# Chapter 37
# Uncertainty Propagation in TELEMAC 2D Dam Failures Modelling and Downstream Hazard Potential Assessment

**Layla Assila, Matthieu Sécher, Thomas Viard, Benoît Blancher, and Cédric Goeury**

**Abstract** This work addresses uncertainty propagation in TELEMAC 2D models with respect to two major types of risks in river hydrodynamics: flood hazard and dam failures. The studied case is a TELEMAC 2D model that extends over approximately 14.4 km$^2$ with a river length of 41 km including 3 major tributaries to the main river and 3 dams. The implementation of the uncertainty propagation approach would not have been feasible and accomplished without the open source platform SALOME-HYDRO and the TELAPY module (PYTHON API) of the TELEMAC-MASCARET SYSTEM. The first step consisted of quantifying uncertain parameters for the acquired hydraulic model and defining adequate probability distributions based on expert judgment and previous specific studies that have been provided by EDF. A sensitivity analysis based on Morris screening method was then carried out to reduce the number of uncertain factors. Uncertainty propagation algorithms such as Monte Carlo and Polynomial Chaos expansion were used to estimate the maximum water depths and velocities, as well as their statistical moments such as the mean and variance and the Sobol indices of the considered parameters. The use of parallelism proved to be necessary to optimize the computation time. The final results are then used to assess the flood casualties and the flood damages. This second estimation is based on the FLOODRISK plugin of QGIS.

**Keywords** Uncertainty · Sensitivity analysis · Dam failures · TELEMAC-MASCARET · TELEMAC 2D · SALOME-HYDRO

L. Assila · M. Sécher (✉) · T. Viard · B. Blancher
EDF Hydro - CIH, Savoie Technolac 4 allée du Lac de Tignes, 73290 La Motte-Servolex, France
e-mail: matthieu.secher@edf.fr

C. Goeury
EDF R&D - LNHE, Quai Watier, 78401 Chatou, France

## 37.1 Introduction

Numerical river hydraulics is based on the discretization of partial differential equations (Saint-Venant or Navier-Stokes) which include simplifying assumptions, input data such as rating curves, bathymetry, hydrographs, and parameters including uncertainties that may influence the results. In the current configuration, most parameters are calibrated a posteriori to ensure a good accuracy and representation of the flow dynamics. However, if the calibration could not be carried out due to lack of data, then the validation of the results is based on expert judgment and is subject to high uncertainties. Thus, uncertainty quantification can prove to be a valuable decision making tool since it can determine confidence intervals and whether model outputs will comply with the regulatory requirements (e.g. design requirements) given the random variation in inputs.

In this thesis, the uncertainty propagation methodology presented in Fig. 37.1 is followed [1]. Three main steps are identified:

- Step A consists in defining the model, the statistical quantity of interest and the corresponding criteria (e.g. criteria on failure probability). The model description is similar to a classical deterministic approach as it defines the inputs and outputs of the model.
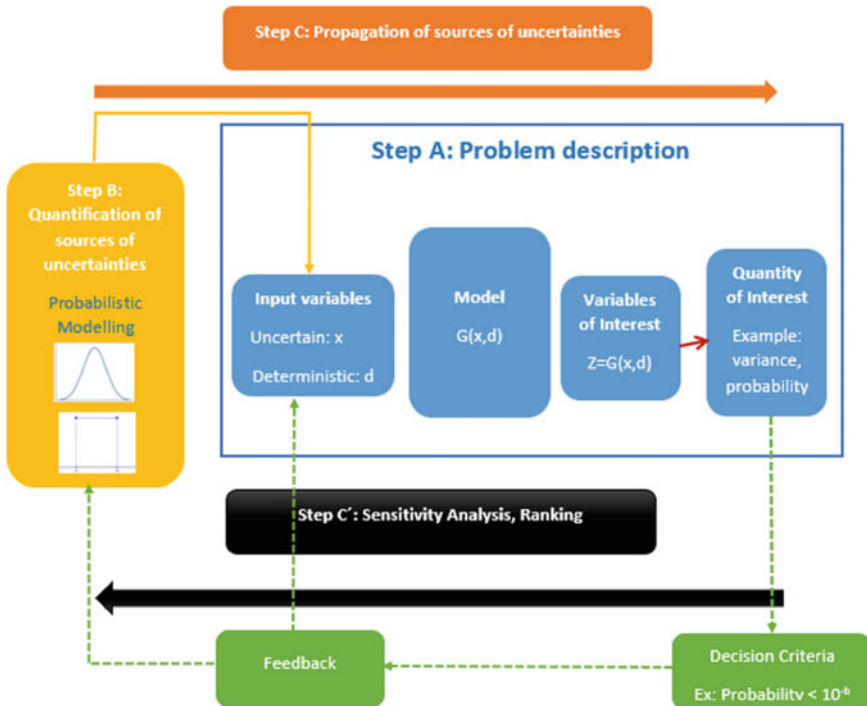


**Fig. 37.1** General framework for uncertainty propagation studies [2]

- Step B consists in quantifying sources of uncertainties on model input parameters which will be described by adequate probability distributions. The result of this step is a random vector of all uncertain variables which is represented by the joint probability distribution of all marginal distributions and a copula that describes the dependence between the variables.
- Step C consists in propagating uncertainties on the input through the model. In most cases, a sensitivity analysis (step C´) is required to assess the influence and the importance of input parameters with respect to the randomness of the output.

One objective of this work is to apply the previously defined steps; adopted by EDF R&D [2], on a TELEMAC 2D hydraulic model that includes dam failures scenarios through the conception of user-friendly PYTHON scripts. The main methods tested in this work are the classical algorithm of Monte Carlo using different sampling techniques such as Quasi-Monte Carlo with low discrepancy sequences, the Morris screening method for sensitivity analysis and the Polynomial Chaos Expansion to build surrogate models.

The model used to test this method of uncertainty propagation is part of an incremental damage study in which submersion waves are simulated for a reference no dam break scenario and an adverse dam break scenario, increments are calculated by comparing the results of both scenarios and flood damages are assessed.

## 37.2  Methodology

### 37.2.1  Hydraulic TELEMAC 2D Model

The geographical location of the modeled river is confidential. The river is delimited by a city downstream which represent a major vulnerability to flood risk in case of dam breaks. The model extends over approximately 14.4 km$^2$ with a river length of 41 km and it includes 3 major tributaries to the main river. The mesh contains approximately 185 000 nodes. The DEM of the model is given in Fig. 37.2.

The simulated discharge correspond to a return period of 5 000 years and it deterministic value is estimated by the Schadex method [3]. A factor 10 is considered by assumption for the return period of the tributaries discharges; i.e. the discharges considered for the tributaries are associated with a return period of 500 years.

The dams are modeled by prescribing rating curves on the upstream boundaries and prescribing discharges on the downstream boundaries. A dam failure correspond to the event of exceeding a dam stability threshold defined based on expert judgment. The rating curves are switched once the break occurs via TELAPY PYTHON script and user FORTRAN file. The solution of Ritter is considered to represent the rating curve associated with the break [4]. The model also includes 7 bridges which are modelled as drag forces in the FORTRAN subroutine Dragfo. The bridges are considered unstable if the water level upstream of the bridge is higher than the bridge deck level.
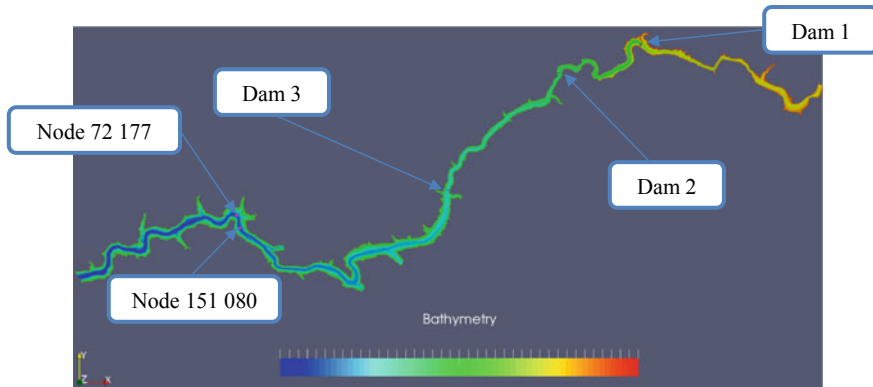
**Fig. 37.2** Bathymetry of the 2D model

In terms of computation time, one run of the model takes approximately 1 h 10 min on 56 processors. This was found to be the optimal number of processors for the studied model.

For this study, the variables of interest are the maximal free surface elevation and the maximal water level (for the flood damages quantification). The statistical quantities of interest are the statistical moments; mainly the mean and the variance, as well as sensitivity measures such as Sobol indices.

### 37.2.2 Quantification of Uncertain Parameters

The quantification of uncertain parameters is carried out based on the categories of the model input data:

- Flow discharges: they correspond to the peak flood discharges at the main river and the 3 tributaries. For this model and based on expert judgment, only the discharges of the main river and one of the tributaries are considered uncertain. Their probability distribution is the truncated normal distribution with a mean equal to the deterministic peak flood discharge estimated with the Schadex method. Variances of approximately 5% and 25% of the means were respectively taken for the main river and the tributary.
- Strickler coefficients: the hydraulic model is divided into 5 areas with different Strickler coefficients. All these coefficients are considered uncertain following a uniform distribution. Since the model was not calibrated, the bounds for these distributions were estimated based on literature values and are given in Table 37.1.
- Dam failure thresholds: the occurrence of a dam break is defined as the event of the hydraulic head upstream of the dam exceeding a specific stability threshold. The latter is considered uncertain following a truncated normal distribution. The parameters of the probability distributions of dam failure thresholds for the 3 dams

**Table 37.1** Probability distribution for Strickler coefficients

| Uncertain variable | Probability distribution | Bounds |
|---|---|---|
| Strickler for urban zones | Uniform | [10, 15] |
| Strickler for forests | Uniform | [5, 12] |
| Strickler for agricultural zones | Uniform | [15, 25] |
| Strickler for meadows | Uniform | [20, 30] |
| Strickler for the riverbed | Uniform | [28, 32] |

included in the model were defined relying on expert judgment and previous EDF studies. The values of these thresholds are confidential.

- Drag coefficients for bridges: based on expert judgment, one of the seven bridges is considered stable since the vulnerabilities are located upstream of the bridge. The other six bridges which are modelled as drag forces, are unstable if the water level exceeds the bridge deck level. The drag coefficients are then considered uncertain since they are empirically estimated based on the shape and material of the bridge. They follow a uniform distribution.
- Dams rating curves coefficients: the equations used to assess the rating curves include several empirical coefficients that are uncertain and that follow uniform distributions.
- Ritter coefficient: the dam break is described by the dam break solution of Ritter with a deterministic value of 0.209 for coefficient of Ritter. It follows a truncated normal distribution with the value 0.3 (spillway overflow coefficient used in the deterministic case to represent dam breaks) as maximal bound.

The quantification step resulted in 27 uncertain parameters that are assumed to be independent.

### 37.2.3  Sensitivity Analysis and Uncertainty Propagation

Given the large number of quantified uncertain parameters, the Morris screening method is tested in order to reduce the problem dimensionality. This method was first introduced in [5] with the aim of identifying the subset of non-influent parameters in a model using a small number of model evaluations (output samples). The input factors are usually classified according to their effects: negligible, linear and uncorrelated, non-linear and correlated. It is based on moving the factors of a sample one at a time (OAT) by a step $\Delta$ in the input physical space. The elementary effects are quantified using (1). The absolute mean and standard deviation of these elementary effects are taken as sensitivity measures.

$$EE_i = \frac{Model(x_1, \ldots, x_i + \Delta, \ldots, x_p) - Model(x_1, \ldots, x_p)}{\Delta} \qquad (1)$$

where $x_1, \ldots, x_p$ are the uncertain factors of the model, and $EE_i$ is the elementary effect of parameter $x_i$.

Sobol indices from a polynomial chaos expansion (PCE) of approximately 7 200 Monte Carlo simulations are computed to compare with the results of the Morris screening method. The PCE is implemented based on a Least Angle Regression Strategy (LARS) using a corrected Leave-One-Out error [6]. The LARS uses a Least Square Regression truncation method of the polynomial decomposition.

The classical algorithm of Monte Carlo is implemented to propagate uncertainties. Convergence of Monte Carlo and Quasi Monte Carlo are studied [7]. The aim is the computation of statistical moments such as the mean and the variance of the maximal water depth and maximal velocity which will be used to assess the flood damages.

### 37.2.4   Dam Break Scenarios

The scenarios considered in this study are:

- Break scenario: total and instantaneous dam breaks triggered by the break of the 1st upstream dam.
- Reference scenario: no dam breaks

### 37.2.5   Implementation Using the APIs and the Clusters

The sensitivity analysis and uncertainty propagation methods were implemented using the C++/PYTHON library OPENTURNS [8] designed for the treatment of uncertainties. It coupling with the hydraulic model (i.e. TELEMAC 2D) is facilitated by the use of the TELAPY module [9] which allows to set and run TELEMAC instances via PYTHON.

Since a large number of simulations are going to be executed, the optimization of the computation time using the available EDF clusters is deemed necessary to accomplish this study. If the TELEMAC model only is parallelized, a minimal time of approximately 48 days 14 h is required for 1 000 simulations. However, if the 1 000 simulations are also parallelized according to Fig. 37.3, then a minimal time of approximately 23 h 20 min is sufficient for all the simulations.
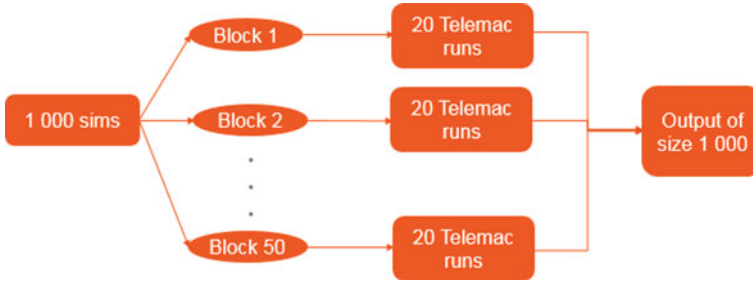
**Fig. 37.3** Parallelization scheme on Clusters

**Table 37.2** Cases ran for the Morris screening method

| Number of trajectories | Number of simulations | Minimal calculation time on cluster |
|---|---|---|
| 20 | 560 | $\sim 20$ h |
| 40 | 1 120 | $\sim 38$ h |
| 60 | 1 680 | $\sim 60$ h |
| 100 | 2 800 | $\sim 4$ days |

## 37.3 Results

### 37.3.1 Morris Screening Method Results

For the Morris screening method, several numbers of trajectories were tested as shown in Table 37.2.

The results for the Morris screening method were not coherent with the expert judgment. In fact, the dam break thresholds and the upstream discharge that were expected to have the major influence on the results based on expert judgment were found to have minimal to zero influence. This is due to the constant delta that is chosen for all parameters even though their values and their types differ significantly. A possible solution to this problem would be to perform an iso-probabilistic transformation on the set of input parameters before generating the samples. Thus, to reduce the set of uncertain parameters, we finally used expert judgment.

### 37.3.2 Uncertainty Propagation and Sensitivity Analysis

#### 37.3.2.1 Convergence

Given the incoherent results of the Morris method, the Monte Carlo algorithm is performed using random sampling method and Quasi Monte Carlo sampling method based on low discrepancy Sobol sequences.

First, the convergence of the dispersion coefficient ($\sigma/\mu$) and the mean of Monte Carlo is graphically analyzed. The mean is bounded by its 95% confidence interval. The dispersion coefficient and the mean of the maximal surface elevation estimated at node 151 080 (node displayed in Fig. 37.2) and shown in Fig. 37.4 and Fig. 37.5 suggests that the convergence of Monte Carlo is obtained at approximately 7 000 simulations. Quasi Monte Carlo converges more rapidly at approximately 3 500 simulations.

Second, a convergence study based on a criteria set on the coefficients of variation of the mean and the variance of the Monte Carlo samples is performed. The criteria were defined a priori (based on the precision deemed acceptable for the model at hand) and are given in (2) and (3).

$$Criterion\,1 : CV_{mean} \leq 0.00005 \tag{2}$$

$$Criterion\,2 : CV_{variance} \leq 0.01 \tag{3}$$

The results displayed in Fig. 37.6 and Fig. 37.7 and estimated on 2 nodes: node 151 080 and node 72 177 located on the downstream agricultural floodplain (Fig. 37.2), confirm that the convergence of Monte Carlo is obtained for approximately 7 000 simulations.
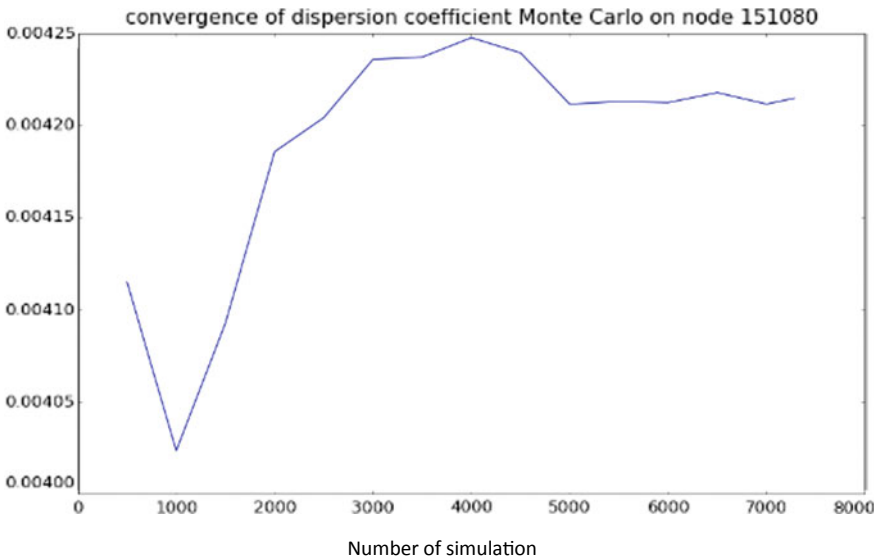


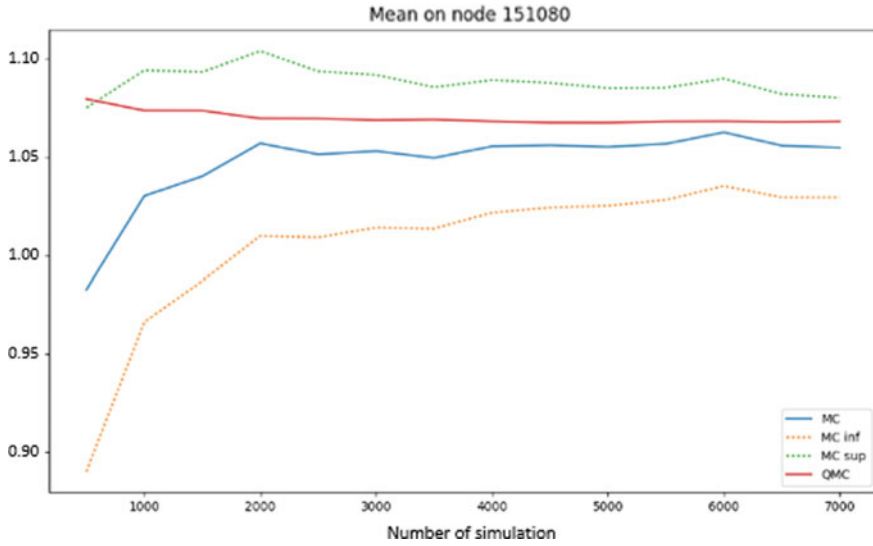**Fig. 37.4** Convergence of the dispersion coefficient of Monte Carlo

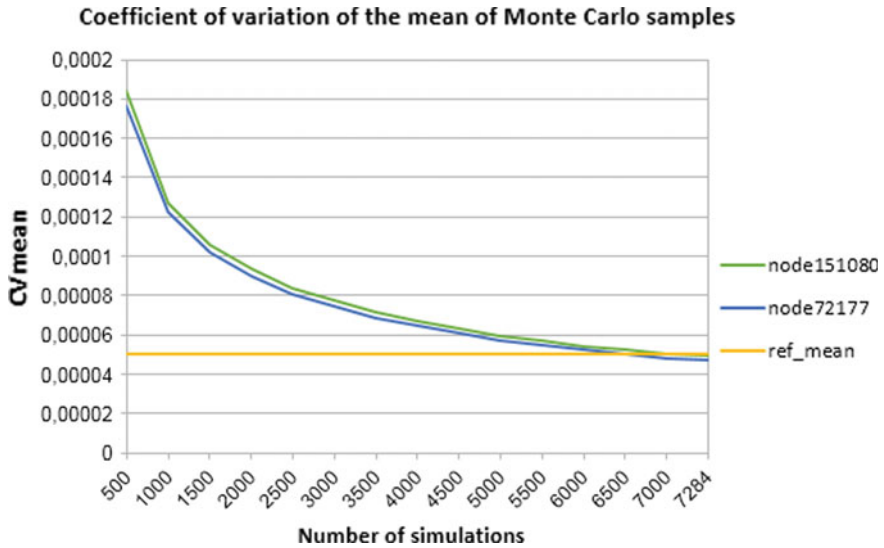**Fig. 37.5** Convergence of the mean of the surface elevation for Monte Carlo and Quasi-Monte Carlo sequence



**Fig. 37.6** Convergence of the coefficient of variation of the Monte Carlo samples variance using criterion 1
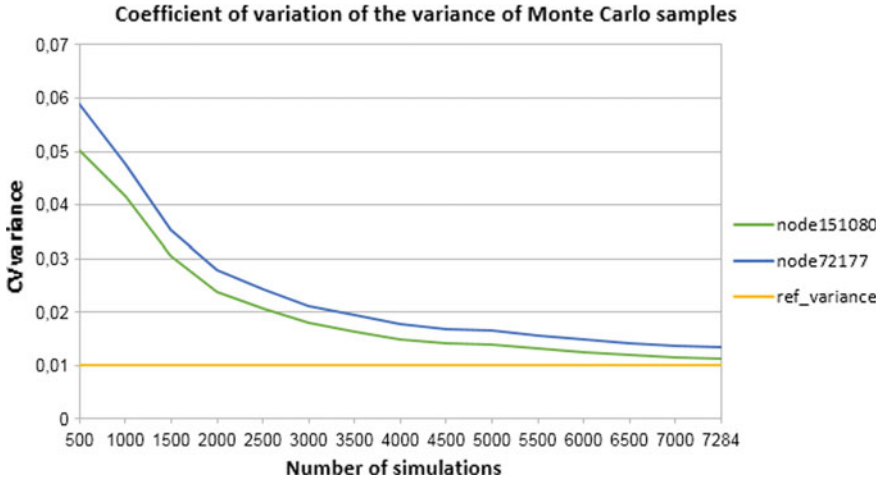
**Fig. 37.7** Convergence of the coefficient of variation of the Monte Carlo samples variance using criterion 2

### 37.3.2.2 Polynomial Chaos Expansion Based on LARS

This method is tested here since it allows the computation of Sobol indices with smaller samples than Monte Carlo or Saltelli algorithm [10]. The theoretical number of simulations required to construct a surrogate model of degree 4 is given by:

$$N_{th} = \binom{p+d}{d} = \frac{(p+d)!}{p!d!} = 31465\,simulations \tag{4}$$

where $p$ is the number of uncertain parameters and $d$ the degree of the polynom.

This number implies a computation time of approximately 47 days on the EDF Eole cluster.

Although the available number of Monte Carlo simulations (~7 200) is inferior to the theoretical number required for a PCE (4), this method was still tested with this sample. A cross-validation was then performed on the constructed surrogate model using a validation Monte Carlo sample of size 1 000.

First, all 7 200 simulations were used to construct surrogate models of different degrees in order to find the optimum precision. The reference values are the mean and variance of the 7 200.

The differences between the means and the variances are approximately $7.10^{-3}$ and $8.10^{-3}$ respectively (as shown in Fig. 37.8). These errors are acceptable and the degree 4 is thus retained for an accuracy study based on the approximation accuracy coefficient $Q_2$ computed using (5) (Fig. 37.9).
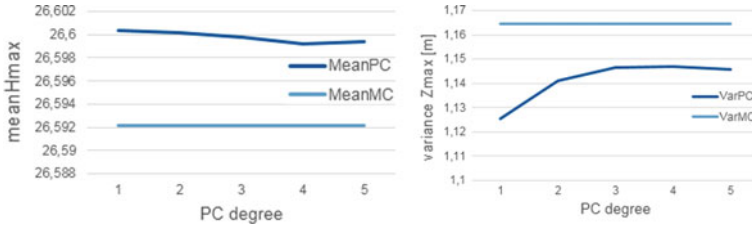
**Fig. 37.8** On the left, estimated mean of the surface elevation on node 151 080 for PCE using 7 200 Monte Carlo simulations; on the right, estimated variance of the surface elevation on node 151 080 for PCE using 7 200 Monte Carlo simulations

**Fig. 37.9** Approximation accuracy coefficient of a degree 4 PCE using different sizes for the Monte Carlo sample
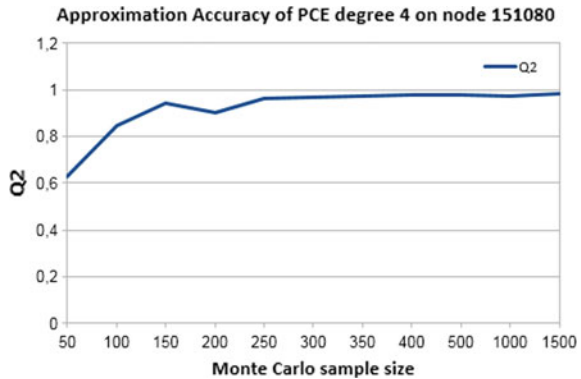


$$Q_2 = 1 - \varepsilon_{LOO} \tag{5}$$

Where $\varepsilon_{LOO}$ is the Leave-One-Out error.

This error is a special case of K-fold error estimate where the number of folds is chosen equal to the cardinality N of the experimental design $X$. Let's denote $\hat{M}$ the surrogate model of the real model M, $\hat{M}^{-i}$ the surrogate model built from the experimental design $X \backslash \{x^i\}$ with the i-th sample $x^i$ being set aside, and $cov(Y)$ the empirical covariance of the response sample Y. The Leave-One-Out error can be calculated using (6).

$$\varepsilon_{LOO} = \frac{\frac{1}{N} \sum_{i=1}^{N} M\left(x^i\right) - \hat{M}^{-i}\left(x^i\right)}{cov(Y)} \tag{6}$$

A cross validation is performed for the PCE surrogate model of degree 4 constructed with 400 Monte Carlo simulations using 1 000 samples. Figure 37.10 indicates that the surrogate model gives a good approximation even if the sample size is less than the required theoretical number of simulations calculated in (4).
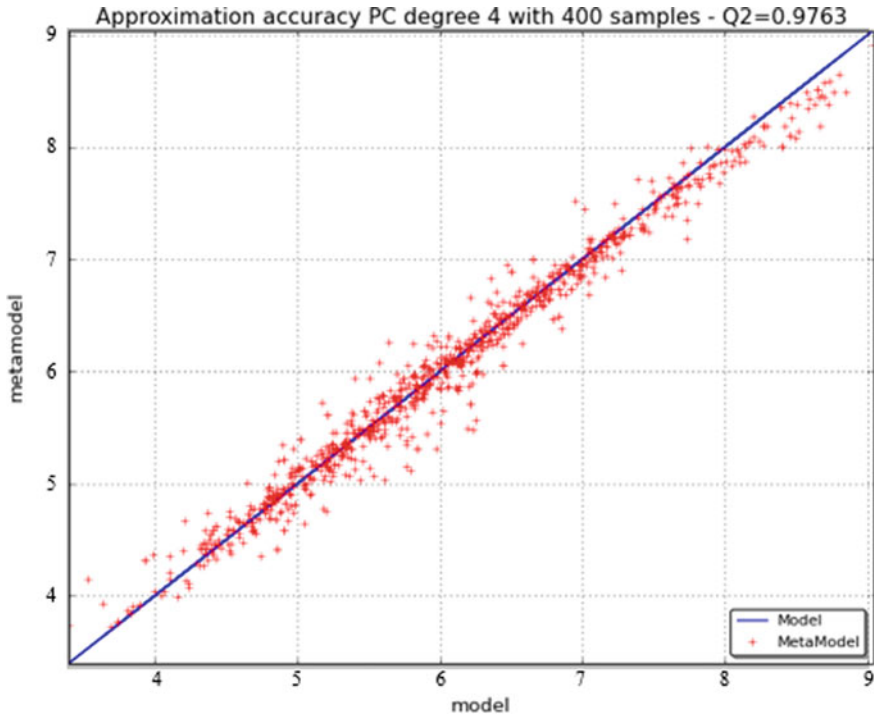
**Fig. 37.10** Cross validation of the degree 4 PCE using 1 000 Monte Carlo samples for validation

### 37.3.2.3 Sobol' Indices Using PCE Surrogate Models

The surrogate models that have been built using polynomial chaos expansion can also be used to perform a sensitivity analysis by computing Sobol indices. In fact, first order Sobol indices have been calculated using a PC model of degree 4 built with all 7 200 Monte Carlo samples. In Fig. 37.11, the conveyance coefficients of dam j are denoted $QZ_{ji}$, $i = 1, \ldots, n_j$, where $n_j$ is the number of conveyance coefficients of dam j. The drag coefficients of bridges are denoted $drag$, the Strickler coefficients are denoted $CF_i$, $i = 1, \ldots, 5$ and the surface elevation dam failure thresholds are denoted $Break_i$, $i = 1, 2, 3$. $Qmain$ and $Qtrib$ refer respectively to the discharges of the main river and the tributary. The results shown in Fig. 37.11 are more coherent with the expert judgment than those of the Morris method. The Strickler coefficient of forest areas stands out as the most influential parameter. This could be explained by the location of the node used for the computation of these indices (node 151 080) or by the fact that the surrogate model still needs to be refined.
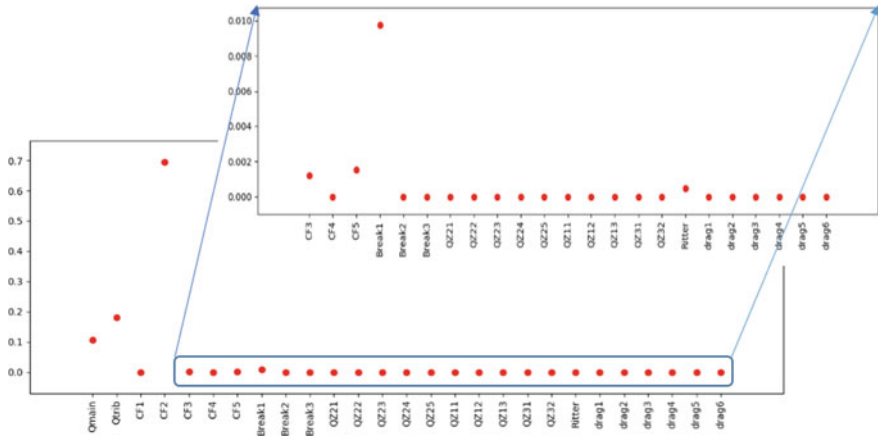
**Fig. 37.11** First order Sobol' indices estimated from a degree 4 PCE

#### 37.3.2.4   Post-treatment and Damages Assessment

Statistical moments such as the mean and the variance of the maximal water depth of the Monte Carlo output sample have been computed. For visualization purposes, these moments are integrated in a result MED file using HERMES APIs, MEDCOUPLING and MEDLOADER [11]. An example of the addition of the mean of the maximal surface elevation from 7 200 Monte Carlo simulations to the results file is given in Fig. 37.12.
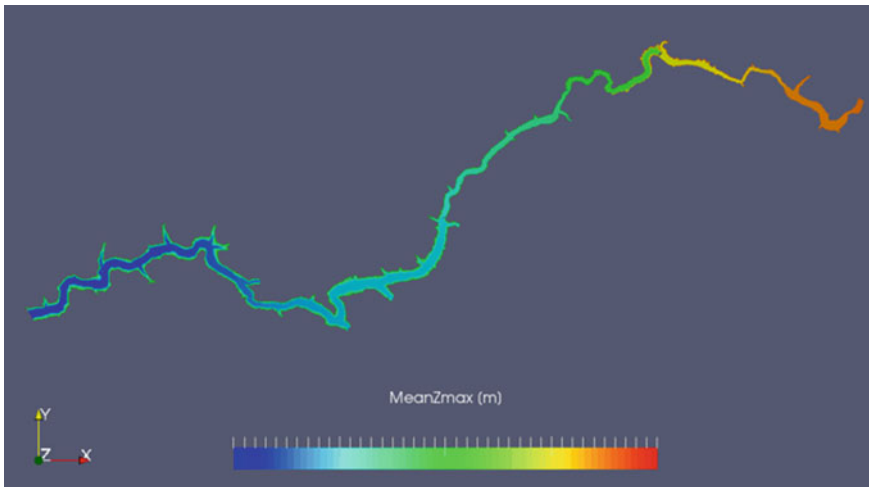


**Fig. 37.12** Integration of the mean of surface elevation to the MED result file

Downstream flood damages were quantified using 2 methods:

- Ramsbottom or Flood risk to People method [12]: used to quantify casualties based on a danger factor.
- FLOODRISK plugin of QGIS [13]: used to evaluate economic damages based on depth-damage curves.



**Fig. 37.13** Vulnerabilities map at the downstream area for total instantaneous dam failures scenario

The FLOODRISK plugin was used to quantify the damages downstream of the last dam using 7200 simulations. The resulting vulnerability map is given in Fig. 37.13. The vulnerabilities are mainly concentrated downstream of the last dam.

## 37.4   Perspectives and Limitations

The results presented here remain indicatory of the prospects uncertainty propagation and sensitivity analysis methods can provide. In fact, a 2nd scenario describing total and consecutive dam breaks (any dam can be the first to fail) is considered and its results still need to be post-treated. Monte Carlo was not exploited yet to calculate Sobol indices since it requires a larger number of simulations [10]. Other nodes, for which the same methods can be applied, located along the streamline or the floodplain can be taken into account. Hence, maps of Sobol indices can be created. Another perspective is the quantification of casualties using the Ramsbottom method, as well as automating the FLOODRISK method without using the QGIS interface.

As for the limitations of the study, the convergence criteria were defined for the specific model at hand. Hence, one should modify and adapt these criteria based on expert judgment and regulatory requirements.

## 37.5   Conclusions

The implementation of the uncertainty propagation approach would not have been feasible and accomplished without the open source platform SALOME-HYDRO,

OPENTURNS and the TELAPY module (PYTHON API) of the TELEMAC-MASCARET SYSTEM. The main difficulty - which is generally common to probabilistic models that treat uncertainty propagation - remains the optimization of computation time with respect to the number of simulations considered and the run-time of the TELEMAC case; whether it is in parallel or sequential mode.

The main objective of this study; which were the application of the uncertainty propagation methods on a hydraulic model at an engineering scale was accomplished. From an engineering point of view, the quantification of sources of uncertainties and their representation with suitable probability distributions is the step that takes the longest time to complete since it mostly relies on expert judgment.

Uncertainty propagation method allowed to criticize deterministic results and permit to explore extensively the surface response of a model. A probabilistic model is then more reliable. In the case presented in this paper, the results showed that deterministic model was more pessimistic than the mean probabilistic result.

# References

1. Sudret B (2007) Uncertainty propagation and sensitivity analysis in mechanical models. Contributions to structural reliability and stochastic spectral methods. Rapport d'Habilitation à Diriger des Recherches
2. Goeury C, David T, Ata R, Boyaval S, Audouin Y, Goutal N, Popelin A-L, Couplet M, Baudin M (2015) Uncertainty quantification on a real case with Telemac-2D. In: TELEMAC-MASCARET user conference proceedings, pp 44–51
3. Paquet E, Lang M, Carré JC (2012) Schadex method for extreme flood estimation: overview, applications and perspectives. In: 24th ICOLD Congress, Kyoto, Japan, June 2012, p 29
4. Castro-Orgaz O, Chanson H (2017) Ritter's dry-bed dam-break flow: positive & negative wave dynamics. Environ Fluid Mech 17(4)
5. Morris MD (1991) Factorial sampling plans for preliminary computational experiments. J Technometrics 33:161–174
6. Blatman G, Sudret B (2011) Adaptive sparse polynomial chaos expansion based on least angle regression. J Comput Phys 230(6):2345–2367
7. Caflisch RE (1998) Monte carlo and quasi-monte carlo methods. Cambridge University Press
8. Baudin M, Dutfoy A, Iooss B, Popelin A-L (2015) OPENTURNS: an industrial software for uncertainty quantification in simulation. In: Handbook of uncertainty quantification, pp 2001–2038
9. Goeury C, Audouin Y, Zaoui F (2017) TelApy user manual. V7P3 TELEMAC-MASCARET documentation
10. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global sensitivity analysis: the primer. Wiley
11. MEDCoupling and MEDLoader API Documentation, Salome platform version 8.4.0

12. HR Wallingford, Flood Hazard Research Centre, Middlesex University, Risk & Policy Analysts Ltd. (2006). The Flood Risks to People Methodology: PHASE 2 FD2321/TR1. Department for Environment, Food and Rural Affairs, Pb no. 11545, March 2006
13. Mancusi L, Albano R, Sole A (2015) FloodRisk: a QGIS plugin for flood consequences estimation. In: Geomatics Workbooks no 12 – FOSS4G Europe Como