HENRY

Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

Conference Paper, Published Version

Mouradi, R. S.; Audouin, Yoann; Goeury, Cédric; Claude, Nicolas; Tassi, Pablo; El Kadi Abderrezzak, Kamal

Sensitivity analysis and uncertainty quantification in 2D morphodynamic models using a newly implemented API for TELEMAC2D/SISYPHE

Zur Verfügung gestellt in Kooperation mit/Provided in Cooperation with:
TELEMAC-MASCARET Core Group

Verfügbar unter/Available at: https://hdl.handle.net/20.500.11970/104533

Vorgeschlagene Zitierweise/Suggested citation:
Mouradi, R. S.; Audouin, Yoann; Goeury, Cédric; Claude, Nicolas; Tassi, Pablo; El Kadi Abderrezzak, Kamal (2016): Sensitivity analysis and uncertainty quantification in 2D morphodynamic models using a newly implemented API for TELEMAC2D/SISYPHE. In: Bourban, Sébastien (Hg.): Proceedings of the XXIIIrd TELEMAC-MASCARET User Conference 2016, 11 to 13 October 2016, Paris, France. Oxfordshire: HR Wallingford. S. 153-162.

Sensitivity analysis and uncertainty quantification in 2D morphodynamic models using a newly implemented API for TELEMAC2D/SISYPHE

R. S. Mouradi[1,2], Y. Audouin [2], C. Goeury [2], N. Claude [2], P. Tassi [2,3] and K. El kadi Abderrezzak [2,3]

[1]École Nationale Supérieure d'Électronique, Informatique, Télécommunications, Mathématiques
et Mécanique, France (ENSEIRB-MATMECA)

[2]EDF R&D National Laboratory for Hydraulics and Environment (LNHE)

[3]Saint-Venant Laboratory for Hydraulics

Chatou, France

Email: remsophia.mouradi@gmail.com

Abstract—In this study, the modules TELEMAC-2D and SISYPHE of the Telemac-Mascaret Modelling System (TMS) have been used in combination with the OpenTURNS library (www.openturns.org) to perform an uncertainty quantification analysis of two-dimensional morphodynamic problems. Open-TURNS is a scientific library usable as a Python module dedicated to uncertainties treatment.

A recently implemented API (Application Program Interface) allowed the communication between OpenTURNS and TELEMAC-2D/SISYPHE, and therefore the efficient implementation of Monte-Carlo like algorithms. In this problem, each uncertain sedimentological parameter, e.g. inlet mean diameter, Shields parameter, etc. has been associated to a statistical distribution, defined with OpenTURNS. A number of TELEMAC-2D/SISYPHE simulations has been proposed regarding the pre-defined random entries in order to guarantee the convergence of the Monte Carlo-like algorithms.

This work allowed the implementation of uncertainty quantification analysis of computationally intensive morphodynamic simulations in the TMS. Thanks to the access to computer resources and optimized software, we were able to perform the uncertainty quantification analysis with a large set of variables, and therefore push the study further with the correlations effects analysis.

Keywords: Uncertainty quantification, Morphodynamic modelling, API, Monte-Carlo like algorithms, Sensitivity analysis

I. INTRODUCTION

Morphodynamic simulations have been increasingly used in the last few decades to model the bed evolution in rivers, coasts and estuaries. In this context, most of the equations are empirical and the parameters involved in the calculations are generally poorly defined in literature. The impact of the uncertainties related to those parameters remains unknown.

In order to quantify the impact of inputs uncertainties on simulations results, an uncertainty study is conducted. Ranking the variables in terms of influence allows to orientate the investigations when performing measurements or calibrating parameters for the simulations. In this study, the uncertainty quantification is applied to SISYPHE [16], a sediments transport module, coupled with TELEMAC2D [15] for hydrody-namics, that integrate the TELEMAC-MASCARET modelling system.

The Monte Carlo method is used to propagate the uncertainties through SISYPHE. This approach requires random generation of several configurations of inputs, using their probability distributions. Successive deterministic model simulations are then submitted to generate a set of responses that correspond to the set of inputs.

In order to have total control over the simulation's parameters and therefore conduct an efficient uncertainty study, an API (Application Program Interface) is developed for SISYPHE. This interface, when coupled with TELEMAC2D's available API, makes running hundreds of cases simultaneously possible through a cluster, taking total benefit from the available processors. Running an optimal number of cases guarantees the statistics convergence.

Finally, The pre-processing of uncertain data, as well as the post-processing of the results, are done using the OpenTURNS uncertainty library [2].

This paper is organized as follows: a description of the context and general goals, followed by the present study objectives are given. Section III deals with the API's implementation and coupling with TELEMAC2D. Section V discusses the uncertainty quantification steps and alternates theory and results for each of these. In this section, a sensitivity analysis followed by an uncertainty propagation are investigated. Correlations between variables are also studied using copulas and an ANCOVA (ANalysis of COVAriance) sensitivity analysis is applied. In the last section, conclusions and perspectives are drawn.

II. CONTEXT AND GOALS

This study is set out in the context of EDF numerical tools development. EDF's R&D National Laboratory for Hydraulics and Environment department (LNHE) uses the TELEMAC-MASCARET system to simulate complex hydro-environmental phenomenons (such as dam breaks and flooding risks) in order to anticipate the risks related to electrical production.

TELEMAC-MASCARET results are therefore expected to produce highly reliable results. However, a great number of parameters used in these studies, specifically in morphodynamic simulations, can be set by the user and are uncertain. In order to determine the uncertain parameters impact on the system's result, an uncertainty quantification study is necessary.

In this context, an uncertainty quantification study is conducted in the morphodynamic simulation module SISYPHE. To make this study possible, an API (Application Program Interface) is implemented and coupled to TELEMAC2D, to guarantee the inter-operability with SISYPHE.

III. API'S DEVELOPMENT

The work here focused on the implementation of SISYPHE's API and its coupling to TELEMAC2D's already available API via FORTRAN modules. The API's main goal is to have control on a simulation while running a case. For example, it must allow the user to stop the simulation at any time step, retrieve some variables values and change them if necessary. SISYPHE's API is contained in the source folder of the TMS, as shown in Figure 1, and can therefore call all SISYPHE's subroutines to conduct morphodynamic simulations.
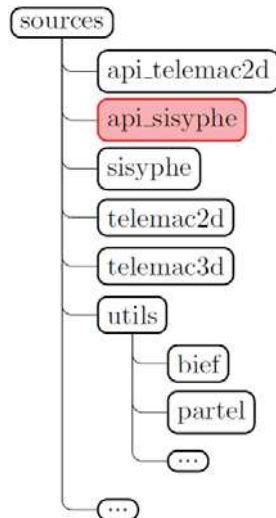

Fig. 1: Extract of TELEMAC-MASCARET sources folders

In order to make this possible, a FORTRAN structure called instance was developed in the API. It contains a list of variables declared as pointers (memory addresses [8]) that are pointing to SISYPHE's variables. This gives direct access to the physical memory of variables, and allows therefore to retrieve their values, and modify them. Furthermore, modifications have been made in SISYPHE's main subroutines to make morphodynamic cases execution possible time step by time step. Finally, parallel runs have also been treated.

In addition to this, to make running coupled cases via the API possible, a communication interface is developed in FORTRAN. This interface contains communication subroutines that send TELEMAC2D's variables to SISYPHE and vice-versa. It also contains subroutines that manage coupled cases and take into consideration the coupling period.

A number of modifications in TELEMAC2D and SISYPHE sources were necessary. These modifications, along with the API and the coupling developments, were validated using three different compilers (NAG, IFORT, GFORTRAN), on classical SISYPHE cases and coupled TELEMAC2D-SISYPHE cases, available in the system.

IV. UNCERTAINTY TREATMENT LIBRARY OPENTURNS

OpenTURNS (Open source initiative to Treat Uncertainties, Risks'N Statistics) [2] is an open source C++ Library for uncertainty treatment used through python scripts. It is co-developed since 2005 by EADS IW, EDF R&D and PHIMECA Engineering. Various statistical methods are implemented in this library and allow to follow the uncertainty study steps [1] represented in Figure 2.
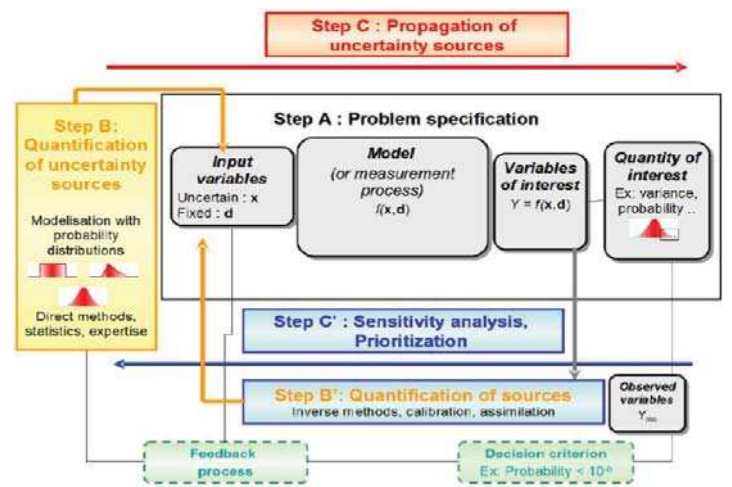

Fig. 2: Steps for an uncertainty study

V. UNCERTAINTY QUANTIFICATION

A. Problem specification

1) Hydrodynamic model: The hydrodynamic phenomenons in two-dimensional fields for free-surface and shallow water flows are modelled using the Saint-Venant equations, which are an integrated form of Navier Stokes 2D equations [9]. Saint-Venant equations are the following:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \nabla \cdot (hu) = 0 \\[2mm] \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} = -g\dfrac{\partial Z_s}{\partial x} + F_x + \dfrac{1}{h}\nabla \cdot (h\nu_t \nabla u) \\[2mm] \dfrac{\partial v}{\partial t} + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} = -g\dfrac{\partial Z_s}{\partial y} + F_y + \dfrac{1}{h}\nabla \cdot (h\nu_t \nabla v) \end{cases}$$
(1)

Where u and v are the velocity components, h the water depth, Z_s the free surface elevation, F_x and F_y are source terms (for example the friction) and ν_t the eddy viscosity.

2) Morphodynamic model: The main goal of a morphodynamic simulation is to predict the bed evolution of a given domain considering the flow conditions. Two types of sediment transport exist [6]:

- **Bed-load** : sediments are transported without losing contact with the bed. Their speed is therefore lower than the flow velocity.

- **Suspension** : corresponds to sediment transport in the absence of direct contact with the bottom. Sediments are here transported with a speed equal to the velocity of the flow.

In this study, only bed-load uncertain parameters are investigated. The main equation for this phenomenon is Exner's equation 2, that calculates the evolution of the bottom elevation Z_f.

$$(1 - \lambda)\frac{\partial Z_f}{\partial t} + \nabla \cdot \overrightarrow{q_b} = 0 \qquad (2)$$

Where λ is the porosity of the sediment, $\nabla\cdot$ the divergence operator and $\overrightarrow{q_b} = q_b \times (\cos \alpha_t \overrightarrow{e_x} + \sin \alpha_t \overrightarrow{e_y})$, α_t being the angle between the flow direction and the sediment transport direction and q_b the bed-load transport rate per width.

The transport rate q_b can be calculated using several empirical formulas. In this study, we focus on Meyer-Peter and Müller formula [13], as written below:

$$\frac{q_b}{\sqrt{g(s-1)d^3}} = \alpha_{MPM}(\theta - \theta_c)^{3/2} \qquad (3)$$

- α_{MPM} being a transport coefficient specified by the user;

- d the sediments mean diameter;

- $s = \rho_s/\rho_f$ the relative density, where ρ_s is the sediments density and ρ_f the fluid density;

- θ and θ_c resp. the Shields number and its critical value indicating the movement's threshold.

The Shields number θ is the dimensionless shear stress τ calculated as in equation 4. The threshold θ_c is given as a parameter of the simulation by the user.

$$\theta = \frac{\tau}{g(\rho_s - \rho)d} \qquad (4)$$

3) Summary of the uncertain parameters: The porosity λ, the mean diameter d, the Shields critical parameter θ_c and five other parameters, described in [16], are considered as uncertain for the current study. A summary of these variables and their definitions are given in Table I Section V-B.

4) Study cases: The uncertainty quantification is conducted for two distinct cases.

Experimental case - Channel : The channel case is a numerical reproduction of Ashida and Michiue's experiment [18] in which the erosion downstream of a dam is studied. The case, illustrated in Figure 3, is modelled in 2D with boundary conditions on flow discharge (Red boundary in Figure 3 - $Q = 0.0314\ m^2 \cdot s^{-1}$) and imposed water depth (Blue boundary in Figure 3 - $H = 0.06\ m$), as well as walls on the remaining boundaries. The channel has a $20\ m$ length, $0.8\ m$ width and $0.2\ m$ slope. An experimental duration of 10 hours is simulated using a time step $\Delta t = 0.1\ s$ within 1 minute of sequential TELEMAC2D/SISYPHE calculations on an Intel-Xeon(R) 3.40GHz core.
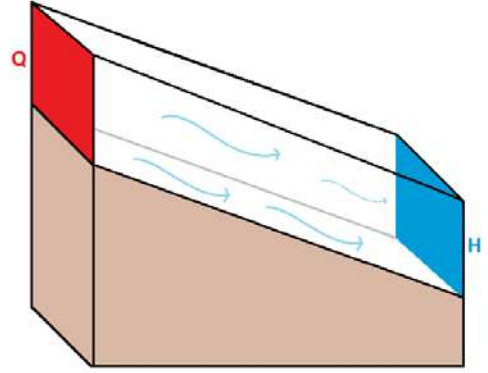

Fig. 3: Ashida et Michiue channel experiment representation

Real case - Bifurcation : A real case of a bifurcation between the Colastiné river in Argentina and a channel access to the Santa Fe harbour is also studied [14], as illustrated in Figure 4. For this case, a flow discharge $Q_1 = 2416.42\ m^2 \cdot s^{-1}$ and two water depths $H_1 = 13.053\ m$ and $H_2 = 13\ m$ are set as boundary conditions. A real duration of 10 days is simulated, with a time step of $\Delta t = 20\ s$, within 1 hour of sequential calculations on an Intel-Xeon(R) 3.40GHz core.
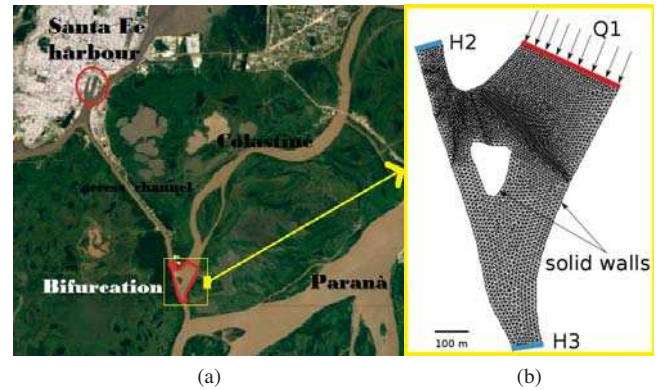

(a)   (b)

Fig. 4: Study domain satellite image (a) and mesh (b)

B. Uncertainty sources quantification

The goal here is to define variation intervals and probability density functions (PDF) for each uncertain parameters. We remind that, for a random variable X defined on an interval $[a,b]$, a PDF $f(x)$ is defined as follows :

$$\mathbb{P}(X \in [a,b]) = \int_a^b f(x)dx \qquad (5)$$

For all the variables from 2 to 8 defined in Table I, variation intervals were found in literature, without further information about their probabilities. Consequently, using the principle of maximum entropy [3], uniform probability density functions are chosen for those variables.

For the experimental channel case, only one mean diameter measure is available and is subject to measurements errors. The errors interval is considered as a variation interval and a uniform PDF is applied for the mean diameter on its measurement interval. For the real bifurcation case, several samples of sediments are extracted in different sections of

the river. Different values of the mean diameter are possible. In order to take all of them into consideration, a PDF that corresponds to the sample is selected and validated via the QQ-Plot method [2].

The summarized variables and their probability density functions are presented in Table I.

Variable	Definition	PDF
d	Sediments mean diameter	Log-Normal(Field Data) / Uniform(Measure errors)
α_{MPM}	Meyer-Peter and Müller transport coefficient	Uniform[5,15]
θ_c	Shields critical parameter	Uniform[0.03,0.06]
λ	Porosity	Uniform[0.25,0.4]
α_S	Skin friction coefficient	Uniform[1.0,6.6]
Φ_S	Angle of repose of the sediments (Slope effect on transport direction)	Uniform[30,80]
β_2	Deviation parameter (Slope effect on transport amplitude)	Uniform[0.1,5.0]
α_c	Secondary currents coefficient	Uniform[0.75,1.0]

TABLE I: Uncertain parameters and their PDFs

C. Sensitivity analysis

1) Monte-Carlo Sampling and statistical estimations: 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 $x^i = (x_1, ..., x_V)^i$ represents a possible configuration for the coupled hydro-morphodynamic simulation. Corresponding realizations of the output are generated by successive deterministic simulations with each configuration of the inputs. Statistical estimators of the response $Y = (Y_1, ..., Y_N) = (M(x^i))_{i \in \{1,...,N\}}$ can therefore be computed from the output as follows :

Mean :
$$E[Y] = \mu_Y = \frac{1}{N} \sum_{i=1}^{N} M(x^i) \qquad (6)$$

Variance :
$$Var(Y) = \frac{1}{N-1} \sum_{i=1}^{N} [M(x^i) - \mu_Y]^2 \qquad (7)$$

Standard deviation : $\quad \sigma_Y = \sqrt{Var(Y)} \qquad (8)$

These statistical moments are useful for both the uncertainty sensitivity analysis and uncertainty propagation.

The convergence order of the Monte-Carlo sampling method is given by the Central Limit Theorem [7] as $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. The estimated statistics are also random quantities and are impacted by estimation uncertainties. Confidence intervals on estimators should therefore be calculated. The non-parametric "bootstrap" method provides information about the statistics uncertainties given few hypothesis [11]. Let $x = (x_1, ..., x_N)$ denote a sample of N independent and identically distributed realizations according to a probability density function $f(x)$. The statistical moment $\theta = T(F)$ (mean, variance, etc.), is estimated by $\hat{\theta} = T(\hat{F})$, where $\hat{F}$

is the empirical cumulative density function that gives equal probability $\frac{1}{N}$ to each observed value x_i defined by :

$$\hat{F}(x) = \frac{1}{d} \sum_{i=1}^{N} 1_{x_i \leq x} \qquad (9)$$

The idea of the non-parametric bootstrap is to simulate data from the empirical cumulative density function. Given that $\hat{F}$ is build upon equal probability for the observations $(x_1, ..., x_N)$, a sample of same size N from $\hat{F}$ would simply be a selection from $(x_1, ..., x_N)$ with repeated values. A number of B samples are generated following this strategy, and estimators properties can therefore be deduced as shown in Figure 5.
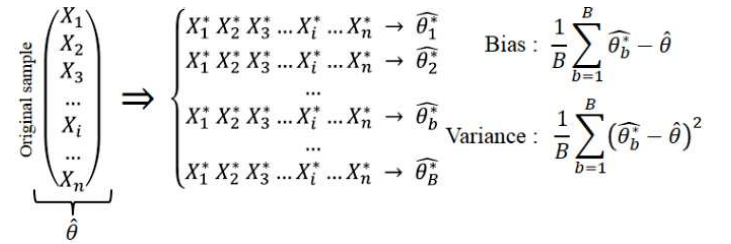

Fig. 5: Bootstrap algorithm [11]

The confidence interval I_θ is then estimated as follows :

$$I_\theta = [0.025 - q_{\hat{\theta}_b}, 0.975 - q_{\hat{\theta}_b}] \qquad (10)$$

Where $\alpha - q_X$ is the α-quantile of a variable X defined as :

$$\mathbb{P}(X \leq q_X(\alpha)) = \alpha \quad \forall \alpha \in [0, 1] \qquad (11)$$

2) Analysis of variance: The main goal of a sensitivity analysis is to rank the uncertain parameters according to their influence. In order to do so, a definition of ranking indices is necessary. The indices used here are called Sobol Indices [10].

The definition of Sobol Indices is a result of the ANOVA (Analysis Of VAriance) variance decomposition. In fact, given a set of V **independent** uncertain parameters $X = (X_1, ..., X_V)$, the variance of a response $Y = M(X)$ can be calculated, using the total variance theorem [4], as follows [10]:

$$Var[Y] = \sum_{i=1}^{V} \mathcal{V}_i(Y) + \sum_{i<j} \mathcal{V}_{ij}(Y) + ... + \mathcal{V}_{12..V}(Y) \qquad (12)$$

Where $\mathcal{V}_i(Y) = Var[E(Y|X_i)]$ and $\mathcal{V}_{ij}(Y) = Var[E(Y|X_i X_j)] - \mathcal{V}_i(Y) - \mathcal{V}_j(Y)$, etc.
$E(Y|X_i)$ is Y's conditional expectation with the condition that X_i remains constant.

One can therefore calculate first order sensitivity indices that estimate the influence of a variable X_i without its interactions with other variables:

$$S_i = \frac{\mathcal{V}_i(Y)}{Var[Y]} = \frac{Var(E[Y|X_i])}{Var[Y]} \qquad (13)$$

And total indices that estimate the global influence of a variable (including interactions):

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i, k \neq i, j<k} S_{ijk} + ... = 1 - \frac{\mathcal{V}_{-i}(Y)}{Var[Y]} \qquad (14)$$

$\mathcal{V}_{-i}(Y)$ being conditional expectations variances that do not involve X_i.

There are two ways to estimate the Sobol indices defined in equations 13 and 14 :

SALTELLI method [7]: Step 1: Two independent samples A and B of size N are generated for the V uncertain variables. For example, sample A can be written as follows:

$$A = \begin{pmatrix} x_1^{A,1} & x_2^{A,1} & ... & x_V^{A,1} \\ x_1^{A,2} & x_2^{A,2} & ... & x_V^{A,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{A,N} & x_2^{A,N} & ... & x_V^{A,N} \end{pmatrix} \quad (15)$$

A new sample C is created using columns of B except for column i that is replaced with data from A:

$$C = \begin{pmatrix} x_1^{B,1} & ... & x_i^{A,1} & ... & x_V^{B,1} \\ x_1^{B,2} & ... & x_i^{A,2} & ... & x_V^{B,2} \\ \vdots & & \vdots & & \vdots \\ x_1^{B,N} & ... & x_i^{A,N} & ... & x_V^{B,N} \end{pmatrix} \quad (16)$$

Simulations using the samples A, B and C result with :

$$\begin{cases} y_k^A = M((x_1^{A,k}, ..., x_V^{A,k})) & k = \{1, N\} \\ y_k^B = M((x_1^{B,k}, ..., x_V^{B,k})) & k = \{1, N\} \\ y_k^C = M((x_1^{C,k}, ..., x_V^{C,k})) & k = \{1, N\} \end{cases} \quad (17)$$

Which can be used to estimate Sobol indices as follows:

$$S_i = \frac{\frac{1}{N} \sum_{k=1}^{N} y_k^A y_k^C - (\mu_{Y^A})^2}{\sigma_{Y^A}^2} \quad (18)$$

$$S_{Ti} = 1 - \frac{\frac{1}{N} \sum_{k=1}^{N} y_k^B y_k^C - (\mu_{Y^B})^2}{\sigma_{Y^B}^2} \quad (19)$$

Overall, for a given sample size N, $(V + 2) \times N$ simulations are necessary to estimate the first order and total Sobol indices for each variable X_i.

Polynomial chaos method (PCE): The models response can be approached by an analytical function :

$$M(X) = M_0 + \sum_{i=1}^{V} M_i(X_i) + \sum_{1 \le i < j \le V} M_{i,j}(X_i, X_j) \\ + ... + M_{1,...,V}(X_1, ..., X_V) \quad (20)$$

$M_i(X_i)$ represents the variable X_i's contribution to the result of the simulation. The variance can therefore be calculated as follows [4]:

$$Var[Y] = \sum_{u \subseteq \{1,...,V\}} [Var[M_u(X_u)] + \\ \sum_{v \subseteq \{1,...,V\}, v \cap u = \emptyset} Cov[M_v(X_v), M_u(X_u)]] \quad (21)$$

In this particular case of independent variables, the covariance term $Cov[M_v(X_v), M_u(X_u)]$ vanishes, and the decomposition becomes then equal to ANOVA. Sobol indices can therefore be estimated as:

$$S_i = \frac{Var[M_i(X_i)]}{Var[Y]} \quad (22)$$

$$S_{Ti} = \frac{\sum_{u \subseteq \{1,...,V\}, i \in u} Var[M_u(X_u)]}{Var[Y]} \quad (23)$$

The contributions $M_i(X_i)$ can be calculated by estimating the models response using a polynomial chaos expansion, which can, in a simplified way, be written as:

$$M(X) = \sum_{|\alpha| \le P} a_\alpha \Psi_\alpha(X) \quad (24)$$

Where $\{\Psi_\alpha, \alpha \in \mathbb{N}^V\}$ is a multivariate polynomial basis and a_α adequate coefficients for the estimation of the model's response, that can be determined using projection methods [4].

The X_i-univariate polynomials shares are the exact contribution of X_i to the polynomial expansion, and are therefore an estimation of $M_i(X_i)$.

3) Results:

Channel's results: For a sample of size $N = 30000$, a number of $(2+8) \times 30000$ simulations are necessary to estimate Sobol indices via a SALTELLI method. Up to 500 cores are used to launch the simulations, for a calculation time of 20 hours overall. The observed variances of Z_f in m^2 and some interest points (1, 86 and 150) are given in Figure 6.
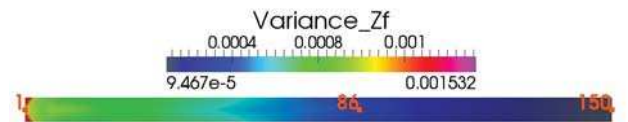

Fig. 6: Channel case - Z_f variances implied by 8 uncertain parameters

First order and total Sobol indices are therefore calculated using the SALTELLI method, and their confidence intervals estimated via Bootstrap. The results are given for the point 86 in Figure 7. Similar results are observed for the other interest points.
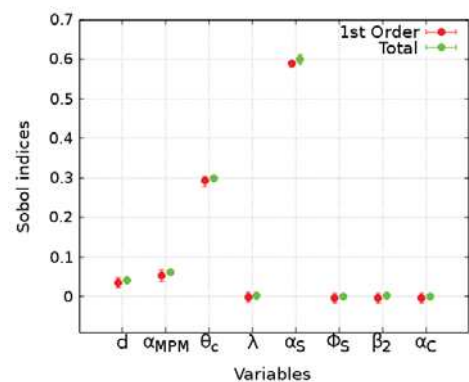

Fig. 7: Channel case - Sobol indices for Point 86

The skin rugosity coefficient α_S is by far the most influencing variable, followed by θ_c the Shields critical parameter. The influence of other variables can be considered negligible.

The SALTELLI method and the PCE method for the estimation of Sobol indices are compared in the interest points for the same sample size $N = 30000$, as shown in Figure 8.
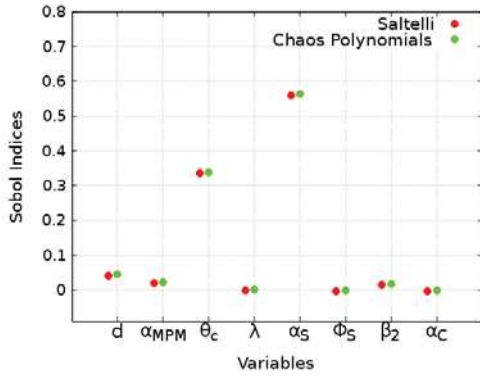
Fig. 8: Channel case - Comparison of Sobol indices calculation using SALTELLI and PCE methods for Point 1

Both methods give equal estimations of Sobol indices for $N = 30000$. Given that the PCE method needs a total of N runs to estimate Sobol indices, which is 10 times less than the cost of SALTELLI method, it will be used from now on.

Bifurcation's results: For the bifurcation application, a sample of size $N = 12800$ is used . An overall 19 hours of simulation using 800 cores is necessary. The observed variances are given in Figure 9.
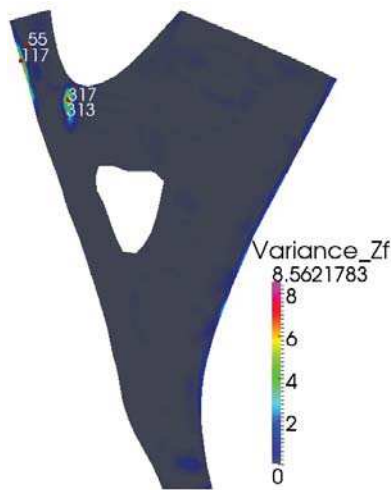

Fig. 9: Bifurcation - Variances of Z_f

Sobol indices are estimated in interest points like 317 (Figure 9) using the PCE method, as shown in Figure 10.
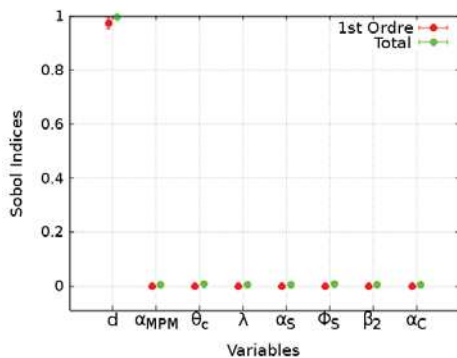

Fig. 10: Bifurcation - Sobol indices for point 317

For this case, the observed variances seem to result only from the diameters uncertainty. All the other uncertain variables seem to be none influencing.

D. Uncertainty propagation

In this section, the impact of the parameters uncertainty on the model's response will be investigated using statistics defined in section V-C1, like the mean and the variance.

1) Channel's results: The sensitivity analysis has shown a great difference between the observed results for an experimental and a real case. In fact, for a real case, the mean diameter of sediments seems to be the only influencing parameter, while it is nearly meaningless for an experimental case. In order to investigate the influence of the sediment diameter, a new set of $N = 30000$ uncertainty runs is launched, considering as an only uncertain parameter the mean diameter (other parameters are constant). The observed statistics (mean in m and variance in m^2) are given in Figure 11.
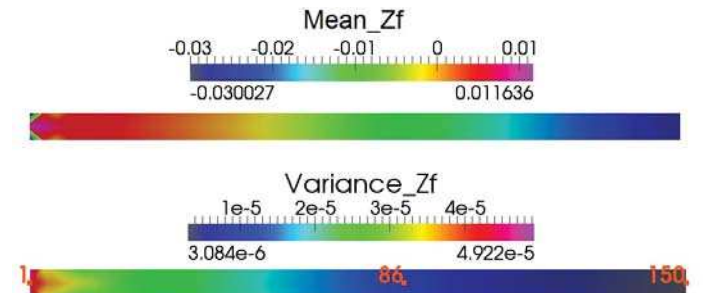

Fig. 11: Channel case - Statistics of Z_f for $N = 30000$

Uncertainties seem to be more important in the upstream direction, which corresponds to a higher erosion upstream and therefore bigger chances of bottom variations. This could also be confirmed through comparing the statistical mean of the final bottom elevations to a deterministic result with the statistical mean of sediment diameters on the channel's center line, as shown in Figure 12. The statistical mean of final bottom elevation is closer to the final deterministic result in the downstream direction.
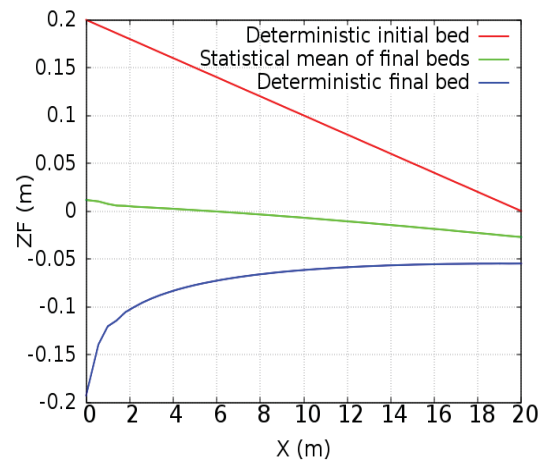

Fig. 12: Evolution of the bottom in the channel's center line

To quantify the uncertainty propagation, a statistical number called variation coefficient is defined in equation 25.

$$c_v(Z_f) = \left| \frac{\sqrt{Var(Z_f)}}{\mu(Z_f)} \right| = \left| \frac{\sigma(Z_f)}{\mu(Z_f)} \right| \quad (25)$$

Variation coefficients c_v are evaluated in interest points with high and low variances (points 86 and 150 in Figure 11), and their convergence investigated with the Bootstrap method, as shown in Figures 13 and 14.
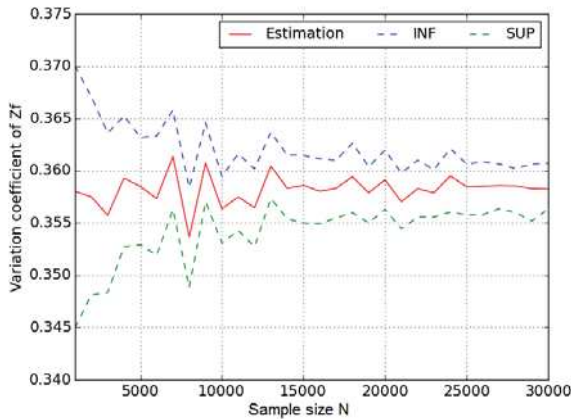

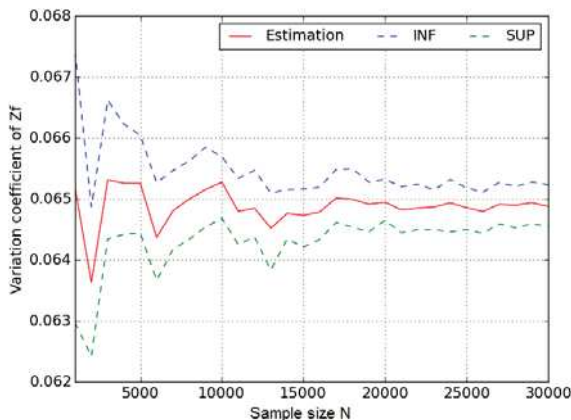Fig. 13: Variation coefficient for point 86

Fig. 14: Variation coefficient for point 150

In comparison with a variation coefficient of $c_v = 0.07$ for the diameter's sample, the uncertainty seems to be highly propagated in some interest points as point 86. This in fact confirms that even if the diameter seems to be none influencing while conducting the sensitivity analysis, its uncertainty is still propagated. In other interest points like 150, the variation coefficient of the output is lower than that of the input. In both cases (low and high c_v) the results have converged at about $N = 16000$. The confidence interval (INF and SUP in Figures 13 and 14) for $c_v = 0.36$ is ±0.005 and for $c_v = 0.06$ is ±0.001. The convergence of the statistical estimations and the narrow confidence intervals indicate that our calculations can be trusted for the chosen size $N = 30000$.

2) Bifurcation's results: For the Bifurcation case, there is no need to investigate the diameter as an only uncertain

parameter, as it is influencing the result at about 100%. The same calculations as in sensitivity analysis have been exploited to estimate the statistics in Figure 15.
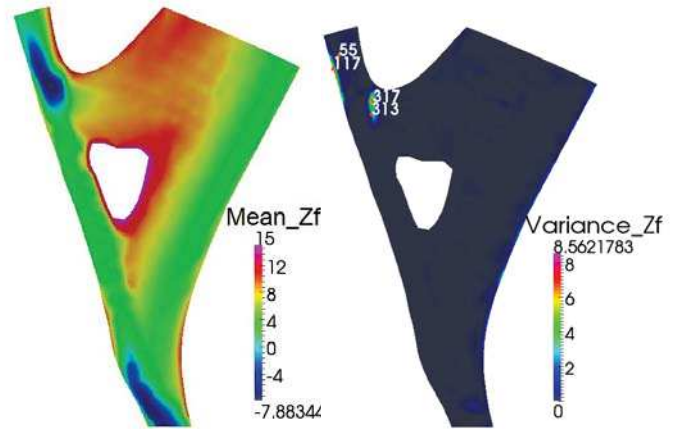

Fig. 15: Bifurcation - Statistics for Z_f - N = 12800

The highest variances seem to be concentrated in specific areas (for example, interest points 317 and 313 in Figure 15). This could be interpreted as the result of higher movements in this areas, which is demonstrated in Figure 16, where the initial and final bottom elevations have been shown for point 317 with a diameter of $d = 0.149mm$. This diameter corresponds to the minimum value of the studied sample. For the maximum value, no bed evolution has been observed.
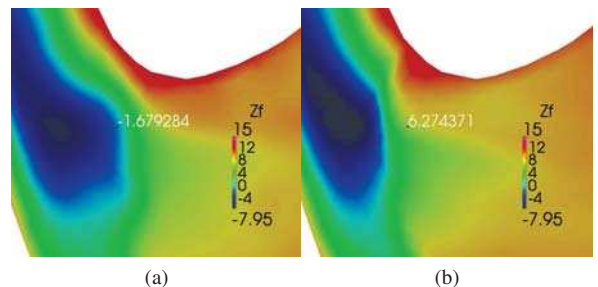

(a)   (b)

Fig. 16: Z_f in point 317 for d=0.149mm - First (a) and last iteration (b)

This difference of movement is explained by the threshold defined by the Shields parameter, which most certainly doesn't allow the maximum diameters sediments to be transported. For $d = 0.149mm$, a sedimentation area around point 317 is observed, which is confirmed by the bottom elevation values, going from $Z_f = -1.68m$ in the beginning of simulation to $Z_f = 6.27m$ at the end.

Nevertheless, interpreting the variances using its minimum-maximum scale can give a false idea about where variances shall appear. In fact, one can interpret that there are no other variances areas than close to points 317 and 313, which is not true and can be proved with rescaling the variances as shown in Figure 17.

Furthermore, the variation coefficients c_v have also been estimated for this case in low and high variance points and their convergence studied. The same conclusions can be drawn.

Fig. 17: Bifurcation - reduced scale of variations for Z_f

For a variation coefficient of the diameter being $c_v = 0.42$, uncertainty is more or less propagated in different points. For example, in point 317 (Figure 18), $c_v = 1.879$ which corresponds to a high uncertainty propagation. For point 55 (Figure 19), $c_v = 0.137$ which is far less than the variation coefficient of the diameter's sample and corresponds to an uncertainty that is not propagated.
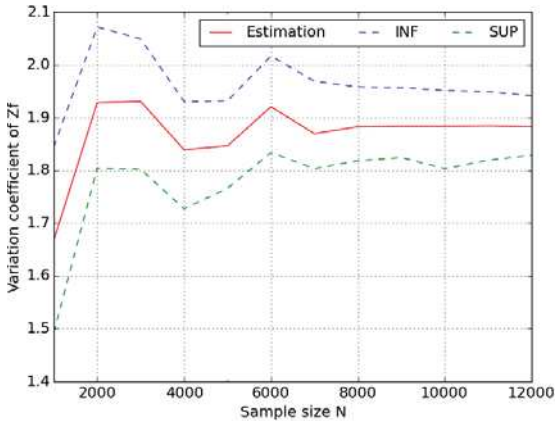

Fig. 18: Bifurcation -Variation coefficient point 317

The variation coefficients converge at about $N = 8000$ for both cases, which signifies that the results of this study can be
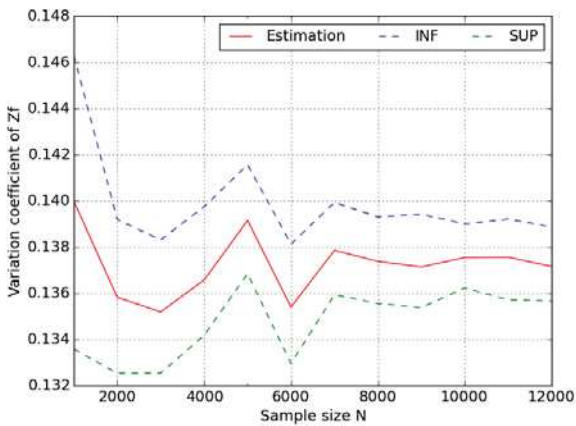

Fig. 19: Bifurcation -Variation coefficient point 55

trusted as $N = 12800$. Furthermore, the confidence intervals are ± 0.1 for $c_v = 1.879$ and ± 0.002 for $c_v = 0.137$ and are tightening with the increasing sample size.

E. Correlations impact

Several physical relationships can exist between the considered uncertain parameters. These correlations between the variables can impact the uncertainty study, given that the parameters are no longer sampled independently. In order to model the correlations, copulas are introduced in section V-E1. For the sensitivity study, the ANOVA method can no longer be used, because of the independent parameters hypothesis it implies. A new method called ANCOVA (ANalysis of COVAriance) is introduced in section V-E2. The uncertainty study is conducted with the correlations consideration on the channel's case. In fact, given that the only influencing parameter in the Bifurcation is the diameter, it has been concluded that the correlation study for this case would be useless.

1) Copulas for correlations modelling: A copula is a function that defines a dependency structure between a set of variables [5]. In fact, it links the multivariate probability density function of random set of variables $(X_1, ..., X_V)$ to their univariate probability density functions.

A copula is a V-dimensional function C defined on $[0, 1]^V$ that verifies:

- $\forall u \in [0, 1]^V \ \forall i \in [|1 : V|]$, if $u_i = 0$ then $C(u) = 0$

- $\forall i \in [|1 : V|]$ and $u_i \in [0, 1]$, $C(1, ..., 1, u_i, 1, ..., 1) = u_i$

- $\forall u, v \in [0, 1]^V$ verifying $\forall i \in [|1 : V|] \quad u_i \leq v_i$ then $V_C([u, v]) \geq 0$

Where $V_C([u, v])$ is the C-volume of the space $[u_1, v_1] \otimes \ldots \otimes [u_V, v_V]$ defined as follows :

$$V_C([u, v]) = \Delta_{u_V}^{v_V} ... \Delta_{u_1}^{v_1} C(\mathbf{w})$$

$\Delta_{u_i}^{v_i}$ being the ith finite differential:

$$\Delta_{u_i}^{v_i} C(\mathbf{w}) = C(w_1, ...w_i, v_i, w_{i+1}, ..., w_V) \\ - C(w_1, ...w_i, u_i, w_{i+1}, ..., w_V)$$

The Sklar theorem [5] allows to define a relation between the multivariate PDF f_X of the set $X = (X_1, ..., X_V)$ and the univariate probability density functions f_i of X_i as follows :

$$f_X(x_1, ..., x_V) = c(F_1(x_1), ..., F_V(x_V)) \times \prod_{i=1}^{V} f_i(x_i) \quad (26)$$

Where F_i are the univariate cumulative distribution functions of X_i associated to the probability density functions f_i. On the other hand, c is the probability density function of the copula C defined as follows:

$$\forall \mathbf{u} \in [0, 1]^V \quad c(u_1, ..., u_V) = \frac{\partial^V C}{\partial u_1 ... \partial u_V}(u_1, ..., u_V) \quad (27)$$

In this study, a classical Gaussian copula is used [2]. It requires the calculation of a correlation Matrix using Spearman

indices [2] from the relationships between variables. The relationships modelled here are the following:

- The empirical relation between the Meyer-Peter and Müller coefficient and the Shields number, that was introduced by Wiberg and Smith [17] as follows:

$$\alpha_{MPM} = 9.64 \times \theta^{0.166} \tag{28}$$

Which implies a correlation between α_{MPM}, α_S and d.

- The relationship between the Shields critical parameter and the dimensionless sediments diameter [16].

- The modified Komura porosity formula [19]:

$$\lambda = 0.13 + \frac{0.21}{(d + 0.002)^{0.21}} \tag{29}$$

- The following relationship between the deviation parameter and the sediments diameter [12]:

$$\beta_2 = 9 \left(\frac{d}{H} \right)^{0.3} \tag{30}$$

2) Analysis of covariance: For dependent variables, it is possible to calculate the variance with the ANCOVA decomposition as follows:

$$Var[Y] = \sum_{u \subseteq \{1, \ldots, V\}} [Var[M_u(X_u)]$$
$$+ \sum_{v \subseteq \{1, \ldots, V\}, v \cap u = \emptyset} Cov [M_v(X_v), M_u(X_u)]] \tag{31}$$

New sensitivity indices can be defined as:

$$\begin{cases} S_i^U = \dfrac{Var[M_i(X_i)]}{Var[Y]} \\[2mm] S_i^C = \dfrac{Cov\left[M_i(X_i), \sum_{v \subseteq \{1,\ldots,V\}, v \cap \{i\} = \emptyset} M_v(X_v)\right]}{Var[Y]} \\[2mm] S_i = S_i^U + S_i^C = \dfrac{Cov[M_i(X_i), Y]}{Var[Y]} \end{cases} \tag{32}$$

Where S_i is the total influence of the variable X_i, S_i^U the uncorrelated part of influence and S_i^C the correlated part.

ANCOVA indices can be negative because of the covariance term. In order to interpret the signification of negative indices, their absolute values are compared. In fact, if $|S_i^C|$ has a high value, this means that S_i^U is close to S_i, which signifies that correlations of the variable X_i have weak influence on the result. Inversely, if it has a high value, this means that correlations of X_i have great impact on the simulation's result.

Lastly, as show in section V-C2, the terms $M_i(X_i)$ can be estimated via the polynomial chaos expansion. However, in order to guarantee the orthogonality of the polynomial chaos basis, it is necessary to estimate the coefficients of the expansion using uncorrelated entries X. The M_i terms are estimated afterwards using the correlated values of the entries.

3) Results: The ANCOVA indices are compared to the Sobol (ANOVA) ones in order to quantify the impact of correlations on the sensitivity analysis, as shown in Figure 20.
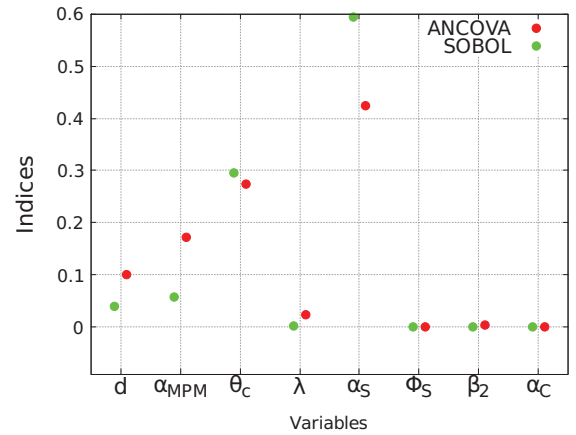

Fig. 20: Sensitivity indices for point 86

In this case, the global order of influence doesn't change. The gap between the variables in terms of influence is reduced, for example between the variables α_{MPM} and θ_C. Furthermore, some of the variables that seemed to be of weak influence at first are now considerably influencing, for example α_{MPM}. Other variables for which no correlation was considered didn't change their sensitivity indices (eg. α_C). In order to estimate the part of the correlations influence in the total ANCOVA influence, the comparative Figure 21 is drawn.
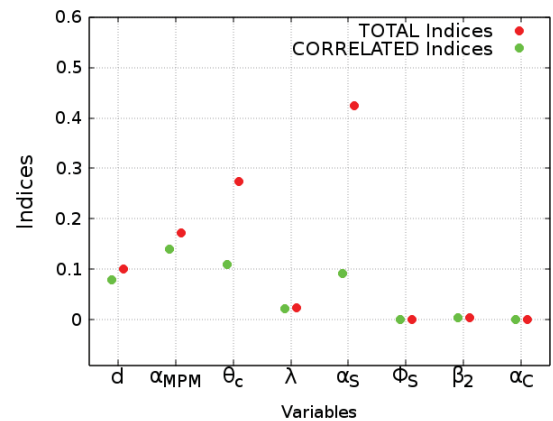

Fig. 21: ANCOVA Indices Vs correlated ANCOVA indices

It can be observed that variables for which the sensitivity indices have considerably increased (d and α_{MPM}) owe all their new influence to the correlations.

Other variables that were initially influencing also have part of their influence that is due to the correlations (θ_C and α_S).

Furthermore, the variances with independent uncertain parameters and correlated uncertain parameters are compared as shown in Figure 22.

It is clearly noticed that variances considerably increase when adding correlations in this case. It can be explained by observing the new variation interval of α_{MPM} for example.
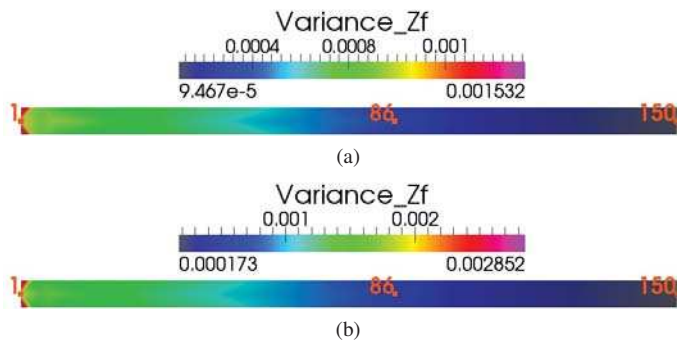
Fig. 22: Uncertainty propagation without (a) and with (b) correlations for the channel's case

In fact, the transport coefficient of the Meyer-Peter and Müller formula initially varies between the values 5 and 15. When correlations are added, α_{MPM} varies between 9.5 and 11, which implies that we statistically observe more movement (erosion) as compared to cases with no correlations.

VI.   CONCLUSION AND PERSPECTIVES

In this study, an uncertainty quantification of a morphodynamic problem has been proposed.

In a sensitivity analysis step, differences were observed between a real case and an experimental case. In fact, for the real bifurcation case, the diameter was the only influencing parameter, which is not the same for the channel.

In order to analyse the influence of sediments diameter on the model's response, an uncertainty propagation study was conducted, considering as an only uncertain parameter the sediments diameter. This study has shown that the diameter has highly propagated uncertainties where there is movement. In fact, high variances were observed in maximum erosion points for the channel, and in a deposition zone for the bifurcation.

Finally, correlations were added and increased the variances significantly. An ANCOVA method was implemented in order to conduct a sensitivity analysis. 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) became of considerable influence when adding correlations.

This study can be generalized to other applications, such as the use of different sediment transport formulas, the study of suspended sediment transport or the influence of different physical phenomenons, for example waves (TOMAWAC module in the TELEMAC-MASCARET system).

REFERENCES

[1] A. PASANISI, M. COUPLET and A.H DUTFOY LEBRUN, "Guide méthodologique pour le Traitement des Incertitudes," *Note EDF-MRI, reference H-T57-2013-02207-FR*, 2013.

[2] Airbus-EDF-IMACS-Phimeca, "Reference Guide - OpenTurns 1.7," 2016.

[3] B. SUDRET, "Uncertainty propagation and sensitivity analysis in mechanical models. Contributions to structural reliability and stochastic spectral methods," *Accreditation to supervise research report*, 2007.

[4] Y. CANIOU, "Global sensitivity analysis for nested and multiscale modelling. PhD thesis. Blaise Pascal University Clermont II France," 2012.

[5] G. CHASTAING, "Indices de Sobol généralisés pour variables dépendantes," 2010.

[6] M. GARCIA, "Sedimentation engineering : processes, measurments, modeling, and practice ASCE," 2008.

[7] C. GOEURY and N. GOUTAL, "Analyse de sensibilité et propagation d'incertitude dans les modèles numériques : Application à l'hydraulique," 2014.

[8] H. DELOUIS and P. CORDE, "Cours Fortran 95 - IDRIS," 2007.

[9] J. HERVOUET, "Hydrodynamics of free surface flows, Wileys," 2007.

[10] B. IOSS, "Revue sur l'analyse de sensibilité globale de modèles numériques," 2010.

[11] M. KADIRI-OTTMANI, "Traitement statistique d'un échantillon," *Technical report of CEA*, 2002.

[12] M. VAN ORMONDT, C. BRIERE, D.J.R. WALSTRA, L.C. VAN RIJN and A.M. TALMON, "FThe effects of bed slope and wave skewness on sediment transport and morphology," 2006.

[13] E. MEYER-PETER and R. MÜLLER, "Formulas for Bed-Load Transport," 2004.

[14] P. TASSI, M.I. MORELL and C.A VIONNER, "Flow pattern at a river diffluence at the alluvia system of the Paraná River," *Journal of Applied Water Engineering and Research*, 2014.

[15] R. ATA, C. GOEURY, P. LANG, J. DESOMBRE, J.M. HERVOUET, "TELEMAC MODELLING SYSTEM - Release 7.0 - USER MANUAL," 2014.

[16] P. TASSI and C. VILLARET, "SISYPHE USER'S MANUAL - Release 6.3 - USER MANUAL," 2014.

[17] P. WIBERG and J. SMITH, "Model for calculating bed load transport of sediment," *Journal of Hydraylic Engineering*, 1989.

[18] W. WU, "Depth-averaged two-dimensional numerical modeling of unsteady flow and uniform sediment transport in open channels," *Journal of Hydraylic Engineering*, 2004.

[19] W. WU and S. WANG, "Formulas for sediment porosity and settling velocity," *Journal of Hydraylic Engineering*, 2006.