# A non-intrusive reduced-order model for uncertainty quantification in numerical solution of one-dimensional free-surface water flows over stochastic beds

Alia Alghosoun[1,*] Nabil El Mocayd[2,†] Mohammed Seaid[1,‡]

[1] *Department of Engineering, University of Durham, South Road, Durham DH1 3LE, UK*
[2] *International Water Research Institute, University Mohammed VI Polytechnic, Hay Moulay Rachid, Benguerir, Morocco*

**Abstract**

Free-surface water flows over stochastic beds are complex due to the uncertainties in topography profiles being highly heterogeneous and imprecisely measured. In the present study, the propagation and influence of several uncertainty parameters are quantified in a class of numerical methods for **one-dimensional** free-surface flows. The governing equations consist of both single-layer and two-layer shallow water equations on either flat or non-flat topography. For this purpose, the free-surface profiles are computed for different realizations of the random variables when the bed is excited with sources whose statistics are well defined. Many research studies have been dedicated to the development of numerical methods to achieve some order of accuracy in free-surface flows. However, little concern was given to examine the performance of these numerical methods in presence of uncertainty. The current work addresses this specific area in computational hydraulics with regards to the uncertainty generated from bathymetric forces. As numerical solvers for the **one-dimensional** shallow water equations, we implement four finite volume methods. To reduce the required number of samples for uncertainty quantification, we combine the proper orthogonal decomposition method with the polynomial chaos expansions for efficient uncertainty quantifications of complex hydraulic problems with large number of random variables. Numerical results are shown for several test examples including dam-break problems for single-layer and two-layer shallow water flows. The problem of flow exchange through the Strait of Gibraltar is also solved in this study. The obtained results demonstrate that in some hydraulic applications, a highly accurate numerical method yields an increase in its uncertainty and makes it very demanding to use in an operational manner with measured data from the field. On the other hand, when the complexity of the physics increases, these highly accurate numerical methods display less uncertainty compared to the low accurate methods.

**Keywords.** Uncertainty quantification; Polynomial chaos expansion; Proper orthogonal decomposition; Shallow water equations; Stochastic beds; Finite volume methods

## 1 Introduction

Free-surface models in hydraulics have gained an increasing interest during last decades. Ranging from flood forecasting to monitoring hydraulic infrastructures such as dams, hydraulic simulations have served for several purposes to support decision making. Water free-surface flows under the influence of gravity can be modeled using the well-established shallow water equations subject to the assumption that the vertical depth is much

*alia.r.al-ghosoun@durham.ac.uk
†nabil.elmocayd@um6p.ma
‡m.seaid@durham.ac.uk

smaller than any typical horizontal length within the hydraulic system. The governing equations in these models are derived by depth-averaging the incompressible Navier-Stokes equations, see for example [44]. However, the main drawback resulting from these assumptions remains the lack of capturing some physical dynamics in the vertical motion of the water flow. Therefore, during the recent years, multi-layer shallow water models have attracted enormous attention and have became very useful tools to solve many hydraulic problems such as rivers, estuaries, bays and other nearshore regions where water flows interact with the bed topography and wind shear stresses, see for instance [21, 2, 26]. From a numerical view point, the most challenging features of the single-layer shallow water equations are the fact that they admit discontinuous and smooth solutions. Even for problems in which the initial data is smooth can lead to discontinuous solutions in a finite time. The hyperbolic nature of the system and its nonlinearity are also among the difficulties that arise when solving numerically these equations. On the other hand, numerical treatment of the two-layer shallow water system often presents difficulties due to their nonlinear form, presence of the non-conservative product terms, lost of hyperbolicity, and difficulty to explicitly obtain its eigenstructure, see for example [26]. Many numerical tools have been proposed in the literature in order to overcome these difficulties. In the present study, we only consider the well-established finite volume methods including Lax-Friedrichs, Rusanov and Roe schemes, see for instance [29]. We also adapt the finite volume modified method of characteristics [3] for solving the two-layer shallow water equations. Despite this continuous effort to improve the accuracy of hydraulic simulations, uncertainty remains ubiquitous for different reasons as discussed in [36] among others. Hence, the numerical methods to solve the **one-dimensional** shallow water systems discretize the continuous equations in space and time which will introduce numerical uncertainties. Furthermore, the empirical modelling of the bed bathymetry and the friction coefficients introduce additional errors. In practice, these parameters are estimated by means of calibration (or inverse problems) where data for a specific event is used, see for example [46].

Nowadays, uncertainty quantification in hydraulics is gradually growing regarding its important in the process of decision making, see for example [34, 30, 47, 14] and further references are therein. Several works in the literature have shown the appealing advantages of including uncertainty quantification in order to approve knowledge of different river, lake and ocean flows [18]. **Uncertainty propagation of dam break flow has also been investigated in [1] using a stochastic non-intrusive method.** Boundary conditions describing the flow are often pointed out as the responsible for uncertainties in hydraulic calculations [37]. This is mainly because these conditions suffer from both epistemic and stochastic errors which are often indistinguishable [32]. However, it turns out that other works including [8, 20] have highlighted that the uncertainty quantification in hydraulics should also include the friction terms, as for some specific flow configurations the uncertainty in hydraulics is more sensitive to these terms than the upstream flow conditions. In fact, the uncertainty expressed around the Manning coefficient is great and its accuracy would significantly impact the overall computational results obtained for the hydraulic simulations [8]. The bathymetry constitutes also an important source of uncertainty in the numerical simulation of hydraulics [10]. For practical hydraulic applications, its description is performed through interpolation of several scarce topographic measurements. For many realistic hydraulic applications, these data are neither available nor functional which make the hydraulic simulations even more uncertain, see for instance [11]. Recent studies have emphasized the importance of estimating the bathymetry in hydraulics specially with the upcoming spatial Surface Water and Ocean Topography (SWOT) mission. In fact, one method to alleviate the aforementioned problem could be achieved by means of Data Assimilation (DA). However, the main challenging part in the DA approaches remains the difficulties to account for uncertainties of the hydraulic simulation expressed by the numerical model, see for example the discussions reported in [22, 33]. The accuracy of uncertainty quantification is considerably increased when using a polynomial surrogate model with a reasonable computational cost, compare [20] and further references are therein. For these reasons, the Polynomial Chaos Expansion (PCE) is chosen to quantify the uncertainty in free-surface water flows. The focus of the present study is on quantifying uncertainties related to hydraulic computations resulting from bathymetric parameters (Manning coefficient and the field of bathymetry) using different numerical schemes.

When considering an uncertainty quantification problem using the PCE, one needs to sample the input parameters following a probability density function. As discussed in [28, 50], assessing stochastic processes such as bathymetry using each value of the vector representing the process as a random variable might be very

demanding. Here, the Karhunen-Loève (KL) expansion is considered to generate different realizations of the stochastic bed. This method has been widely used in the framework of uncertainty quantification, see [39, 43] for a detailed review. The PCE is becoming increasingly used in hydraulics community with its intrusive and non-intrusive forms. The intrusive formulation appears to be less appealing because of the hyperbolic nature of the shallow water equations, whereas the classical non-intrusive PCE struggles with the physical shock occurring in open-channel hydraulics. A recent study reported in [41] overcomes this issue by using a wavelet basis instead of the classical orthogonal polynomial basis. The main advantage of the method lies on the fact that the uncertainty of the hydraulic state is obtained in a straightforward way. However, the main drawback related to the *curse of dimensionality* as a sparse basis is hardly achievable with these methods. On the other hand, non-intrusive PCE has been successfully implemented for open-channel hydraulics in numerous studies, see for example [20, 25]. The appealing advantage of this method relies mainly in the possibility to use the adaptive PCE introduced by [5]. However, the computational cost of the method could be impacted by the discretization level of the hydraulic state. This issue has also been treated in previous studies using methods like the Proper Orthogonal Decomposition (POD), see for instance [19]. The idea in this approach is to decompose the hydraulic state using a POD technique. The uncertainty is then expressed on the eigenvalues of this decomposition. In the present study, the uncertainty of the bathymetric field is propagated through the hydraulic model using four different numerical solvers.

The aim in this study is to quantify the uncertainty of hydraulic computations over stochastic beds using different numerical schemes. Numerical results are presented for simple tests such as dam-break problems and lock-exchange flows and a realistic test of the exchange flow through the Strait of Gibraltar subject to a stochastic bathymetry. To the best of our knowledge, quantifying uncertainties in finite volume methods for two-layer shallow water flows is presented for the first time. Numerical results presented in the current study demonstrate high resolution of the proposed techniques and confirm their capability to provide efficient uncertainty quantification for simulation of free-surface water flows over stochastic beds. This paper is organized as follows. The equations for free-surface water flows are presented in Section 2. In Section 3 we present the numerical methods considered in the present work for solving the **one-dimensional** shallow water systems. The formulation of the finite volume modified method of characteristics for the numerical solution of two-layer models is also presented in this section. The general methodology used to address the problem of uncertainty quantification is described in section 4. This section includes the reduction methodology along with the PCE and the POD techniques. Section 5 is devoted to numerical results for several test examples of free-surface water flows over stochastic topography. Our new approach is shown to enjoy the expected accuracy as well as the robustness. Section 7 contains concluding remarks.

## 2 Governing equations for free-surface water flows

Shallow water equations have widely been used in the literature to model free-surface flows in rivers and coastal areas as well as to study a wide variety of phenomena in hydraulics and oceanography. The main feature in these equations is related to the fact that vertical effects can be neglected compared to the horizontal ones with a good approximation by replacing the vertical momentum equation by the hydrostatic pressure distribution. Indeed, the shallow water equations have been derived by integrating the three-dimensional incompressible Navier-Stokes equations along the vertical direction based on the assumptions that the vertical motion is ignored and the pressure distribution is hydrostatic, see [44] among others. Thus, in terms of the dependent variables water height $h(t,x)$ and water velocity $u(t,x)$, the single-layer shallow water equations read

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} &= 0, \\ \frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{1}{2}gh^2\right) &= -gh\frac{\partial Z}{\partial x} - ghM_b^2\frac{u \mid u \mid}{h^{4/3}}, \end{aligned} \tag{1}$$

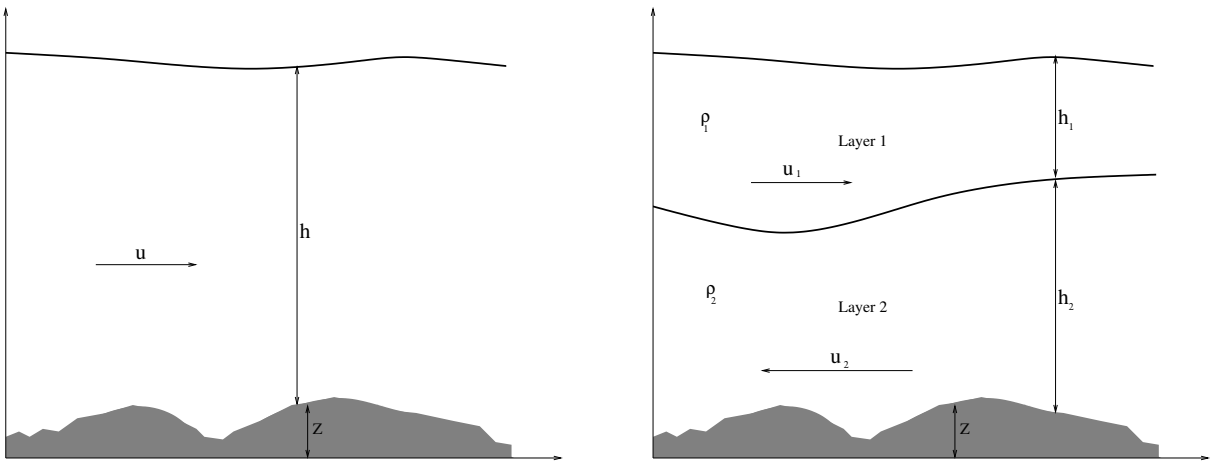


Figure 1: Schematic illustration of a single-layer shallow water flow (left plot) and a two-layer shallow water flow (right plot) over a given non-flat topography.

where $Z(x)$ is the function characterizing the bottom topography, $g$ the acceleration due to gravity, $M_b$ the Manning roughness coefficient at the bed, see the left plot in Figure 1 for an illustration. Notice that the system (1) is strictly hyperbolic with real and distinct eigenvalues given by

$$\lambda_1 = u - \sqrt{gh}, \qquad \lambda_2 = u + \sqrt{gh}. \tag{2}$$

In the current study, we also consider the two-layer shallow water equations for modelling free-surface water flows. For this situation, the governing equations are written in a conservative form as

$$\begin{aligned}
\frac{\partial h_1}{\partial t} + \frac{\partial \left(h_1 u_1\right)}{\partial x} &= 0, \\
\frac{\partial \left(h_1 u_1\right)}{\partial t} + \frac{\partial}{\partial x}\left(h_1 u_1^2 + \frac{1}{2} g h_1^2\right) &= -g h_1 \frac{\partial Z}{\partial x} - g h_1 \frac{\partial h_2}{\partial x}, \\
\frac{\partial h_2}{\partial t} + \frac{\partial \left(h_2 u_2\right)}{\partial x} &= 0, \\
\frac{\partial \left(h_2 u_2\right)}{\partial t} + \frac{\partial}{\partial x}\left(h_2 u_2^2 + \frac{1}{2} g h_2^2\right) &= -g h_2 \frac{\partial Z}{\partial x} - g h_2 \frac{\rho_1}{\rho_2} \frac{\partial h_1}{\partial x} - g h_2 M_b^2 \frac{u_2 \mid u_2 \mid}{h_2^{4/3}},
\end{aligned} \tag{3}$$

where the subscripts 1 and 2 represent respectively, the upper and lower layers in the hydraulic system, see the right plot in Figure 1 for an illustration. Here, $\rho_j$ is the water density of the $j$th layer, $h_j(t,x)$ is the water height of the $j$th layer and $u_j(t,x)$ is the local water velocity for the $j$th layer, $j = 1, 2$. It should be stressed that the exact calculation of eigenvalues associated with the two-layer system (3) is not trivial, see for instance [26]. Indeed, the four eigenvalues $\lambda_k$ $(k = 1, \ldots, 4)$ of the system are the zeros of the characteristic polynomial

$$P(\lambda) = \left(\lambda^2 - 2u_1\lambda + u_1^2 - g h_1\right)\left(\lambda^2 - 2u_2\lambda + u_2^2 - g h_2\right) - g^2 r h_1 h_2, \tag{4}$$

where the density ratio $r = \rho_1/\rho_2$. It should be stressed that, in many applications for free-surface flows, the density ratio $r \approx 1$ and the velocities $u_1 \approx u_2$. In these situations, a first-order approximation of the eigenvalues

can be obtained by expanding (4) in terms of $1-r$ and $u_2-u_1$ as

$$\begin{aligned}
\lambda_1 &\approx U_m - \sqrt{g\,(h_1+h_2)},\\
\lambda_2 &\approx U_m + \sqrt{g\,(h_1+h_2)},\\
\lambda_3 &\approx U_c - \sqrt{(1-r)g\frac{h_1h_2}{h_1+h_2}\left(1-\frac{(u_2-u_1)^2}{(1-r)g\,(h_1+h_2)}\right)},\\
\lambda_4 &\approx U_c + \sqrt{(1-r)g\frac{h_1h_2}{h_1+h_2}\left(1-\frac{(u_2-u_1)^2}{(1-r)g\,(h_1+h_2)}\right)},
\end{aligned} \tag{5}$$

with

$$U_m = \frac{h_1u_1+h_2u_2}{h_1+h_2}, \qquad U_c = \frac{h_1u_2+h_2u_1}{h_1+h_2}.$$

Note that, depending on the values of the density ratio $r$, the eigenvalues in (5) may become complex. In this case, the system is not hyperbolic and it yields the so-called Kelvin-Helmholtz instability at the interface separating the two layers. Therefore, a necessary condition for the two-layer system (3) to be hyperbolic is

$$(u_2-u_1)^2 < (1-r)g\,(h_1+h_2)\,. \tag{6}$$

The system (3) can be rearranged in a vector form as

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{W})}{\partial x} = \mathbf{Q}(\mathbf{W}) + \mathbf{R}(\mathbf{W}), \tag{7}$$

where $\mathbf{W}$ is the vector of conserved variables, $\mathbf{F}$ the vector of flux functions, $\mathbf{Q}$ and $\mathbf{R}$ the vectors of source terms defined as

$$\mathbf{W} = \begin{pmatrix} h_1 \\ h_1u_1 \\ h_2 \\ h_2u_2 \end{pmatrix}, \qquad \mathbf{F}(\mathbf{W}) = \begin{pmatrix} h_1u_1 \\ h_1u_1^2 + \frac{1}{2}gh_1^2 \\ h_2u_2 \\ h_2u_2^2 + \frac{1}{2}gh_2^2 \end{pmatrix},$$

$$\mathbf{Q}(\mathbf{W}) = \begin{pmatrix} 0 \\ -gh_1\frac{\partial}{\partial x}\,(h_2+Z) \\ 0 \\ -gh_2\frac{\partial}{\partial x}\,(rh_1+Z) \end{pmatrix}, \qquad \mathbf{R}(\mathbf{W}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -gh_2M_b^2\frac{u_2 \mid u_2 \mid}{h_2^{4/3}} \end{pmatrix}.$$

Similarly, the equations (1) can be reformulated in the vector form (7) with

$$\mathbf{W} = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \mathbf{F}(\mathbf{W}) = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \end{pmatrix}, \quad \mathbf{Q}(\mathbf{W}) = \begin{pmatrix} 0 \\ -gh\frac{\partial Z}{\partial x} \end{pmatrix}, \quad \mathbf{R}(\mathbf{W}) = \begin{pmatrix} 0 \\ -ghM_b^2\frac{u \mid u \mid}{h^{4/3}} \end{pmatrix}.$$

Note that we solve the equations (7) for free-surface flows in a time interval and a spatial domain equipped with given initial and boundary conditions with fixed topography and constant Manning roughness. In many hydraulic applications, these parameters are not deterministic and may be hard to estimate. In the present study, we also account for stochastic input parameter $\boldsymbol{\omega}$ in the free-surface equations (7) such that the topography $z(x,\boldsymbol{\omega})$ is a function of space and stochastic variables, and the Manning coefficient $M_b(\boldsymbol{\omega})$ depends on the stochastic inputs.

## 3 Numerical simulation of free-surface water flows

For the numerical solution of the free-surface equations (7) we consider a class of finite volume methods. We discretize the spatial domain into control volumes $[x_{i-1/2}, x_{i+1/2}]$ with uniform size $\Delta x = x_{i+1/2} - x_{i-1/2}$ and divide the temporal domain into subintervals $[t_n, t_{n+1}]$ with stepsize $\Delta t$. Here, $t_n = n\Delta t$, $x_{i-1/2} = i\Delta x$ and $x_i = (i + 1/2)\Delta x$ is the centre of the control volume. Integrating the equation (7) with respect to space over the control volume $[x_{i-1/2}, x_{i+1/2}]$, we obtain the following semi-discrete equations

$$\frac{d\mathbf{W}_i}{dt} + \frac{\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}}{\Delta x} = \mathcal{Q}_i + \mathcal{R}_i, \tag{8}$$

where $\mathbf{W}_i(t)$ is the space average of the solution $\mathbf{W}$ in the control volume $[x_{i-1/2}, x_{i+1/2}]$ at time $t$ *i.e.*,

$$\mathbf{W}_i = \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{W}(t, x)\, dx,$$

and $\mathcal{F}_{i\pm1/2} = \mathbf{F}(\mathbf{W}_{i\pm1/2})$ are the numerical fluxes at $x = x_{i\pm1/2}$ and time $t$. In (8), $\mathcal{Q}_i$ and $\mathcal{R}_i$ are the difference approximations of the discretized source terms $\mathbf{Q}(\mathbf{W}_i)$ and $\mathbf{S}(\mathbf{W}_i)$ in (7), respectively. To integrate the system (8) in time we consider an operator splitting method consisting first of the predictor step

$$\mathbf{W}_i^{n+1/2} = \mathbf{W}_i^n + \Delta t \mathcal{R}_i^n, \tag{9}$$

followed by the corrector step

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^{n+1/2} - \Delta t \frac{\mathcal{F}_{i+1/2}^{n+1/2} - \mathcal{F}_{i-1/2}^{n+1/2}}{\Delta x} + \Delta t \mathcal{Q}_i^{n+1/2}. \tag{10}$$

It should be pointed out that as with all explicit time stepping methods, the theoretical maximum stable time step $\Delta t$ is specified according to the Courant-Friedrichs-Lewy (CFL) condition

$$\Delta t = Cr \frac{\Delta x}{\max\limits_k(|\lambda_k^n|)}, \tag{11}$$

where $Cr$ is a constant (Courant number) to be chosen less than unity and $\lambda_k$ are the eigenvalues given by (2) or (5) for single-layer or two-layer shallow water equations, respectively. The spatial discretization of the equation (10) is complete when a reconstruction of the numerical fluxes $\mathcal{F}_{i\pm1/2}^{n+1/2}$ and source terms $\mathcal{Q}_i^{n+1/2}$ is chosen. In general, the reconstruction of the numerical fluxes requires a solution of Riemann problems at the interfaces $x_{i\pm1/2}$, see for example [29, 38]. In general this step can be carried out using any finite volume method developed in the Literature for solving hyperbolic systems of conservation laws. In the current study we consider the following reconstructions:

**Lax-Friedrichs method:** The numerical fluxes $\mathcal{F}_{i+1/2}^n$ are defined by

$$\mathcal{F}_{i+1/2}^n = \frac{1}{2}\Big(\mathbf{F}\left(\mathbf{W}_{i+1}^n\right) + \mathbf{F}\left(\mathbf{W}_i^n\right)\Big) + \frac{\Delta x}{2\Delta t}\left(\mathbf{W}_i^n - \mathbf{W}_{i+1}^n\right). \tag{12}$$

**Rusanov method:** The numerical fluxes $\mathcal{F}_{i+1/2}^n$ are defined by

$$\mathcal{F}_{i+1/2}^n = \frac{1}{2}\Big(\mathbf{F}\left(\mathbf{W}_{i+1}^n\right) + \mathbf{F}\left(\mathbf{W}_i^n\right)\Big) + \frac{\lambda}{2}\left(\mathbf{W}_i^n - \mathbf{W}_{i+1}^n\right), \tag{13}$$

where $\lambda$ is the Rusanov speed defined by $\lambda = \max\limits_k(\lambda_k^n)$, with $\lambda_k$ are the eigenvalues given by (2) or (5) for single-layer or two-layer shallow water equations, respectively.

**Roe method:** The numerical fluxes $\mathcal{F}^n_{i+1/2}$ are defined by

$$\mathcal{F}^n_{i+1/2} = \frac{1}{2}\left(\mathbf{F}(\widehat{\mathbf{W}}^n_{i+1}) + \mathbf{F}(\widehat{\mathbf{W}}^n_i)\right) + \frac{1}{2}\mathbf{A}\left(\widehat{\mathbf{W}}^n_{i+1/2}\right)\left(\widehat{\mathbf{W}}^n_i - \widehat{\mathbf{W}}^n_{i+1}\right), \tag{14}$$

where the averaged state $\widehat{\mathbf{W}}^n_{i+1/2}$ is calculated as

$$\widehat{\mathbf{W}}^n_{i+1/2} = \begin{pmatrix} \dfrac{\widehat{h}^n_i + \widehat{h}^n_{i+1}}{2} \\ \dfrac{\sqrt{\widehat{h}^n_i}\widehat{u}^n_i + \sqrt{\widehat{h}^n_{i+1}}\widehat{u}^n_{i+1}}{\sqrt{\widehat{h}^n_i} + \sqrt{\widehat{h}^n_{i+1}}} \end{pmatrix},$$

and $\mathbf{A} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{-1}$ is the Roe matrix defined by

$$\mathbf{R} = \begin{pmatrix} 1 & 1 \\ \widehat{\lambda}_1 & \widehat{\lambda}_2 \end{pmatrix}, \qquad \mathbf{\Lambda} = \begin{pmatrix} \widehat{\lambda}_1 & 0 \\ 0 & \widehat{\lambda}_2 \end{pmatrix},$$

where $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ are the eigenvalues in (2) calculated at the averaged state. Note that the extension of the Roe scheme to the two-layer system is not possible since (5) is only an approximation of the eigenvalues associated with the system and no explicit exact formulation is provided for these eigenvalues. Therefore, we also consider a finite volume modified method of characteristics (FVC) in the current study. The FVC method is simple, easy to implement, and accurately solves the equations (7) without relying on Riemann problem solvers. This method has been proposed in [3] for solving the single-layer shallow water equations (1) and its extension to the two-layer shallow system (3) is presented in the present work.

### 3.1 Finite volume modified method of characteristics

To reconstruct the numerical fluxes $\mathcal{F}^n_{i\pm1/2}$ in (10), we consider the method of characteristics applied to the advective version of the system (3). In general, the advective form of the two-layer system (3) is built such that the non-conservative variables are transported with the same velocity field associated with each layer. Here, the two-layer shallow water equations (3) are reformulated in an advective form as

$$\begin{aligned} \frac{\partial \mathbf{U}_1}{\partial t} + u_1\frac{\partial \mathbf{U}_1}{\partial x} &= \mathbf{S}_1, \\ \frac{\partial \mathbf{U}_2}{\partial t} + u_2\frac{\partial \mathbf{U}_2}{\partial x} &= \mathbf{S}_2, \end{aligned} \tag{15}$$

where

$$\mathbf{U}_1 = \begin{pmatrix} h_1 \\ u_1 \end{pmatrix}, \quad \mathbf{U}_2 = \begin{pmatrix} h_2 \\ u_2 \end{pmatrix}, \quad \mathbf{S}_1 = \begin{pmatrix} -h_1\partial_x u_1 \\ -g\dfrac{\partial}{\partial x}\left(Z + h_1 + h_2\right) \end{pmatrix}, \quad \mathbf{S}_2 = \begin{pmatrix} -h_2\partial_x u_2 \\ -g\dfrac{\partial}{\partial x}\left(Z + rh_1 + h_2\right) \end{pmatrix}.$$

The fundamental idea of the method of characteristics is to impose a regular grid at the new time level and to backtrack the flow trajectories to the previous time level. At the old time level, the needed quantities are evaluated by interpolation from their known values on a regular grid, for more discussions we refer the reader to [45, 40] among others. Thus, the characteristic curves associated with the equations (15) are solutions of the initial-value problem

$$\begin{aligned} \frac{dX_{j,i+1/2}(\tau)}{d\tau} &= u_{j,i+1/2}\Big(\tau, X_{j,i+1/2}(\tau)\Big), \quad \tau \in [t_n, t_{n+1}], \\ X_{j,i+1/2}(t_{n+1}) &= x_{i+1/2}, \qquad j = 1, 2. \end{aligned} \tag{16}$$

Note that $X_{j,i+1/2}(\tau)$ is the departure point at time $\tau$ of a particle that will arrive at point $x_{i+1/2}$ in time $t_{n+1}$. The method of characteristics does not follow the flow particles forward in time, as the Lagrangian schemes do, instead it traces backward the position at time $t_n$ of particles that will reach the points of a fixed mesh at time $t_{n+1}$. To compute the solutions of (16) we use an explicit second-order Runge-Kutta scheme. Once the departure points $X_{j,i+1/2}(t_n)$ are known, a solution at the cell interface $x_{i+1/2}$ is reconstructed as

$$\mathbf{U}^n_{j,i+1/2} = \mathbf{U}_j\left(t_{n+1}, x_{i+1/2}\right) = \widetilde{\mathbf{U}}_j\left(t_n, X_{j,i+1/2}(t_n)\right), \tag{17}$$

where $\widetilde{\mathbf{U}}_j\left(t_n, X_{j,i+1/2}(t_n)\right)$ is the solution at the characteristic foot computed by interpolation from the grid-points of the control volume where the departure point $X_{j,i+1/2}(t_n)$ resides. In the present work we consider a quadratic Lagrange interpolation to perform this stage.

Applied to the equations (15), the characteristic solutions are given by

$$\begin{aligned}
h^n_{1,i+1/2} &= \widetilde{h}^n_{1,i+1/2} - \frac{\nu}{2}\widetilde{h}^n_{1,i+1/2}\left(u^n_{1,i+1} - u^n_{1,i}\right),\\
u^n_{1,i+1/2} &= \widetilde{u}^n_{1,i+1/2} - \frac{\nu}{2}g\Big((Z + h^n_1 + h^n_2)_{i+1} - (Z + h^n_1 + h^n_2)_i\Big),\\
h^n_{2,i+1/2} &= \widetilde{h}^n_{2,i+1/2} - \frac{\nu}{2}\widetilde{h}^n_{2,i+1/2}\left(u^n_{2,i+1} - u^n_{2,i}\right),\\
u^n_{2,i+1/2} &= \widetilde{u}^n_{2,i+1/2} - \frac{\nu}{2}g\Big((Z + rh^n_1 + h^n_2)_{i+1} - (Z + rh^n_1 + h^n_2)_i\Big),
\end{aligned} \tag{18}$$

where $\nu = \frac{\Delta t}{\Delta x}$ and

$$\widetilde{h}^n_{1,i+1/2} = h_1\left(t_n, X_{1,i+1/2}(t_n)\right), \qquad \widetilde{u}^n_{1,i+1/2} = u_1\left(t_n, X_{1,i+1/2}(t_n)\right),$$

$$\widetilde{h}^n_{2,i+1/2} = h_2\left(t_n, X_{2,i+1/2}(t_n)\right), \qquad \widetilde{u}^n_{2,i+1/2} = u_2\left(t_n, X_{2,i+1/2}(t_n)\right),$$

are the solutions at the characteristic foot computed by interpolation from the gridpoints of the control volume where the departure points $X_{1,i+1/2}(t_n)$ and $X_{2,i+1/2}(t_n)$ belong. The numerical fluxes $\mathcal{F}_{i\pm1/2}$ in (8) are calculated using the intermediate states $\mathbf{W}^n_{i\pm1/2}$ recovered accordingly from the characteristic solutions $\mathbf{U}^n_{j,i\pm1/2}$ in (17). Hence, the FVC method (10) reduces to

$$\begin{aligned}
h^{n+1}_{1,i} &= h^n_{1,i} - \nu\left((h_1u_1)^n_{i+1/2} - (h_1u_1)^n_{i-1/2}\right),\\
q^{n+1}_{1,i} &= q^n_{1,i} - \nu\left(\left(h_1u_1^2 + \frac{1}{2}gh_1^2\right)^n_{i+1/2} - \left(h_1u_1^2 + \frac{1}{2}gh_1^2\right)^n_{i-1/2}\right) - \frac{1}{2}\nu g\hat{h}^n_{1,i}\left((Z + h_2)_{i+1} - (Z + h_2)_{i-1}\right),\\
h^{n+1}_{2,i} &= h^n_{2,i} - \nu\left((h_2u_2)^n_{i+1/2} - (h_2u_2)^n_{i-1/2}\right),\\
q^{n+1}_{2,i} &= q^n_{2,i} - \nu\left(\left(h_2u_2^2 + \frac{1}{2}gh_2^2\right)^n_{i+1/2} - \left(h_2u_2^2 + \frac{1}{2}gh_2^2\right)^n_{i-1/2}\right) - \frac{1}{2}\nu g\hat{h}^n_{2,i}\left((Z + rh_1)_{i+1} - (Z + rh_1)_{i-1}\right),
\end{aligned} \tag{19}$$

where $q_1 = h_1u_1$ and $q_2 = h_2u_2$ are the water discharges associated with the upper layer and lower layer, respectively. In our FVC method, the reconstruction of the term $\hat{h}^n_{1,i}$ and $\hat{h}^n_{2,i}$ in (19) is carried out such that the discretization of the source terms is well balanced with the discretization of the flux gradients using the concept of C-property as detailed in [3]. Hence, if the source terms $\hat{h}^n_{1,i}$ and $\hat{h}^n_{2,i}$ in the stage of (19) are discretized as

$$\hat{h}^n_{1,i} = \frac{1}{4}\left(h^n_{1,i+1} + 2h^n_{1,i} + h^n_{1,i-1}\right), \qquad \hat{h}^n_{2,i} = \frac{1}{4}\left(h^n_{2,i+1} + 2h^n_{2,i} + h^n_{2,i-1}\right), \tag{20}$$

then the proposed FVC method satisfies the C-property. In summary, the implementation of the FVC algorithm to solve the two-layer shallow water equations (3) is carried out in the following steps. Given $\left(h^n_{1,i}, q^n_{1,i}, h^n_{2,i}, q^n_{2,i}\right)$, we compute $\left(h^{n+1}_{1,i}, q^{n+1}_{1,i}, h^{n+1}_{2,i}, q^{n+1}_{2,i}\right)$ via:

**Step 1.** *Compute the departure points* $X_{1,i+1/2}(t_n)$ *and* $X_{2,i+1/2}(t_n)$ *using the explicit second-order Runge-Kutta scheme for solving (16).*

**Step 2.** *Compute the approximations*

$$\widetilde{h}^n_{1,i+1/2} = h_1\left(t_n, X_{1,i+1/2}(t_n)\right), \qquad \widetilde{u}^n_{1,i+1/2} = u_1\left(t_n, X_{1,i+1/2}(t_n)\right),$$

$$\widetilde{h}^n_{2,i+1/2} = h_2\left(t_n, X_{2,i+1/2}(t_n)\right) \quad \textit{and} \quad \widetilde{u}^n_{2,i+1/2} = u_2\left(t_n, X_{2,i+1/2}(t_n)\right),$$

*employing a quadratic Lagrange interpolation procedure.*

**Step 3.** *Evaluate the intermediate states* $h^n_{1,i+1/2}$, $u^n_{1,i+1/2}$, $h^n_{2,i+1/2}$ *and* $u^n_{2,i+1/2}$ *from the predictor stage* (18).

**Step 4.** *Update the solutions* $h^{n+1}_{1,i}$, $q^{n+1}_{1,i}$, $h^{n+1}_{2,i}$ *and* $q^{n+1}_{2,i}$ *using the corrector stage* (19).

Note that the application of the FVC scheme for solving the single-layer shallow water equations (1) is straightforward and its formulation is omitted here.

# 4 Uncertainty quantification methods

In general, the purpose of uncertainty quantification is to identify the main sources of uncertainty in a physical model (*e.g.* parameters, external forcing, boundary conditions, initial conditions) and to quantify their impact on the quantities of interest simulated by the numerical model (prognostic variables, probability of exceeding the threshold). This allows to associate every forecast with a level of certainty since the accuracy of a simulation significantly depends on both the quantity and the quality of the input data. Therefore, to better understand the results of numerical simulations it is necessary to take into account these uncertainties in the simulations. In addition, a problem of quantification of uncertainties is a problem which aims at estimating uncertainty on the outputs of a numerical simulation according to uncertainties on the knowledge of its input parameters. Because of the random nature of uncertainty, the probabilistic approach to deal with a problem of uncertainty quantification is to consider the uncertain data of the model as random variables or random processes, and to reconsider the real deterministic numerical model as a stochastic model. This section presents the techniques used in the current study for uncertainty quantification. First, methods of generating different independent realizations of the bathymetric field are presented. Then, the classical PCE method is used as a surrogate tool to alleviate the computational cost. Finally, as the model output (*i.e.* the hydraulic state) would depend on the space and in order to avoid building as many surrogate models as the number of gridpoints, we suggest to reduce the dimension of the problem using the POD and to compute a PCE only for the associated nonphysical variables.

## 4.1 Karhunen-Loève expansion for stochastic process

A stochastic process is defined by the means of an indexation set $X$, $E$ is a measurable space and $(\Omega, \mathcal{F}, \mathcal{P})$ is a $\sigma$-algebra representing the probability space [49]. The stochastic process $\mathbf{z}$ is then defined as a collection $\{\mathbf{z}_x, x \in X\}$ which the random values are in the state space $E$ and described statistically by the probabilistic space $(\Omega, \mathcal{F}, \mathcal{P})$

$$\mathbf{z}_x : \Omega \longrightarrow E. \tag{21}$$

Consequently, one is able to define a realization $\omega \in \Omega$ for a stochastic process as

$$\mathbf{z}(\omega) : \begin{array}{ccc} X & \longrightarrow & E \\ x & \longmapsto & \mathbf{z}_x(\omega) \end{array} \tag{22}$$

The Karhunen-Loève expansion (KL) allows to model a random process based on a spectral decomposition of its spatial covariance matrix $\mathcal{C}(x, x')$, see for example [31]. Note that, by construction the covariance function

is real, symmetric and positive definite. Consequently, the set of eigenfunctions form a complete orthogonal basis for the space in which the process belongs. Hence, a stochastic process can be defined as

$$\boldsymbol{z}(x,\boldsymbol{\omega}) = \overline{\boldsymbol{z}}(x) + \sum_{i=0}^{\infty} \omega_i l_i \phi_i(x), \tag{23}$$

where $x$ is the curvilinear abscissa, $\overline{\boldsymbol{z}}$ the mean of the random process, $\boldsymbol{\omega} = \{\omega_1, \ldots, \omega_\infty\}$ a set of independent random variables, $l_i$ the eigenvalues of the covariance function, and $\phi_i(x)$ the eigenfunctions. In practice and for computational reasons, the equation (23) is truncated at a certain degree $d$. The choice of this latter is often determined by a threshold $\epsilon$ from which the eigenvalues could be neglected *i.e.*

$$\frac{\sum_{i=0}^{d} l_i}{\sum_{i=0}^{\infty} l_i} > 1 - \epsilon.$$

Thus, the expansion (23) is replaced by

$$\mathbf{z}(x,\boldsymbol{\omega}) = \overline{\mathbf{z}}(x) + \sum_{i=0}^{d} \omega_i l_i \phi_i(x). \tag{24}$$

It should be stressed that the major limitation of the KL decomposition is the *a priori* knowledge of the covariance matrix. In the literature, this parameter is estimated empirically and therefore could be subject to the uncertainty itself. This uncertainty is not considered in the present work and the reader may be referred to [43] for more discussions on this topic. This expansion is commonly used to model the uncertainty of stochastic input parameters for two reasons: firstly the mean-square error of the finite representation (24) is minimized such that the equation (23) converges following the $\ell_2$-norm. The second reason is related to the generation of random samples. In fact, making a realization of $\boldsymbol{z}(x,\boldsymbol{\omega})$ amounts to randomly draw the different $\omega_i$ following a defined probability density function. Those parameters are considered independent random variables. Therefore, the KL decomposition offers a good representation of the input parameters when the covariance is known. In our study, The covariance matrix $\mathcal{C}(x,x')$ is supposed to be exponential. Indeed, [51] showed that the uncertainty in the description of bathymetric fields followed this spatial correlation function. Thus, the correlation matrix is defined as

$$\mathcal{C}(x,x') = \sigma^2 \exp\left(-\frac{|x-x'|}{\theta}\right), \tag{25}$$

where $\sigma$ and $\theta$ are the hyper-parameters of the covariance function with $\sigma$ is the standard deviation of the process and $\theta$ is the correlation length. Note that, calculating the eigenvalues and eigenvectors is well known by the Fredholm problem. There exits many algorithms that aim to solve this latter given a well defined matrix, see for example [27] for more details. Once, $l_i$ and $\phi_i(x)$ are known, the KL expansion is implemented in a straightforward manner. Therefore, the bathymetry could be written as described in (24) and could be sampled in order to consider its uncertainty in hydraulic computations. Numerical values should be given hereafter.

### 4.2 Polynomial chaos expansions

The PCE has been intensively used as a surrogate model in the context of uncertainty quantification, see [25, 20] among others. It aims to reproduce the global behavior of the considered shallow water models. Supposing that the input parameters of this model are represented by $M$ independent random variables $\boldsymbol{\zeta} = \{\zeta_1, \zeta_2, \ldots, \zeta_M\}$ with finite variance well defined in a probabilistic space, the response $\mathbf{W}$ of this model is also random. Note that $\mathbf{W}$ is a vector-valued response as it represents the spatial variability of the hydraulic states. Considering

that the expectation $\mathbb{E}\left[\|\mathbf{W}\|^2\right]$ is finite, the behavior of $\mathbf{W}$ could be reproducible following a polynomial decomposition namely PCE [42] as

$$\mathbf{W}(x,\boldsymbol{\zeta}) = \sum_{i\in\mathbb{N}} \alpha_i(x)\Psi_i(\boldsymbol{\zeta}), \tag{26}$$

where $\Psi_i$ are the multivariate polynomials that form the basis are chosen in such way they are orthonormal with respect to the associated probability density function of $\boldsymbol{\zeta}$, *i.e.* $\mathbb{E}\left[\Psi_i(\boldsymbol{\zeta})\Psi_j(\boldsymbol{\zeta})\right] = \delta_{ij}$ with $\delta_{ij}$ is the Kronecker symbol, compare for example [48]. In (26), $\alpha_i$ are the unknown spectral coefficients of the decomposition to be determined. Again, the sum in (26) is truncated to a finite series as

$$\mathbf{W}(x,\boldsymbol{\zeta}) \approx \sum_{i\in\mathcal{I}} \alpha_i(x)\Psi_i(\boldsymbol{\zeta}), \tag{27}$$

where $\mathcal{I} \subset \mathbb{N}$ is the finite set of indices. The determination of a PCE is therefore conditioned by the estimation of the spectral coefficients $\alpha_i$. There are many methods used in the literature to achieve this step and we refer to [50] for a review on these methodologies. In the current work, only the regression method is used and this choice comes from the fact that these methods coupled with compressed sensing techniques are advantageous when the stochastic dimension $M$ of the problem is high (which is more likely to be the case here after the use of the KL decomposition). The regression method is based on solving a least-square minimization problem in some $\ell_2$-norm to estimate the coefficients $\alpha_i$, see for instance [4]. In practice, we begin by defining an error $\boldsymbol{\epsilon}$ as the distance between the model and the PCE for a finite set of randomly sampled input variables of size $N_{ls}$ as

$$\boldsymbol{\epsilon} = \left\| \mathbf{W}(x,\boldsymbol{\Xi}) - \sum_{i\in\mathbb{N}} \alpha_i(x)\Psi(\boldsymbol{\Xi}) \right\|_2 \equiv \boldsymbol{\mathcal{W}} - \boldsymbol{\alpha}^\top \boldsymbol{\Psi}, \tag{28}$$

where $\boldsymbol{\Xi} = \left(\boldsymbol{\zeta}^{(1)}, \boldsymbol{\zeta}^{(2)}, \ldots, \boldsymbol{\zeta}^{(N_{ls})}\right)^\top$ is the set of realizations for the stochastic input variables $\boldsymbol{\zeta}$ in (27) and $\boldsymbol{\mathcal{W}} = \left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \ldots, \mathbf{W}^{(N_{ls})}\right)^\top$ the vector of associated model outputs. We also define $\boldsymbol{\alpha} = \left(\alpha_0, \alpha_1, \ldots, \alpha_{N_{PC-1}}\right)^\top$ as the vector of the $N_{PC} = Card(\mathcal{I})$ unknown coefficients and $\boldsymbol{\Psi}$ is the matrix of size $N_{PC} \times N_{ls}$ assembling the values of all orthonormal polynomials at the stochastic input realizations values $\boldsymbol{\Psi}_{ik} = \Psi_i(\boldsymbol{\zeta}^{(k)})$, with $i = 0, 1, \ldots, N_{PC-1}$ and $k = 1, 2, \ldots, N_{ls}$. Following the ordinary least-square solver (28), the estimation of the set of coefficients $\boldsymbol{\alpha}$ is equivalent to minimize the following function

$$\mathrm{J}(\boldsymbol{\alpha}) = \boldsymbol{\epsilon}^\top\boldsymbol{\epsilon} = \left(\boldsymbol{\mathcal{W}} - \boldsymbol{\alpha}^\top\,\boldsymbol{\Psi}\right)^\top \left(\boldsymbol{\mathcal{W}} - \boldsymbol{\alpha}^\top\,\boldsymbol{\Psi}\right), \tag{29}$$

which leads to a standard well-known linear algebraic solution as

$$\boldsymbol{\alpha} = \left(\boldsymbol{\Psi}^\top\,\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}^\top\,\boldsymbol{\mathcal{W}}. \tag{30}$$

Here, the input space exploration is fulfilled using a Monte-Carlo sampling-based approach [6, 20]. It is worth mentioning that the number of coefficients $N_{PC}$ needed is directly affected by the stochastic dimension $M$ as well as the polynomial degree $p$. As a consequence, the Monte-Carlo size $N_{ls}$ will also increase significantly with $M$ and $p$. This is a classical problem of PCE also known as *the curse of dimensionality*. The adaptive Least Angle Regression (LAR) method [16] has been introduced and used to overcome this specific problem [6]. This method introduced in the context of compressed sensing has made it possible to recover accurately the solution with fewer model simulations. It will be used in the present work as it has been demonstrated to be very efficient in this framework. The idea behind this method is to select with an iterative manner an optimal sparse basis among the original one and then to compute a limited number of coefficients using a standard regression method. In the context of PCE, a hybrid LAR method is used. The method consists on using LAR to evaluate the best set of predictors among the full basis elements with the help of a cross-validation method. These coefficients are estimated using the classical least square method, see [6] for more details. The infinite expansion (26) describing the PCE converges with respect to the standard $\ell_2$-norm known by the mean-squared convergence, see for example [50]. However, due to the errors of truncation and spectral coefficient estimation,

the accuracy of this expansion must be evaluated in the same error norm. There are many different error metrics that allow to assess the accuracy of the PCE, see [20, 50] among others. In the present work, all the PCEs are assessed using the Leave-One-Out ($LOO$) error. This method avoids integrating the model over another set of validation samples and therefore the computational cost is optimized. It has been introduced in this context with the introduction of sparse PCE [4] and the $LOO$ technique required the formulation of several surrogate models [6]. Each surrogate model is built excluding one point out of the input sample and the accuracy of the surrogate model is then quantified at this particular point. Following this theory, the error $\epsilon_{LOO}$ is defined by

$$\epsilon_{LOO} = \frac{\sum_{k=1}^{N_{ls}} \left(\mathbf{W}^{(k)} - \mathbf{W}^{(-k)}\right)^2}{\sum_{k=1}^{N_{ls}} \left(\mathbf{W}^{(k)} - \overline{\mathbf{W}}\right)^2}, \tag{31}$$

where $\overline{\mathbf{W}}$ denotes the sample-averaged model simulations and $\mathbf{W}^{(-k)}$ stands for the evaluation of the PCE at $\boldsymbol{\zeta}^{(k)}$ when the surrogate has been built using an experimental design in which $\boldsymbol{\zeta}^{(k)}$ was excluded. It should be stressed that, in the context of PCE, it has been proven that the LOO error could be computed analytically using all the samples **from the Monte-Carlo simulations**. In this case, a corrected LOO error is computed using only a full PCE rather than several decompositions, compare [5]. Furthermore, it has been proven that in the case of sparse PCE, the LOO error is robust and conservative, see [6]. The relative error and the LOO error have also been compared in [20] for the case of the PCE applied to hydraulic uncertainties and the results reported in this reference have shown a good agreement between the two metrics. **In the present work, the determination of the optimal polynomial degree is performed using an iterative procedure. Thus, a PCE is computed for different degrees varying from 1 to 20 and the optimal degree is determined based on the value of the corresponding $\epsilon_{LOO}$ error. For the same value of a given error, the lowest value of the degree is retained in our simulations.**

Note that, when dealing with numerical models with spatial or temporal dependency, the classical way used in numerical simulations is to discretize the physical domain and resolve the governing equations for each control volume. This procedure leads to multiple decompositions to ensure the numerical convergence. As a result, uncertainty quantification becomes computationally very demanding not only because of the stochastic dimensions but also because of the spatial or temporal dimensions of the output variables. For these reasons, many methods have been introduced in the literature in order to reduce the dimensionality of the output vector, see for instance [35, 7, 12]. Among all these different techniques, the POD have proven to be efficient in the context of physical fields with rapid variability as in the case of hydraulics [35].

### 4.3 Stochastic proper orthogonal decomposition

The proper orthogonal decomposition (POD) is a well-established technique that allows a high-dimensional system to be approached by a low-dimensional one, compare [13] among others. This method is based on determining a basis of orthogonal eigenvalues representative of the simulated physics. The eigenvectors are obtained by solving the integral of Fredholm problem whereas, the kernel of this integral is constructed from a set of simulations constructed using a set of experiments. The interesting property of this representation lies in the fact that the eigenfunctions associated with the problem are optimal in the sense of the energetic representation which makes it possible to use them to construct a reduced representation of the physics under study. Notice that the POD is used in the uncertainty quantification to reduce the size of a random vector at the output of the model. The uncertainty is therefore carried out over each direction defined by the eigenvectors $\phi_i(x)$. The idea is based on projecting the solution $\mathbf{W}$ of the model into a finite and orthonormal basis $\{\phi_i,\ i \in \mathcal{I}_{POD}\}$, where $\mathcal{I}_{POD}$ is a discrete finite set of indices. Thus, the process $\mathbf{W}(x, \boldsymbol{\zeta})$ is decomposed as

$$\mathbf{W}(x, \boldsymbol{\zeta}) = \sum_{i \in \mathcal{I}_{POD}} \widehat{\lambda}_i(\boldsymbol{\zeta}) \phi_i(x). \tag{32}$$

The estimation of the finite basis $\{\phi_i,\ i \in \mathcal{I}_{POD}\}$ is performed by decomposing the covariance matrix as

$$\mathbf{C} = \frac{1}{N_{ls}} \mathbf{W}\mathbf{W}^{\top}. \tag{33}$$

Indeed, the literature is rich of techniques that aim to decompose a covariance matrix. One of the most known methods is the Singular Value Decomposition (SVD) algorithm [15]. Hence, we define a POD-truncated error $\epsilon$ such as only the most $k$ invaluable eigenvectors are retained as

$$\frac{\sum_{i=0}^{k} \lambda_i}{\sum_{i=0}^{N_{ls}} \lambda_i} > 1 - \epsilon, \tag{34}$$

where $\widehat{\lambda}_i$ is the mean value of $\lambda_i(\boldsymbol{\zeta})$. In summary, the stochastic POD procedure can be implemented using the following steps:

**Step 1.** *Define the covariance matrix* (33).

**Step 2.** *Expand the matrix* $\mathbf{C}$ *using an SVD algorithm to determine* $\lambda_i$ *and* $\phi_i(x)$.

**Step 3.** *Retain only the first* $k$ *eigenvalues and eigenvectors in the expansion using the condition* (34).

**Step 4.** *Reconstruct the stochastic solutions* $\mathbf{W}(x, \boldsymbol{\zeta})$ *using* (32).

It is worth remarking that the selection of convergence criterion is problem-dependent and therefore the selection of $\epsilon$ for test examples in the present investigation is discussed in section 5 where numerical examples are described.

### 4.4 The POD-PCE surrogate model

Once the stochastic POD is reconstructed, the eigenvalues are considered as stochastic independent variables (since the eigenvectors form a basis). This means that we can define a PCE for each eigenvalue following the same manner as described in the previous section on the polynomial chaos expansions leaving the spatial dependence described by the eigenvectors $\phi_i(x)$ as

$$\lambda_i(\boldsymbol{\zeta}) = \sum_{j=0}^{N_{PC}} \gamma_j \Psi_j(\boldsymbol{\zeta}), \tag{35}$$

**where $\gamma_j$ are the corresponding spectral coefficients computed following the methodology described in Section 4.2**. Hence, using (35), the equation (27) reduces to

$$\mathbf{W}(x, \boldsymbol{\zeta}) = \sum_{i \in \mathcal{I}_{POD}} \left( \sum_{j=0}^{N_{PC}} \gamma_j \Psi_j(\boldsymbol{\zeta}) \right) \phi_i(x). \tag{36}$$

Figure 2 summarizes both algorithms that we consider in the present work for the quantification of uncertainty. Here, the first step consists of developing the KL decomposition to be able to sample the bathymetric fields. Obviously, this step depends mainly on the definition of a spatial correlation associated with the problem under consideration. In practice, the covariance matrix hyper-parameters could be inferred from observational data, compare for example [43]. For all our simulations carried out in this study, the correlation matrix is assumed to be exponential with given correlation length for each test problem and the associated eigenvalues and eigenvectors are obtained analytically as in [28]. The second step in our algorithm for uncertainty quantification

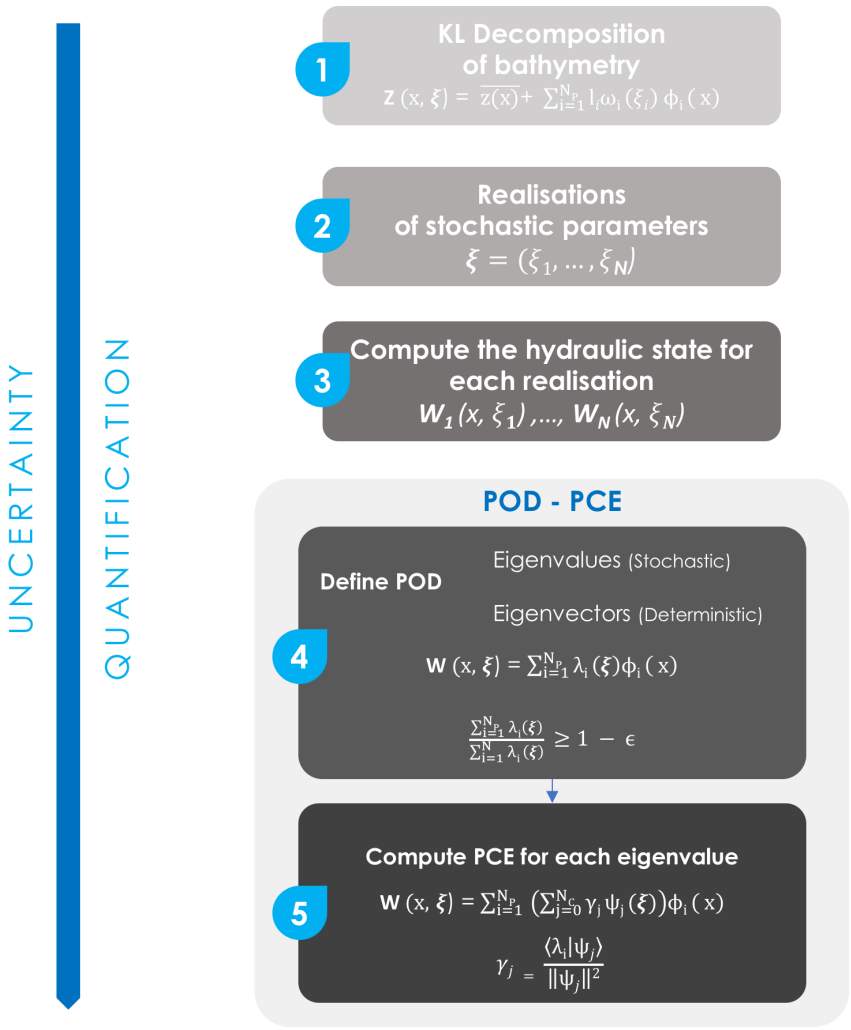


Figure 2: Schematic representation of the uncertainty quantification strategy used in the present study using a POD-PCE based surrogate model.

is the sampling of different random variables using a Monte Carlo approach which include the eigenvalues of the KL decomposition and the Manning coefficient following a well defined probabilistic distribution. For each realization, the hydraulic state is computed using step 3 in the algorithm shown in Figure 2. Next, a resulting hydraulic state is obtained for each realization which allows to perform a POD over the hydraulic state. Note that the eigenvectors represent the spatial variability while, the eigenvalues are responsible for the uncertainty as explained in [19]. The final step in our algorithm is to carry out an adaptive spare PCE over the eigenvalues of the POD. It is evident that, using this surrogate model, one is able to quantify the uncertainty for each numerical solver.

## 5 Numerical results and examples

We present numerical results for several test problems of shallow water flows using single-layer and two-layer shallow water equations over flat and non-flat beds. The main goal of this section is to illustrate the numerical performance of the techniques described above. In all the computations reported herein, the Courant number is set to $Cr = 0.8$ and the time stepsize $\Delta t$ is adjusted at each time step according to the CFL stability condition (11). We present numerical results obtained using the Lax-Friedrichs, Rusanov, Roe and FVC methods.

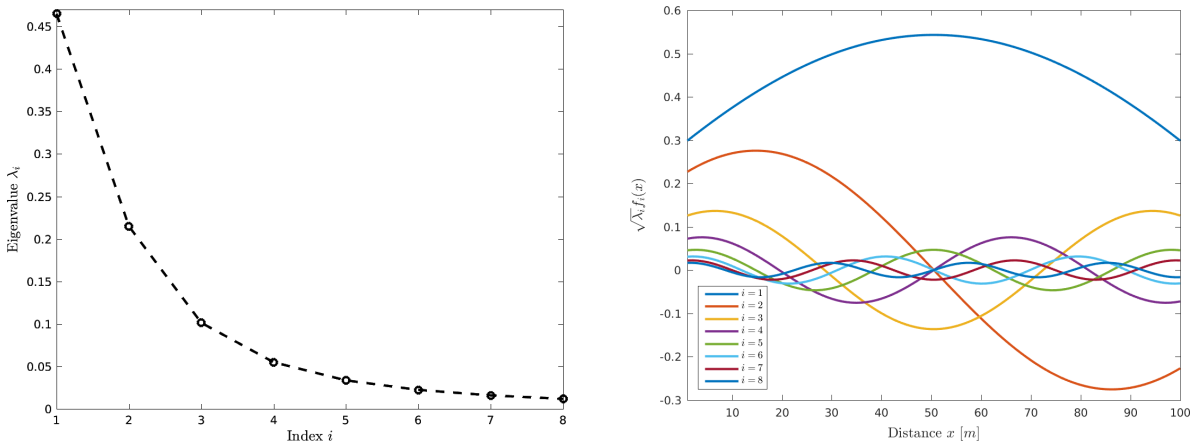


Figure 3: Eigenvalues (left plot) and associated eigenfunctions (right plot) of the correlation matrix in the KL decomposition for the dam-break problem over a flat bed.

## 5.1 Results for single-layer shallow flows

In this class of flow problems, we investigate the uncertainty using the proposed finite volume methods for the single-layer shallow water equations (1). More precisely we quantify the uncertainty related to the bathymetry displayed by the considered numerical methods. The aim here is to rank these finite volume methods regarding the uncertainty on the computed water height and water velocity variables when the knowledge about the bathymetry is uncertain. Here, the POD-PCE meta-model is built using 1000 forward simulations. In all cases considered in this section, we follow the general methodology described in section 5. Thus, we first decompose the bathymetry field using the KL decomposition, then we run Monte-Carlo simulations to obtain a set of fields for both water height and water velocity. These fields are then reduced using the POD approach and are reduced further by using a truncation threshold in the PCE over each POD modes.

### 5.1.1 Dam-break problem over a flat bed

For this test example, the uncertainty quantification is performed for a dam-break flow problem over a stochastic frictionless flat bed (*i.e.* $Z = 0$ and $M_b = 0$). The channel is of length 30 $m$ and subject to the following initial conditions

$$h(0,x) = \begin{cases} 1\ m, & \text{if} \quad x \leq 15\ m, \\ 0.3\ m, & \text{if} \quad x > 15\ m, \end{cases} \qquad u(0,x) = 0\ m/s.$$

At time $t = 0$ the dam collapses and the flow problem consists of a shock wave traveling downstream and a rarefaction wave traveling upstream. In this case example, **the variation coefficient in the bathymetry $CV_b = 0.05$ and the Manning coefficient is kept deterministic**. The computed results are presented at time $t = 2\ s$.

Figure 3 presents the eigenvalues and eigenvectors obtained for the bathymetry process. This latter is supposed to be a Gaussian process described with an exponential spatial autocorrelation function. Given a threshold value $\epsilon = 10^{-2}$, 8 eigenvalues are required to correctly sample the bathymetric field. In other words, the stochastic dimension of this problem is set at 8. For each of the four numerical methods tested in this study, a 1000 Monte-Carlo simulations are run while the POD is then used to reduce the output space. In this case, only two modes are retained to correctly represent the water height and the water velocity fields. The PCE is built over each POD mode corresponding to 4 PCEs (2 for the water height and 2 for the water velocity) instead of $2 \times 100$ decompositions in the conventional approach. An iterative procedure is used to

Table 1: Best polynomial degree along with the LOO error for the two POD modes in the water height and water velocity using different numerical schemes for the dam-break problem over a flat bed.

| | First POD mode | | | |
|---|---|---|---|---|
| | FVC | Lax-Friedrichs | Roe | Rusanov |
| Optimal polynomial degree ($h$) | 3 | 3 | 3 | 3 |
| Optimal polynomial degree ($u$) | 3 | 3 | 3 | 3 |
| LOO error ($h$) | $1.5 \times 10^{-9}$ | $1.8 \times 10^{-9}$ | $4.5 \times 10^{-11}$ | $3.8 \times 10^{-11}$ |
| LOO error ($u$) | $10^{-7}$ | $1.6 \times 10^{-6}$ | $2.3 \times 10^{-8}$ | $2.2 \times 10^{-9}$ |

| | Second POD mode | | | |
|---|---|---|---|---|
| | FVC | Lax-Friedrichs | Roe | Rusanov |
| Optimal polynomial degree ($h$) | 3 | 3 | 3 | 3 |
| Optimal polynomial degree ($u$) | 3 | 3 | 3 | 3 |
| LOO error ($h$) | $4.8 \times 10^{-9}$ | $3.6 \times 10^{-9}$ | $0.75 \times 10^{-11}$ | $0.4 \times 10^{-11}$ |
| LOO error ($u$) | $8.8 \times 10^{-7}$ | $2 \times 10^{-6}$ | $1.3 \times 10^{-8}$ | $5.8 \times 10^{-9}$ |

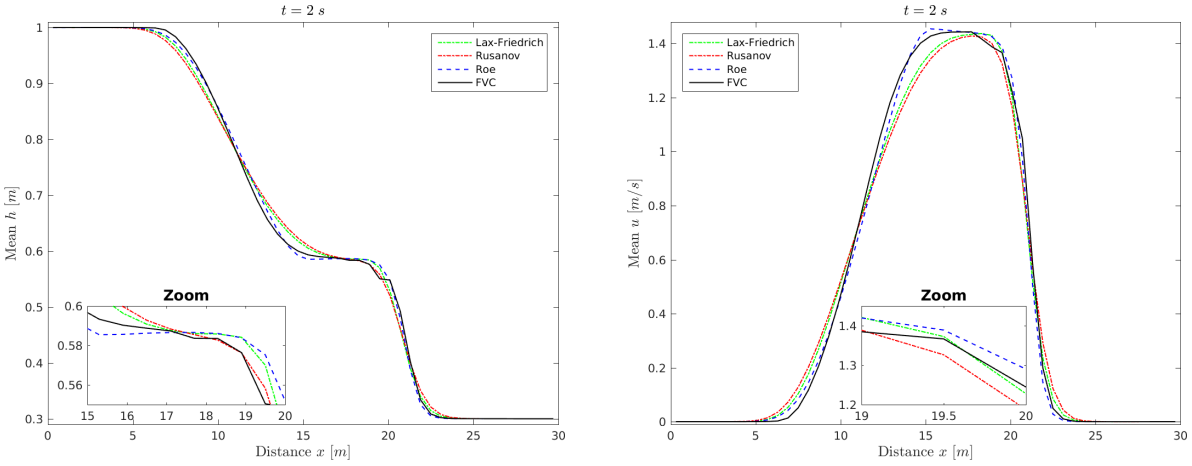


Figure 4: Mean solutions obtained using Lax-Friedrichs, Roe, Rusanov and FVC methods for water height (left plot) and water velocity (right plot) for the dam-break problem over a flat bed.

determine the best polynomial degree along with the LOO error that is used to assess the robustness of the considered PCE.

Table 1 summarizes the results obtained for the first and the second POD modes. It is evident that a polynomial of degree 3 is enough to correctly reproduce the uncertainty translated from errors in the bathymetry in the four methods considered in this study. Depending on the numerical method and the output field (water height or water velocity), the LOO error varies from $10^{-11}$ to $10^{-6}$ which makes the surrogate model very reliable for uncertainty quantification in this class of dam-break problems. The results obtained for the mean and variance fields are depicted in Figure 4 and Figure 5, respectively. As can be seen from these results, the mean fields highlight some well-established results on the accuracy of the numerical methods, *i.e.* the numerical diffusion is more pronounced for the Lax-Friedrichs method than the FVC method. However, the variance fields show noticeable differences in these numerical methods particularly where the hydraulic shock. It is also clear that the FVC and Roe methods are more sensitive to changes in the bathymetry than the Rusanov and Lax-Friedrichs methods. Note that the FVC and Roe methods, being more accurate than the Rusanov and Lax-Friedrichs methods, are more sensitive to the stochastic inputs. Therefore, these numerical

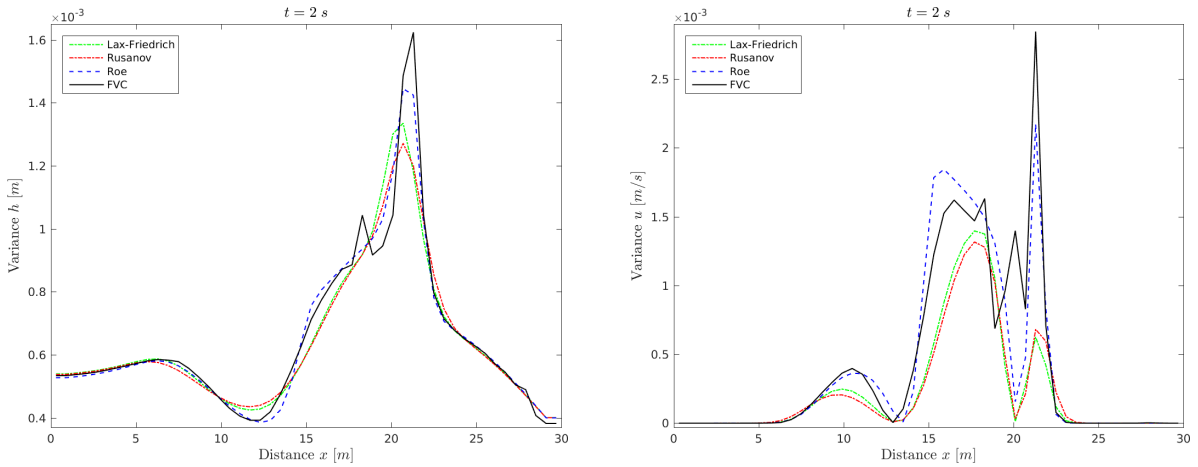


Figure 5: Variance solutions obtained using Lax-Friedrichs, Roe, Rusanov and FVC methods for water height (left plot) and water velocity (right plot) for the dam-break problem over a flat bed.

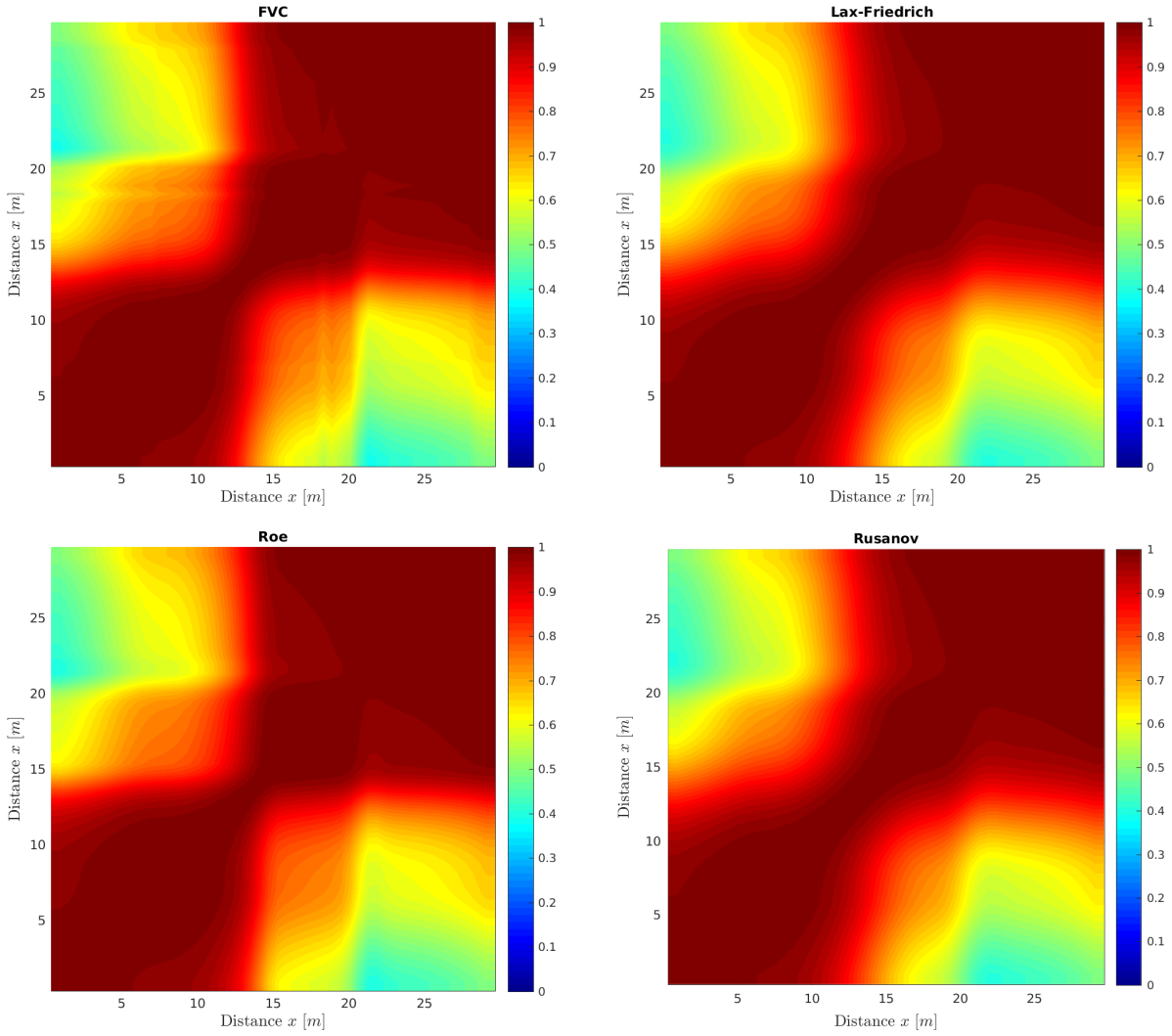


Figure 6: Spacial correlation field for water height using Lax-Friedrichs, Roe, Rusanov and FVC methods for the dam-break problem over a flat bed.

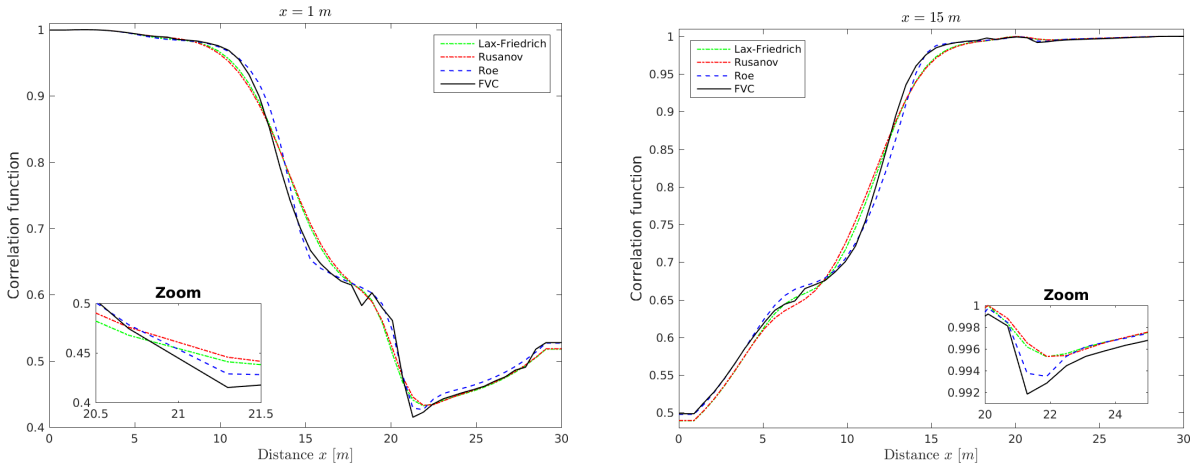


Figure 7: Correlation functions for water height at $x = 1\ m$ (left plot) and $x = 15\ m$ (right plot) using Lax-Friedrichs, Roe, Rusanov and FVC methods for the dam-break problem over a flat bed.

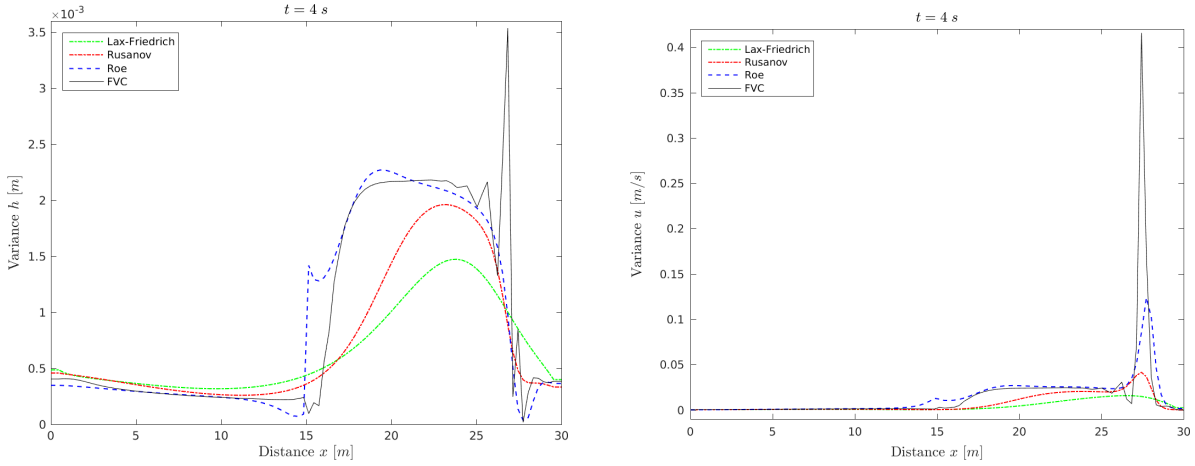


Figure 8: Same as Figure 5 but at time $t = 4\ s$.

methods display more uncertainty in both water height and water velocity solutions, especially around the discontinuity. Furthermore, regarding their importance in the context of data assimilation for bathymetric corrections, the correlation matrices and functions are presented for this test example. Figure 6 illustrates the correlation matrices using the four numerical methods whereas, Figure 7 represents the correlation function in the spatial points located at $x = 1\ m$ and $15\ m$ in the computational domain. There are no obvious differences in these statistical informations for this test case.

Next, we present numerical results for later time at $t = 4\ s$ to quantify the propagation of uncertainty in the hydraulic model. **It should be noted that at each time step, another surrogate model is built for this test case.** Figure 8 displays the obtained results at this simulation time. Comparing these results with those shown in Figure 5, clearly the uncertainty is very sensitive to the simulation time and therefore to the hydrodynamic itself. This confirms that the uncertainty quantification should be addressed carefully regarding the considered physics and the results for one simulation time are far to be generalized to another simulation time. Regarding the numerical methods considered in this study, the results are quite similar to the previous case and the high accurate methods such as FVC and Roe methods, display more uncertainties around the discontinuity location compared to the other methods, and this uncertainty tends to propagate spatially within the simulation time.

Next we compare sensitivity results of the uncertainty quantification for two different meshes with 100 and

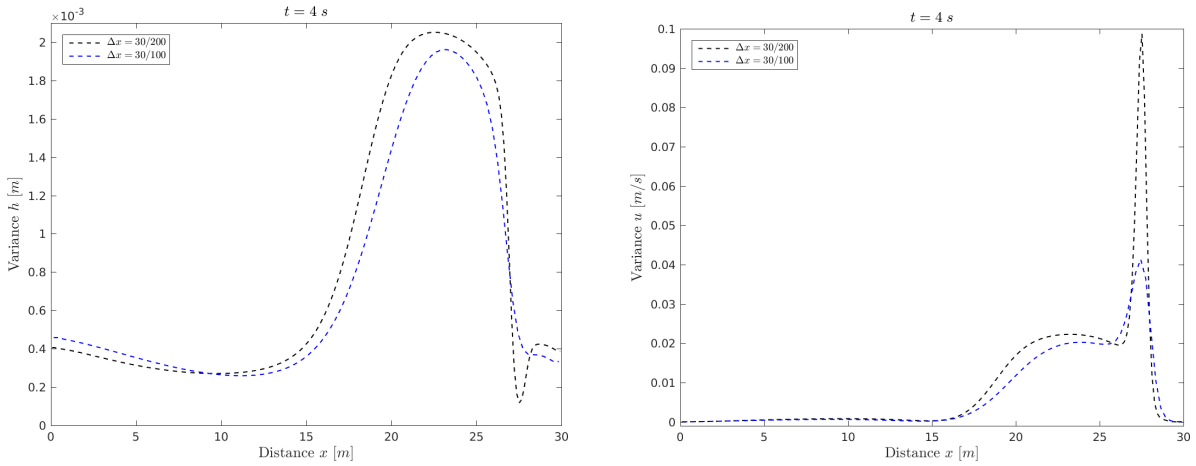


Figure 9: Variance solutions obtained using Rusanov method for water height (left plot) and water velocity (right plot) at time $t = 4\ s$ for the dam-break problem over a flat bed using two different meshes with 100 and 200 control volumes.

Table 2: Variation coefficients in the bathymetry $CV_b$ and in the Manning $CV_m$ used in our simulations for the dam-break problem over a non-flat bed.

| | Test 1 | Test 2 | Test 3 | Test 4 |
|---|---|---|---|---|
| $CV_m$ | 0.05 | 0.01 | 0.01 | 0.05 |
| $CV_b$ | 0.05 | 0.05 | 0.1 | 0.1 |

200 control volumes. For brevity in the presentation, only results obtained using the Rusanov method are taken under consideration since this case is representative of all the other methods. Figure 9 shows the results of the uncertainty and its sensitivity to the spatial discretization. Note that refining the mesh will ultimately improve the accuracy of the model and intuitively the uncertainty should decrease. This statement is only true for the upstream region and it is especially highlighted for the water height. However, around the area where the shock occurs, refining the spatial discretization will lead to more uncertainty mainly because of the complex physics occurring in this area.

### 5.2 Dam-break problem over a non-flat bed

In this test example we consider a dam-break problem over a non-flat bed defined as

$$Z(x) = \frac{1}{5} \exp\left(-\frac{(x-15)^2}{20}\right).$$

The aim here is to study effects of the supposed uncertainty in both the bathymetry and the Manning coefficient on the hydraulic state over a non-flat bed. Using a mean value of 0.025 for the Manning coefficient, a mesh with 100 control volumes and a simulation time of $t = 2\ s$. Four different tests, as described in Table 2, are considered in this study. We follow the same strategy to quantify the uncertainty as discussed in section 3. As described in Table 3, two POD modes are enough to represent all the uncertainty displayed by the hydraulic state. In general, a polynomial of degree 3 is able to correctly study the uncertainty as the LOO error ranges from $10^{-11}$ to $10^{-6}$ depending on which numerical method is used.

The variance of both water heights and velocity fields using the considered numerical methods and for the cases listed in Table 2 are displayed in Figure 10. It is notable that our comments in the previous section are still

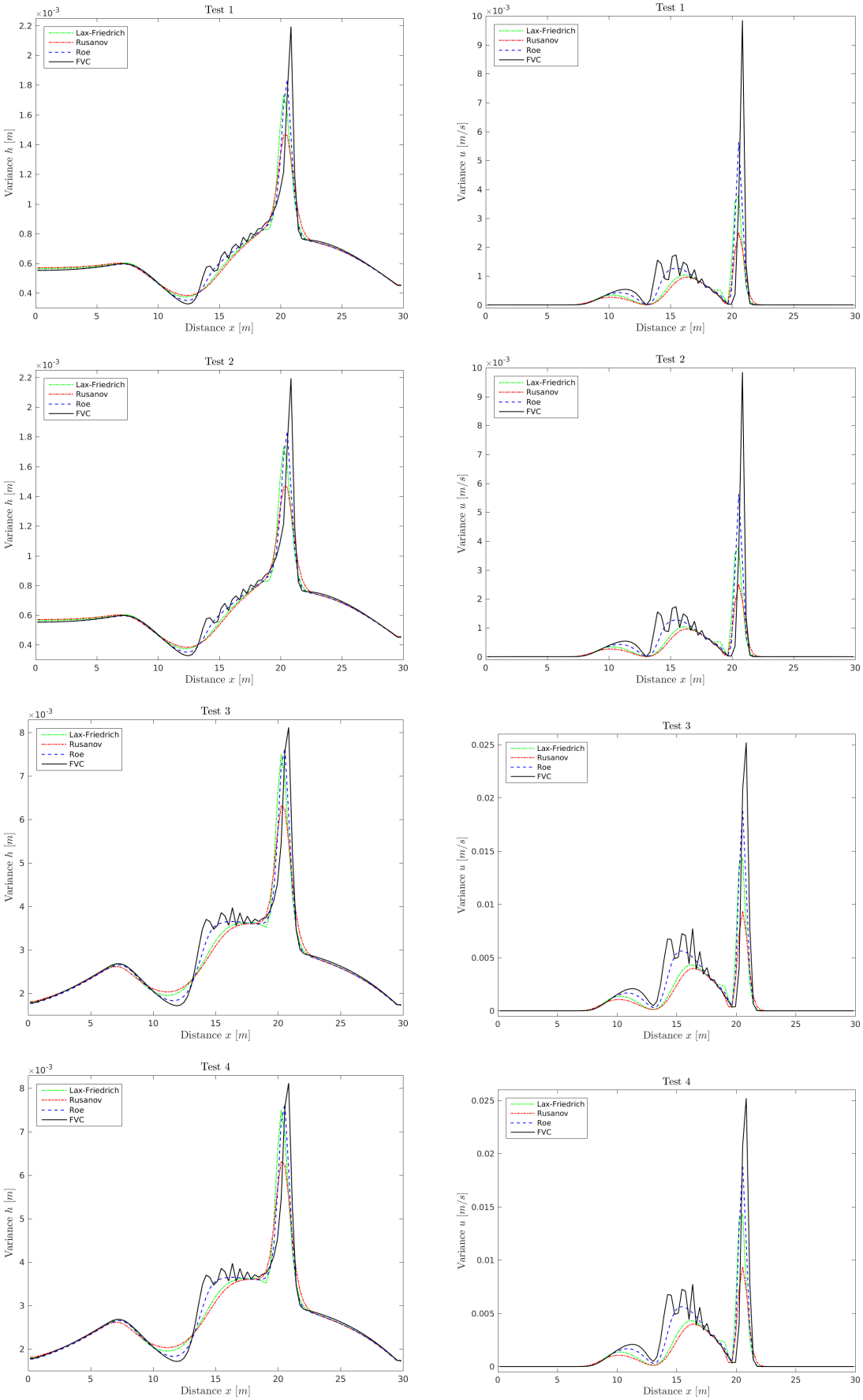


Figure 10: Variance solutions for the water height $h$ (left plots) and water velocity $u$ (right plots) obtained for the dam-break problem over a non-flat bed at time $t = 2\ s$ in the four considered tests listed in Table 2.

Table 3: Best polynomial degree along with the LOO error for the two POD modes in the water height and water velocity using different numerical schemes for the dam-break problem over a non-flat bed.

| | First POD mode | | | |
|---|---|---|---|---|
| | FVC | Lax-Friedrichs | Roe | Rusanov |
| Optimal polynomial degree ($h$) | 3 | 3 | 3 | 3 |
| Optimal polynomial degree ($u$) | 3 | 3 | 3 | 3 |
| LOO error ($h$) | $0.2 \times 10^{-10}$ | $0.5^{-9}$ | $2.3^{-11}$ | $2.2^{-11}$ |
| LOO error ($u$) | $10^{-8}$ | $1.6^{-6}$ | $0.8^{-8}$ | $1.1^{-8}$ |

| | Second POD mode | | | |
|---|---|---|---|---|
| | FVC | Lax-Friedrichs | Roe | Rusanov |
| Optimal polynomial degree ($h$) | 3 | 3 | 3 | 3 |
| Optimal polynomial degree ($u$) | 3 | 3 | 3 | 3 |
| LOO error ($h$) | $1.3\ 10^{-10}$ | $2.4\ 10^{-9}$ | $3.7\ 10^{-11}$ | $1.5\ 10^{-11}$ |
| LOO error ($u$) | $1.1\ 10^{-8}$ | $0.6\ 10^{-6}$ | $1.6\ 10^{-8}$ | $0.7\ 10^{-8}$ |

valid for these simulations here as well. The variance does not significantly change from a numerical method to another however, near the hydraulic jump, the uncertainty increases and therefore the calibration data need to be more accurate when using an accurate method. It is clear that the variance around the hydraulic jump for the FVC method attends the highest values. Figure 10 also points out that the hydraulic state is more sensitive to the bathymetry than the Manning coefficient. For instance, the variance of the two last tests in Table 2 is four times the variance of the first two tests regardless of the uncertainty taken in the Manning coefficient.

# 6 Results for two-layer shallow water flows

We present numerical results for the two-layer shallow water equations (3) subject to different hydraulic conditions. We consider two test examples including a lock-exchange flow problem and the problem of flow exchange in the Strait of Gibraltar. We focus mainly on assessing the uncertainty of the hydraulic states composed by the hydraulic layers and two velocities related to these layers with respect to the bathymetry. Notice that, in absence of exact expressions for the eigenvalues in the system (3), the Roe method can not be applied and only results obtained using the Lax-Friederich, Rusanov and FVC methods are presented in this section. Since the purpose here is to quantify the uncertainty regarding the numerical method used, a small variation of the bathymetry around the mean is considered to guaranty stability in the considered numerical methods.

## 6.1 Lock-exchange flow problem

In the first example we consider the lock-exchange problem proposed in [21]. The bottom topography is considered to be a Gaussian function defined as

$$Z\left(x\right) = \exp\left(-x^2\right) - 2. \tag{37}$$

The two layers are initially separated and the lighter water is on the left while the heavier one is on the right *i.e.*,

$$\left(h_1(x,0), h_2(x,0)\right)^\top = \begin{cases} \left(-Z\left(x\right), 0\right)^\top, & \text{if} \quad x \leq 0, \\ \left(0, -Z\left(x\right)\right)^\top, & \text{elsewhere,} \end{cases} \qquad u_1(x,0) = u_2(x,0) = 0. \tag{38}$$

Table 4: Best polynomial degree along with the LOO error for the POD mode in the water height and water velocity using different numerical schemes for the lock-exchange flow problem.

| | FVC | Lax-Friedrichs | Rusanov |
|---|---|---|---|
| Optimal polynomial degree ($h_1$) | 2 | 2 | 3 |
| Optimal polynomial degree ($h_2$) | 2 | 2 | 2 |
| Optimal polynomial degree ($u_1$) | 2 | 2 | 2 |
| Optimal polynomial degree ($u_2$) | 2 | 2 | 2 |
| LOO error ($h_1$) | $2 \times 10^{-5}$ | $1.9 \times 10^{-5}$ | $8.8 \times 10^{-6}$ |
| LOO error ($h_2$) | $5.8 \times 10^{-6}$ | $4.7 \times 10^{-6}$ | $1.2 \times 10^{-8}$ |
| LOO error ($u_1$) | $8.8 \times 10^{-6}$ | $10^{-5}$ | $1.6 \times 10^{-8}$ |
| LOO error ($u_2$) | $1.1 \times 10^{-5}$ | $1.2 \times 10^{-5}$ | $1.1 \times 10^{-8}$ |

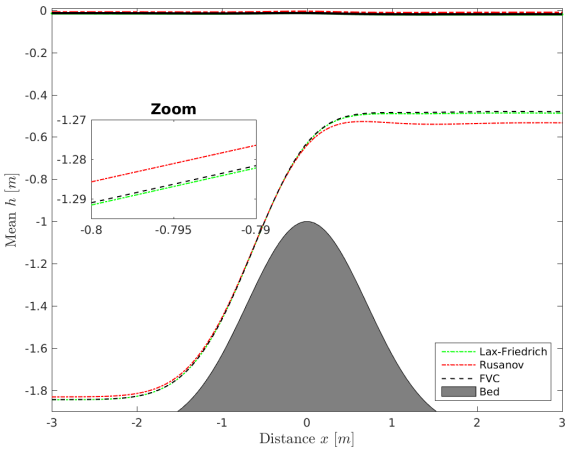


Figure 11: Mean value of the hydraulic state for the lock-exchange flow problem.

The density ratio is $\rho_1/\rho_2 = 0.98$, the computational domain is $[-3, 3]$ and the boundary conditions are imposed on the water discharges $q_1 = h_1 u_1$ and $q_2 = h_2 u_2$ as $q_1 = -q_2$ at each end of the domain. The variation coefficients in the bathymetry and in the Manning are $CV_b = 0.01$ and $CV_m = 0.01$, respectively. In this hydraulic problem, the heavier water propagates to the left while the lighter one moves to the right. The solution is expected to converge to a smooth steady state and we compute the numerical steady-state solution on a mesh with 100 control volumes. Following the same methodology described in section 2, the uncertainty of the hydraulic state is quantified. For this flow problem, first the assessment of the surrogate model is performed. The obtained results show that given the uncertainty of the bathymetry, only one POD mode is needed. Indeed, the perturbation in the bathymetry highly affects the numerical stability and therefore, the domain of randomness of this parameter has been narrowed which would also translate to a narrow domain of uncertainty in the hydraulic state. Table 4 summarizes the results of the surrogate model built for each component of the hydraulic state and for each numerical method. Here, a polynomial degree of 2 is enough to correctly estimate the uncertainty and the LOO error ranges from $10^{-8}$ to $10^{-5}$. This makes the proposed surrogate model very reliable for the purpose of uncertainty quantification.

Figure 11 illustrates the mean values of the water free-surface and water interface corresponding to the two water layers whereas, Figure 12 reports the mean value of the water velocities for the three numerical methods considered in the present study. It is clear that there are no large differences between the results obtained for the water free-surface solutions. For the water velocities, the Rusanov method is underestimating the mean solutions compared to the Lax-Friedrichs and FVC methods, but the differences are also very small. It is also worth noting that it could be corrected using observations. Figure 13 and Figure 14 display the results obtained

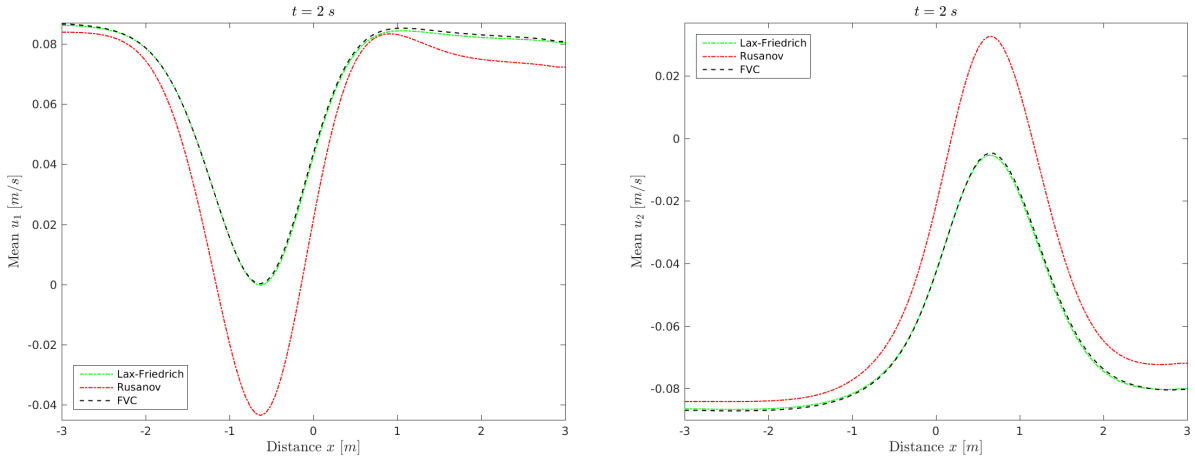


Figure 12: Mean values for water velocity $u_1$ (left plot) and $u_2$ (right plot) for the lock-exchange flow problem.

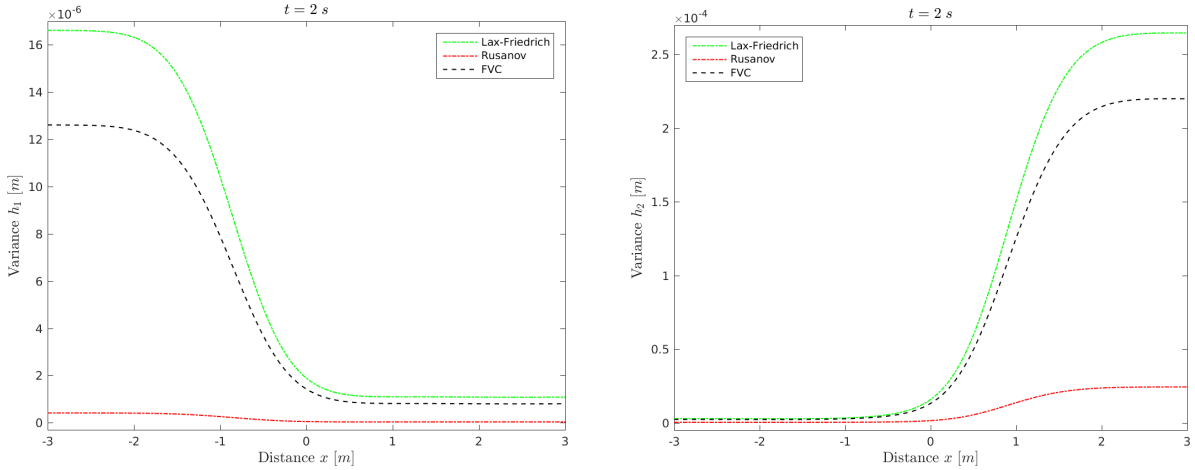


Figure 13: Variance in water height $h_1$ (left plot) and $h_2$ (right plot) for the lock-exchange flow problem.

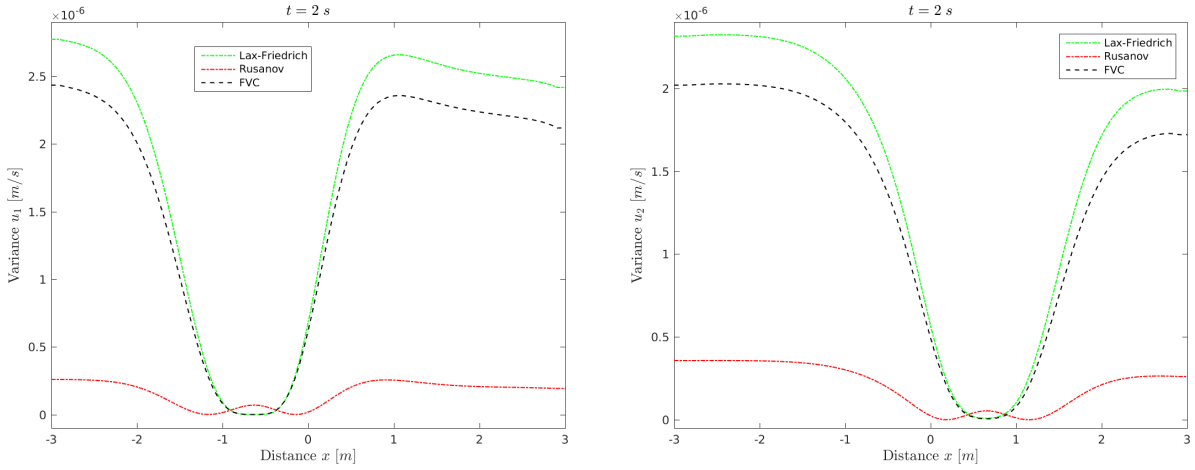


Figure 14: Variance in water velocity $u_1$ (left plot) and $u_2$ (right plot) for the lock-exchange flow problem.

for the variance in water height and water velocity, respectively. In contrary to the mean value, these results are very different. It is clear that the Lax-Friedrichs method generates the largest amount of the uncertainty

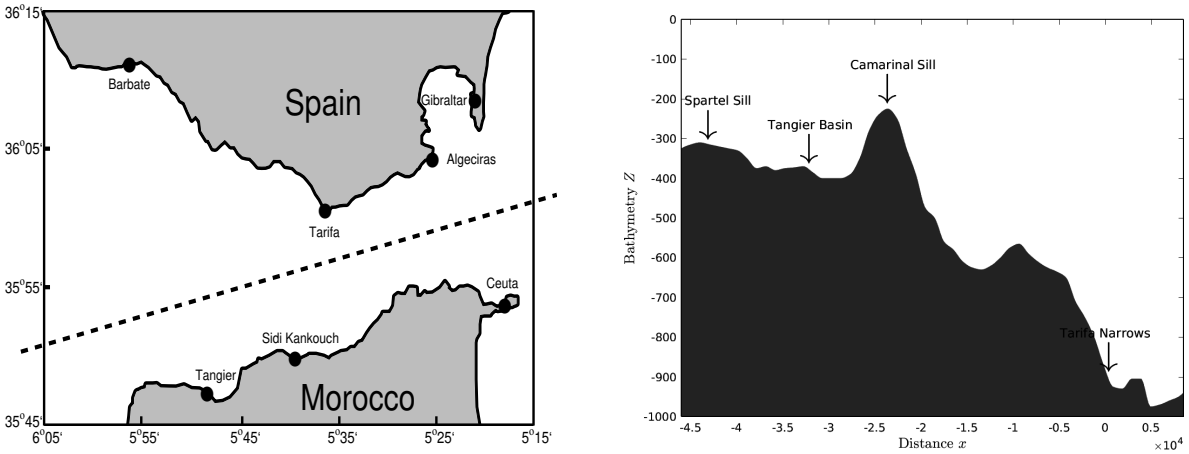


Figure 15: Schematic map of the Strait of Gibraltar along with relevant locations (left plot) and the bathymetry used in our simulations (right plot).

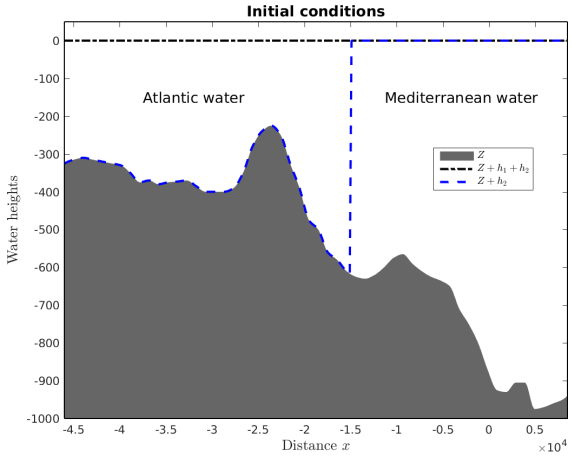


Figure 16: Initial conditions for the flow exchange through the Strait of Gibraltar.

compared to the other methods. This large amount of uncertainty is mainly driven by the model itself as the numerical diffusion in this case is also very important compared to the single-layer shallow water model. The hydraulic calculations using the FVC method are also less uncertain compared to the Rusanov method.

### 6.1.1 Flow exchange through the Strait of Gibraltar

Our last example consists of solving the problem of flow exchange through the Strait of Gibraltar. This hydraulic problem is selected because it presents a real example of two-layer shallow water flows for two major reasons. Firstly, the domain of the Strait of Gibraltar is a large-scale domain including high gradients of the bathymetry and well-defined shelf regions. Secondly, the Strait contains two water bodies with different densities, which present a challenge in the shallow water modelling. Indeed, the basic Oceanic circulation in the Strait of Gibraltar consists of an upper layer of cold, fresh surface Atlantic water (with density $\rho_1 = 1027\ kg/m^3$) and an opposite deep current of warmer and salty outflowing Mediterranean water (with density $\rho_2 = 1029\ kg/m^3$), compare for example [23]. A schematic map of the Strait of Gibraltar along with relevant locations is depicted in the left plot of Figure 15. The system is bounded to the north and south by the Iberian and African continental forelands, respectively, and to the west and east by the Atlantic Ocean and the Mediterranean sea. In geographical coordinates, the Strait of Gibraltar is $35^o45'$ to $36^o15'$ N latitude and $5^o15'$ to $6^o05'$

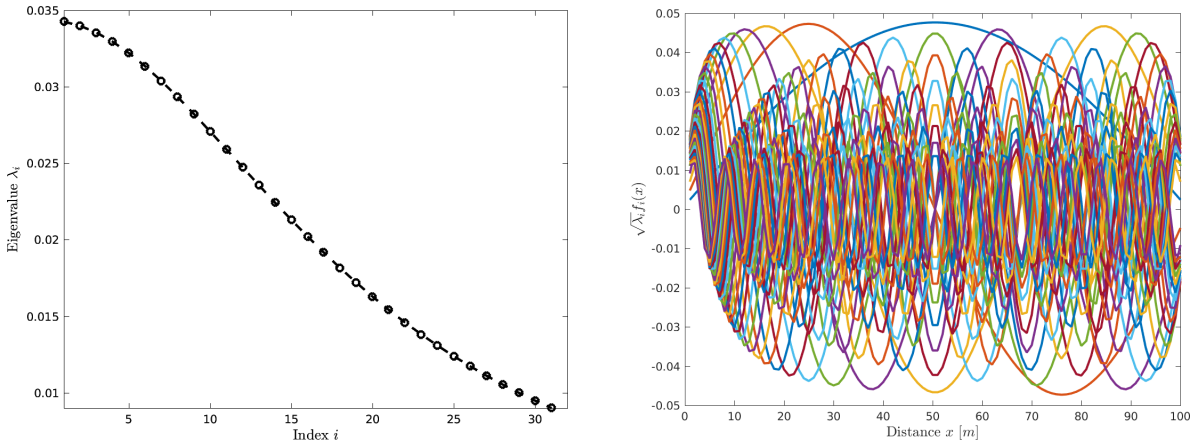


Figure 17: Eigenvalues (left plot) and associated eigenfunctions (right plot) used in the KL decomposition to sample the bathyemtric field over the Strait of Gibraltar using an exponential Kernel with a correlation length of 10 $km$.

W longitude. Here, we consider a one-dimensional cross-section along the Strait (obtained by a longitudinal section along the dashed line in the left plot of Figure 15). The computational domain and the associated bathymetry are displayed in the right plot of Figure 15. This restricted domain has also been considered in [26, 24] among others.

This hydraulic problem was experimentally studied in [9] and numerically solved in [17] using a semi-Lagrangian scheme and in [26] using a discontinuous Galerkin method. The flow parameters of this experiment are those used in [9, 17, 26]. Thus, an artificial dam is included in the Strait separating the water bodies from the Atlantic Ocean and the Mediterranean sea. On the boundaries, the ratio between the water discharges is set to 1, *i.e.*, $q_1 = -q_2$ at each boundary node. Here, the computational domain is covered with 100 control volumes and Figure 16 exhibits the initial water heights $h_1$ and $h_2$. We evaluate the uncertainty displayed by the three numerical methods used to identify the flow in this region. The variation coefficients in the bathymetry and in the Manning are $CV_b = 0.005$ and $CV_m = 0.01$, respectively. The aim is to quantify the uncertainty generated by an error in the reconstruction of the Strait bathymetry using the three different numerical methods. First, a KL decomposition is carried out in order to represent the stochastic aspect of the bathymetric field. **The spatial correlation matrix was supposed to have an exponential Kernel as proposed in [51] and the correlation length is set to a value of** 10 $km$**, which is the minimum distance existing between two different observations stations**. Figure 17 displays the results of this decomposition. It is evident that the stochastic dimension of our hydraulic problem is very high as more than 30 modes are needed for the LOO error to be under the threshold error of $10^{-2}$. This is mainly attributed to the size of the physical domain used in the modelling of flow exchange through the Strait of Gibraltar.

Following the methodology discussed in Section 4, the fields of water height and velocity are decomposed using a POD. Then, the uncertainty is supposed to be described by the eigenvalues of the POD approach. For the considered hydraulic conditions and given the threshold used in the POD, only one mode is enough to represent correctly the uncertainty. Again, the considered variance in the bathymetry has been selected reasonably small in order to maintain the stability of the numerical schemes. Table 5 describes the assessment of the PCE for this mode in all the three numerical methods used in this study. Overall, a polynomial degree of 3 is able to correctly resolve the uncertainty with an error estimated to be less than $10^{-5}$ for the four fields considered in a two-layer shallow water model and for the three numerical methods considered for its numerical simulations.

Figure 18 represents the mean fields of the water free-surface $h_1$ and the water interface $h_2$ obtained at the steady-state time. As can be seen, the Mediterranean water moves downstream and exits the Strait resulting in the formation of strong and weak shocks. Under actual flow conditions, it is clear that a hydraulic jump

Table 5: Best polynomial degree along with the LOO error for the POD mode in the hydraulic states using different numerical schemes for the problem of flow exchange through the Strait of Gibraltar.

| | FVC | Lax-Friedrichs | Rusanov |
|---|---|---|---|
| Optimal polynomial degree ($h_1$) | 3 | 3 | 3 |
| Optimal polynomial degree ($h_2$) | 3 | 3 | 3 |
| Optimal polynomial degree ($u_1$) | 3 | 3 | 3 |
| Optimal polynomial degree ($u_2$) | 3 | 3 | 3 |
| LOO error ($h_1$) | $3.2 \times 10^{-6}$ | $2.9 \times 10^{-6}$ | $3.1 \times 10^{-6}$ |
| LOO error ($h_2$) | $2.6 \times 10^{-5}$ | $2.1 \times 10^{-5}$ | $2.3 \times 10^{-5}$ |
| LOO error ($u_1$) | $1.9 \times 10^{-7}$ | $1.5 \times 10^{-7}$ | $1.5 \times 10^{-7}$ |
| LOO error ($u_2$) | $2.1 \times 10^{-7}$ | $1.7 \times 10^{-5}$ | $1.75 \times 10^{-7}$ |

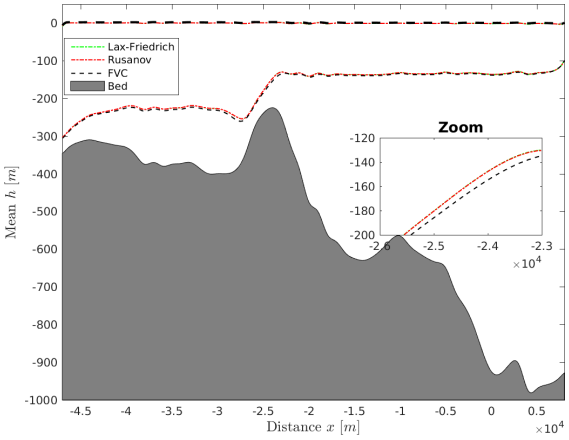


Figure 18: Mean hydraulics for the problem of flow exchange through the Strait of Gibraltar.

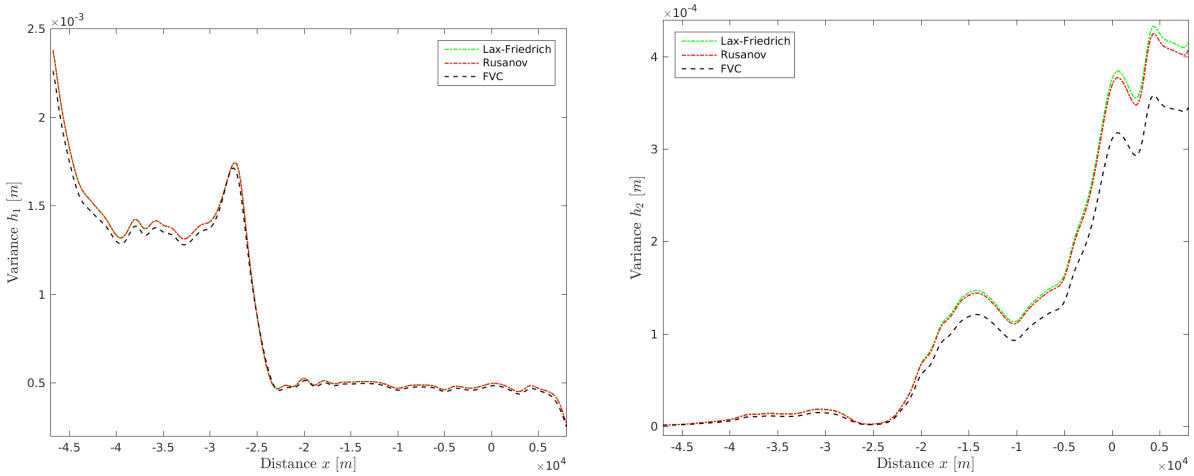


Figure 19: Variance in water height $h_1$ (left plot) and $h_2$ (right plot) for the problem of flow exchange through the Strait of Gibraltar.

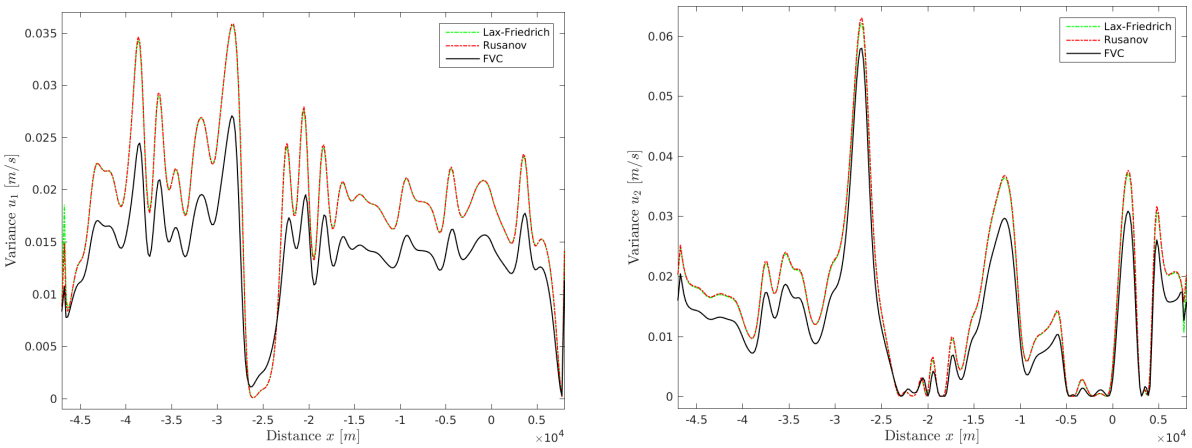


Figure 20: Variance in water velocity $u_1$ (left plot) and $u_2$ (right plot) for the problem of flow exchange through the Strait of Gibraltar.

is detected in the water interface near the Camarinal Sill. The exchange discharge predicted by our method is $q_1 = -q_2 = 0.856\ sv$. There is an excellent agreement between the numerical results obtained by the considered methods and those from previous studies reported in [9, 17, 26]. It is worth remarking that the exact solutions to these examples are not available, but the computed solutions using the proposed methods seem to converge to the physically relevant solutions in all selected test cases. The numerical methods used in this case capture the shock accurately and do not diffuse the fronts or exhibit oscillations near the steep gradients in the computational domain. Overall, there is a good convergence between the results of the three numerical methods considered for this exchange flow problem. This agreement is also observed in the mean velocity field of the upper layer for the three numerical methods. However, for the water interface, there is a slight difference between the mean fields obtained using the FVC method and those obtained using the other considered numerical methods for this flow problem. These discrepancies are also observed in the variance fields, see Figure 19 and Figure 20. Hence, the FVC method being numerically more accurate, it displays less uncertainty than the diffusive methods. This case example demonstrates that when the physical hydraulics is very complex, the use of a highly accurate numerical method is necessary and the uncertainty quantified by this method is expected to be well predicted.

# 7 Conclusions

In the current study, we have introduced a new methodology to quantify uncertainties generated by bathymetric forces on the hydraulic states. This methodology allows to overcome many known problems in the uncertainty quantification for computational hydraulics and its combines: (i) a Karhunen-Loève decomposition to sample stochastic processes independently of the required number of control volumes in the mesh to achieve reasonable accuracy, (ii) a sparse polynomial chaos expansion to overcome the well known problem of curse of dimensionality risen by the Karhunen-Loève decomposition at a reasonable computational cost, and (iii) a proper orthogonal decomposition for hydraulic states to increase further the efficiency of the proposed method. Furthermore, we have examined the uncertainty quantification in a class of finite volume methods for free-surface flows. The models consist of both single-layer and two-layer shallow water equations with stochastic bathymetry and Manning roughness coefficients. Numerical results are presented for dam-break problems for single-layer models and for exchange flows for two-layer models including the problem of flow exchange through the Strait of Gibraltar. It has been shown that the uncertainty in the hydraulic calculations is very sensitive to both the mathematical models and the numerical methods selected for their approximations. It has been found that the uncertainty depends on the numerical scheme used in the solution procedure. While accurate numerical schemes may increase the precision of the deterministic solution, they need continuous data

to make them usable in the operational context. In addition, the simulation time has also been highlighted to be a subject of uncertainty such that the uncertainty tends to increase with the simulation time. Efforts on correcting the bathymetric fields would be very relevant to reduce the uncertainty of such models. In the case of two-layer shallow water flows such as flow through the Strait of Gibraltar, highly accurate finite volume schemes are shown to be the best numerical tools to quantify uncertainties in their involved parameters. **The present study demonstrated the need to accurately choose the appropriate numerical tools with respect to the uncertainty depicted in the bathymetry since, when the uncertainty goes beyond a certain level, some numerical schemes tend to loose their stability. For instance, highly accurate finite volume methods exhibit oscillatory behavior at the hydraulic jumps.** Although, we have studied only the case of one-dimensional problems in hydraulics, the extension to two-dimensional problems would be an encouraging next step and requires an in-depth study on high-order finite volume methods for solving stochastic shallow water equations in unstructured meshes.

**Acknowledgments.** This work was partly performed while the first author was visiting IWRI at University Mohammed VI Polytechnic Benguerir. Financial support provided by IWRI is gratefully acknowledged.

## References

[1] A. Abdedou, A. Soulaïmani, and G.W. Tchamen. Uncertainty propagation of dam break flow using the stochastic non-intrusive B-splines Bézier elements-based method. *Journal of Hydrology*, 590:125342, 2020.

[2] E. Audusse, F. Benkhaldoun, S. Sari, M. Seaid, and P. Tassi. A fast finite volume solver for multi-layered shallow water flows with mass exchange. *Journal of Computational Physics*, 272:23–45, 2014.

[3] F. Benkhaldoun and M. Seaid. A simple finite volume method for the shallow water equations. *J. Comp. Applied Math*, 234:58– 72, 2010.

[4] M. Berveiller, B. Sudret, and M. Lemaire. Stochastic finite element: A non intrusive approach by regression. *European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique*, 15:81–92, 2006.

[5] G. Blatman and B. Sudret. An adaptive algorithm to build up sparse polynomial Chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25(2):183–197, 2010.

[6] G. Blatman and B. Sudret. Adaptative sparse polynomial Chaos expansion based on Least Angle Regression. *Journal of Computational Physics*, 230:2345–2367, 2011.

[7] G. Blatman and B. Sudret. *Sparse polynomial Chaos expansions of vector-valued response quantities*. CRC Press/Balkema, 2013.

[8] S. Bozzi, G. Passoni, P. Bernardara, N. Goutal, and A. Arnaud. Roughness and discharge uncertainty in 1D water level calculations. *Environmental Modeling & Assessment*, 20(4):343–353, 2015.

[9] H. Bryden, J. Candela, and T. Kinder. Exchange through the Strait of Gibraltar. *Progress in Oceanography*, 33(3):201–248, 1994.

[10] L. Cea and J. French. Bathymetric error estimation for the calibration and validation of estuarine hydrodynamic models. *Estuarine, Coastal and Shelf Science*, 100:124–132, 2012.

[11] E. Chávarri, A. Crave, M. Bonnet, A. Mejía, J. Da Silva, and J. Guyot. Hydrodynamic modelling of the amazon river: Factors of uncertainty. *Journal of South American Earth Sciences*, 44:94–103, 2013.

[12] M. Chevreuil and A. Nouy. Model order reduction based on proper generalized decomposition for the propagation of uncertainties in structural dynamics. *International Journal for Numerical Methods in Engineering*, 89(2):241–268, 2012.

[13] R. Crisovan, D. Torlo, R. Abgrall, and S. Tokareva. Model order reduction for parametrized nonlinear hyperbolic problems as an application to uncertainty quantification. *Journal of Computational and Applied Mathematics*, 348:466–489, 2019.

[14] T.D. Dao, Q. Serra, S. Berger, and E. Florentin. Error estimation of polynomial chaos approximations in transient structural dynamics. *International Journal of Computational Methods*, 17(10):2050003, 2020.

[15] Z. Drmac. A posteriori computation of the singular vectors in a preconditioned Jacobi SVD algorithm. *IMA journal of numerical analysis*, 19(2):191–213, 1999.

[16] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of Statistics*, 32:407–499, April 2004.

[17] M. El-Amrani and M. Seaid. An essentially non-oscillatory semi-Lagrangian method for tidal flow simulations. *International Journal for Numerical Methods in Engineering*, 81(7):805–834, 2010.

[18] N. El Moçayd. La décomposition en polynômes du chaos pour lámélioration de lássimilation de données ensembliste en hydraulique fluviale. *PhD Université de Toulouse, INP*, 2017.

[19] N. El Moçayd, S. Mohamed, D. Ouazar, and M. Seaid. Stochastic model reduction for polynomial Chaos expansion of acoustic waves using proper orthogonal decomposition. *Reliability Engineering & System Safety*, 195:106733, 2020.

[20] N. El Moçayd, S. Ricci, N. Goutal, M. Rochoux, S. Boyaval, C. Goeury, D. Lucor, and O. Thual. Polynomial surrogates for open-channel flows in random steady state. *Environmental Modeling & Assessment*, pages 1–23, 2017.

[21] D. Farmer and L. Armi. Maximal two-layer exchange over a sill and through a combination of a sill and contraction with barotropic flow. *J. Fluid Mech.*, 164:53–76, 1986.

[22] I. Gejadze and P. Malaterre. Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model. *International Journal for Numerical Methods in Fluids*, 83(5):405–430, 2017.

[23] F. Gómez. The role of the exchanges through the Strait of Gibraltar on the budget of elements in western mediterranean sea: Consequences of humain-induced modifications. *Marine Pollution Bulletin.*, 46:685–694, 2003.

[24] C. González, E. Reyes, Ó. Álvarez, A. Izquierdo, M. Bruno, and R. Mañanes. Surface currents and transport processes in the Strait of Gibraltar: Implications for modeling and management of pollutant spills. *Ocean & Coastal Management*, 179:104869, 2019.

[25] N. Goutal, C. Goeury, R. Ata, S. Ricci, N. El Mocayd, M. Rochoux, H. Oubanas, I. Gejadze, and P. Malaterre. Uncertainty quantification for river flow simulation applied to a real test case: The Garonne valley. In *Advances in Hydroinformatics*, pages 169–187. Springer, 2018.

[26] N. Izem, M. Seaid, and M. Wakrim. A discontinuous Galerkin method for two-layer shallow water equations. *Mathematics and Computers in Simulation*, 120:12–23, 2016.

[27] A. Kendall. A survey of numerical methods for the solution of Fredholm integral equations of the second kind. 1976.

[28] O. Le Maıtre, M. Reagan, H. Najm, R. Ghanem, and O. Knio. A stochastic projection method for fluid flow: Ii. Random process. *Journal of computational Physics*, 181(1):9–44, 2002.

[29] R. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, Cambridge, 2002.

[30] T. Li, S. Deng, K. Zhang, H. Wei, R. Wang, J. Fan, J. Xin, and J. Yao. A non-intrusive parametrized reduced-order model for periodic flows based on extended proper orthogonal decomposition. *International Journal of Computational Methods*, 2021.

[31] M. Loève. Elementary probability theory. In *Probability Theory I*, pages 1–52. Springer, 1977.

[32] B. Merz and A. Thieken. Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, 309(1-4):114–132, 2005.

[33] H. Oubanas, I. Gejadze, P. Malaterre, and F. Mercier. River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model. *Journal of hydrology*, 559:638–647, 2018.

[34] F. Pappenberger and K. Beven. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5), 2006.

[35] M. Raisee, D. Kumar, and C. Lacor. A non-intrusive model reduction approach for polynomial Chaos expansion using proper orthogonal decomposition. *International Journal for Numerical Methods in Engineering*, 103(4):293–312, 2015.

[36] J. Refsgaard, J. Sluijs, A. Højberg, and P. Vanrolleghem. Uncertainty in the environmental modelling process–a framework and guidance. *Environmental modelling & software*, 22(11):1543–1556, 2007.

[37] M. Ricchiuto, P. Congedo, and A. Delis. *Runup and uncertainty quantification: sensitivity analysis via ANOVA decomposition.* PhD thesis, INRIA, 2014.

[38] P. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, 43:357–372, 1981.

[39] C. Schwab and R. Todor. Karhunen–Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100–122, 2006.

[40] M. Seaid. Semi-Lagrangian integration schemes for viscous incompressible flows. *Comp. Methods in App. Math.*, 4:392–409, 2002.

[41] J. Shaw, G. Kesserwani, and P. Pettersson. Probabilistic Godunov-type hydrodynamic modelling under multiple uncertainties: robust wavelet-based formulations. *Advances in Water Resources*, page 103526, 2020.

[42] C. Soize and R. Ghanem. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing*, 26(2):395–410, 2004.

[43] I. Sraj, O. Le Maître, O. Knio, and I. Hoteit. Coordinate transformation and polynomial Chaos for the bayesian inference of a Gaussian process with parametrized prior covariance function. *Computer Methods in Applied Mechanics and Engineering*, 298:205–228, 2016.

[44] J. Stoker. *Water waves.* Interscience Publishers, Inc, New York, 1986.

[45] C. Temperton and A. Staniforth. An efficient two-time-level Galerkin-characteristics semi-implicit integration scheme. *Quart. J. Roy. Meteor. Soc.*, 113:1025–1039, 1987.

[46] J. Wang and Z. Zhang. Evaluating riparian vegetation roughness computation methods integrated within HEC-RAS. *Journal of Hydraulic Engineering*, 145(6):04019020, 2019.

[47] Z.H. Wang, C. Jiang, X.X. Ruan, Y.Q. Zhang, Z.L. Huang, C.S. Wang, and T. Fang. Uncertainty propagation analysis of T/R modules. *International Journal of Computational Methods*, 16(07):1850105, 2019.

[48] N. Wiener. The homogeneous Chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

[49] D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1, Foundations*. Wiley, 1979.

[50] D. Xiu. *Numerical methods for stochastic computations: a spectral method approach*. Princeton university press, 2010.

[51] Y. Yoon, M. Durand, C. Merry, E. Clark, K. Andreadis, and D. Alsdorf. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of hydrology*, 464:363–375, 2012.