



HAL
open science

Feedback from Uncertainties Propagation Research Projects Conducted in Different Hydraulic Fields: Outcomes for Engineering Projects and Nuclear Safety Assessment

Vito Bacchi, Claire-Marie Duluc, Lise Bardet, Nathalie Bertrand, Vincent
Rebour

► **To cite this version:**

Vito Bacchi, Claire-Marie Duluc, Lise Bardet, Nathalie Bertrand, Vincent Rebour. Feedback from Uncertainties Propagation Research Projects Conducted in Different Hydraulic Fields: Outcomes for Engineering Projects and Nuclear Safety Assessment. SIMHYDRO 2017 - Choosing The Right Model in Applied Hydraulics, Jun 2017, Nice, France. pp.221-241, 10.1007/978-981-10-7218-5_15. hal-04215770

HAL Id: hal-04215770

<https://hal.science/hal-04215770>

Submitted on 9 Oct 2023

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

FEEDBACK FROM UNCERTAINTIES PROPAGATION RESEARCH PROJECTS CONDUCTED IN DIFFERENT HYDRAULIC FIELDS; OUTCOMES FOR ENGINEERING PROJECTS AND NUCLEAR SAFETY ASSESSMENT

Vito Bacchi¹

Phone: 00 33 (0) 1 58 35 7193, Fax: 00 33 (0) 1 58 35 8130, E-Mail: vito.bacchi@irsn.fr

Claire-Marie Duluc, Lise Bardet, Nathalie Bertrand, Vincent Rebour

Mailing address

claire-marie.duluc@irsn.fr, lise.bardet@irsn.fr, nathalie.bertrand@irsn.fr, vincent.rebour@irsn.fr

KEY WORDS

Nuclear safety, flood risk, uncertainty quantification, global sensitivity analysis, Promethee Environment.

ABSTRACT

Hydraulic river models are applied for various purposes such as safety against flooding, navigation or ecological rehabilitation. Much effort has been put into the development of sophisticated numerical model systems. These numerical models are based on a deterministic approach and the results are presented in terms of measurable quantities (water depths, flow velocities, etc.). However, the modelling of river processes involves numerous uncertainties associated both to the numerical structure of the model, to the knowledge of the physical parameters which force the system and to the randomness inherent to the natural phenomena. As a consequence, dealing with uncertainties can be a difficult task for both practitioners [1] and new guidance [2].

In the context of nuclear safety, the Institute for Radioprotection and Nuclear Safety (IRSN) assesses studies conducted by the operators for different reference flood situations (local rain, small or large watershed flooding, sea levels, etc...), in agreement with the recommendation reported by the guide ASN n°13 [2]. The guide provides some recommendations to deal with uncertainties, by proposing a specific conservative approach to cover hydraulic modelling uncertainties. Especially, the most influencing parameter of the numerical model is identified and an unfavorable value is taken in order to cover a whole set of parameters. Depending of the situation, the influencing parameter might be the Strickler coefficients, levee behaviors, simplified topographic assumptions, etc. Obviously, identifying the most influencing parameter and giving it a penalizing value is challenging and usually questionable.

In this context, IRSN conducted cooperative (Compagnie Nationale du Rhone, I-CiTy laboratory of Polytech'Nice, Atomic Energy Commission) research activities since 2011 in order to investigate feasibility and benefits of Uncertainties Analysis (UA) and Global Sensitivity Analysis (GSA) when applied to hydraulic modelling. A specific methodology, presented in section 2, was tested by using the computational environment Promethee, which allows carrying out uncertainties propagation study. This methodology was applied with various numerical models and in different contexts (section 3), as river flooding on the Rhône River [3] and on the Garonne River (in the context of the Garonne river test case), for the studying of local rainfall [4] or for tsunami generation, in the framework of the ANR-research project TANDEM. The feedback issued from these previous studies is analyzed (technical problems, limitations, interesting results, etc...) in section 4 and the perspectives and a discussion on how a probabilistic approach of uncertainties should improve the actual deterministic methodology for risk assessment (also for other engineering applications) is finally given.

¹ Corresponding author

1. INTRODUCTION

Flooding is considered to be one of the most destructive among many natural disasters, bringing catastrophic and costly damages to human lives, infrastructure and the environment all over the world. As the most frequently occurring natural disaster impacting more people worldwide than any other natural disasters [5], understanding floods and assessing the risks associated to it have become more and more important nowadays [5-8]. With this aim, today, much effort has been put into the development of sophisticated numerical model systems. These numerical models are based on a deterministic approach and the results are presented in terms of measurable quantities (water depths, flow velocities, *etc.*). However, the modelling of river processes involves numerous uncertainties associated both to the numerical structure of the model, to the knowledge of the physical parameters which force the system and to the randomness inherent to the natural phenomena. As a consequence, dealing with uncertainties is a difficult but central task for both practitioners [9] and new guidance [2].

The French regulation requires that flooding hazard due to external causes must be taken into consideration in the demonstration of nuclear safety of basic nuclear installations (BNI). Therefore, the ASN guide n°13 [2] details the recommendations concerning the external flooding hazard, a method to quantify them and a list of recommendations for defining protective equipments are also provided. External flood refers to flooding coming from sources outside to the structures, areas or buildings of the BNI accommodating systems or components to be protected, whatever the cause(s) of that flooding (rainfall, river spates, storms, pipes failures, *etc.*). The water sources that could initiate or contribute to a flood affecting the study site are the rainfall (1), the groundwater (2), the sea and oceans (3), the watercourses (rivers and canals) (4), the natural (lakes, glaciers) (5) and the man-made (storage dams, tanks, water towers, pipes, *etc.*) reservoirs. The events or combinations of events that could cause a flood hazard for the installation, for each of the identified water sources, are also listed in the guide. A particular "event" is usually characterized by a physical quantity that defines its intensity (volume, height, flow rate, *etc.*) and when it's possible, by a probability or frequency of exceedance of that intensity, and a duration. The guide also defines as "Reference Flood Situation" (RFS) an event, or a combination of events, whose characteristics may be increased if necessary (unfavorable combination or margin to compensate for the limits of current knowledge). The eleven RFS to consider are introduced in the section 2.3.2 of the guide [2].

Today, uncertainties are taken into account in a deterministic and simplified way in the flooding guide. The aim of the recommendations concerning the consideration of combinations of events, initial states, or more broadly all the uncertainties, is to introduce conservative measures. For instance, in the case of the large watershed flooding, the RFS is characterized by a reference flow rate, a reference water level and the associated flood plain. The reference flow rate is the instantaneous peak flow rate associated with the thousand-year return period flood taking into consideration the upper bound limit of the 70% confidence interval increased by 15%. The guide pays attention to the definition of the influencing parameter, which is a parameter whose variations have a significant impact on the calculation results. Depending on the model, the influencing parameter can be the roughness coefficient (Strickler coefficient), the behavior of a dyke (which can resist or not during an event), the boundary conditions of the model and so on. To cover all uncertainties related to hydraulic modelling, the guide recommends performing a single simulation, with penalized values for one or two influencing parameters. Moreover, the number of sensitivity studies can be limited by identifying the most influencing parameter and taking the most unfavourable value in order to cover the uncertainties for a whole set of parameters.

However, identifying the most influencing parameter and giving it a penalizing value is challenging and usually questionable and it seems of interest to develop a rigorous and objective methodology to define this parameter with the associated uncertainty. To deal with uncertainties, computer experiment methodologies based upon statistical techniques have already been developed [10]. Especially, authors agree in the fact that a rigorous way for validating models and selecting the influencing parameters is to perform Uncertainty Analysis (UA) and Global Sensitivity Analysis (GSA) in deterministic models [9, 11, 12].

UA and GSA have started to be used [12, 13] and become broadly applied for a wide range of environmental modelling problems [14, 15] and hydraulic studies [3, 4, 16]. UA consists in the propagation of uncertainty sources through model, and focuses on the quantification of uncertainties in model outputs to check their robustness [17]. GSA methods are invaluable tools as they allow studying how the uncertainty in the output

of a model can be apportioned to different sources of uncertainty in the model inputs [11, 12]. Moreover, the method can be used to determine the most contributing input variables to an output behavior as the non-influential inputs, or ascertain some interaction effects within the model. Objectives of GSA are numerous; *i*) model verification, *ii*) understanding, *iii*) model simplification and *iv*) factor prioritization. Finally, GSA is an aid in the validation of a computer code, guidance research efforts, or to justify system design safety for example. Both UA and SA are essential to analyze complex systems [17, 18], as study of uncertainties related to input parameters is of prime interest for applied practitioners willing to decrease uncertainties in their models results [1].

In this context, IRSN conducted cooperative (Compagnie Nationale du Rhone, I-CiTy laboratory of Polytech'Nice, Atomic Energy Commission) research activities since 2011 in order to investigate feasibility and benefits of UA and GSA when applied to hydraulic modelling. IRSN has developed this kind of approach for nuclear and criticality issues but it can as well be applied for every kind of model, as soon as numerous simulations can be performed. A specific methodology, presented in section 2, was tested by using the computational environment Promethee, which allows carrying out uncertainties propagation study. This methodology was applied with various numerical models and in different contexts (section 3), a flood case on the Rhône River [3], on the Var River in an urban area [4] and finally on tsunami (wave) propagation along the French coasts (in the framework of the ANR-research project TANDEM). The feedback issued from these previous studies is analyzed (technical problems, limitations, interesting results, etc...) in section 4 and perspectives and discussion on how a probabilistic approach of uncertainties should improve the actual deterministic methodology for risk assessment (also for other engineering applications) is finally given.

2. METHODOLOGY

2.1 Global methodology for uncertainty studies and numerical tool employed by IRSN

The proposed methodology is based on the classical uncertainty study steps [11, 12, 19, 20], as represented in Figure 1. The first step (Step A) of an uncertainty study can be roughly described as "the definition of the problem". This may seem obvious, but starting an uncertainty study requires an analysis of some key issues and ensuring that physical problem we aim to study has been correctly translated in mathematical terms. This step consists in the definition of the input variables of the model, the choice of the numerical model and of the output variables of interest (as, for instance, the water height at a given location).

Once step A has been carried out, the next step (Step B) consists in the identification of the inputs affected by uncertainty, and consequently the definition of the variation intervals and the probability density functions (PDF) for each uncertainty parameters. We remind that, for a random variable X defined on an $[a, b]$ interval, a PDF $f_X(X)$ is defined as follows:

$$P(X \in [a, b]) = \int_a^b f_X(X) dx \quad (1)$$

The stochastic inputs can be operating conditions, geometry, as well as empirical parameters which define the physical models. Most of the time, the probability distributions of random inputs is unknown. Hence, expert's knowledge or the experimental evidence available is used as the basis for the PDF definition.

The next step (Step C) requires propagating the uncertainty defined in the previous step in the model. The objective is to compute the PDF $f_y(Y)$ of the quantities of interest y . This step is most complex and computationally demanding. There are various methods to deal with this step, and some of them will be presented in the next paragraph.

Finally, in a probabilistic framework, a better understanding of uncertainties can be achieved by analyzing the contribution of the different uncertainty sources to the uncertainty of the variables of interest (Step D). For each couple "criteria of the study / propagation method used in Step C", post-treatment procedures are used in order to rank the uncertainty sources. It is important to note that an uncertainty study rarely stops after a first processing of steps A, B, C and D, and the last step then plays a crucial role. Indeed, the ranking results highlight the variables that truly determine the relevancy of the final results of the study. If the uncertainty model of some of these variables has been chosen a bit roughly in Step B *e.g.* because of time

constraints or any practical difficulties, collecting further information on these meaningful sources can be a relevant upgrade to refine the analysis.

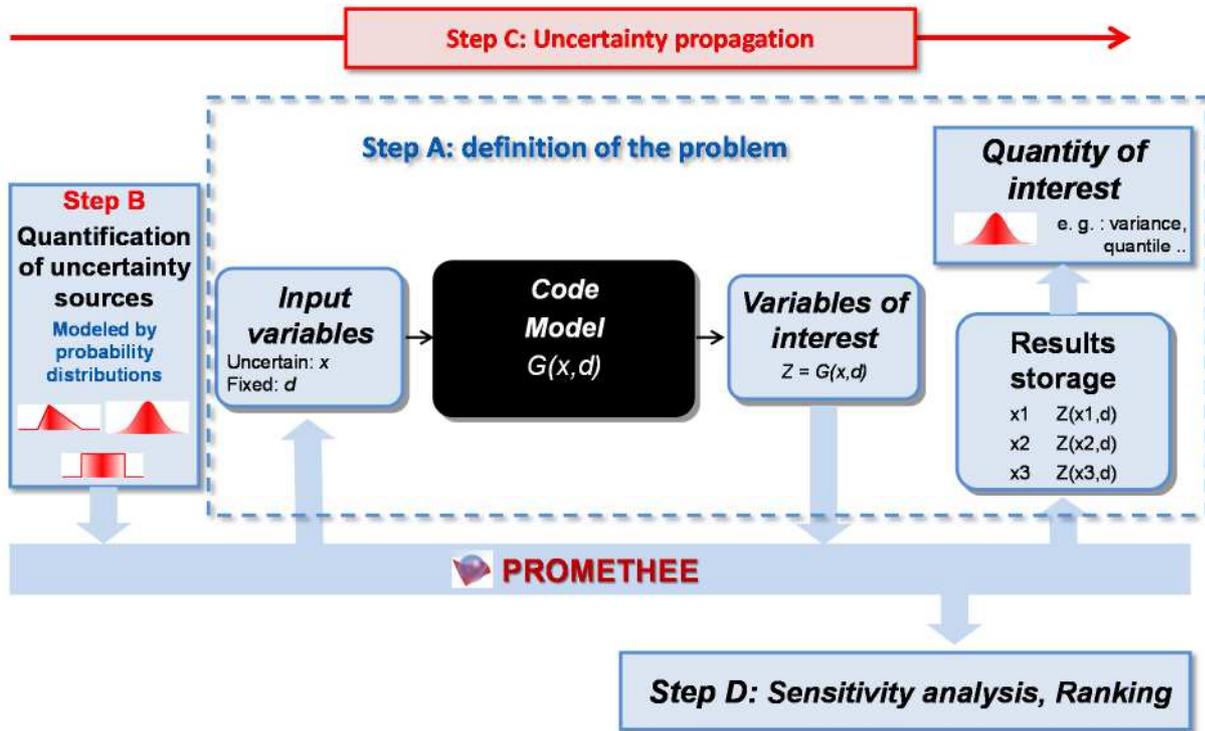


Figure 1 : Uncertainty and Sensitivity analysis steps driven by Promethee Environment.

In the examples reported in the next section, Step A and Step B are strictly related to the objectives of the study (e.g. Tsunami Hazard Assessment) and will be presented for each case. However, for Step C and D we choose to explore a methodology essentially based on the coupling between the IRSN Promethee-Environment and a hydrodynamic model adapted to the objectives of the study. Especially, Promethee is an environment for parametric computation that allows carrying out uncertainties propagation study, when coupled (or warped) to a code. This software is freely distributed by IRSN (<http://promethee.irsn.org/doku.php>). Promethee allows the parameterization with any numerical code and is optimized for intensive computing. Moreover, statistical post-treatment, such as UA and GSA can be performed using Promethee as it integrates R statistical computing environment [21]. The methods employed for Step C and D are described in the next paragraphs.

2.2 The Monte-Carlo method for uncertainty Propagation (Step C)

The Monte-Carlo method requires random generation of input variables from their probability distributions. The resulted sampling of a given size N is a $N \times V$ matrix, V being the number of uncertain parameters. Each row of the matrix $x^i = (x_1, \dots, x_V)^T$ represents a possible configuration of the input parameters of the model. Corresponding realizations of the output are generated by successive deterministic simulations with each configuration of the inputs. Statistical estimators of the response $Y = (Y_1, \dots, Y_N) = (G(x^i))_{i \in \{1, \dots, N\}}$ can therefore be computed from the output as follows:

$$E[Y] = \mu_Y = \frac{1}{N} \sum_{i=1}^N G(x^i) \quad (2)$$

$$\text{Var}(Y) = \frac{1}{N-1} \sum_{i=1}^N [G(x^i) - \mu_Y]^2 \quad (3)$$

$$\sigma_Y = \sqrt{\text{Var}(Y)} \quad (4)$$

Where $E[Y] = \mu_Y$, $\text{Var}(Y)$ and σ_Y are, respectively, the mean, the variance and the standard deviation of the response Y given by the model G . These statistical moments are useful for both the uncertainty propagation and the uncertainty sensitivity analysis (see next paragraph).

The convergence order of the Monte-Carlo sampling method is given by the Central Limit Theorem [11] as

$O\left(\frac{1}{\sqrt{N}}\right)$. As a consequence, a large amount of simulations are necessary to obtain convergent statistics. The

number of simulations increase with the increasing of the uncertainty parameters. As a consequence, other techniques have been developed to reduce the number of simulations by preserving the quality of the statistical results (*i.e.* quasi Monte-Carlo screening methods, meta-models methods, and so on).

2.3 Sensitivity Analysis (Step D)

The role of the sensitivity analysis is to determine the strength of the relation between a given uncertain input parameters and the model outputs [20]. In this study, we focused our attention on Global Sensitivity Analysis approaches which rely on sampling based methods for uncertainty propagation, willing to fully map the space of possible model predictions from the various model uncertain input parameters and then, allow to rank the significance of the input parameter uncertainty contribution to the model output variability [22]. GSA approaches are well suited to be applied with models having nonlinear behaviour and when interaction among parameters occurs [23]. These approaches going through an intensive sampling are computationally demanding, as they most often rely on Monte-Carlo (MC) approach, even though some more parsimonious sampling method such as Latin hypercube or pseudo-Monte Carlo are sometimes applied (see Helton et al. [18] for a review). The GSA approaches we focused on for our explorative studies are:

- Screening methods, such as Morris method [24];
- Sobol indices computation, that considers the output hyperspace (x) as a function ($Y(x)$) and performs a functional decomposition ([1, 19] or a Fourier decomposition (FAST method) of the variance.

2.3.1 The Morris Screening Method

Screening methods are based on a discretization of the inputs in levels, allowing a rapid exploration of the code behavior. These methods are adapted to a large number of inputs; practice has often shown that only a small number of inputs are influential [19]. The aim of this type of method is to identify the non-influential inputs with a small number of model calls while making realistic hypotheses on the model complexity. The model is therefore simplified before using other sensitivity analysis methods, more complex but more costly.

The most engineering-used screening method is based on the so-called “One At a Time” (OAT) design, where each input is varied while fixing the others. Our choice has been to use the Morris method [24], which is the most complete and the most costly one [19]. The method of Morris, as fully detailed in [19], allows classifying uncertain input parameters in three categories: (*i*) those that have negligible effect, (*ii*) those that have linear effect and (*iii*) those that have nonlinear effects or effects in interaction with other input parameters. The method consists in discretizing the input space for each variable, then performing a given number of OAT design. Such designs of experiments are randomly chosen in the input space, and the variation direction is also random. The repetition of these steps allows the estimation of elementary effects for each input. From these effects, sensitivity indices are derived. Let us denote r the number of OAT designs (Saltelli et al. [20] propose to set parameter r between 4 and 10). If one discretizes the input space in a d -dimensional grid with n levels by input. Let us denote the elementary effect of the j -th variable obtained at the i -th repetition, defined as:

$$E_j^{(i)} = \frac{f(X^{(i)} + \Delta e_j) - f(X^{(i)})}{\Delta} \quad (5)$$

where Δ is a pre-determined multiple of $\frac{1}{(n-1)}$ and a e_j vector of the canonical base. Indices are obtained as follows:

$$\mu_j^* = \frac{1}{r} \sum_{i=1}^r |E_j^{(i)}| \quad (\text{mean of the absolute value of the elementary effects}) \quad (6)$$

$$\sigma_j = \sqrt{\frac{1}{r} \sum_{i=1}^r \left(E_j^{(i)} - \frac{1}{r} \sum_{i=1}^r E_j^{(i)} \right)^2} \quad (\text{standard deviation of the elementary effects}) \quad (7)$$

so, μ_j^* is a measure of influence of the j -th input on the output. The larger μ_j^* is, the more the j -th input contributes to the dispersion of the output. σ_j is a measure of non-linear and/or interaction effects of the j -th input. If σ_j is small, elementary effects have low variations on the support of the input. Thus, the effect of a perturbation is the same all along the support, suggesting a linear relationship between the studied input and the output. On the other hand, the larger σ_j is, the less likely the linearity hypothesis is. Thus a variable with large σ_j will be considered having non-linear effects, or being implied in an interaction with at least one other variable. These methods perform well to discriminate influencing parameters on output variability with 2D flood modelling studies [25].

2.3.1 Functional decomposition of the variance for the computation of the Sobol's indices: FAST-method [26, 27]

With respect to the Morris method [24], GSA approaches relying on Sobol index computation go one step further, allowing to quantify the contribution to the output variance of the main effect of each input parameters [23, 27-29]. The definition of Sobol Indices is a result of the ANOVA (ANalysis Of VAriance) variance decomposition. To be able to write the variance decomposition, some mathematical hypothesis are required (see i.e. Saltelli et al. [12] or Faivre et al. [11]). Especially, let us consider an integrable model G on the domain $A = A_1 \times \dots \times A_v$. It exists an unique decomposition of the model:

$$G(Y) = f_0 + f_1(Y_1) + \dots + f_m(Y_m) + f_{1,2}(Y_1, Y_2) + \dots + f_{v-1}(Y_{v-1}, Y_v) + \dots + f_{1,\dots,v}(Y_1, \dots, Y_v) \quad (8)$$

Where all the functions f_i are mutually orthogonal, implying that $\langle f_i | f_j \rangle = 0$ if $i \neq j$, with the scalar product defined as follow:

$$\langle f | g \rangle = \int_A f(x)g(x)\pi(x)dX \quad (9)$$

Where π is a probability distribution defined on A . Under these assumptions, given a set of V independent uncertain parameters X , the variance of a response $Y = G(X)$ can be calculated, using the total variance theorem, as follows :

$$Var(Y) = \sum_{i=1}^V D_i(Y) + \sum_{i<j}^V D_{ij}(Y) + \dots + D_{12..V}(Y) \quad (10)$$

Where $D_i(Y) = Var[E(Y|X_i)]$ and $D_{ij}(Y) = Var[E(Y|X_i X_j)] - D_i(Y) - D_j(Y)$ and so on for higher order interactions. $E(Y|X_i)$ is the Y conditional expectation with the condition that X_i remains constant. The so-called ‘‘Sobol indices’’ are obtained as follows:

$$S_i = \frac{D_i(Y)}{\text{Var}(Y)}, \quad S_{ij} = \frac{D_{ij}(Y)}{\text{Var}(Y)}, \quad \dots \quad (11)$$

These indices express the share of variance of Y that is due to a given input or input combination. For instance, S_1 is the first order sobol index. The number of indices grows in an exponential way with the number V of dimensions. So, for computational time and interpretation reasons, Homma and Saltelli [30] introduced the so-called “total indices” or “total effects” that write as follows:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots = \sum_{l \in \neq i} S_l \quad (12)$$

Where $\neq i$ are all the subsets of $\{1, \dots, V\}$ including i .

To estimate Sobol' indices, Monte Carlo sampling based methods have been developed: Sobol [29] for first order and interaction indices and Saltelli [26] for first order and total indices. For our applications, we used the so-called "extended-FAST" method [27] (or FAST99). This method allows the estimation of first order and total Sobol' indices for all the factors “ p ” at a total cost of “ $n \times p$ ” simulations, by improving the original FAST method (for Fourier Amplitude Sensitivity Test, from Cukier et al. method [31]), that not allows the computation of the higher-order indices. For the computation of the Sobol indices, we used the open source sensitivity-package R [32] or redeveloped R scripts to adapt the method to our specific cases. For more details about the methodology, see [26, 27, 29, 33].

3. EXAMPLES OF APPLICATION

3.1 Modelling the Rhône River with a 1D hydraulic model

3.1.1 Description and objectives of the study: Steps A and B

Between 2012 and 2014, IRSN conducted some cooperative researches activities with CNR (Compagnie Nationale du Rhone), in order to study the uncertainty associated to the numerical modelling of a large watershed flooding. The results we present in this section were already discussed in a previous paper [3] and we limit here to report a short synthesis of this work. This is the first UA and GSA analysis conducted by IRSN in a hydraulic context and its purpose was mainly methodological and explorative. Especially, the main objectives of this research were:

- to develop a methodology for UA and GSA studies in hydraulic fields;
- to test this methodology with a real test-case (a numerical model currently used for the management of some sections of a river);
- to have a feedback concerning the technical problems issued by this kind of study, the methodology employed and the results obtained in comparison with some deterministic simulation.

The 1D-hydrodynamic model Crue9 (developed by the CNR) was coupled with the open-source Promethee environment developed by IRSN. We focused our attention on the *Vieux Rhône Bourg-lès-Valence* section of the CNR numerical model of *Beauchastel* (France), delimited by the *Roche de Glun* dam (Bourg-lès-Valence, France), upstream, and the *Charme* dam (Charme-sur-Rhône, France) downstream. The numerical model was composed by 63 nodes, 327 profiles and 81 numerical branches.

For the purposes of the study, the model was decomposed in three homogenous morphological areas and UA and GSA were focused on the water height computed at one selected location of each area. Fifteen variables of interest were chosen, grouped in 9 global categories. The probability distributions of the variables of interest (Step A) were “designed” around their nominal values, corresponding to the values of the calibrated model (*i.e.* the Strickler coefficients or the weirs flow coefficients), to the geometrical data used for the construction of the model (*i.e.* the bottom topography or the levels of the weirs connecting the floodplain), to the statistical reference flood (in this case, the 1,000-years return period hydrogram), or to the validated numerical parameters and coefficients (*i.e.* the dam flow coefficient of the De Marchi law, the time step and the weighting coefficient of the Preissmann implicit numerical scheme).

Nearly 3,000 simulations were performed for the UA study, by using the Monte-Carlo method. Both the Morris-screening method and the FAST99 Method were used for GSA.

3.1.2 Results and feed-back

An illustration of the numerical results obtained is reported in Figure 2. From a methodological point of view, this study allowed to show how powerful was the coupling tool Promethee/Crue9. This coupling permitted to realize an important number of simulations by stochastically parameterizing the modelling files, which was not possible manually, but also allowing the synthesis and statistic treatments of the results. Simple UA allowed to calculate the variation ranges of model outputs, and to evaluate the criticality of the results (by examining parameters combinations that gave “critical” outputs). Besides hydraulics intuitions confirmation (upstream directed by downstream), the studies that were carried out showed on the one hand the non-linearity of the system response, and on the other hand, that the input parameters hypotheses are essential. GSA results showed a major contribution obtained of the input discharge “q” and of the Strickler coefficients “kmi”, as reported in Figure 2. Indeed, it is necessary to ensure a great expertise of the most influencing parameters, for variation ranges and statistical distributions, which determine the repartition (shape and amplitude) of calculated outputs. It is of interest to outline that this study permitted also to explore the sensitivity of the model to the numerical scheme implemented in Crue9. Moreover, the coupled tool is actually used by CNR in technical and research activities in the hydraulic context, as for example the study of levees behavior [34].

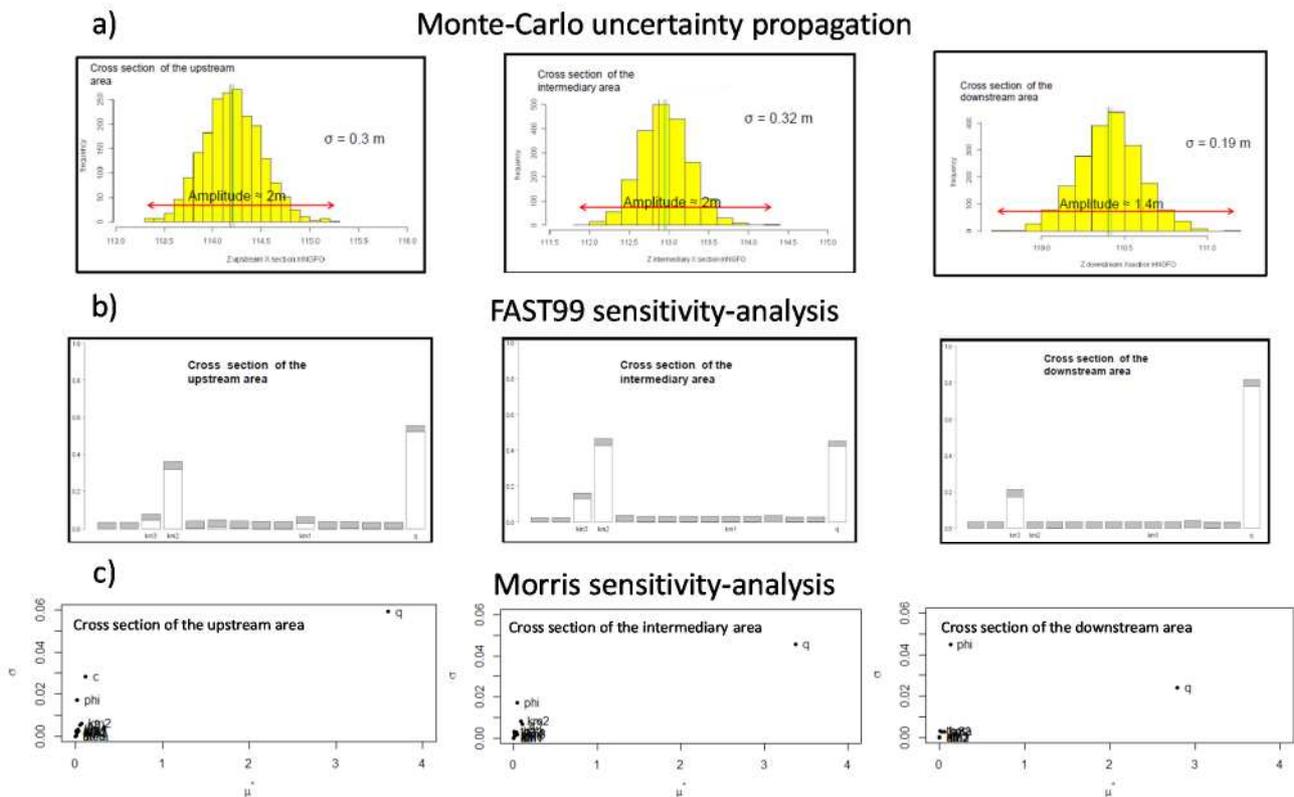


Figure 2 : a) Monte-Carlo Uncertainty Propagation (frequency distribution of the water height at three selected location); b) 1st order (in white) and total order (in grey) computed with FAST99 method; c) Morris results. Figures issued from Nguyen et al. [3].

3.2 Studying uncertainty related to topography with a 2D-model

3.2.1 Description and objectives of the case: Steps A and B

Between 2013 and 2016, IRSN conducted some cooperative researches activities with the I-CiTy laboratory of Polytech'Nice, in order to analyze uncertainties in 2D urban flood related to the use of High Resolution classified topographic data (HR) in an urban area. With this aim, the main objectives of the study were:

- to develop a numerical tool able to perform UA and GSA study with a 2D hydrodynamic model solving Saint-Venant Equations, in a context where spatial variability and interactions are likely to occur;
- to perform UA and GSA on two categories of uncertain parameters (measurement errors and uncertainties related to operator choices) relative to the use of HR in a 2D urban flood model having spatial inputs and outputs;
- to investigate the benefit of producing sensitivity maps based on Sobol index computation for results analysis in this context of spatial inputs and outputs.

Full details are reported in Abily [35] and Abily *et al.* [4, 36-38] and we limit here to report a synthesis. The study area is a 17.8 km² domain that represents the last five highly urbanized downstream kilometers of the low Var valley (Nice, France) where flooding events occurred in the last decades [39]. For the study, a 3-Dimensional classified data-set of topography data was created from photointerpretation procedure [4, 35]. This data-set was used for generating HR DEMs (Digital Elevation Models), one for each simulation. So, only the input DEM changes from one simulation to another and the hydraulic parameters of the model were set identically for the simulations. Three inputs parameters related to uncertainties when willing to use HR 3D classified data in 2D hydraulic models were identified: (i) the topographic input error (called var. E), (ii) the categorical order parameter (called var. S) having values representing the level of above ground features details impacting flow directions included in DEM (S1 is a Digital Terrain Model only, S2 is S1 combined with buildings elevation inclusion, S3 is S2 completed with walls, and S4 in S3 plus fine concrete structures, as sidewalks, road-curbs, *etc.*) and (iii) the choices made by the modeler (called Var. R) concerning the computational grid cells resolution in the model (in this study, five values of cells resolutions from 1 m to 5 m were used). So, a total of 2,000 DEMs were generated and used in the implementation of the GSA (1st order Sobol index were computed). UA was carried out using a MC approach to randomly sample in the database.

Promethee has been coupled with FullSWOF_2D (a 2D free surface modelling code) over a High Performance-Computing (HPC) structure [4, 38]. FullSWOF_2D (for Full Shallow Water equation for Overland Flow in 2 dimensions) is a code developed by the Orléans University as free software based on 2D SWE [40]. On the HPC structure (Interactive Computation Centre of Nice Sophia Antipolis University), up to 1,152 CPU are available and up to 30 simulations can be launched simultaneously using Promethee-FulSWOF_2D. A database of flood maps results has been produced using a total of 400,000 CPU hours. The required unitary computation time is two hours over 64 CPU, for simulations using the finest resolution grid size (1 m), which has 17.8 millions of computational points. At the coarsest resolution (5 m), the grids size is decreased to 712,000 computational points and using 64 CPU, the computational time decreases to few minutes.

3.2.2 Results and feed-back

A database of 1,500 simulations of a river flood event scenario over a densely urbanized area described with a HR classified topographic dataset has been built. A random sampling on the produced results database was performed to follow a Monte-Carlo sampling. After checking the convergence [4], a UA and a variance based functional decomposition GSA have been performed over the output of interest. Output of interest being the maximal overland flow water depth (Y(x)) reached at every point of the computational grids.

The feasibility of spatial GSA approach for HR 2D flood modelling has been achieved by this proof of concept test study. Important requirements are involved when implementing UA and GSA (i) to establish/specify the method/the problem and (ii) to characterize input parameters as complexity of this step increases to consider spatial variability of the input parameters and can involve an important pretreatment phase (*e.g.* for DEMs generation). Eventually accuracy of HR topographic dataset spatial information might

not be available. In that case, basic assumption attributing regionally different characteristics to PDF could be relevant. However, this part of the process is “subjective” and it can be time consuming. Its application in dedicated tools (such as Promethee-FullSWOF_2D) might not be straight forward.

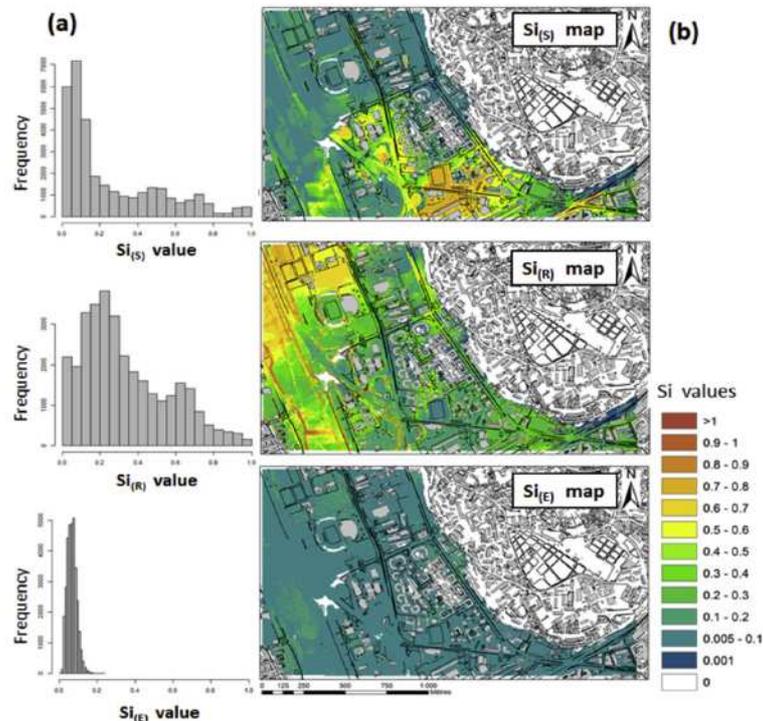


Figure 3: a) Distribution of computed Si and b) details of Si maps (from [4]).

The UA has allowed to quantify uncertain parameters impacts on output variability and to describe the spatial pattern of this variability. The spatial GSA has allowed the production of Sobol index (S_i) maps (an illustrative example is reported in Figure 3) over the area of interest, enhancing the relative weight of each uncertain parameter on the variability of calculated overland flow. The spatial distribution of (S_i) illustrates the major influence of the modeler choices, when using the HR topographic data in 2D hydraulic models (var. S and var. R) with respect to the influence of HR dataset accuracy (var. E), related to the choice of the mesh resolution. The fact that var. S is the most contributing parameter in densely urbanized areas is not surprising. Indeed, in that case, a change in var. S highly influences the representation in the model of physical properties of the urban environment, therefore impacting model results. Var. R indirectly impacts quality of small scale elements representation as well. Nevertheless var. E assumes a spatially uniform RMSE and does not take into consideration errors in photointerpretation. Therefore, uncertainties related to HR measurement are probably locally underestimated or overestimated in this study.

In conclusion, this research allows to underline (with a robust numerical methodology) that, in urban area, it is not necessary to focused on the roughness coefficients of the model or the mesh resolution if the urban structures (buildings, walls, sidewalks, etc...) are not correctly represented. This research also permitted to provide a schematic overview of the high-resolution topographic data use possibilities for flood modelling studies [35]. Especially, a reflection on the *good-practices* was proposed [35], in order to helps modellers to have a better understanding of theirs model results, and to provide some effective strategies for model improvement throw GSA and UA analysis. Nevertheless, it must be noted that this kind of approach does not permit to clearly quantify the effects of a given topographic zone on another zone (as the flow driven by particular topographic characteristics). Moreover, computational cost of this type of approach remains a clear limitation for standard practical applications.

3.3 Tsunamigenic potential of AZGBR seismic zone

3.3.1 Description and objectives of the case: Steps A and B

In order to assess the global tsunamigenic potential of the Azores-Gibraltar Fracture Zone (AGFZ), an uncertainty propagation study was carried out by IRSN between 2014 and 2016, in association with the “Commissariat à l’Energie Atomique et aux Energies Alternatives” (CEA). This ongoing research project was realized in the context of the French ANR-TANDEM research project (<http://www-tandem.cea.fr/>). Structurally, the AGFZ region is complex and not well understood. Moreover, some of its faults produce earthquakes with significant vertical slip, of a type that can generate tsunamis events. In this context, the objectives of this research are:

- to develop a numerical tool able to cover the whole AGFZ tsunamigenic potential through UA approach;
- to investigate the interest of producing “uncertainty” maps in this context of large scale simulations (Atlantic scale);
- to investigate the effects of the dependences between the input parameters through the analysis of sorted numerical results.

To manage uncertainties on tsunami parameters, CEA tsunami numerical code has been coupled with IRSN environment Promethee, in order to perform a large number of simulations. Numerical simulations were realized on HPC cluster of the “Laboratoire de Géologie de l’Ecole normale supérieure” (ENS). The seismic source input parameters were represented with uniform probability distributions and a Monte Carlo approach was used to sample the parameters in the chosen ranges of variation, leading to many simulations and results. Ten seismic parameters were modelled by uniform distributions: latitude [degrees], longitude [degrees], depth [km], strike [degrees], rake [degrees], dip [degrees], length L [km], width W [km], slip D [m] and rigidity μ [GPa]. The input variables were sampled randomly and independently. The magnitude M_w and its probability distribution were obtained from the seismic moment “ M_0 ”, with $M_0 = \mu LWD$. The same bathymetric grid, with a space resolution of 2’ (~3.6 km), was used for the simulations of the tsunami propagation.

3.3.2 Results and feed-back

Large variability of input data and its independent sampling give numerous combinations of seismic source parameters. A data-base of more than 20,000 simulations were performed with the Promethee/CEA-code tool. Each row (or scenario of the data base) contains the input seismic parameters and the associated maximum wave height and arrival time computed by the tsunami-call. For the purposes of the analysis, we sorted the data base with some “physical” filters (which are characterized by a chosen relation between the input variables). In this way, we consider that this kind of analysis could permit to roughly consider the dependence between the input parameters, even if the key hypotheses adopted for UA was that the input variables were independent (see section 2). A first filter (the “geophysical filter”) based on the relation between the fault length and the fault width has been applied. A second filter (the “focal filter”), based on the focal mechanisms of the seismic source, was also applied. Such filters considerably reduce the number of tsunami scenarios of the database: after the two filters, about 1,000 remaining scenarios can be exploited.

Some results are reported in Figure 4, showing the aggregating maps of the tsunami maximum wave heights, for the whole database or after the different filters: ‘mean+1 σ ’ values of the simulations are represented. These previous results show the impact that “physical filters” could have on results from UA and also confirm the interest of a spatial visualization of the numerical results. A coming paper is in progress to present this work in details. It will describe the used methodology, with a focus on the two applied filters, and the tsunami maximum wave heights will be analyzed with respect to the different filters employed.

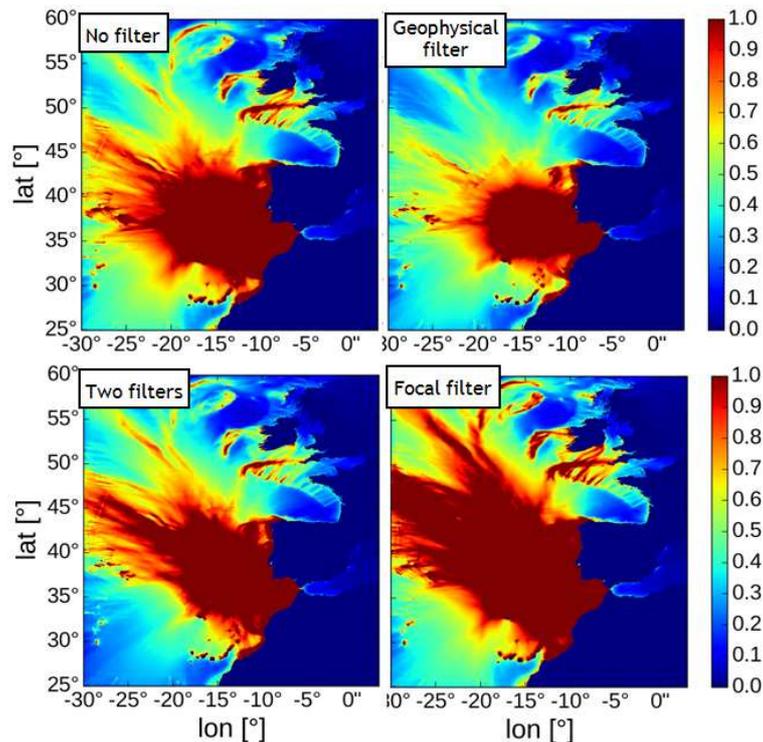


Figure 4: illustration of the maximum wave heights (one standard deviation over mean values) extracted from the AGFZ tsunami database, in view of different filters on the seismic source. Water height scale was limited to 1.0 m for clarity, but values higher than 1.0 were computed.

4. DISCUSSION & CONCLUSIONS

As reported in the introduction, since 2011 IRSN conducted explorative researches activities in order to investigate the feasibility of UA and GSA in hydraulic studies. A specific methodology, based on the coupling of the computational environment Promethee with an adapted numerical tool, was applied in different contexts, providing a first feed-back. The different coupling tools developed allowed to run an important number of simulations by stochastically parameterizing the modelling files, which was not possible manually, but also allowing the synthesis and statistic treatments of the results [3]. Even if for each application case some technical problems had to be solved (especially for the treatment of the model results, as in [4, 35]), a first conclusion is the flexibility of the use of Promethee with different codes developed or used by different operators (CNR, Orléans University, CEA, University of Nice).

Generally, we consider that the quantification of uncertainty through UA goes in the direction of improvement of state of the art, compared to quantification of uncertainty based on expert opinion only. Investigations on the UA can lead to deeper understanding of mechanisms leading to variable of interest variability. GSA by ranking uncertain parameters allows practical approaches to better understand issues in the model and to improve the model. Indeed, even if GSA results can vary from one approach to another at least, it helps modeller to have a better understanding of its model results, and provides effective strategies for model improvement.

In this section, we focused mainly on the engineering applications related to our studies, on the perspectives and, finally, on the role of uncertainty analysis in the research activities conducted by IRSN.

4.1 Interesting results for engineering applications

Besides hydraulics intuitions confirmation (*i.e.* upstream directed by downstream in [3]), the studies that were carried out showed some physical interesting results, as the non-linearity of the system response [3, 4]. Another interesting result is that GSA use to spatially rank uncertain parameters effects can give a valuable insight to modeler [4]: it can help reducing variability in the output putting effort on improving knowledge

about a given parameter or helps for optimization (*e.g.* to define relevant areas where spatial discretization is important prior to non-structured mesh use).

In the context of the nuclear safety, the feed-back of these studies opens various interesting perspectives. Indeed, the ANS guide N°13 [2] is focused on the identification and the penalization of the most influencing parameter of the numerical model. Depending on the reference flow situation (watershed flooding, waves, *etc.*), the influencing parameter can be difficult to define. As a consequence, UA and GSA study could permit to identify the influencing parameters in a rigorous way.

More generally, this methodology allow quantifying the uncertainty of the model results related to the input parameters and also to better define a critical flow situation. Indeed, UA permits to totally (thousands of possible combinations of the input parameters) explore a numerical model. As a consequence, the uncertainty propagation might lead to some rare combination of critical flooding case (*e.g.* the “cliff effects”) that would have not been identified with an expert opinion. This perspective should be of interested to deal with uncertainty related to structure (as the levee) or equipment’s (as valves, weirs, and so on) behavior.

4.2 Limitations and improvements of the methodology employed for UA and GSA

The methods employed for UA and GSA (described in the section 2) are very time consuming and require a large capacity for data storages. This is particularly true for the spatial analysis of the numerical results (as for [4]). For instance, the data-base developed in the frame of the ANR-TANDEM project is of about 400 Go, requiring an adapted procedure for data analysis and post-treatment. Moreover, only Monte-Carlo method was employed for UA and the methods of Morris and FAST99 for GSA [32]. These methods have the advantage to be robust even if numerically expensive.

From a practical point of view, in order to reduce the computational time and the volume of the data stored, a possible solution is to focus the analysis on some selected point (local analysis). However, in this way the advantages of a spatial analysis disappeared [4, 35]. So, it could be of interest to test other techniques for UA and GSA in order to reduce the number of computation (and, more important, the size of stored data) without reducing the convergence of the numerical results. Especially, to overcome to computational challenges, a meta-modelling approach could be helpfully. Replacing the numerical model (geometry file, result file, input file) by a function can really reduce the time necessary for the study and the volume of the stored data. For instance, the Gaussian Process Emulator (GPE), already tested in a coastal context [41] or the Polynomial chaos method (PCE), already tested in a complex case of coupling between an hydrodynamic and a morphodynamic numerical code [42], permit to approach the hydrodynamic model response with an adapted function. Considering the number of existing methods for GSA and UA [11, 12], the open-source numerical tools already available [32], and the evidence that, as reminded in Pappenberger et al. [43], depending in the method GSA might produce different results, furthers research activities on these topics are necessary for improve the methodology presented in this study.

Moreover, several physical relationships can exist between the considered uncertain parameters. These correlations can impact the uncertainty study, given that the parameters are no longer sampled independently. With this assumption, the ANOVA method can no longer be used, because of the independent parameters hypothesis it implies. A new method called ANCOVA (ANalysis of COVariance) can be used, instead. The dependency structure between a set of variables can be defined by a function called Copula [44], which links the multivariate probability density function of random set of variables ($X_1; \dots; X_v$) to their univariate probability density functions. This approach was recently tested in the hydraulic field on a simple test-case [42] but giving very promising results. Especially, it was found that when correlations were added they increased the variances significantly. The study [42] shows that with the ANCOVA method the gap between the variables influences decreased and variables that seemed first none-influencing (sediments diameter and transport coefficient of the Meyer-Peter and Müller formula in [42]) became of considerable influence when adding correlations. Until now, in the context of the ANR-TANDEM research project, we explore the question of the dependence of the input variables by introducing, in the post-processing phase of the study, some physical filters linking the independent input parameters. However, even if some previous interesting results have been obtained (Figure 4), this approach is not mathematically rigorous and should be of interest to test a rigorous approach based on the Copula function on a real-case.

Finally, the good choice of the distribution of the input parameters of the model is of primary importance because its influence on the results of the study [3]. Indeed, it is necessary to ensure, at least, a great expertise of the most impacting parameters, for variation ranges and statistical distributions, which determine the repartition (shape and amplitude) of calculated outputs.

4.3 The role of uncertainty analysis in the research activities conducted by IRSN

The research efforts provided since 2011 by IRSN confirm the possibility to introduce UA and GSA as an engineering tool for exploring hydraulic numerical models in different situations. The technical constraints related to the large amount of data do not seem a real limitation, considering the increasing evolution of the HPC capacity and the possibility of using other techniques for the analysis (as GPE or PCE). However, some theoretical constraints still persist, as the definition of the PDF of the input variables, the treatment of the dependences or the influence of the methodology employed (for GSA or UA) on the results. These scientific topics are a relevant part of our research activities.

Moreover, in the frame of the flow hazard studies, UA and GSA approaches can be both exploited in other contexts, as the determination of the uncertainty related to levees failure, partially explored by Bertrand et al. in the framework of the Bench-mark Garonne, or integrated in a more general context of quantification of the Probability Flood Hazard Assessment.

ACKNOWLEDGEMENTS

The authors are thankful to the “Compagnie Nationale du Rhône”, the “University of Nice” and the “Commissariat à l’Energie Atomique et aux Energies Alternatives” for the precious methodological and technical discussions that permitted to testing, analyzing and continuous improving this ongoing research activities on uncertainty studies in hydraulic fields. The authors are also thankful to the “Bureau de Recherches Géologiques Minières” for the exchanges and discussions in the framework of the ANR-TANDEM research project.

6. REFERENCES

- [1] Iooss, B. (2011). Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Société Française de Statistique*, **152**(1), p. 1-23.
- [2] ASN (2013). *Protection of Basic Nuclear Installations against external flooding*.
- [3] Nguyen, T.-m. & al. (2015). Propagation des incertitudes dans les modèles hydrauliques 1D. *La Houille Blanche*, (5), p. 55-62.
- [4] Abily, M. & al. (2016). Spatial Global Sensitivity Analysis of High Resolution classified topographic data use in 2D urban flood modelling. *Environmental Modelling & Software*, **77**, p. 183-195.
- [5] Banks, J.C., Camp, J.V. & Abkowitz, M.D. (2014). Adaptation planning for floods: a review of available tools. *Natural Hazards*, **70**(2), p. 1327-1337.
- [6] Dale, M. & al. (2012). Probabilistic flood forecasting and decision-making: an innovative risk-based approach. *Natural Hazards*, **70**(1), p. 59-72.
- [7] Santos, P.P.d. & Tavares, A.O. (2015). Basin flood risk management: a territorial data-driven approach to support decision-making. *Water*, **7**(2), p. 480-502.
- [8] Son, C.H. & al. (2015). The effects of mitigation measures on flood damage prevention in Korea. *Sustainability*, **7**(12), p. 16866-16884.
- [9] Iooss, B. (2013). Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Société Française de Statistique*, **152**(1), p. 1-23.
- [10] Marrel, A. & al. (2009). Calculations of the sobol indices for the gaussian processes metamodel. *Reliability Engineering and System Safety*, **94**, p. 742-751.
- [11] Faivre, R. & al. (2013). *Analyse de sensibilité et exploration de modèles. Application aux sciences de la nature et de l'environnement*. Ed. Quae.
- [12] Saltelli, A. & al. *Global Sensitivity Analysis. The Primer*.
- [13] Saltelli, A., Chan, K. & Scott, E.M. (2000). *Sensitivity Analysis*.
- [14] Refsgaard, J.C. & al. (2007). Uncertainty in the environmental modelling process - a framework and guidance. *Environmental Modelling & Software*, **22**(11), p. 1543-1556.
- [15] Uusitalo, L. & al. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, **63**, p. 24-31.
- [16] Ranzi, R. & al. (2013). Levee Breaches Statistics, "Geotechnical Uncertainty", residual risk in flood hazard mapping. In *Proceedings of the 35th IAHR world congress, September*, p. 8-13.
- [17] Saint-Geours, N. & al. (2014). Multi-scale spatial sensitivity analysis of a model for economic appraisal of flood risk management policies. *Environmental Modelling & Software*, **60**, p. 153-166.
- [18] Helton, J.C. & al. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, **91**(10), p. 1175-1209.
- [19] Iooss, B. & Lemaître, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty Management in Simulation-Optimization of Complex Systems*, p. 101-122. Springer.
- [20] Saltelli, A. & al. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. John Wiley & Sons.
- [21] R Core Team, 2016 - "R : A langage and environment for statistical computing", R Foundation for Statistical Computing. Available from <http://www.R-project.org>.
- [22] Baroni, G. & Tarantola, S. (2014). A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. *Environmental Modelling & Software*, **51**, p. 26-34.

- [23] Saint-Geours, N. (2012). *Sensitivity analysis of spatial models: application to cost-benefit analysis of flood risk management plans*. Phd thesis, Université Montpellier II-Sciences et Techniques du Languedoc.
- [24] Morris, M.D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, **33**(2), p. 161-174.
- [25] Willis, T.D. (2014). *Systematic analysis of uncertainty in flood inundation modelling*. Phd thesis, University of Leeds.
- [26] Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, **145**(2), p. 280-297.
- [27] Saltelli, A., Tarantola, S. & Chan, K.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, **41**(1), p. 39-56.
- [28] Sobol, I.M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, **55**(1), p. 271-280.
- [29] Sobol, I.M. (1993). Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments*, **1**, p. 407-414.
- [30] Homma, T. & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, **52**(1), p. 1-17.
- [31] Cukier, R. & al. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of chemical physics*, **59**(8), p. 3873-3878.
- [32] Pujol, G. & al., 2016 - "Sensitivity: Global Sensitivity Analysis of Model Outputs". Available from <https://CRAN.R-project.org/package=sensitivity> [R package version 1.13.0].
- [33] Cukier, R., Levine, H. & Shuler, K. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of computational physics*, **26**(1), p. 1-42.
- [34] Alliau, D. & al. (2015). Étude du risque d'inondation d'un site industriel par des crues extrêmes: de l'évaluation des valeurs extrêmes aux incertitudes hydrologiques et hydrauliques. *La Houille Blanche*, (2), p. 67-74.
- [35] Abily, M. (2015). *High-resolution modelling with bi-dimensional shallow water equations based codes : high-resolution topographic data use for flood hazard assessment over urban and industrial environments*. Phd thesis, Université Nice Sophia Antipolis. Available from <https://tel.archives-ouvertes.fr/tel-01288217>.
- [36] Abily, M. & al. (2016). Use of 3D classified topographic data with FullSWOF for high resolution simulation of a river flood event over a dense urban area. *arXiv preprint arXiv:1603.07463*.
- [37] Abily, M. & al. (2016). High-resolution Modelling With Bi-dimensional Shallow Water Equations Based Codes–High-Resolution Topographic Data Use for Flood Hazard Assessment Over Urban and Industrial Environments. *Procedia Engineering*, **154**, p. 853-860.
- [38] Abily, M. & al. (2016). Global sensitivity analysis with 2D hydraulic codes: application on uncertainties related to high-resolution topographic data. *Advances in Hydroinformatics*, p. 301-315.
- [39] Guinot, V. & Gourbesville, P. (2003). Calibration of physically based models: back to basics? *Journal of Hydroinformatics*, **5**(4), p. 233-244.
- [40] Delestre, O. & al. (2014). FullSWOF: A software for overland flow simulation. In *Advances in Hydroinformatics*, p. 221-231. Springer.
- [41] Malde, S. & al. (2016). Applying emulators for improved flood risk analysis. In *FLOODrisk 2016 - 3rd European Conference on Flood Risk Management*, Lyon.
- [42] Mouradi, R.S. & al. (2016). Sensitivity analysis and uncertainty quantification in 2D morphodynamic models using a newly implemented API for TELEMAC2D/SISYPHE. In *23rd Telemac-Mascaret User Club*, Paris.

- [43] Pappenberger, F. & al. (2008). Multi-method global sensitivity analysis of flood inundation models. *Advances in water resources*, **31**(1), p. 1-14.
- [44] Chastaing, G. (2013). *Indices de Sobol généralisés pour variables dépendantes*. Phd thesis, Mathématiques appliquées, Grenoble, Université de Grenoble, 218 p. Available from <https://tel.archives-ouvertes.fr/tel-00930229>.