fréjus. This region has been chosen for its strong exposure to flash flood events: two major events occurred in recent years. In June 2010 and in November 2011, the flash floods of the Argens River and its tributaries respectively account for EUR 550 and 250 million in insurance losses (CCR estimates in 2013). On the map in Fig. 2, the official flood-prone area (the highest known floods) is overlaid with the exceedance probability maps generated by the probabilistic flood model. The flood areas on the map are the Argens and tributaries floodplains (Reyran, Grande Garonne).

In the aggregate, the probabilistic river overflow fits with the official flood-prone areas, especially for short return periods (< 50 years). Nevertheless, the modeled flood zone covers a larger territory. On paper, the modeled flood zones could

**Table 5.** Description of the F1 and F2 model results: per event and annual loss distributions (EUR million) for individual risks for the Macif portfolio.

Losses	Per event		Annual
Distribution parameters	F1 model	F2 model	Combined F1&F2 models
Mean	7.4	0.5	29.9
90th percentile	19.1	1.3	60.8
99.9th percentile	258.5	34.7	583.3

cover a larger expanse than the highest known floods. But, between two high return periods, e.g., 100 and 150 years, there are only small differences between the levels of water (centimeter scale). The modeled flood zone should remain close to the highest known flood zone. Important differences are probably due to modeling uncertainties. The differences can be explained by uncertainties in the generated water level or in the elevation model (1 m elevation resolution).

The map of Fig. 2 reveals the important geographical exposure to the surface water runoff, for territories outside the Argens floodplain. For example, Saint-Raphaël, Boulouris-sur-mer, Saint-Aigulf, Les Issambres and the northeast of Fréjus.

To validate the model results, the historical flood claims, geocoded at the “street number” precision, have been overlaid with the official flood-prone areas (Fig. 3) and the probabilistic flood map (Fig. 4).

Figure 3 shows that many claims are located outside the floodplains. A large majority of flood claims in this region is concentrated in the urban centers, outside the flood zone. These claims account for 46 % of the total losses for this region. They represent 87 % by loss count. The first conclusion is that the most expensive claims are located in the floodplains, since 13 % of the data count for more than half the total losses. The river overflow inside the floodplains, in this region, explains only the half of global flood losses. Our assumption is that surface runoff in the urbanized areas



**Table 6.** Deterministic model results in a selection of historical flood event sets. These model results are compared with the CCR estimations for the event losses (EUR million). These estimations are based on the extrapolation of insurance claims.

Losses	Simulated		Historical
	Inf.	Sup.	Best Estimate
Historical events			
Southeast September 2009	48.4	66.4	53.6
Center-east November 2008	94.3	171.3	125.6
Rhône Alps storms September 2008	45.3	85.0	41.8
Southwest storms May 2007	4.1	11.4	19.4
Meurthe and Moselle October 2010	25.6	59.7	51.5
Gard September 2005	120.1	277.6	83.8
Arles and Rhône December 2003	422.4	964.9	834.4
Gard September 2002	297.2	600.5	609.1
Var June 2010	138.5	610.1	462.7
Var June 2011	387.8	448.3	266.6

**Table 7.** Loss distribution for a historical event (southeast in September 2009) simulated with the deterministic model for the insurance market (in EUR million). The error distribution is calibrated on a selection of the 1995–2010 event set by comparing modeled losses with real claims.

Modeled loss distribution	10th perc.	30th perc.	70th perc.	90th perc.
Historical event Southeast September 2009	40.6	48.4	66.4	93.6

has generated these claims located outside the Argens major riverbed. Indeed, in the Nat Cat scheme, the insurance expert has to verify if the claims are due to water rising from the ground (considered as natural disaster) or coming down from the roof (not considered as a natural disaster). We assume here that all flood claims are natural disaster claims. The results of the probabilistic model seem to confirm this theory.

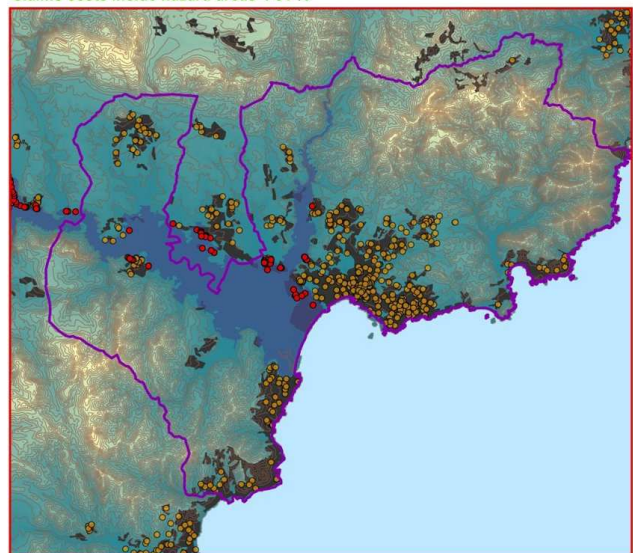
This confirms the importance of modeling the surface water runoff in a comprehensive flood model. On the map of Fig. 4, when combining river overflow with surface water runoff, a significantly higher ratio of claims (81 %) and losses (96 %) are located inside the flood areas. The remaining claims located outside the modeled flood areas can be explained by uncertainties in DEM data coupled with uncertainties in the hazard model, which are a subject for more investigations. For example, the model does not take into account the sewer network and the risk of sewer overflow inside the urban areas. Furthermore, the model does not take into account the street network and its effect on the water velocity. The use of a more precise DEM for elevation resolution ( $< 1$  m) and spatial resolution ( $< 50$  m) should also enhance the simulation results. In the short term, a 5m DEM will be used in the urbanized areas for rainfall–runoff and flood propagation computing. Some claims are also generated by

### Claims due to flooding || Highest known floods

Côte d'Azur : Fréjus / Saint-Raphaël / Roquebrune

Rate of claims inside hazard areas : 13 %

Claims costs inside hazard areas : 54 %

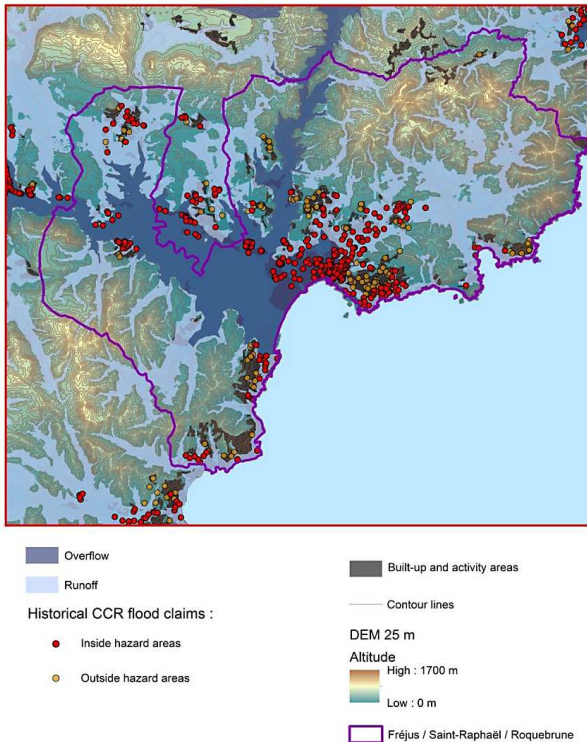
**Figure 3.** Official flood-prone areas (highest known water level) map overlayed with street number precision geocoded market claim data. Focus on the downstream floodplain of the Argens River near Fréjus, Saint-Raphaël and Roquebrune.

### Claims due to flooding || CCR Probabilistic Hazard

Côte d'Azur : Fréjus / Saint-Raphaël / Roquebrune

Rate of claims inside hazard areas : 81 %

Claims costs inside hazard areas : 96 %



**Figure 4.** Exceedance probability map overlaid with street number precision geocoded market claim data. Focus on the downstream floodplain of the Argens River near Fréjus, Saint-Raphaël and Roquebrune.

other perils, such as sea surge, and cannot be explained by river overflow and surface water runoff.

#### 4.2.2 Metropolitan France probabilistic flood map

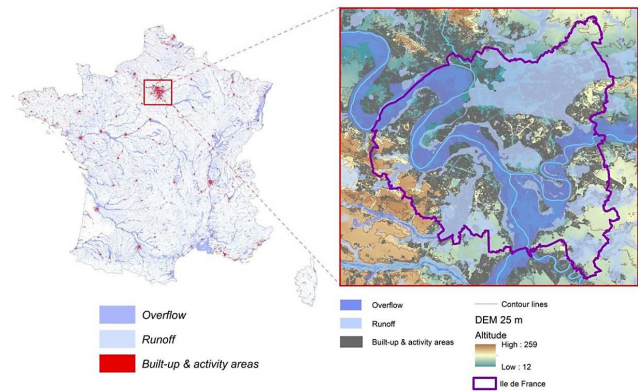
Figure 5 presents the results of the exceedance probability mapping for the whole French metropolitan territory with a focus on Île-de-France. At the French territory scale, 74 % of the flood claims are located inside the modeled areas, whereas only 29 % are inside the official flood-prone areas (Atlas des Zones Inondables). Among the claims located inside the modeled areas, 29 % are located inside the floodplains and 45 % in the surface runoff hazard areas.

Counting the flood claims inside the flood areas is necessary but not sufficient for the validation of the flood model. If an urbanized area, located inside the flood zones, has never been concerned by any flood claim, either the possible flood has not yet occurred, or the flood zone is a model error.

The ratio of the number of claims to the number of insurance policies inside and outside the flood areas has been

Probabilistic floods in France

Probabilistic floods in Ile de France

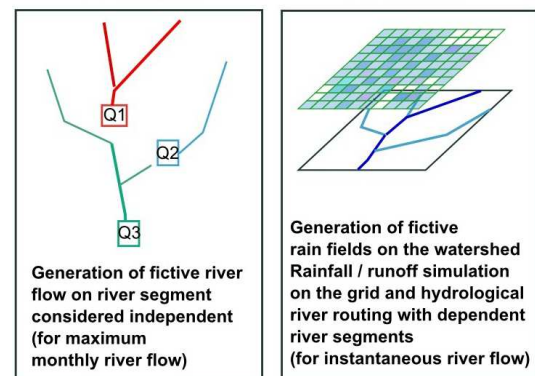


**Figure 5.** Exceedance probability map for the French territory combining the river overflow and the surface water runoff. This map is overlaid with the build-up and activity areas. A focus on Île-de-France overlaid with a 25 m DEM is displayed.

### Simulation of fictive flood events

F1 model

F2 model



**Figure 6.** Hydrological method to simulate river flow. Comparison between F1 and F2 model on a fictive small river catchment.

computed. This ratio is the so-called “claim frequency”. It is calculated at the level of the whole French metropolitan territory. The model is validated if the claim frequency is optimized inside the flood zones and minimized outside. Inside the flood zone, we expect the modeled claim frequency to be close to the historical claim frequency, which is usually between 1 and 10 % depending on the hazard intensity. Outside the flood zone, the claim frequency must be close to zero.

To validate the modeled flood zones, we need to verify if the claim frequency is significantly higher inside the modeled areas than outside. For this purpose, we calculate a specific ratio, the validation criterion  $k$ , with the following method:

$$k = \frac{\frac{S}{P}_{in}}{\frac{S}{P}_{out}}, \quad (12)$$

with  $S$  the number of claim data and  $P$  the number of risks inside (in) and outside (out) the hazard areas. The validation



criteria  $k$  is thus a claim frequency ratio. We can calculate  $k$  for every hazard map type: official, modeled or real event maps.

The criterion  $k$  value for flood-prone areas (AZI) is estimated at 3.1. In the model results,  $k$  has a value of 3.2. Our model identifies more claims than the official flood-prone areas, but with a similar precision. We can conclude that mapping both river overflow and surface runoff is mandatory to identify a majority of claims. The increase of precision in our modeling approach is still an important issue for our works.

The probabilistic flood map shown on Fig. 5 is the first flood map in France at a national scale, with a homogeneous method combining two perils: surface water runoff and river overflow. The areas exposed to floods represent 52 374 km<sup>2</sup>, i.e., 9.8 % of the French metropolitan territory. River overflow areas and surface water runoff respectively account for 6.1 and 3.7 % of the metropolitan territory. The part of the territory exposed to < 50 years, < 100 and > 100 years return period flood is respectively 4.6, 5.4 and 9.8 %. The surface-water-runoff-exposed areas represent 30 to 40 % of the modeled flood areas, depending on the return period.

Due to the large-scale approach for the generation of our probabilistic map, it must be used for large-scale exposure studies. The best scale to use it on is the insurance portfolio. In the mid-term, the improvements of the methodology will allow us to improve the map and study the exposure at a smaller scale (commune for example).

## 5 Conclusion, limits and perspectives

The objective of this paper is to estimate the financial exposure to floods for the French insurance market. Two steps are identified to achieve this goal: first, to develop and validate a deterministic model, ARTEMIS, for the flood damage estimation on a single event. Second, to link this deterministic model to a probabilistic flood event set.

ARTEMIS combines a flood hazard and a damage estimation model. It aims to account for the two perils representing the major part of the flood losses in France: river overflow and surface water runoff due to heavy rainfall, particularly in the southern regions.

The modeling scale for the hazard and the damages has to be consistent (Apel et al., 2009). Given our objectives which concern the entire French territory, the hazard model is simplified. Thus, the damage model, even if based on an individual risk database, follows a large number statistical method. This approach is more consistent on individual risks than on professional risks, which are too heterogeneous.

This ARTEMIS model combines many uncertainties: in the hazard model, in the vulnerability and in the damage model. The uncertainties are calculated at the end of the whole process: error between real loss value and model estimates to determine the deterministic model precision.

For the hazard model part, major sources of uncertainties are due to the simplified river routing model for the estimation of river discharge and DEM precision for the estimation of the flooded area. The enhancement of this model part is one of the main perspective for the short term. The use of a variable velocity model could significantly improve the river discharge calculation. More simply, the calibration of the stream flow celerity on every river from the hydrographic network will have a significant impact on the results (the celerity is calibrated at the event scale).

The lack of knowledge on river characteristics (precise width, flow gauges network) and flood protections will generate many modeling errors. But the large-scale modeling does not allow, by definition, a high precision for topographic and hydrographic description. Nevertheless, the hazard model is validated by comparison with the historical river flow (Banque Hydro) or with available satellite images of the flooded area, for the historical event set. Whilst not shown in this paper, these results constitute the main calibration of the deterministic hydrological modeling.

The uncertainties in the vulnerability and damage models come from the lack of precision in the address location of the insurance policies and risks, the estimation of the insured values for the different types of risks (individual, commercial, agricultural) can be an important source of uncertainties if not present in the insurance database. Furthermore, in the claim database, the lack of major industrial claim data makes it difficult to calibrate a damage model for these risks, and thus the statistical approach on the industrial risks reaches its limits.

The probabilistic flood event set combines a river flow database with a rain field database to simulate the flood events. The results of this probabilistic flood model are studied based on the Macif portfolio, a French insurance company and specifically on the individual risk types. The average and extreme annual flood losses are calculated. High differences appear between model estimation and historical observations. Indeed, the flood event set takes into account extreme events and the distribution tail will have significant impact on the mean annual loss. These results are the first application of our modeling system. Due to many uncertainties explained above, an significant amount of work is still in progress to strengthen results. We can list the main perspectives which are in the mid-term (1 or 2 years):

- the enhancement of the river routing model: using the kinematic wave with variable velocity or calibrate the constant velocity for each stream
- the use of historical data to optimize the distribution on the historical water discharges for the probabilistic event set F1 (Hosking and Wallis, 1986; Payrastré et al., 2011)

- the simulation of fictive rain fields at a large scale to allows for rainfall–runoff simulations of major catchment floods.

In spite of these limitations and the prospects of improvement in the methodology, we present in this paper our multi-peril exceedance probability flood map for the entire French territory, combining river overflow and surface water runoff with a homogeneous approach. This map is not available at small scale since it is only designed for large-scale studies. The flood exposure map is a CCR internal tool for exposure quantification and should not be used as an official flood map since it is an intermediary result of our model.

**Acknowledgements.** The authors would like to thank Jean-Michel Soubeyroux for his precious advice on the use of meteorological data. We also like to thank Laurent Montador and Pierre Michel for supporting this CCR project since 2010.

We would like to thank our two anonymous reviewers for their precious advice and Maria-Carmen Llasat, our editor.

Edited by: M.-C. Llasat

Reviewed by: two anonymous referees

## References

- Aon Benfield (Ed.): Global Catastrophe Recap, Impact Forecasting, 2013.
- Apel, H., Aronica, T., Kreibich, H., and Thielen, A. H.: Flood risk analyses - how detailed do we need to be?, *Nat. Hazards*, 49, 79–98, 2009.
- Arnaud, P. and Lavabre, J.: Estimation de l'aléa pluvial en France métropolitaine, Guide pratique, Editions QUAE, 2010.
- Banque Hydro, Banque nationale de données pour l'hydrométrie et l'hydrologie, Ministère de l'Environnement et du Développement Durable, Paris, available at: <http://www.hydro.eaufrance.fr/accueil.html>, 2006.
- Bates, B. C., Kundzewicz, Z. W., Wu, S., and Palutikof, J. P. (Eds.): Climate change and water, Technical paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp., 2008.
- Bouilloud, L., Chancibault, K., Vincendon, B., Ducrocq, V., Habets, F., Saulnier, G.-M., Anquetin, S., Martin, É., and Noilhan, J.: Coupling the ISBA land surface model and the TOPMODEL hydrological model for Mediterranean flash-flood forecasting : Description, calibration and validation, *J. Hydrometeorol.*, 11, 315–333, 2009.
- Bouvier, C. and DelClaux, F.: ATHYS: a hydrological environment for spatial modelling and coupling with GIS. IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 235, 19–28, 1996.
- Bradbrook, K., Waller, S., and Morris, D.: National Floodplain Mapping: Datasets and Methods – 160,000 km in 12 months, *Nat. Hazards*, 36, 103–123, 2005.
- de Moel, H., van Alphen, J., and Aerts, J. C. J. H.: Flood maps in Europe – methods, availability and use, *Nat. Hazards Earth Syst. Sci.*, 9, 289–301, doi:10.5194/nhess-9-289-2009, 2009.
- Dráb, A. and Říha, J.: An approach to the implementation of European Directive 2007/60/EC on flood risk management in the Czech Republic, *Nat. Hazards Earth Syst. Sci.*, 10, 1977–1987, doi:10.5194/nhess-10-1977-2010, 2010.
- Durand, Y., Brun, E., Merindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E.: A meteorological estimation of relevant parameters for snow schemes used with atmospheric models, *Ann. Glaciol.*, 18, 65–71, 1993.
- Edijatno, M. C., Nascimento, N. O., Yang, X., Makhoulouf, Z., and Michel, C.: GR3J: a daily watershed model with three free parameters, *Hydrol. Sci. J.*, 44, 263–277, 1999.
- Estupina-Borrel, V.: Vers une modélisation hydrologique appliquée à la prévision opérationnelle des crues éclair. Application à de petits bassins versants du sud de la France, Thèse de doctorat de l'Université Paul Sabatier Toulouse, 3, 254 pp., 2009.
- EXCIMAP (European Exchange Circle on Flood Mapping): Handbook on good practice on flood mapping in Europe, available at: <http://floods.jrc.ec.europa.eu/flood-risk/excimap-handbook-on-flood-mapping>, 2007.
- Falter, D., Vorogushyn, S., Lhomme, J., Appel, H., Gouldby, B., and Merz, B.: Hydraulic model evaluation for large-scale floor risk management, *Hydrol. Process.*, 27, 1331–1340, 2013.
- Gomez, E.: Modélisation intégrée du transfert de nitrates à l'échelle régionale dans un système hydrologique. Application au bassin de la Seine, Thèse de doctorat de l'École Nationale Supérieure des Mines de Paris, 2009.
- Green, W. H. and Ampt, G.: Studies of soil physics part I – the flow of air and water through soils, *J. Agr. Sci.*, 4, 1–24, 1911.
- Habets, F., Gascoin, S., Korkmaz, S., Thiéry, D., Zribi, M., Amraoui, N., Carli, M., Ducharme, A., Leblois, E., Ledoux, E., Martin, E., Noilhan, J., Ottilé, C., and Vienne, P.: Multi-model comparison of a major flood in the groundwater-fed basin of the Somme River (France), *Hydrol. Earth Syst. Sci.*, 14, 99–117, doi:10.5194/hess-14-99-2010, 2010.
- Hosking, J. R. M. and Wallis, J. R.: The Value of Historical Data in Flood Frequency Analysis, *Water Resour. Res.*, 22, 1606–1612, doi:10.1029/WR022i011p01606, 1986.
- Krige, D. G.: A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of Chemical, Metallurg. Mining Soc. South Afr.*, 52, 119–139, 1951.
- Kohonen, T.: Self-organizing maps, Springer Series in Information Sciences, Vol. 30, Springer Verlag, Springer, Berlin, Heidelberg, New York, 3rd Edn., 1995.
- Leblois, E. and Creutin, J.-D.: Space-time simulation of intermittent rainfall with prescribed advection field: adaptation of the turning band method, *Water Resour. Res.*, 49, 3375–3387, 2013.
- Legifrance.gouv.fr: (Ed), Loi no 82–600 du 13 juillet 1982 relative à l'indemnisation des victimes de catastrophes naturelles, 1982.
- Lloyd's market association: Catastrophe modelling, Guidance for non-catastrophe modellers, download available at: <http://www.lmalloyds.com/CMDownload.aspx?ContentKey=5f8ba7c4-40f9-4e94-b212-f039f24fe761&ContentItemKey=894a1f1e-fff7-4940-9f61-441b309e1a73>, June 2013.
- Manning, R.: On the flow of water in open channels and pipes, *Transactions, Inst. Civ. Engrs, Ireland*, 20, 161–207, 1891.
- Matheron, G.: The intrinsic random functions and their applications, *Adv. Appl. Probability*, 5, 439–468, 1973.

- Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article “Assessment of economic flood damage”, *Nat. Hazards Earth Syst. Sci.*, 10, 1697–1724, doi:10.5194/nhess-10-1697-2010, 2010.
- Messner, F., Penning-Rowsell, E., Green, C., Meyer, V., Tunstall, S., and van der Veen, A.: Evaluating flood damages: guidance and recommendations on principles and methods, FLOODsite-Report T09-06-01, 2007.
- Moncoulon, D.: Proposition d’une méthode d’estimation de l’exposition financière aux inondations pour le marché de l’assurance en France. Modélisation hydrologique et économique probabiliste spatialisée, Thèse de Doctorat de l’Université Paul Sabatier Toulouse 3, 4 June, 2014.
- Moncoulon, D. and Quantin, A.: Modélisation des événements extrêmes d’inondation en France métropolitaine, *La Houille Blanche*, *Revue Internationale de l’Eau*, I, 2013.
- Nelsen, R. B.: An introduction to Copulas, *Lectures Notes in Statistics*, Vol. 139, Springer Verlag, New York, 1999.
- Payrastre, O., Gaume, E., and Andrieu, H.: Usefulness of historical information for flood frequency analyses: Developments based on a case study: HISTORICAL INFORMATION FOR FLOOD FREQUENCY ANALYSES, *Water Resour. Res.*, 47, W08511, doi:10.1029/2010WR009812, 2011.
- Poulard, C. and Leblois, E.: Towards objective design of dry dams at watershed scale: how to take into account the spatial structure of the rainfall and its variability. *Proceedings: hydrological extremes in small basins: 12th Biennial International Conference of the Euromediterranean Network of Experimental and Representative Basins (ERB)*, Krakow, Poland, 18–20 September 2008, available at <http://unesdoc.unesco.org/images/0018/001828/182846e.pdf>, 2008.
- Quantin, A.: Modélisation du péril inondation: passage d’une approche déterministe à une approche probabiliste en vue d’une tarification à l’exposition des couvertures de réassurance dans le cadre du régime des Catastrophes Naturelles. *Rapport confidentiel. Mémoire d’actuaire*, Centre d’Etudes Actuarielles, p. 164, 2011.
- Sklar, A.: Fonctions de répartition à  $n$  dimensions et leurs marges, *Publications de l’Institut de Statistique de Paris*, 8, 229–231, 1959.
- Swiss Re: Flood – an underestimated risk. *Inspect, inform, insure*, Swiss Re Productions, Zurich, 2012.
- Swiss Re: SIGMA, natural catastrophes and man-made disasters in 2013: large losses from floods and hail; Haiyan hits the Philippines, 1, 2014.