Méthodes stochastiques non-intrusives pour l’analyse de la propagation d’incertitudes: Application aux écoulements des ruptures de barrages

par

Azzedine ABDEDOU

THÈSE PAR ARTICLES PRÉSENTÉE À L’ÉCOLE DE TECHNOLOGIE SUPÉRIEURE COMME EXIGENCE PARTIELLE À L’OBTENTION DU DOCTORAT EN GÉNIE
Ph.D.

MONTRÉAL, LE 16 MARS 2021

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

PRÉSENTATION DU JURY

CETTE THÈSE A ÉTÉ ÉVALUÉE

PAR UN JURY COMPOSÉ DE:

M. Azzeddine Soulaïmani, directeur de thèse
Département de génie mécanique, École de technologie supérieure

M. Saad Bennis, président du jury
Département de génie de la construction, École de technologie supérieure

M. Antoine Tahan, membre du jury
Département de génie mécanique, École de technologie supérieure

M. Serge Prudhomme, examinateur externe indépendant
Département de mathématiques et de génie industriel, Polytechnique Montréal

ELLE A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE “05 MARS 2021”

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Le présent travail a été effectué au sein du laboratoire GRANIT de l'École de Technologie Supérieure, sous la direction du Professeur Azzeddine Soulaïmani. Qu'il trouve ici l'expression de ma profonde gratitude de m'avoir accueilli au sein de son groupe de recherche, pour son soutien constant et ses précieux conseils durant toutes ces années. Je vous suis reconnaissant.

Je remercie les membres du jury : Prof. Saad Bennis, Prof. Antoine Tahan et Prof. Serge Prudhomme qui m'ont fait l'honneur d'accepter d'évaluer ce travail.

Je remercie également Hydro-Québec pour leur soutien tout au long de ce projet, en particulier Dr Georges Tchamen (Hydro-Québec) pour son étroite collaboration.

Mes remerciements s'adressent aussi à : Amine Ben Hadj Ali, Jean-Marie Zokagoa, Pierre Jacquier, Vincent Delmas et tous mes collègues au sein du laboratoire GRANIT.

À mes filles : Lydia et Lynda ;

À ceux et celles qui me sont chers ;

Méthodes stochastiques non-intrusives pour l'analyse de la propagation d'incertitudes: Application aux écoulements des ruptures de barrages

Azzedine ABDEDOU

RÉSUMÉ

La présente thèse est une contribution à l'analyse de la propagation des incertitudes à travers les modèles numériques des écoulements de ruptures de barrages. La complexité de ces phénomènes et le temps de calcul relativement élevé requis par les schémas numériques rendent l'application des méthodes d'échantillonnage telles que Monte Carlo ou Latin Hypercube Sampling (LHS) très onéreuse en temps de calcul au regard de la taille considérable des échantillons que ces méthodes exigent pour une convergence satisfaisante.

Dans ce contexte, une nouvelle approche non-intrusive dénommée B-Splines Bézier Elements based Method (BSBEM) est proposée comme un outil efficace pour l'analyse de la propagation d'incertitudes pour des problèmes avec un comportement hyperbolique. Le caractère générique de l'approche proposée permet sa généralisation à d'autres problèmes de l'ingénierie. L'efficacité de la technique BSBEM est évaluée et comparée, dans un premier temps, avec la méthode des polynômes du chaos (PCE) et la méthode de Monte Carlo (considérée comme solution de référence), à travers une série d'exemples numériques reproductibles, et aussi en utilisant la solution analytique de Stoker décrivant une rupture idéalisée d'un barrage hydraulique avec la présence d'une discontinuité dans la réponse de sortie. La méthode proposée est par la suite appliquée à l'analyse de la propagation d'incertitudes d'une hypothétique rupture d'un barrage avec des données réelles (Rivière de Batiscan, province du Québec). Les paramètres d'entrée tels que, le débit amont, le coefficient de frottement de Strickler et la largeur moyenne de la brèche sont considérés comme des variables aléatoires à partir desquelles des incertitudes peuvent émaner et se propager à travers les modèles numériques. Les résultats obtenus montrent la capacité de l'approche BSBEM à prédire avec efficacité les moments statistiques et les densités de probabilité des réponses de sortie, représentées en termes d'hydrogramme, du niveau d'eau et du temps d'arrivée du front de l'onde de submersion avec des profiles lisses, contrairement aux prédictions des polynômes du chaos qui présentent un fort comportement oscillatoire.

Une autre contribution de la présente thèse concerne l'introduction d'une nouvelle approche non-intrusive de réduction de modèles appelée 'proper orthogonal decomposition-based B-Splines Bézier Elements Method (POD-BSBEM)'. La méthode adopte l'approche de la décomposition orthogonale aux valeurs propres (POD) à deux niveaux pour l'extraction de la base réduite. Un troisième niveau de POD est ensuite appliqué sur les données de projection résultantes afin de découpler la dépendance entre les modes temporels et les coefficients paramétriques, qui sont exprimés comme une décomposition en utilisant les fonctions de bases B-Splines locales dans chaque élément de Bézier. Les performances de la technique proposée sont évaluées et comparées avec d'autres approches notamment la technique du modèle réduit basée sur les réseaux de neurones artificiels (POD-ANN) et l'approche classique des polynômes du chaos (Full-PCE), et ce à travers des exemples numériques stochastiques et instationnaires. L'approche POD-BSBEM est ensuite appliquée pour l'analyse de la propagation des incertitudes paramétriques à travers un modèle numérique décrivant l'onde de submersion générée par une rupture hypothétique d'un barrage le long d'une rivière avec une bathymétrie complexe. Les principaux résultats obtenus montrent la fiabilité des prédictions de la méthode proposée dans l'approximation des moments statistiques avec un facteur de réduction significatif de l'effort de calcul dans les deux phases (apprentissage et prédictions) en comparaison avec les autres techniques.

Mots-clés: Propagation d'incertitudes, B-Splines, éléments de Bézier, polynômes du chaos, écoulements de ruptures de barrages, décomposition orthogonale aux valeurs propres, modèle réduit.

Non-intrusive stochastic methods for uncertainty propagation analysis : Application to dam-break flows

Azzedine ABDEDOU

ABSTRACT

This thesis is a contribution to the uncertainties propagation analysis through numerical models of dam-break flows. The complexity of these phenomena and the relatively high computation time required by numerical solvers render the application of classical sampling methods such as Monte Carlo, Latin Hypercube Sampling (LHS), computationally cumbersome due to the high sample size that require such techniques to reach a satisfactory convergence.

In this context, a novel non-intrusive approach named B-Splines Bézier Elements based Method (BSBEM) is proposed as an efficient tool for the uncertainty propagation analysis in physical hyperbolic problems. The generic character of the proposed approach allows it to be implemented for several engineering fields. The predictive efficiency of the BSBEM is first assessed and compared to the polynomial chaos expansion (PCE) and Monte Carlo (MC) methods using benchmark numerical examples and Stoker's analytical solution for an idealized dam-break flow with discontinuous outputs. The proposed methodology is then applied to analyze the uncertainty propagation of a hypothetical failure of an actual dam (Batiscan River) located in the province of Québec (Canada). Three parameters, the input discharge, the Strickler roughness coefficient of the main channel, and the average breach width are considered as input random variables from which uncertainties may occur and propagate through the hydraulic numerical models. The obtained results reveal the ability of BSBEM to efficiently predict the statistical moments and probability distributions of the output quantities of interest represented in terms of the downstream discharge, water level, and front wave arrival time at different locations of the studied reach with smooth profiles, unlike those from the polynomial chaos expansion that show oscillatory behavior.

Another contribution of this thesis concerns the proposal of a novel non-intrusive reduced order model technique named proper orthogonal decomposition-based B-Splines Bézier Elements Method (POD-BSBEM) for the uncertainty propagation analysis of stochastic time-dependent problems. The method uses a two-step proper orthogonal decomposition (POD) technique to extract the reduced basis. A third POD level is then applied to separate the time-dependent modes from the stochastic parametrized coefficients, which are approximated in the stochastic parameter space using B-splines basis functions defined in the corresponding Bézier element. The accuracy and the efficiency of the proposed method are assessed and compared to reduced-order model-based artificial neural network (POD-ANN) and the full-order model-based polynomial chaos expansion (Full-PCE) using benchmark steady-state and time-dependent problems. The POD-BSBEM is then applied to analyze the uncertainty propagation through a food wave flow stemming from a hypothetical dam break within a river with a complex bathymetry. The results confirm the ability of the POD-BSBEM to accurately predict the statistical moments of the output quantities of interest with a substantial speed-up for both offline and online stages compared to other techniques.

Keywords: Uncertainty propagation, B-splines, Bézier elements, polynomial chaos, dam-break flows, Proper orthogonal decomposition, reduced-order model.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

Page

LISTE DES ALGORITHMES

Page

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

MC	Monte Carlo
LHS	Latin Hypercube Sampling
PCE	Polynomial Chaos Expansion
Pcol	Point de collocation
PCM	Probabilistic Collocation Method
NIPC	Non-Intrusive Polynomial Chaos
IGA	IsoGeometric Analysis
BSBEM	B-Splines Bézier Elements Method
POD	Proper Orthogonal Decomposition
ROM	Reduced Order Model
SVD	Singular Value Decomposition
ANN	Artificial Neural Network
SWE	Shallow Water Equations
PDF	Probability Density Function
CDF	Cumulative Density Function
CI	Confidence Interval
GPR	Gaussian Process Regression
RBF	Radial Basis Functions
CPU	Central Processing Unit
GPU	Graphics Processing Unit
CMM	Communeauté Métropolitaine de Montréal
Full-PCE	Full Polynomial Chaos Expansion
POD-ANN	Proper Orthogonal Decomposition based Artificial Neural Network
POD-BSBEM	Proper Orthogonal Decomposition based B-Splines Bézier Elements Method

LISTE DES SYMBOLES ET UNITÉS DE MESURE

x, y	Distance logitudinale et transversale
z	Bathyétrie
h	Hauteur d'eau
t	Temps
g	Accélération gravitationnelle
n	Coefficient de friction de Manning
$\mathbf{S}_f$	Coefficient de friction
m	Nombre de variables aléatoire
N_s	Taille de l'échantillon
N_t	Nombre de d'intervalles de temps
N_e	Nombre d'éléments du maillage
$\mathcal{M}$	Modèle deterministe
Y	Réponse du modèle
$\mathbf{Y}$	Vecteur de réponses de sortie globales
$\mathbf{Y}^e$	Vecteur de réponses de sortie locales
M	Nombre total de termes du développement polynomial
p	Ordre des bases polynomiales
P	Mesure de probabilité
$\lvert\mathbf{q}\rvert$	Multi-indices
i, j, k, r, s, ℓ, e	Indices
n_p	Rapport de suréchantillonnage
$\boldsymbol{\eta}$	Vecteur de variables aléatoires dans le domaine physique
I_e	Éléments de Bézier

nx_i	Nombre d'intervalle correspondant à la variable η_i
N_{elt}	Nombre d'éléments de Bézier
n_b	Nombre de fonctions de bases dans chaque élément
K_s	Coefficient de friction de Strickler
Q_{up}	Débit amont
B_{av}	Largeur de la brèche du barrage
h_{up}	Hauteur d'eau en amont du barrage
h_{ds}	Hauteur d'eau en aval du barrage
H_i	Nombre de neurones constituant la couche cachée
$N_{i,p}$	Fonctions B-spiles d'ordre p
$B_{i,p}$	Fonctions de Beirstein d'ordre p
$\mathbf{C}^e$	Opérateur d'extraction de Bézier
L	Nombre de modes propores
Re	Nombre de Reynolds
μ	Moyenne
σ	Écart-type
Cv	Coefficients de variation
$\boldsymbol{\alpha}$	Vecteur des coefficients globaux
$\boldsymbol{\alpha}^e$	Vecteur des coefficients locaux
$\boldsymbol{\xi}$	Vecteur de variables aléatoires dans le domaine paramétrique
Θ	Espace d'événnements
ϵ_s, ϵ_t	Tolérances paramétrique et temporelle pour la sélécetion des bases réduite
ϱ_i	Fonction de densité de probabitlité de la variable η_i
$\psi(\xi)$	Fonction de base univariée
$\Psi(\boldsymbol{\xi})$	Fonction de base multivariée

$\boldsymbol{\Psi}^e$	Matrice de conception (design) local
$\boldsymbol{\Psi}$	Matrice de conception (design) global
$\mathbf{A}$	Opérateur d'assemblage
$\mathcal{N}$	Distribution Gaussienne
$\boldsymbol{\mathcal{U}}$	Matrice des snapshots
$\boldsymbol{\Phi}$	Matrice des bases réduites
$\mathcal{B}$	Matrice des données de projection de la matrice des snapshots sur la base réduite

INTRODUCTION

0.1 Généralités et problématique de recherche

Les barrages hydrauliques sont régulièrement exposés à des phénomènes naturels et conditions météorologiques extrêmes (séisme, intempéries, cyclones, glissement de terrains, . . . etc) qui, dans certains cas, sont à l'origine de leur rupture causant ainsi des inondations dont les conséquences sont dévastatrices. La recrudescence de ces événements s'est considérablement accentuée ces dernières années, à la fois en fréquence et en intensité, menant à des situations désastreuses tant sur le plan humain que matériel à cause de leur caractère instantané et imprévisible. La compréhension de la dynamique régissant ces phénomènes s'avère primordiale pour établir des plans d'urgence permettant d'éviter, ou du moins d'atténuer, les effets des sinistres qui en découlent.

De ce fait, les autorités réglementaires ont renforcé l'arsenal juridique mettant en exergue la sécurité des barrages, en exigeant des propriétaires l'établissement de plans d'urgence en tenant compte de tous les scénarios éventuels. Ainsi, pour chaque barrage à forte contenance, la loi sur la sécurité des barrages du Québec (S-3.1.01, r. 1, mise à jour du 01 janvier 2017) exige qu'une carte d'inondation soit produite indiquant le temps d'arrivée de l'onde de submersion ainsi qu'une description détaillée des zones qui seront potentiellement affectées. Ces cartes d'inondations doivent être établies à la fois pour le cas de rupture en conditions normales et pour le cas d'une crue de sécurité, considéré comme un cas exceptionnel. En se basant sur les informations fournies par ces cartes, des mesures d'urgence sont ensuite élaborées afin de délimiter les zones qui seront affectées, et qui requièrent un plan d'évacuation. Les dispositifs réglementaires exigent aussi la prise en considération de toutes les variabilités pouvant affecter la fiabilité des cartes d'inondation et par conséquent l'efficacité des mesures d'urgence.

Ainsi, afin d'apporter des éléments de réponse à toutes ces exigences à la fois techniques et réglementaires, des efforts considérables de recherche ont été déployés, ces dernières décennies, pour développer des modèles numériques capables de reproduire la dynamique des écoulements à surface libre. Ces modèles de prédiction, dont l'efficacité ne cesse de s'améliorer avec l'évolution de la puissance de calcul, servent comme solution alternative, et dans certains cas comme solutions complémentaires, pour des approches traditionnelles expérimentales dont le coût de mise en œuvre est souvent très onéreux. La fiabilité des prédictions résultant de ces modèles est souvent tributaire des différents paramètres pris en compte lors des calculs (paramètres physiques d'entrée, conditions initiales, conditions aux limites. . . etc), qui sont, dans la plupart des cas, entachés d'incertitudes dont il faut impérativement tenir compte afin d'atteindre un niveau de précision des résultats obtenus le plus optimal possible.

0.2 Objectifs de la thèse

La présente thèse est une contribution à l'analyse de la propagation d'incertitudes à travers les modèles numériques des écoulements de ruptures de barrages hydrauliques. Les techniques classiques dites d'échantillonnage représentent un défi majeur pour l'analyse de la propagation d'incertitudes au regard de la taille des échantillons nécessaires pour une précision adéquate. À cela s'ajoute les temps de calcul que demandent les modèles numériques traitant des écoulements à surface libre dans le cas de grands domaines de calcul avec des maillages de l'ordre de millions de nœuds. Devant de telles configurations, l'application des méthodes classiques pour le traitement stochastique s'avère très onéreuse en termes de coût de calcul, d'où la nécessité de recourir à d'autres méthodes, existantes ou nouvelles, avec une précision au moins équivalente, mais avec un effort de calcul considérablement réduit. C'est dans cette optique que s'inscrit la présente thèse dont les objectifs sont énumérés comme suit :

1. Effectuer une revue de littérature, aussi exhaustive que possible, des différentes méthodes d'analyse de propagation d'incertitudes paramétrique.

2. Implémenter les méthodes d'échantillonnage telles que Monte Carlo (MC) et Latin Hypercube Sampling (LHS) pour la propagation d'incertitudes paramétrique à travers les modèles numériques de l'hydraulique des ruptures de barrages, et qui serviront aussi de solutions de référence pour la vérification d'autres techniques.
3. Implémenter les méthodes non-intrusives du développement stochastique, à l'instar de la méthode des polynômes du chaos, pour le traitement stochastique de la réponse de sortie des modèles de l'hydraulique des ruptures de barrages.
4. Proposer de nouvelles approches non-intrusives pour l'analyse de propagation des incertitudes paramétriques qui comblent les limitations des autres techniques, particulièrement pour les problèmes à fort comportement hyperbolique ou discontinus, comme c'est le cas des écoulements de rupture de barrage.
5. Proposer de nouvelles approches non-intrusives de modèles réduits adaptées au traitement stochastique des réponses de sortie de modèles numériques des écoulements à surface libre.
6. Exploiter les différentes approches implémentées et développées pour l'analyse de propagation des incertitudes paramétriques à travers des applications sur des cas de barrages avec des données réelles, permettant l'établissement de carte d'inondations probabilistes fiables avec un effort de calcul considérablement réduit.

0.3 Contributions originales

Le travail réalisé dans le cadre de ce travail de recherche a permis d'apporter des contributions conséquentes à l'analyse de la propagation d'incertitudes paramétriques à travers des modèles numériques de l'hydraulique des ruptures de barrages. Ainsi, les contributions et originalités réalisées dans la présente thèse sont :

1. Les différentes techniques d'échantillonnage à l'instar de Monte Carlo et LHS ont été implémentées pour l'analyse de la propagation des incertitudes à travers des modèles

numériques hydrauliques. Ces méthodes ont été aussi utilisées comme solutions de référence pour la validation d'autres techniques.

2. La méthode des polynômes du chaos a été implémentée et intégrée dans l'analyse de propagation d'incertitudes. Elle a permis l'estimation des moments statistiques de la réponse de sortie moyennant un nombre limité d'appels du modèle déterministe (solveur numérique). Elle offre aussi la possibilité de construire des modèles de substitution stochastiques (métamodèles) à moindre coût et dont le temps d'exécution est considérablement réduit. Néanmoins, ces techniques présentent des solutions avec un comportement oscillatoire dans certaines applications spécifiques avec un fort comportement hyperbolique ou discontinu.

3. Comme contribution originale de la présente thèse, une nouvelle méthode non-intrusive dénommée B-splines Bézier Elements Methods (BSBEM) a été introduite afin de remédier aux limitations de de la méthode des polynômes du chaos. Cette nouvelle approche, qui s'inspire du concept de l'analyse iso-géométrique (Isogeomtric analysis, IGA), combine les fonctions locales de type B-splines avec l'aspect multi-éléments de l'opérateur d'extraction de Bézier pour une représentation stochastique locale dans le domaine paramétrique. Ce travail a fait l'objet d'une publication dans la revue :
   Journal of Computer Methods in Applied Mechanics and Engineering.
   https ://doi.org/10.1016/j.cma.2018.10.047.

4. En second lieu, une application de l'approche proposée (BSBEM) a été réalisée sur un cas d'un barrage hydraulique avec des données réelles du site et celles relatives aux incertitudes des paramètres d'entrée en considérant les différents scénarios de rupture possibles. Une analyse approfondie des performances de la méthode a été effectuée sur des aspects fondamentaux de l'analyse stochastique à travers l'estimation de la variabilité des réponses de sortie telles que le niveau de la surface d'eau, l'hydrogramme, et le temps d'arrivé de l'onde de submersion, moyennant la quantification des moments statistiques et

les fonctions de distribution de probabilité. Cette contribution a été publié dans la revue : *Journal of Hydrology*. https ://doi.org/10.1016/j.jhydrol.2020.125342.

5. Dans la même optique, une nouvelle méthode non-intrusive de réduction de modèles dénommée POD-BSBEM a été récemment proposée dans le cadre de ce travail de recherche. Elle combine les notions de l'approche BSBEM et celles de la décomposition orthogonale aux valeurs propres pour l'analyse de la propagation d'incertitudes avec un coût de calcul considérablement réduit, ce qui s'avère une solution adéquate pour les écoulements de ruptures de barrages où les domaines d'étude sont excessivement larges. Les performances de la méthode proposée sont comparées avec celles du modèle réduit basé sur les réseaux de neurones artificiels (POD-ANN). Une troisième publication, découlant de cette nouvelle approche pour les modèles réduits, a été soumise pour évaluation à la revue :
Journal of Computational Physics.
6. Enfin, le développement de toutes ces techniques d'analyse stochastique permet l'établissement de prédictions probabilistes fiables qui tiennent compte de toutes les éventuelles variabilités dans les paramètres d'entrée. Ces outils représentent aussi un atout majeur d'aide à la décision dans des situations critiques nécessitant des mesures d'urgence.

0.4 Organisation de la thèse

La présente thèse est présentée sous le format d'une thèse par articles qui contient quatre chapitres. Un premier chapitre est consacré à la revue de littérature, deux articles déjà publiés forment les chapitres 2 et 3, et le contenu d'un troisième article soumis pour évaluation est présenté dans le chapitre 4.

Ainsi, le chapitre 1 présente une revue de littérature des écoulements à surface libre, des différentes approches d'analyse de propagation des incertitudes paramétriques et des techniques de réduction de modèles. Quelques notions complémentaires sur les concepts fondamentaux des méthodes utilisées dans le cadre du présent travail sont aussi abordées.

Dans le chapitre 2, les fondements mathématiques de la nouvelle méthode non-intrusive, en l'occurrence BSBEM, sont abordés. La fiabilité et la performance de la méthode sont démontrées à travers une comparaison des résultats obtenus avec ceux de la méthode des polynômes du chaos et les méthodes d'échantillonnage, moyennant une série d'exemples numériques reproductibles de référence, mais aussi à travers une application sur une hypothétique rupture d'un barrage hydraulique.

Le troisième chapitre est consacré à l'application de la méthode proposée sur un cas de rupture de barrage factuel avec des données réelles, utilisées dans des études de ruptures déterministes. L'efficacité de l'approche BSBEM a été poussée davantage sur des configurations beaucoup plus complexes, en examinant ses capacités à estimer avec précision les fonctions de densité de probabilité, en plus des moments statistiques, de la réponse de sortie avec un fort comportement hyperbolique entourant l'onde de submersion, contrairement aux méthodes du chaos polynomial qui présentent des solutions oscillatoires.

Une nouvelle approche non-intrusive de réduction de modèles, dénommée POD-BSBEM, pour des problèmes instationnaires et paramétriques, est présentée dans le chapitre 4, où le cadre global qui combine à la fois les notions de l'approche BSBEM et la technique POD est détaillée. Cette nouvelle approche utilise, d'une part, la procédure de la double POD pour la génération de la base réduite, et d'autre part, un troisième niveau de POD, pour découpler la dépendance paramétrique et temporelle des coefficients de projection, dont le comportement stochastique est décrit par des approximations locales dans le domaine paramétrique, en fonctions des variables aléatoires, en utilisant les concepts de la méthode BSBEM.

CHAPITRE 1

REVUE DE LITTÉRATURE ET CONCEPTS FONDAMENTAUX

1.1 Introduction

Dans ce chapitre, une revue de littérature de la problématique des incertitudes en modélisation numérique de l'hydraulique des ruptures de barrages est présentée. Les dernières évolutions des différentes approches adoptées dans l'analyse de la propagation des incertitudes sont passées en revue, en mettant l'accent tant sur l'aspect fiabilité des prédictions que sur le coût de calcul associé.

Les concepts fondamentaux relatifs aux méthodes de quantification d'incertitudes telles que : les méthodes d'échantillonnage classiques, les méthodes du développement stochastiques non-intrusives, ainsi que les nouvelles approches de réduction de modèles pour les problèmes paramétriques instationnaires, sont abordés dans la suite de ce chapitre.

1.2 Écoulements à surface libre

Les écoulements à surface libre sont d'un intérêt grandissant auprès des chercheurs à cause des risques accrus d'inondations et le besoin pressant d'assimiler leur dynamique. Différents travaux ont été menés dans ce sens en développant des modèles numériques de prédiction permettant de définir les zones inondables afin de mieux estimer les risques inhérents (De Almeida & Bates, 2013 ; Guinot, 2012). Ainsi, les progrès réalisés dans l'accroissement de la puissance de calcul a entrainé une amélioration de la capacité de ces modèles numériques à travers l'implémentation de nouveaux schémas (Bellos & Tsakiris, 2015 ; Lai, Guo, Lin & Tan, 2010 ; Loukili & Soulaimani, 2007) capable de simuler des configurations de plus en plus complexes, permettant d'avoir des prédictions sur des domaines de très grandes tailles, se rapprochant, aussi fidèlement que possible, des données réelles disponibles, récoltées à la suite d'événements antérieurs.

La dynamique complexe des écoulements à surface libre est souvent modélisée par les équations de Saint- Venant, appelées aussi équations d'eaux peu profondes (Shallow Water Equations, SWE), qui sont caractérisées par un comportement hyperbolique, et dont l'expression mathématique est donnée par (Leendertse & Gritton, 1971 ; Ying, Khan & Wang, 2004 ; Zokagoa, 2011) :

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{H}(\mathbf{U})}{\partial y} = \mathbf{S}(\mathbf{U}) \tag{1.1}$$

Où :

$$\mathbf{U} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \ \mathbf{G}(\mathbf{U}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix}, \ \mathbf{H}(\mathbf{U}) = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}, \ \mathbf{S}(\mathbf{U}) = \begin{bmatrix} 0 \\ gh(S_{0_x} - S_{f_x}) \\ gh(S_{0_y} - S_{f_y}) \end{bmatrix}$$

avec :

$$S_{0_x} = -\frac{\partial z}{\partial x}, \ S_{0_y} = -\frac{\partial z}{\partial y}, \ S_{f_x} = \frac{n^2 u\sqrt{u^2+v^2}}{h^{4/3}}, \ S_{f_y} = \frac{n^2 v\sqrt{u^2+v^2}}{h^{4/3}}$$

u et v sont les composantes de la vitesse moyenne dans les directions x et y, respectivement. h représente la hauteur d'eau, qui est la différence entre la surface libre η et la bathymétrie z telle que $(h = \eta - z)$, comme le montre la figure 1.1. n et $\mathbf{S}_f$ désignent le coefficient de friction de Manning et le terme source de friction, respectivement.

Les équations de conservation décrites ci-dessus sont aussi utilisées pour la modélisation des écoulements à surface libre découlant d'une rupture soudaine de barrages hydrauliques, bien que l'accélération verticale soit non négligeable. Ces écoulements présentent un défi majeur en termes de modélisation du fait des interactions avec certains aspects tels que le transport des sédiments, le changement continue de la bathymétrie du lit de rivières (Saikia & Sarma, 2010), mais aussi de leur comportement fortement instationnaire, et de la présence de discontinuités au voisinage du front de l'onde de submersion (Cannata & Marzocchi, 2012 ; Xiong, 2011). Le caractère inopiné de ces écoulements se traduit par des conséquences désastreuses tant sur le plan humain, matériel qu'environnemental (Dai, Lee, Deng & Tham, 2005 ; Lempérière, 2017).

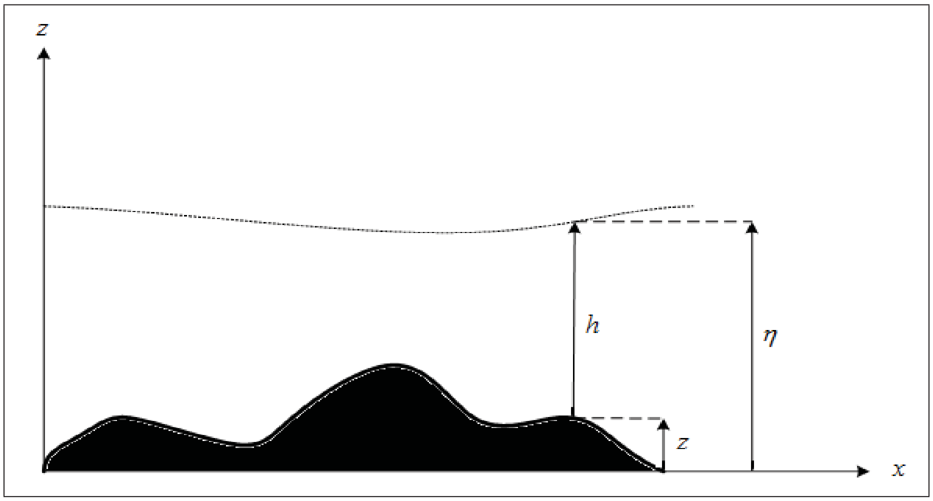


Figure 1.1 Représentation schématique de la hauteur d'eau (h), de la surface libre (η) et de la bathymétrie (z) (Zokagoa, 2011).

L'utilité de développer des modèles numériques qui sont en mesure de reproduire la dynamique de ces phénomènes est plus que jamais nécessaire afin d'éviter, ou, dans une moindre mesure, d'amenuiser l'ampleur des dommages générés. Une multitude de chercheurs s'est autosaisie afin de faire de ces axes de recherche une priorité où des efforts considérables ont été consentis ces dernières décennies afin d'améliorer la compréhension de la dynamique régissant l'onde de submersion (Erpicum, Dewals, Archambeau & Pirotton, 2010 ; Haltas, Elçi & Tayfur, 2016 ; Seyedashraf & Akhtari, 2017 ; Zhang, Xia, Yuan & Jiang, 2014).

Les études portant sur les écoulements de rupture de barrages ont été menées sur différents aspects : analytique, numérique et expérimental. La complexité du phénomène physique associé à ce type d'écoulements rend la recherche de solutions analytiques très ardue. Néanmoins, certaines d'entre elles ont été proposées pour des configurations simplifiées où un certain nombre d'hypothèses ont été admises permettant ainsi d'idéaliser le processus du rupture de barrage unidimensionnel avec un fond plat et sans présence de frottement, à l'instar des solutions de Ritter (Ritter, 1892), et de Stoker (Stoker, 1957), qui décrivent l'évolution du niveau d'eau et la vitesse de l'onde de choc générée par une rupture de barrage sur un lit sec et mouillé, respectivement. Malgré le caractère simplifié de ces solutions analytiques, elles ont souvent servi comme référence pour tester et valider les différents modèles numériques qui sont développés pour des configurations complexes proches de cas réels.

Les modèles numériques qui sont utilisés pour la modélisation des écoulements de rupture de barrages sont basés sur la discrétisation des équations gouvernantes en utilisant, principalement, les schémas de différence finis, volumes finis, ou éléments finis. Différents outils, basés sur ces techniques, ont été développés pour simuler la dynamique de l'onde de crue avec notamment des modèles 1D (Butt, Umar & Qamar, 2013 ; Goutal & Maurel, 2002 ; Petaccia & Natale, 2013), qui sont largement utilisés pour résoudre la forme unidimensionnelle des équations de Saint-Venant qui permettent un gain substantiel en temps de simulation pour des configurations avec des domaines d'étude, représentés par une série de sections de passage, s'étalant sur de longues distances. L'amélioration des capacités de calculs a contribué à développer des modèles 2D, et récemment des modèles 3D (Ferrari, Dumbser, Toro & Armanini, 2009 ; Marsooli & Wu, 2014 ; Munoz & Constantinescu, 2020 ; Ozmen-Cagatay & Kocaman, 2011), qui reflètent mieux la réalité physique en remédiant aux limitations observées dans les modèles 1D, particulièrement pour des configurations complexes proches de la réalité du terrain. L'introduction de ces modèles dans divers outils de simulation a permis une amélioration substantielle de la qualité des prédictions relatives aux écoulements de rupture de barrage.

Les prédictions émanant de ces modèles numériques sont considérées comme étant déterministes, du fait qu'elles sont obtenues avec des valeurs nominales fixes de paramètres d'entrée et des conditions aux limites comme le coefficient de friction, la bathymétrie, le débit d'entrée, la hauteur initiale et la largeur moyenne de la brèche. Les valeurs nominales de ces paramètres sont estimées, entre autres, par le biais d'observations et de relevés in-situ, ainsi que par des relations empiriques, comme c'est le cas pour la largeur des brèches, où plusieurs corrélations ont été proposées pour prédire leurs caractéristiques géométriques au moment de la rupture (Ahmadisharaf, Kalyanapu, Thames & Lillywhite, 2016 ; Sammen, Mohamed, Ghazali, Sidek & El-Shafie, 2017). Il convient de souligner que des incertitudes majeures peuvent découler de ces processus d'estimation en raison d'un manque de connaissances, d'erreurs de mesure et de la nature aléatoire des paramètres d'entrée (Butt *et al.*, 2013 ; Kim & Sanders, 2016), ce qui peut influencer de manière significative la fiabilité des prédictions. La propagation de ces incertitudes à travers les modèles numériques déterministes peut conduire à des comportements tout à fait inattendus du fait des systèmes fortement non linéaires, tels que les équations régissant les écoulements à surface libre, ce qui, par conséquent, peut impacter la fiabilité des prédictions. Il est donc crucial de prendre en compte la variabilité des paramètres d'entrée afin de quantifier son effet sur la réponse de sortie, permettant ainsi l'établissement de cartes d'inondations probabilistes (Bharath & Elshorbagy, 2018 ; Tsai, Yeh & Huang, 2019), qui sont en mesure de fournir des prévisions crédibles et réalistes.

1.3 Analyse de propagation d'incertitudes

1.3.1 Cadre général

Les principales étapes d'une analyse de propagation d'incertitudes à travers un modèle numérique peuvent être décrites comme suit (De Rocquigny, Devictor & Tarantola, 2008 ; Sudret, 2007) : ***i)*** *quantification des sources d'incertitudes* qui implique une estimation de la variabilité des paramètres d'entrée à travers des lois de distributions appropriées ; ***ii)*** *propagation des incertitudes* à travers le modèle numérique qui transforme la mesure de l'incertitude de l'entrée en mesure de l'incertitude de la réponse de sortie ; ***iii)*** *post-traitement des résultats* par le calcul des moments statistiques et de la fonction de distribution de probabilité de la réponse de sortie. Une analyse de sensibilité pourrait aussi être effectuée afin d'établir les paramètres d'entrée qui contribuent le plus aux incertitudes du système étudié. Une représentation schématique de ces étapes est illustrée sur la figure 1.2.

1.3.2 Méthodes d'analyse de propagation d'incertitudes

L'étude de la propagation des incertitudes paramétriques à travers des modèles de prédiction déterministes peut se faire par plusieurs méthodes. La méthode de Monte Carlo est considérée

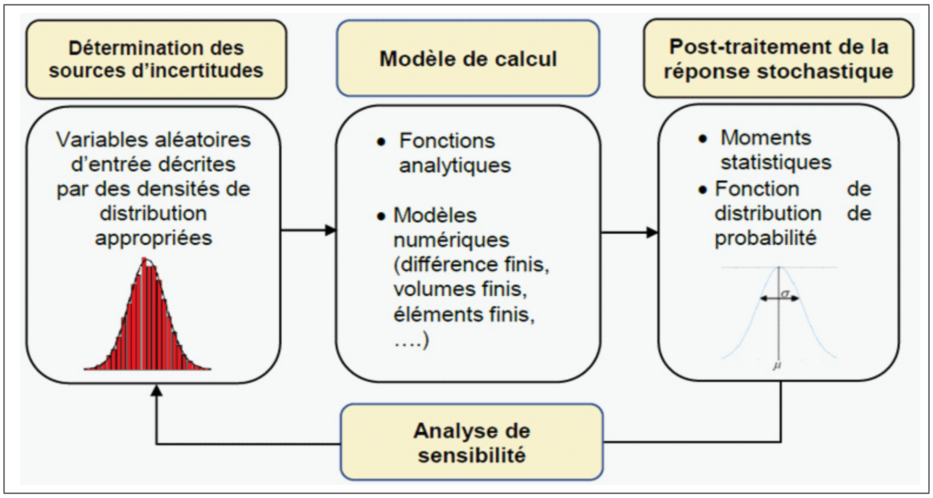


Figure 1.2 Principales étapes d'une analyse de propagation d'incertitudes.

comme l'une des méthodes les plus couramment utilisées pour le traitement des problèmes liés à l'analyse de la propagation des incertitudes paramétriques (Weitz, Blanco, Charon, Dauchet, El Hafi, Eymet, Farges, Fournier & Gautrais, 2016). Le principe de la méthode consiste à générer aléatoirement un échantillon des variables incertaines suivant une distribution de probabilité $\{\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N_s)}\}$, et les injecter, ensuite, dans le modèle déterministe $Y \equiv \mathcal{M}(\boldsymbol{\eta})$. Les valeurs de la solution, ainsi générées, sont traitées en calculant les moments statistiques de la réponse de sortie (Sudret, 2007) :

$$\mu_Y = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathcal{M}\left(\boldsymbol{\eta}^{(i)}\right) \tag{1.2}$$

$$\sigma_Y^2 = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} \left[\mathcal{M}\left(\boldsymbol{\eta}^{(i)}\right) - \mu_Y\right]^2 \tag{1.3}$$

La méthode Monte Carlo présente une attractivité due au fait de sa simplicité à mettre en œuvre et de son caractère non intrusif. Elle est souvent considérée comme solution de référence pour tester la fiabilité d'autres techniques modernes. Néanmoins, la convergence de la méthode reste assez lente, elle de l'ordre $\frac{1}{\sqrt{N_s}}$, où N_s est la taille de l'échantillon étudié (Ballio & Guadagnini, 2004 ; Blatman, 2009). Pour cette raison, et afin d'atteindre le niveau de précision souhaité, des échantillons de très grande taille sont nécessaires. Cette condition peut représenter un inconvénient majeur en particulier pour les problèmes où le temps de calcul du modèle déterministe pour une seule simulation est très conséquent.

Devant une telle contrainte, d'autres approches ont été développées afin d'accélérer la convergence avec des échantillons de tailles beaucoup plus réduites en permettant une meilleure distribution des points d'échantillonnage le long du domaine paramétrique (Hammersley & Handscomb, 1964 ; Sobol & Levitan, 1999). Parmi ces techniques, la méthode d'échantillonnage par l'Hypercube Latin (Latin Hypercube Sampling, LHS) (McKay, Beckman & Conover, 1979 ; Shields & Zhang, 2016) est une méthode d'échantillonnage par stratification qui permet une meilleure représentation des valeurs de l'échantillon généré pour chaque variable aléatoire étudiée. C'est une méthode dont l'usage est largement répandu auprès de la communauté scientifique du fait des différents avantages qu'elle procure. Elle permet, en effet, d'avoir un niveau de précision satisfaisant avec des échantillons de tailles réduites, comparativement à la méthode de Monte Carlo. Le domaine représenté par la densité de probabilité cumulative est divisé en n intervalles équidistants (d'équiprobabilité). Une fois les valeurs des probabilités sont sélectionnées dans chaque strate $p^{(i)}$, les valeurs de la variable aléatoire correspondantes $\eta^{(i)}$ sont obtenues en utilisant la fonction inverse de la densité de probabilité cumulative $x^{(i)} = F^{-1}(p^{(i)})$ (Kwon, Moon & Khalil, 2007). La procédure d'échantillonnage par la méthode LHS est schématisée sur la figure 1.3.

1.3.2.1 Méthode des polynômes du chaos

Comme alternative aux techniques d'échantillonnage, les méthodes des polynômes du chaos (PC), qui appartiennent aux méthodes de développement stochastique, basées sur la théorie du chaos homogène Wiener (1938), ont été introduites pour l'analyse de propagation d'incertitudes (Ghanem & Spanos, 1991). Le principe de ce type de méthodes est basé sur une approximation de la réponse du modèle par un développement en série du type polynômes du chaos en fonction des variables aléatoires d'entrée, i.e. $\widehat{y} \approx \sum_{j=1}^{M} \alpha_j \Psi_j(\boldsymbol{\eta})$. Les bases multivariées $\Psi_j(\boldsymbol{\eta})$ sont

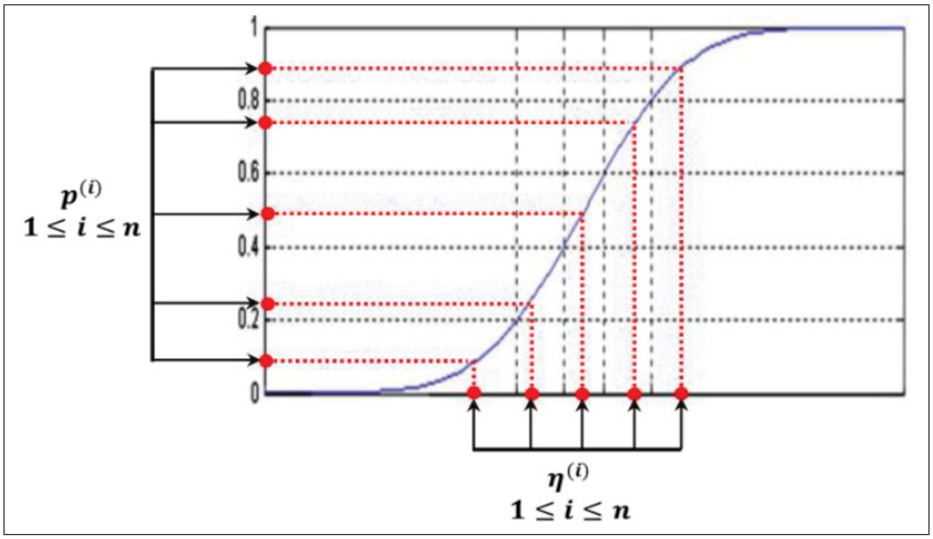


Figure 1.3 Procédure d'échantillonnage par la méthode LHS.

construites à partir des bases univariées $\psi_i(\eta_i)$. Le choix de ces bases polynomiales orthogonales est étroitement lié à la fonction densité de probabilité de la variable aléatoire d'entrée η_i car la convergence en dépend fortement (Xiu & Karniadakis, 2002). Les correspondances entre les distributions de probabilité et le type des polynômes orthogonaux sont indiquées dans le tableau 3.1.

Les coefficients de l'approximation du chaos polynomial sont évalués, dans leurs premières phases, par des techniques dites intrusives basées sur la projection de Galerkin, qui requiert une modification des équations gouvernantes et résoudre ainsi un système d'équations couplées dont les inconnues sont les coefficients α_i (Dinescu, Smirnov, Hirsch & Lacor, 2010 ; Ghanem, 1999). L'avantage de cette méthode réside dans sa bonne convergence, mais son inconvénient majeur est son caractère intrusif qui peut représenter un réel défi dans certaines configurations où des développements conséquents sont nécessaires (Zokagoa, 2011). Afin de remédier à cet inconvénient, d'autres méthodes qui traitent le modèle déterministe comme une boite noire, d'où l'appellation de méthodes non-intrusives, sont développées. L'idée de base de ces méthodes est de pouvoir estimer les coefficients en se basant sur un nombre réduit de solutions émanant du modèle déterministe, c'est-à-dire un ensemble d'évaluations de la réponse du modèle à partir d'un certain nombre de valeurs des variables aléatoires d'entrée (Loeven, Witteveen & Bijl, 2007).

Les méthodes non intrusives peuvent être scindées en deux grandes catégories : les méthodes de projection et les méthodes de régression. Le principe des méthodes de projection repose sur l'exploitation de la propriété de l'orthogonalité des fonctions de bases polynomiales afin d'évaluer les coefficients de l'approximation de la réponse du modèle (Ghiocel & Ghanem, 2002). L'utilisation de cette technique peut s'avérer laborieuse en termes d'effort de calcul, requis pour évaluer l'intégrale multidimensionnelle résultante, particulièrement pour le cas où le nombre de paramètres incertains d'entrée devient important (Blatman & Sudret, 2011). La méthode de collocation consiste à évaluer les coefficients par régression, qui permettent d'approcher au mieux la réponse de la sortie, et ce à travers un nombre limité d'évaluations du modèle déterministe obtenant un système d'équations algébrique, formé d'une part par une matrice dont les membres sont les valeurs des bases polynomiales évaluées aux points de collocation et d'autre part par le vecteur contenant les réponses du modèle (Berveiller, Sudret & Lemaire, 2006).

Le choix des points de collocation est souvent une étape cruciale dans l'application des méthodes non-intrusives basées sur la régression, car l'efficacité de la méthode et son taux de convergence en dépendent fortement. Différentes techniques ont été proposées dans la littérature pour un choix optimal de ces points de collocation et qui dans certains est à l'origine de nouvelles méthodes de collocation. Ainsi, la méthode de collocation probabiliste (PCM) (Loeven *et al.*, 2007 ; Tatang, Pan, Prinn & McRae, 1997) considère les points de collocation comme étant les racines des bases polynomiales orthogonales. Une autre approche basée sur la régression en l'occurrence la méthode du point de collocation (Hosder & Walters, 2010 ; Hosder, Walters & Balch, 2007) est largement utilisée au regard de sa simplicité. Elle combine l'efficacité des différentes techniques d'échantillonnage avec un coefficient de suréchantillonnage approprié afin d'optimiser la distribution des points de collocations à travers le domaine paramétrique.

Différents critères peuvent avoir un effet déterminant sur l'efficacité des approximations stochastiques basées sur les développements en polynômes du chaos à l'instar de l'ordre polynomial p et le nombre de variables aléatoires m, qui déterminent la taille de l'expansion à construire M, donnée par l'équation 3.2. La construction des bases polynomiales multivariées, comme mentionné précédemment, se fait moyennant le produit tensoriel entre les termes univariés de la base orthogonale en introduisant la notion du multi-indice $|\mathbf{q}| = \sum_{i=1}^{m} p_i^j$ (Miller, Berg, Davison, Sudicky & Forsyth, 2018 ; Sudret, 2014).

Tableau 1.1 Polynômes d'Hermite univarié d'ordre 3.

p_i^j	$\psi_{p_1^j}(\xi_1)$	$\psi_{p_2^j}(\xi_2)$
0	1	1
1	ξ_1	ξ_2
2	$\xi_1^2 - 1$	$\xi_2^2 - 1$
3	$\xi_1^3 - 3\xi_1$	$\xi_2^3 - 3\xi_2$

Un exemple illustratif de cette procédure est représenté dans le tableau 1.2 pour le cas d'une base bivariée ($m = 2$), en utilisant des polynômes d'Hermite univariés (qui sont associés avec des variables aléatoires décrites par une distribution normale) dont les expressions sont données dans le tableau 1.1 (Hu, 2015). Les bases polynômiales ainsi obtenues conservent, par construction, la propriété d'orthogonalité.

L'une des caractéristiques les plus attractives des méthodes des polynômes du chaos, en plus de la construction d'un modèle de substitution stochastique, est la facilité qu'elles procurent pour estimer les moments statistiques de la réponse de sortie à partir des valeurs des coefficients, préalablement calculés, sans aucun effort de calcul supplémentaires, et ce en se basant sur la propriété d'orthogonalité des bases générées (Salehi, Raisee, Cervantes & Nourbakhsh, 2017). Les coefficients obtenus sont aussi utilisés pour le calcul direct des indices de sensibilité de Sobol (Homma & Saltelli, 1996 ; Sobol, 2001) qui permettent un classement des variables aléatoires qui contribuent le plus à la variabilité de la réponse de sortie. Ces approches permettent un gain

Tableau 1.2 Construction d'une base polynômiale d'Hermite à deux variables aléatoires.

$\lvert\mathbf{q}\rvert$	(p_1^j, p_2^j)	j	$\psi_{p_1^j}(\xi_1)\psi_{p_2^j}(\xi_2)$	$\Psi_j(\xi_1, \xi_2)$
0	$(0,0)$	0	$\psi_0(\xi_1)\psi_0(\xi_2)$	1
1	$(1,0)$	1	$\psi_1(\xi_1)\psi_0(\xi_2)$	ξ_1
	$(0,1)$	2	$\psi_0(\xi_1)\psi_1(\xi_2)$	ξ_2
2	$(2,0)$	3	$\psi_2(\xi_1)\psi_0(\xi_2)$	$\xi_1^2 - 1$
	$(1,1)$	4	$\psi_1(\xi_1)\psi_1(\xi_2)$	$\xi_1\xi_2$
	$(0,2)$	5	$\psi_0(\xi_1)\psi_2(\xi_2)$	$\xi_2^2 - 1$
3	$(3,0)$	6	$\psi_3(\xi_1)\psi_0(\xi_2)$	$\xi_1^3 - 3\xi_1$
	$(2,1)$	7	$\psi_2(\xi_1)\psi_1(\xi_2)$	$(\xi_1^2 - 1)\xi_2$
	$(1,2)$	8	$\psi_1(\xi_1)\psi_2(\xi_2)$	$\xi_1(\xi_2^2 - 1)$
	$(0,3)$	9	$\psi_0(\xi_1)\psi_3(\xi_2)$	$\xi_2^3 - 3\xi_2$

substantiel dans l'effort de calcul, contrairement aux méthodes d'échantillonnage classique, étant donné qu'elles ne requièrent pas de calculs additionnels pour effectuer l'analyse de sensibilité (Blatman & Sudret, 2010b ; Rahman, 2011).

L'utilisation des méthodes des polynômes du chaos non-intrusives s'est largement répandue dans l'analyse de la propagation d'incertitudes et la construction de modèles de substitution stochastiques qui permettent une estimation efficace des moments statistiques ainsi que de la fonction de distribution de la réponse de sortie avec un faible coût de calcul. Des améliorations continues de ces techniques ont été proposées à travers de nouvelles approches plus efficaces et moins couteuses (Blatman & Sudret, 2010a,1 ; Thapa, Mulani & Walters, 2018). Néanmoins, pour des problèmes avec un fort comportement hyperbolique ou présentant des discontinuités dans le domaine physique ou paramétrique, ces techniques présentent quelques limitations dans l'estimation de la réponse de sortie où un comportement oscillatoire peut être observé dans les profiles des moments statistiques. Ce phénomène est connu sous le terme 'phénomène de Gibbs' (Bijl, Lucor, Mishra & Schwab, 2013). Une panoplie de méthodes a été introduite pour surmonter cette limitation que présente les polynômes du chaos. Certaines de ces techniques sont basées sur une approche multi-éléments qui décomposent le domaine paramétrique en sous-domaines adjacents avec des approximations du type chaos polynomial (Chouvion & Sarrouy, 2016 ; Foo, Wan & Karniadakis, 2008 ; Wan & Karniadakis, 2005), ou d'autres fonctions de base à l'instar des fonctions de base type multi-ondelettes (Le Maître, Knio, Najm & Ghanem, 2004a ; Pettersson, Iaccarino & Nordström, 2013).

1.3.2.2 Fonctions B-splines avec éléments de Bézier

L'utilisation des fonctions de base type B-splines a connu une progression constante du fait des propriétés mathématiques qu'elles procurent. Couplées à la méthode des éléments finis, elles ont donné naissance à l'approche communément appelée *Isogeometric Analysis (IGA)* (Hughes, Cottrell & Bazilevs, 2005) dont le champ d'application s'est étendu à divers domaines de l'ingénierie ces dernières années (Hughes, Reali & Sangalli, 2010 ; Schuß, Dittmann, Wohlmuth, Klinkel & Hesch, 2019). L'introduction des fonctions de base B-splines a permis d'améliorer grandement le processus de discrétisation et d'approximation des problèmes physiques régis par les équations aux dérivées partielles grâce aux multiples propriétés mathématiques attractives qu'elles procurent (da Veiga, Buffa, Sangalli & Vázquez, 2016). Les principales propriétés mathématiques qui caractérisent les fonctions de bases type B-splines sont, entre autres, le support compact, la partition d'unité, la non-négativité, l'indépendance linéaire sur leur support compact et la continuité d'ordre C^{p-1}, où p est l'ordre polynomial (Hughes *et al.*, 2005). Une description détaillée des expressions mathématiques des fonctions B-splines univariées est présentée dans la section 2.3.1.

En plus de l'essor que connait les fonctions B-splines dans les études déterministes, un nombre grandissant de travaux a été mené en introduisant ces fonctions de base dans les analyses stochastiques des différents problèmes physiques où d'autres techniques, se basant sur le développement stochastique, ont montré quelques insuffisances dans l'approximation de la réponse de sortie (Baroth, Bodé, Bressolette & Fogli, 2006 ; Baroth, Bressolette, Chauvière & Fogli, 2007 ; Hien & Noh, 2017). Ces fonctions de bases sont utilisées à la fois comme fonctions d'interpolation, mais aussi comme fonctions de pondérations dans les techniques de projection (Beck, Tamellini & Tempone, 2019 ; Dertimanis, Spiridonakos & Chatzi, 2017). Le potentiel qu'offre les fonctions B-splines dans l'analyse de la propagation et la quantification des incertitudes a fait l'objet de plusieurs investigations. Une comparaison détaillée des performances des fonctions de base B-splines par rapport aux autres fonctions de base de type chaos polynomial a été récemment proposée par Ditkowski, Fibich & Sagiv (2018), où les capacités des B-splines à approximer, avec efficacité, le comportement stochastique de la réponse de sortie, et plus particulièrement la fonction de distribution de probabilité, ont été abordées. Cet aspect représente un défi majeur, au-delà des moments statistiques, pour quelques problèmes typiques instationnaires, fortement non linéaires et dans certains cas discontinus.

Dans ce contexte, les techniques qui disposent d'outil mathématique permettant de procéder à des approximations locales par morceaux peuvent s'avérer plus efficace pour remédier aux insuffisances des autres approches qui sont définies sur le tout le domaine. L'efficacité de ces approximations d'autant plus grande lorsqu'elles sont associées à des approches multi-éléments (Halder, Sanderse & Koren, 2019). Ainsi, au regard de leurs caractéristiques, les fonctions B-splines présentent un champ d'application très prometteur dans le domaine de l'analyse de propagation des incertitudes en particulier pour les phénomènes où il y a présence de discontinuités, tels que les écoulements de ruptures de barrages qui sont fréquemment soumis de fortes discontinuités à travers l'onde de crue. Une approximation locale de ces phénomènes est rendue possible grâce à l'aspect multi-éléments qu'offre la décomposition du domaine paramétrique (domaine de définition des fonctions de base) en sous domaine adjacents nommés éléments de Bézier. Les fonctions de base locales sont ainsi définies dans chaque élément à partir des fonctions de base de Bernstein moyennant l'opérateur d'extraction linéaire de Bézier (Borden, Scott, Evans & Hughes, 2011).

En combinant les caractéristiques mathématiques des fonctions de base B-splines et l'aspect multi-eléments de l'opérateur d'extraction de Bézier, une nouvelle approche nommée *B-Splines Bézier Elements Method (BSBEM)* (Abdedou & Soulaïmani, 2019 ; Abdedou, Soulaïmani & Tchamen, 2020) a été proposée, dans le cadre des axes de recherche menés dans la présente thèse, comme une contribution aux différentes techniques de l'analyse de propagation d'incertitudes à travers les modèles numériques et le traitement stochastique des réponses de sortie qui en résultent. La méthode s'appuie aussi, en plus des aspects évoqués ci-dessus, sur la correspondance entre le domaine paramétrique des fonctions B-splines et le domaine de définitions des fonctions de densité cumulative (CDF) (Barth, 2013) qui offre la possibilité de simplifier considérablement l'évaluation des moments statistiques de la réponse de sortie à travers une série de transformations mathématiques appropriées. La méthode s'inscrit dans le principe des méthodes de collocation basées sur le développement stochastique multi-éléments avec un caractère non-intrusif, qui ne requiert aucune modification des équations gouvernantes pour l'évaluation des coefficients. Il y a lieu de préciser que le cadre fondamental de la méthode est abordé en détail dans les chapitres qui suivent, et uniquement quelques notions complémentaires sont décrites dans la présente section.

Ainsi, pour chaque variable aléatoire η_i définie dans le domaine physique par une distribution de probabilité appropriée, décrivant au mieux sa variabilité, représentant le paramètre d'entrée incertain, une variable correspondante dans le domaine paramétrique lui est associée, $\xi_i \in [0, 1]$, avec $i = 1, \ldots, m$ où m désigne le nombre de variables aléatoires. Ce domaine paramétrique correspondant à la variable ξ_i est ensuite subdivisé en un nombre fini nx_i d'intervalles disjoints. Des fonctions de base polynomiales locales univariées d'ordre p_i, et dont le nombre est de $(n_b = p_i + 1)$, sont définies dans chaque intervalle. Les valeurs p_i et nx_i sont spécifiques à chaque variable aléatoire ξ_i et peuvent être choisies par l'utilisateur indépendamment les unes des autres. Des exemples de ces bases polynomiale locales univariées correspondant aux variable aléatoires ξ_1 et ξ_2 avec un nombre dintervalle $nx_1 = nx_2 = 2$ sont ilustrés dans les tableaux 1.3 et 1.4, respectivement. Le domaine paramétrique de chaque variable aléatoire est divisé, dans cet exemple, en $nx_1 = 2$ intervalles à savoir : $e_{11} = [0, 0.5]$ et $e_{12} = [0.5, 1]$ pour la variable ξ_1, et de la même manière en $nx_2 = 2$ intervalles : $e_{21} = [0, 0.5]$ et $e_{22} = [0.5, 1]$ pour la variable

ξ_2, dans lesquels les fonctions de base locales sont définies.

Tableau 1.3 Exemple de fonctions de base locales univariées d'ordre $p_1 = 2$ de la variable ξ_1 avec un nombre d'intervalles $nx_1 = 2$.

Intervalle e_{11}	**Intervalle** e_{12}
$\psi_1^{e_{11}}(\xi_1) = \frac{(\xi_1-1)^2}{4}$	$\psi_1^{e_{12}}(\xi_1) = \frac{(\xi_1-1)^2}{8}$
$\psi_2^{e_{11}}(\xi_1) = \frac{1}{8}(-3\xi_1^2 + 2\xi_1 + 5)$	$\psi_2^{e_{12}}(\xi_1) = \frac{1}{8}(-3\xi_1^2 - 2\xi_1 + 5)$
$\psi_3^{e_{11}}(\xi_1) = \frac{(\xi_1+1)^2}{8}$	$\psi_3^{e_{12}}(\xi_1) = \frac{(\xi_1+1)^2}{4}$

Tableau 1.4 Exemple de fonctions de base locales univariées d'ordre $p_2 = 2$ de la variable ξ_2 avec un nombre d'intervalles $nx_2 = 2$.

Intervalle e_{21}	**Intervalle** e_{22}
$\psi_1^{e_{21}}(\xi_2) = \frac{(\xi_2-1)^2}{4}$	$\psi_1^{e_{22}}(\xi_2) = \frac{(\xi_2-1)^2}{8}$
$\psi_2^{e_{21}}(\xi_2) = \frac{1}{8}(-3\xi_2^2 + 2\xi_2 + 5)$	$\psi_2^{e_{22}}(\xi_2) = \frac{1}{8}(-3\xi_2^2 - 2\xi_2 + 5)$
$\psi_3^{e_{21}}(\xi_2) = \frac{(\xi_2+1)^2}{8}$	$\psi_3^{e_{22}}(\xi_2) = \frac{(\xi_2+1)^2}{4}$

Dans le cas d'un problème avec plusieurs variables aléatoires $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_m\}$, les fonctions de base locales sont multivariées dont la construction se fait par tensorisation à partir des bases univariées. Pour des raisons de simplification, un exemple typique de construction de bases locales bivariées (ξ_1, ξ_2) dans chaque éléments de Bézier sera détaillé, et ce en utilisant les expressions des bases reportées dans les tableaux 1.3 et 1.4. Pour illustrer davantage le processus de construction, une représentation schématique de la décomposition de l'espace paramétrique bivarié, décrit précédemment, est représenté sur la figure 1.4.

Ainsi, et comme illustré sur la figure 1.4, le domaine paramétrique bidimensionnel est décomposé en éléments de Bézier $I_e = e_{1s} \times e_{2r}$, avec $s = 1, 2$, $r = 1, 2$ et $e = 1, \ldots, 4$. Le nombre total d'éléments dans le cas multidimensionnel est $N_{elt} = \prod_{i=1}^{m} nx_i$ et le nombre de fonctions de base dans chaque élément est $n_b = \prod_{i=1}^{m}(p_i + 1)$. L'expression mathématique permettant la

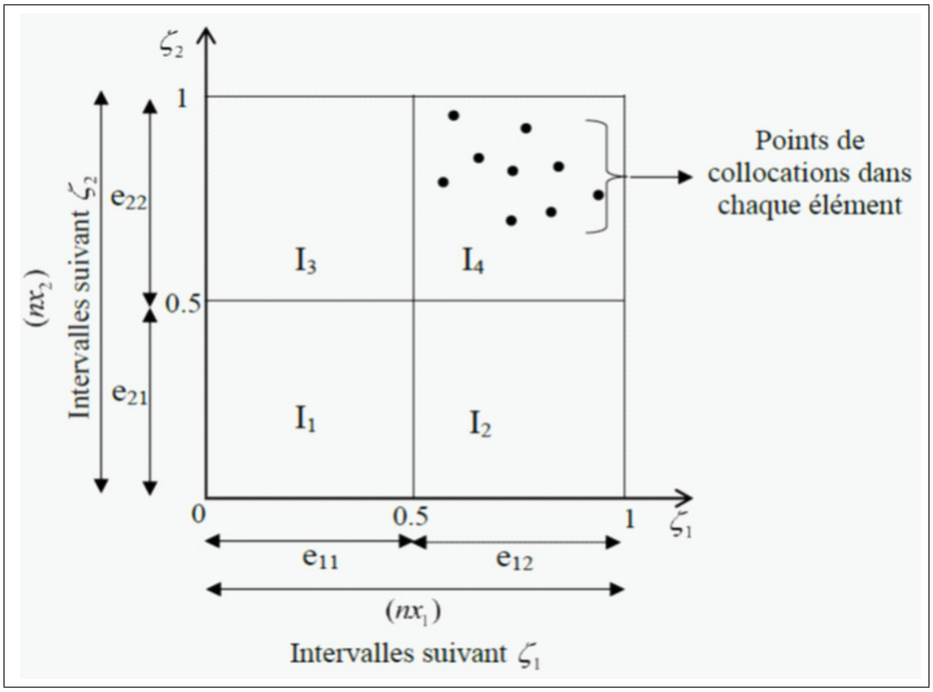


Figure 1.4 Représentation schématique d'une décomposition d'un domaine paramétrique bivarié avec $p_1 = p_2 = 2$ et $nx_1 = nx_2 = 2$.

construction d'une base bivariée dans chaque élément de Bézier est :

$$\Psi_j^e(\xi_1, \xi_2) = \psi_k^{e_{1s}}(\xi_1)) \times \psi_\ell^{e_{2r}}(\xi_2) \tag{1.4}$$

avec $s, r = 1, 2$; $k, \ell = 1, \ldots, (p_1 + 1)(p_2 + 1)$; $e = 1, \ldots, N_{elt}$ et $j = 1, \ldots, n_b$. Pour une meilleure illustration de la procédure de construction des bases bivariées, les principales étapes sont détaillées dans l'algorithme 1.1.

L'autre aspect important qui caractérise l'approche proposée est le processus de sélection des points de collocation. Ainsi, un nombre minimal de points ($n_s^e \geq n_b$) est généré aléatoirement dans l'élément I_e du domaine paramétrique. Par la suite, pour chaque point représentant les variables $\xi_1^{(q)}$, $\xi_2^{(q)}$ lui est associé son point correspondant dans le domaine physique en utilisant la fonction inverse de la distribution cumulative appropriée de sorte : $\eta_1^{(q)} = F_1^{-1}(\xi_1^{(q)})$ et $\eta_{21}^{(q)} = F_2^{-1}(\xi_2^{(q)})$, avec $q = 1, \ldots, n_s^e$. Il y a lieu de mentionner que chaque variable aléatoire est décrite par sa propre fonction de distribution de probabilité. Afin de mieux imager ce processus

Algorithme 1.1 Procédure de construction d'une base locale bivariée

Entrées : $p_1 ; p_2 ; nx_1 ; nx_2 ; \left\{\psi_k^{e_{1s}}(\xi_1)\right\}_{k=1}^{p_1+1} ; \left\{\psi_\ell^{e_{2r}}(\xi_2)\right\}_{\ell=1}^{p_2+1} ; k = 1, \ldots, P_1 + 1$
$\ell = 1, \ldots, P_2 + 1 ; s, r = 1, 2$
Sortie : $\left\{\Psi_j^e(\xi_1, \xi_2)\right\}_{j=1}^{n_b}$
pour $r = 1 : nx_2$ **faire**
  pour $s = 1 : nx_1$ **faire**
    $e = (r-1)nx_1 + s$
    $I_e = e_{1s} \times e_{2r}$
    pour $k = 1 : p_1 + 1$ **faire**
      pour $\ell = 1 : p_2 + 1$ **faire**
        $j = (k-1)(P_1 + 1) + \ell$
        $\Psi_j^e(\xi_1, \xi_2) = \psi_k^{e_{1s}}(\xi_1))\psi_\ell^{e_{2r}}(\xi_2)$
      fin
    fin
  fin
fin

de sélection des points de collocation dans le domaine paramétrique et leurs transformations vers le domaine physique, un exemple illustratif est schématisé dans la figure 1.5 pour le cas d'une distribution normale.

1.3.2.3 Méthodes non-intrusives de réduction de modèles

Les techniques du développement stochastique présentées précédemment, avec leurs différentes variantes, permettent la construction de modèles de substitution avec un coût de calcul largement inférieur comparativement aux techniques classiques d'analyses stochastiques. Ces modèles de substitution sont construits moyennant une combinaison entre les bases polynomiales et les coefficients associés au niveau de chaque point de calcul constituant le maillage et pour chaque temps de simulation dans le cas des problèmes instationnaires. Certains de ces problèmes requièrent des maillages très volumineux, qui sont de l'ordre de millions de nœuds, mais aussi un temps de simulation relativement élevé avec des pas de temps raffinés, ce qui par conséquent peut s'avérer extrêmement couteux en termes de temps de calcul résultant, comme c'est le cas, entre autres, des écoulements de ruptures de barrages (Goutal, Goeury, Ata, Ricci, El Mocayd,

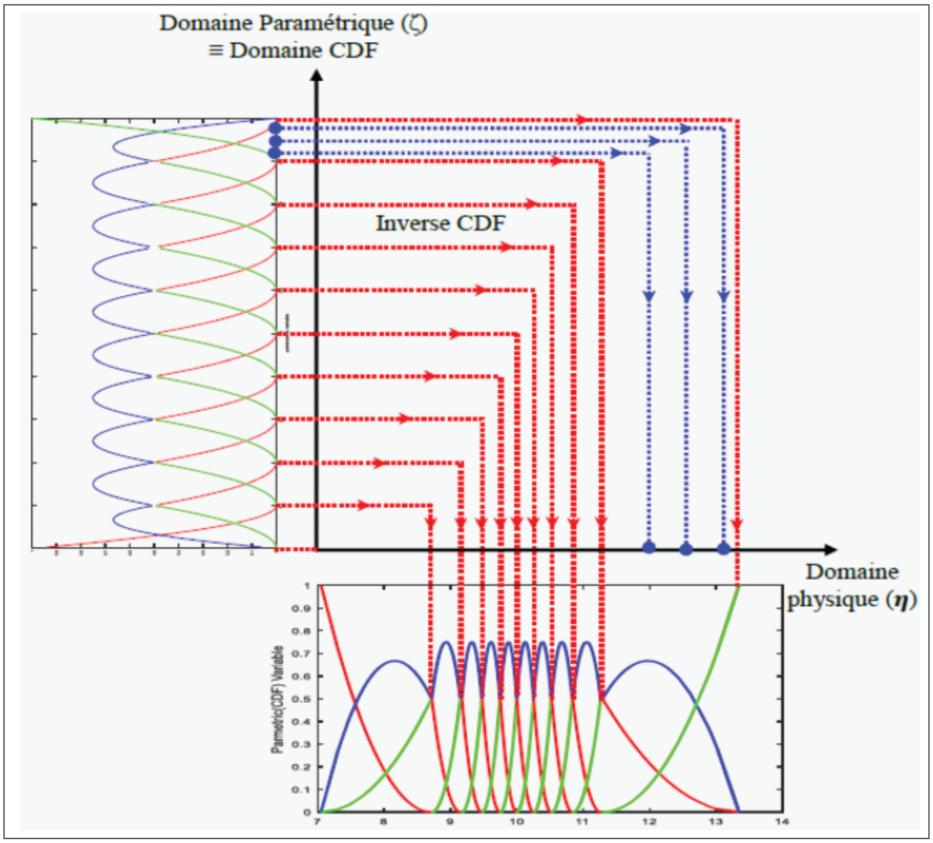


Figure 1.5 Représentation schématique de la procédure d'échantillonnage dans le domaine paramétrique et le transfert vers le domaine physique, avec (p = 2, nx = 10).

Rochoux, Oubanas, Gejadze & Malaterre, 2018). Cette situation est d'autant plus problématique pour les techniques de développement stochastique, particulièrement lorsque le nombre de paramètres incertains devient important. De ce fait, l'introduction des techniques de réduction de modèles s'avère primordiale afin d'apporter des éléments de réponse à la limitation du temps de calcul qu'exigent certaines situations, au regard du potentiel qu'elles offrent pour réduire considérablement l'effort de calcul tout en maintenant une bonne précision dans les approximations (Schilders, Van der Vorst & Rommes, 2008).

Parmi les techniques de réduction de modèles, la décomposition orthogonale aux valeurs propres (POD) (Lumley, 1967) est considérée comme étant la plus couramment utilisée. Elle est aussi connue sous d'autres appellations comme, par exemple, l'analyse par composantes principale (PCA) (Goodall, 1988) ou encore la décomposition de Karhunen-Loeve (Berkooz, Holmes & Lumley, 1993 ; Newman, 1996) pour les processus stochastiques. Elle permet de construire une base réduite optimale de la réponse du modèle en utilisant la décomposition modale. L'introduction de la méthode des snapshots (Sirovich, 1987), qui consiste en une collection de la réponse du modèle objet de l'étude à différents instants, combinée à la technique de décomposition en valeurs singulière (SVD) a contribué à l'essor de l'approche POD dont l'utilisation s'est largement généralisée à des problèmes avec un nombre très élevé de degré de liberté, et à des domaines aussi variés en allant de la mécanique des fluides aux problèmes du traitement et compression d'images (Fukunaga, 2013).

La construction de modèles réduits est basée principalement sur deux grandes catégories d'approches pour l'évaluation des coefficients associés aux modes principaux de la base réduite ; en l'occurrence les méthodes intrusives et les méthodes non-intrusives (Zokagoa, 2011). Les techniques de projection de Galerkin sont parmi les méthodes intrusives les plus répandues. Elles font appel aux équations gouvernantes et nécessitent une modification des codes sources pour générer les systèmes d'équations à résoudre pour le calcul des coefficients (Couplet, Basdevant & Sagaut, 2005 ; Liberge, 2008 ; Zokagoa & Soulaïmani, 2012). Les techniques intrusives permettent dans certaines configurations de donner des résultats remarquables. Néanmoins, leur caractère intrusif représente un inconvénient majeur, particulièrement pour les cas complexes qui nécessitent des modifications importantes des équations gouvernantes, et dans certaines situations l'accès même aux codes sources des solveurs numériques est impossible, comme c'est le cas dans la plupart des codes commerciaux.

De ce fait, et afin de contourner cette limitation, des techniques dites non-intrusives sont apparues et dont l'intérêt ce cesse de grandir du fait qu'elles ne nécessitent aucune modification des équations gouvernantes régissant les phénomènes physiques, et traitent le code source comme une boite noir (Xiao et al., 2016b). Plusieurs techniques non-intrusives ont été proposées ces dernières années qui se basent sur différentes approches d'interpolations pour le calcul des coefficients à l'instar des approches basées sur les fonctions de base radiales (RBF) (Audouze, De Vuyst & Nair, 2013), l'interpolation par la grille creuse de Smolyak (Xiao, Fang, Buchan, Pain, Navon & Muggeridge, 2015b), décomposition du domaine combinée avec les fonctions radiales (Xiao, Heaney, Fang, Mottet, Hu, Bistrian, Aristodemou, Navon & Pain, 2019). La construction des modèles réduits pour les problèmes instationnaires paramétriques représente un défi majeur du fait de la difficulté les bases réduites au regard de la taille des matrices de snapshots générées. L'utilisation de la technique POD standard peut s'avérer insuffisante dans certaines applications qui exigent des capacités en mémoire très importantes. La technique dite double POD a été ainsi introduite (Jacquier, Abdedou, Delmas & Soulaïmani, 2021 ; Wang, Hesthaven & Ray, 2019) permettant de procéder à une décomposition POD à deux niveaux (temporel et paramétrique) pour l'extraction de la base réduite. Les techniques non-intrusives basées sur les réseaux de neurones artificiels (ANN) ont connues un développement considérable où différents algorithmes d'apprentissage machine et d'apprentissage profond ont été proposés (Gao, Wang & Zahr, 2020 ; Guo & Hesthaven, 2019 ; Kani & Elsheikh, 2019), une revue approfondie des différentes variantes ces approches est présentée dans (Yu, Yan & Guo, 2019). Ces techniques ont contribué grandement à la construction de modèles réduits de plus en plus efficace avec l'avantage de la facilité d'implémentation, mais requièrent tout de même, pour certains cas particuliers, un temps d'apprentissage relativement élevé.

Un autre axe de recherche sur lequel se sont penchées les techniques de réduction de modèles est l'analyse de propagation des incertitudes émanant des paramètres d'entrée et quantifier les effets qu'elles génèrent sur la réponse de sortie. Plusieurs méthodes ont été développées dans cette optique comme par exemple les polynômes du chaos qui, une fois combinés, à l'approche POD a donné lieux à une panoplie de variantes (Hijazi, Stabile, Mola & Rozza, 2020 ; Raisee, Kumar & Lacor, 2015). Dans certaines configurations, les bases du chaos polynomial sont utilisées pour extraire des bases orthogonales optimales à partir d'une grille grossière, pour ensuite qu'elles soient utilisées pour les calculs à travers un maillage plus raffiné (Abraham, Tsirikoglou, Miranda, Lacor, Contino & Ghorbaniasl, 2018 ; Raisee, Kumar & Lacor, 2019). D'autres techniques ont utilisé les bases des polynômes du chaos pour construire une régression dans le domaine paramétrique permettant le calcul des coefficients stochastiques associés à la base réduite (El Moçayd, Mohamed, Ouazar & Seaid, 2020 ; Sun, Pan & Choi, 2019). C'est dans cette optique qu'une nouvelle approche est proposée dans le cadre de cette thèse, et qui fait l'objet du chapitre 4, pour construire un modèle réduit stochastique non-intrusif en combinant la technique de la double POD et la méthode BSBEM pour l'évaluation des coefficients associés avec un troisième niveau POD.

1.4 Conclusion

Dans ce chapitre, une revue de littérature portant sur les différents axes de recherche faisant l'objet de la présente thèse a été présentée. Une présentation exhaustive des différents travaux qui ont traité des écoulements à surface libre, des méthodes d'analyse de propagation d'incertitudes et des techniques non-intrusives de réduction de modèles. Quelques notions complémentaires portant sur des concepts fondamentaux des différentes approches ont été présentées, particulièrement les techniques des polynômes du chaos et la nouvelle approche BSBEM qui feront l'objet d'une présentation détaillée dans le prochain chapitre.

CHAPITRE 2

A NON-INTRUSIVE B-SPLINES BÉZIER ELEMENTS-BASED METHOD FOR UNCERTAINTY PROPAGATION

Azzedine Abdedou, Azzeddine Soulaïmani

Department of Mechanical Engineering, Ecole de Technologie Supérieure, 1100 Notre-Dame W., Montreal (QC), Canada H3C 1K3

Article publié dans la revue :
Journal of Computer Methods in Applied Mechanics and Engineering, Volume 345, 13 November 2018
https ://doi.org/10.1016/j.cma.2018.10.047

Abstract

A non-intrusive B-Splines Bézier Elements based Method (BSBEM) is proposed as an efficient tool for uncertainty propagation analysis in physical hyperbolic problems. The model's output response is approximated using a surrogate model whose coefficients are obtained from a set of deterministic calls by means of a regression technique. The accuracy, efficiency and the generality of the proposed approach are assessed using benchmark numerical examples by comparing the convergence of the statistics moments with those of the Polynomial chaos-based point collocation method (Pcol) and the Monte Carlo (MC) method. The generic character of the proposed approach allows it to be implemented for several engineering fields. BSBEM is applied to quantify uncertainty propagation through dam break flows modelled by shallow water equations. The obtained results, which are depicted in terms of water depth and inundation line confidence intervals, show that with a meticulous exploitation of the multi-element aspect and the smoothness feature of the basis functions, the proposed method provides an accurate and smooth approximation of the stochastic output response.

keywords

Uncertainty propagation, B-splines, Bézier elements, polynomial chaos, dam break flows, shallow water equations

2.1 Introduction

The increase of computational power has contributed to the development of advanced numerical models that are an essential and a powerful means to deepen the understanding and improve the prediction of physical phenomena in several fields of engineering sciences. The numerical results and the products derived from these models are usually performed with fixed values of the input parameters without taking into account the numerous uncertainties to which they may be subjected. The theme of uncertainty quantification (UQ) analysis has therefore attracted attention over the past few decades owing to its various engineering applications, including structural design (Capillon, Desceliers & Soize, 2016 ; Chen & Qiu, 2018 ; Gao, Zhou & Huang, 2017), aerospace engineering (Chen, Qiu, Wang, Li & Wang, 2017 ; Weaver, Alexeenko, Greendyke & Camberos, 2011), nuclear safety analysis, thermal engineering (Fajraoui, Fahs, Younes & Sudret, 2017) and hydraulic-environmental studies (Chen, Yang, Wang, Xu & Yu, 2013 ; Faghih, Mirzaei, Adamowski, Lee & El-Shafie, 2017).

The Monte Carlo method is among the most common and straightforward methods for uncertainty propagation due to its robustness and implementation simplicity (Weitz *et al.*, 2016). However, the use of such a method for cases involving complex deterministic models can prove to be prohibitive in computational cost due to the large number of executions required to achieve an acceptable accuracy. Despite this drawback, the Monte Carlo method is often used as a benchmark solution to evaluate the accuracy and reliability of other sampling methods, such as Latin Hypercube Sampling (LHS)(McKay *et al.*, 1979) and Multi-Level Monte Carlo (MLMC)(Mishra, Schwab & Šukys, 2016), which were introduced to compensate for the low convergence rate of the Monte Carlo method.

As an alternative to sampling approaches, the polynomial chaos (PC) expansion proposed by Ghanem and Spanos (Ghanem & Spanos, 1991), which belongs to the stochastic expansion methods, was developed to quantify uncertainty propagation in numerical models. This method, based on the homogenous chaos theory of Wiener (Wiener, 1938), uses Hermite polynomial basis functions to model the uncertainties of the output response resulting from a stochastic process with normally-distributed random input parameters. For non-Gaussian random input variables, an extension of the Wiener-Hermite PC expansion, known as the generalized polynomial chaos expansion (gPCE), was proposed by Xiu and Karniadakis (Xiu & Karniadakis, 2002). In gPCE, orthogonal polynomials are selected from the Askey scheme in accordance with the corresponding probability distributions of the random input variables, thereby allowing for an exponential convergence rate of the statistical moments.

The coefficients of the PCE method are evaluated, in their earliest development, by the intrusive approach based on the Galerkin projection, which requires modification of the governing equations of the deterministic model, thus leading to a system of coupled equations whose unknowns are the coefficients of the PC expansion (Dinescu *et al.*, 2010 ; Ghanem, 1999). This intrusive technique presents a real difficulty in some complex deterministic codes, requiring time-consuming cumbersome developments, which renders it much less attractive. As a result, other methods have been proposed, such as the so-called non-intrusive polynomial chaos expansion (NIPCE), which uses deterministic models as black boxes and thus require no modification of the governing equations. The coefficients of the NIPC expansion are evaluated either with a spectral projection technique that exploits the orthogonality of the polynomial basis, using the quadrature approach to estimate the resulting multidimensional integral (Ghiocel & Ghanem, 2002), or with a regression technique, solving an oversized linear system of equations by using the least square minimization approach (Blatman & Sudret, 2010a). When the number of input random variables increases for a specified order of the polynomial basis, the number of model evaluations that the projection approach requires to evaluate the multidimensional integrals becomes prohibitively large, which leads to the so-called curse of dimensionality (Blatman & Sudret, 2011). As an alternative, the sparse or Smolyak quadrature technique is introduced to alleviate the curse of dimensionality by reducing the number of the collocation points, as reported in (Matthies & Keese, 2005), particularly in stochastic collocation methods (Xiu, 2007). Another scheme, the so-called probabilistic collocation method, which considers the collocations points as roots of the orthogonal polynomials, is proposed as a non-intrusive PCE and allows the number of model evaluations to be reduced (Loeven *et al.*, 2007 ; Tatang *et al.*, 1997). The point collocation method (Hosder & Walters, 2010 ; Hosder *et al.*, 2007), is also considered as an alternative approach to deal with uncertainty quantification with an appropriate accuracy. It combines the gPCE approach with well-known sampling methods (that are also associated with an adequate oversampling technique) to choose the collocation points.

Recently, much work has been devoted into developing approaches that are more efficient for uncertainty propagation analysis and thereby reduce simulation costs. Some of these techniques have deeply investigated in terms of how to optimize sampling procedures and how to reduce the number of polynomial basis functions required to capture the main stochastic features of the output response. These efforts include the least angle regression technique (LAR)(Blatman & Sudret, 2010a,1), adaptive sparse polynomial chaos expansion (Cheng & Lu, 2018 ; Quicken, Donders, van Disseldorp, Gashi, Mees, van de Vosse, Lopata, Delhaas & Huberts, 2016), sliced inverse regression-based sparse polynomial chaos expansions (Pan & Dias, 2017), multi-fidelity non-intrusive polynomial chaos (Palar, Tsuchiya & Parks, 2016 ; Palar, Zuhal, Shimoyama & Tsuchiya, 2018 ; Salehi, Raisee, Cervantes & Nourbakhsh, 2018), and more recently, using polynomial chaos decomposition with a differentiation approach (Thapa *et al.*, 2018).

Thanks to their multiple attractive features, non-intrusive polynomial chaos approximations are widely used in uncertainty quantification analysis to provide accurate statistical moments and to build an efficient surrogate for the complex original model of the output response. Nevertheless, the use of such basis functions presents some limitations when the output response presents hyperbolic behaviour (Xiu, 2010). In those cases, non-intrusive polynomial chaos approaches may fail to reproduce the stochastic behavior of the model's output response. Several methods are proposed in the literature to overcome this limitation, such as the Wiener-Har expansion based on wavelets, which belongs to a multi-element generalized polynomial chaos approach (Le Maître *et al.*, 2004a ; Wan & Karniadakis, 2005), multi resolution analysis (MRA)(Le Maître, Najm, Ghanem & Knio, 2004b) and a piecewise polynomial approximation based on a hybrid global and adaptive polynomial algorithm (Barth, 2013).

The basic concept motivating the present work is to introduce the B-Spline basis functions as an alternative approach to the existing schemes to overcome, as much as possible, the limitation of the polynomial chaos method to describe an output response with strong hyperbolic behaviour. It must be emphasized that B-Spline functions have only just been introduced in the field of uncertainty quantification analysis in recent studies (Baroth *et al.*, 2006,0 ; Hadrich, Zribi & Masmoudi, 2016), in which they were used as weighting functions in the projection phase to evaluate the coefficients of the approximations. More recently, B-spline functions have been introduced as basis functions to investigate the data-driven uncertainty quantification of structural systems (Dertimanis *et al.*, 2017).

In this paper, a non-intrusive regression stochastic method is proposed, wherein the expansion is defined on the cumulative probabilities domain that is decomposed into elements (called Bézier elements), and where the B-splines functions (with the input parameters' sampling) are established locally in the Bézier elements. The use of such interpolations allows smooth statistics of the outputs (mean and variance) to be obtained in the case of discontinuous wave flows. The multi-element aspect of the method and the compact support feature of the BSBE method are explored to increase the efficiency of the sampling procedure and the smoothness of the approximation.

The paper is organized as follows. Section 2.2, presents a review of the polynomial chaos method. The proposed non-intrusive B-Spline Bézier elements method (BSBEM) is presented in section 2.3. Section 2.4 evaluates the accuracy of the BSBE method using analytical test cases with two and three input random variables, and the proposed approach is applied to dam break flow test cases, first over a test bed and then over real topography. The accuracy of the method is compared to both to the PCE approach and to the Monte Carlo reference solutions. Finally, the conclusions of this study and some suggestions for future research are presented in section 2.5.

2.2 Non-intrusive PCE for uncertainty propagation using stochastic collocation

The polynomial chaos expansion is based on the spectral representation of the scalar output response ($Y = f(\mathbf{x})$) of the deterministic model, which is approximated as a truncated sum of the orthonormal multivariate polynomials of a vector of uncertain input parameters $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$:

$$\widehat{Y} = \sum_{i=1}^{M} \alpha_i \Psi_i(\mathbf{x}) \tag{2.1}$$

where m is the number of input random variables and α_i are the deterministic weighting coefficients to be determined. $\Psi_i(\mathbf{x})$ are the multivariate basis functions that describe the stochastic part of the expansion. The multivariate basis functions can be constructed using the tensor product of the univariate polynomials $\psi_{p_i^j}(x_i)$, involving a multi-index p_i^j associated to each input random variable (Eldred & Burkardt, 2009) :

$$\Psi_j(\mathbf{x}) = \prod_{i=1}^{m} \psi_{p_i^j}(x_i) \tag{2.2}$$

The total number of expansion terms M, with the same p-order polynomial chaos expansion (for all m variables), is thus reduced to :

$$M = \frac{(p+m)!}{p!m!} \tag{2.3}$$

The choice of univariate polynomial basis function $\psi_{p_i^j}(x_i)$ is closely related to the probability density function of each input random variable x_i, which ensures an exponential convergence. For instance, Hermite and Legendre polynomials serve as optimal basis functions for normal and uniform distributions, respectively. Further details on common probability distributions and their corresponding optimal polynomials in the Askey scheme are provided in (Xiu & Karniadakis, 2002).

The polynomial chaos expansion coefficients can be computed using the non-intrusive regression technique (Blatman & Sudret, 2010a). This approach is based on a set of deterministic calculations for M vectors $\mathbf{x}_i = (x_1, x_2, \ldots, x_m)_i^T$, $i = 1, 2, \ldots, M$ of the input uncertain parameters, which are known as an experimental design set. The evaluation of the deterministic model at these collocation points leads to the response vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_M)^T$, where its i^{th} component represents the model response corresponding to the i^{th} vector of the input uncertain parameters ($Y_i = f(\mathbf{x}_i)$, $i = 1, \ldots, M$). The selection of collocation points in the experimental design has a profound impact on the accuracy of the regressive methods, such as the point collocation expansion (Hosder *et al.*, 2007), which is considered among the most widely-used methods for UQ. This approach combines different sampling techniques such as random sampling, LHS and Harmemesly sampling with the oversampling concept by introducing the oversampling ratio n_p :

$$N_s = n_p \frac{(p+m)!}{p!m!} = n_p M \tag{2.4}$$

where N_s is the number of the collocation points of the oversampled experimental design, which is a multiple of the minimum required points given by Eq.(2.3). The obtained overdetermined system of equations, whose unknowns are the PC coefficients that are the solutions of the following normal equation :

$$(\mathbf{\Psi}^T \mathbf{\Psi})\alpha = \mathbf{\Psi}^T \mathbf{y} \tag{2.5}$$

where $\mathbf{\Psi}^T \mathbf{\Psi}$ denotes the pseudo inverse of the data matrix $\mathbf{\Psi}$, also called the design matrix, which is built by evaluating the M multivariate polynomials, representing each row, at a particular $\mathbf{x}$

sample, as follows :

$$\mathbf{\Psi} = \left(\Psi_{ij} = \Psi_j(\mathbf{x}_i)\right), \; i = 1, \ldots, n \text{ and } j = 1, \ldots, M \tag{2.6}$$

Once the coefficients of the expansion are computed, the evaluations of the mean μ and the standard deviation σ of the variable of interest become straightforward using the orthogonality property of the basis functions, as in :

$$\mu = \mathbb{E}\left[\widehat{Y}\right] = \int_{\Xi} \left[\sum_{i=1}^{M} \alpha_i \Psi_i(\mathbf{x})\right] \varrho_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \alpha_1 \tag{2.7}$$

$$\sigma = \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}\left[\widehat{Y}\right]\right)^2\right]^{1/2} = \left[\sum_{i=2}^{M} \alpha_i^2 \langle \Psi_i^2\,(\mathbf{x})\rangle\right]^{1/2} \tag{2.8}$$

The joint probability (weight) function $\varrho_{\mathbf{x}}(\mathbf{x})$ is defined as the product of the marginal density probability $\varrho_i(x_i)$ of each single independent random variable x_i, as given below :

$$\varrho_{\mathbf{x}}(\mathbf{x}) = \prod_{i=1}^{m} \varrho_i(x_i) \tag{2.9}$$

The orthogonality property with respect to the inner product $\langle \bullet, \bullet \rangle$ is used to evaluate the integrals in Eq.(2.8) :

$$\langle \Psi_i, \Psi_j \rangle = \int_{\Xi} \Psi_i\,(\mathbf{x})\,\Psi_j\,(\mathbf{x})\,\varrho_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \tag{2.10}$$

2.3 A non-intrusive B-splines Bézier elements' expansion

The use of B-Splines functions has been steadily increasing due to the mathematical properties they provide. They have been coupled with the finite element method to set up a new concept called Isogeometric Analysis, introduced by Hughes et al (Cottrell, Hughes & Bazilevs, 2009 ; Hughes *et al.*, 2005). The use of such basis functions in uncertainty quantification seems to be especially promising for phenomena with hyperbolic behaviour.

2.3.1 B-splines and the Bézier extraction operator

A univariate B-spline basis is defined by a knot vector $V = \left\{\xi_1, \xi_2, \ldots, \xi_{n_b+p+1}\right\}$, which represents a set of non-decreasing coordinates $0 \leqslant \xi_1 \leqslant \xi_2 \leqslant \ldots \leqslant \xi_{n_b+p+1}$, where ξ_i is the i^{th} knot, p is the polynomial order and n_b is the number of B-spline basis functions $N_{i,p}(\xi)$ that are defined recursively using the Cox-de Boor formula (Hughes *et al.*, 2005), starting from the piecewise constant basis function for $p = 0$:

$$N_{i,o}(\xi) = \begin{cases} 1, & \text{if } \xi_i \leqslant \xi \leqslant \xi_{i+1} \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

and for $p \geqslant 1$, the B-spline basis functions are defined by :

$$N_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi) \tag{2.12}$$

In the following, we will use an open knot vector with $\xi_1 = \xi_2 = \ldots = \xi_{p+1}$ and $\xi_{n_b+1} = \xi_{n_b+2} = \xi_{n_b+p+1}$ for which $\xi_i \neq \xi_j$ if $p + 1 < i, j \leq n_b$. The expression in Eq.(2.12) defines n_b B-spline functions with C^{p-1} continuity and $nx = n_b - p$ intervals $I_e = [\xi_e, \xi_{e+1}]$, called Bézier elements. Local piecewise B-spline basis functions can also be obtained using the Bézier decomposition technique by introducing the local Bézier extraction operator, which leads to a mapping from the Bernstein basis functions onto Bézier elements. The local piecewise B-spline functions thus obtained can be defined as (Borden *et al.*, 2011 ; Li, Gao, Wu, Song & Chen, 2018) :

$$\mathbf{N}^e(\xi) = \mathbf{C}^e \mathbf{B}^e(\xi) \tag{2.13}$$

where $\mathbf{N}^e(\xi) = \left(N^e_{i,p}(\xi)\right)^T$, $(i = 1, \ldots, p + 1)$ indicates the vector of the $(p + 1)$ B-spline basis functions defined on the Bézier element e and $\mathbf{C}^e$ is the linear element extraction operator. A deeper insight into the Bézier extraction procedure and the extraction operator algorithm is provided by Borden *et al.* (2011). $\mathbf{B}^e = \left(B^e_{i,p}\right)$ denotes the Bernstein functions of degree p on the e^{th} Bézier element which can be expressed on the parametric domain $[-1, 1]$ as (Malakiyeh, Shojaee & Javaran, 2018 ; Scott, Borden, Verhoosel, Sederberg & Hughes, 2011) :

$$B_{i,p}(\tau) = \frac{1}{2^p}\binom{p}{i-1}(1-\tau)^{p-(i-1)}(1+\tau)^{i+1} \quad \text{for} \quad \tau \in [-1,1] \tag{2.14}$$

where :

$$\binom{p}{i-1} = \frac{p!}{(i-1)!(i+1-p)!}$$

is the binomial coefficient with $1 \leqslant i \leqslant p+1$.

The univariate Bernstein basis function can also be expressed in an explicit recursive form as (Buffa & Sangalli, 2016) :

$$B_{i,p}(\tau) = \frac{1}{2}(1-\tau)\,B_{i,p-1}(\tau) + \frac{1}{2}(1+\tau)\,B_{i-1,p-1}(\tau) \tag{2.15}$$

with $B_{i,0}(\tau) = 1$, and $B_{i,p}(\tau) = 0$ if $i < 1$ or $i > p+1$. Given a value for $\xi \in I^e$, the parametric variable is computed using the following change of variable :

$$\tau = \frac{2\xi - (\xi_e + \xi_{e+1})}{\xi_{e+1} - \xi_e} \tag{2.16}$$

2.3.2 Stochastic framework

Let x be the univariate random variable defined on the probability space (Θ, Σ, P), where Θ is the event space, P the probability measure and Σ is the σ-algebra on Θ. The statistical behaviour of $x(\theta)$, with its support denoted by Ξ, is characterized by a probability density function $\varrho(x)$ such that $dF = \varrho(x)dx$, with the marginal cumulative distribution function (CDF) defined by : $F(x) = \int_{-\infty}^{x} \varrho(x')dx'$ and its image domain is $[0,1]$. In the proposed approach, the CDF domain will be decomposed using an open knot vector of nx intervals (Bézier elements) $I_e = \left[\xi_{e+p}, \xi_{e+p+1}\right]$, with order p and $nx = n_b - p$. It is important to note that ξ will be used as the parametric variable, and that it physically represents a cumulative probability value, i.e : $\xi(x) = F(x) = \int_{-\infty}^{x} \varrho(x')dx'$, which implies $d\xi = \varrho(x)dx$. The non-intrusive B-spline Bézier elements' expansion is an approximation $\widehat{Y}$ of the output stochastic response of the variable of interest $Y = f(x(\theta))$ as the sum of the piecewise polynomials of the parameter ξ on each Bézier element I_e. For the sake of simplicity, the proposed approach will be detailed first for the univariate case :

$$\widehat{Y}^e = \sum_{i=1}^{p+1} \alpha_i^e \psi_i^e(\xi) \tag{2.17}$$

where α_i^e are the unknown coefficients and $\psi_i^e(\xi)$ are the p-order univariate piecewise B-spline basis functions $\psi_i^e = N_{i,p}^e$ (for $i = 1, \ldots, p+1$) over the Bézier element I_e.
The evaluation of each component of the set of basis functions at the n^e collocation points in each element allows the construction of the local design matrix $\mathbf{\Psi}^e$, whose components are $\psi_{ij}^e = N_{j,p}^e(\xi_i^e)$ (with $i = 1, \ldots, n^e$ and $j = 1, \ldots, p+1$), which can be further expressed as :

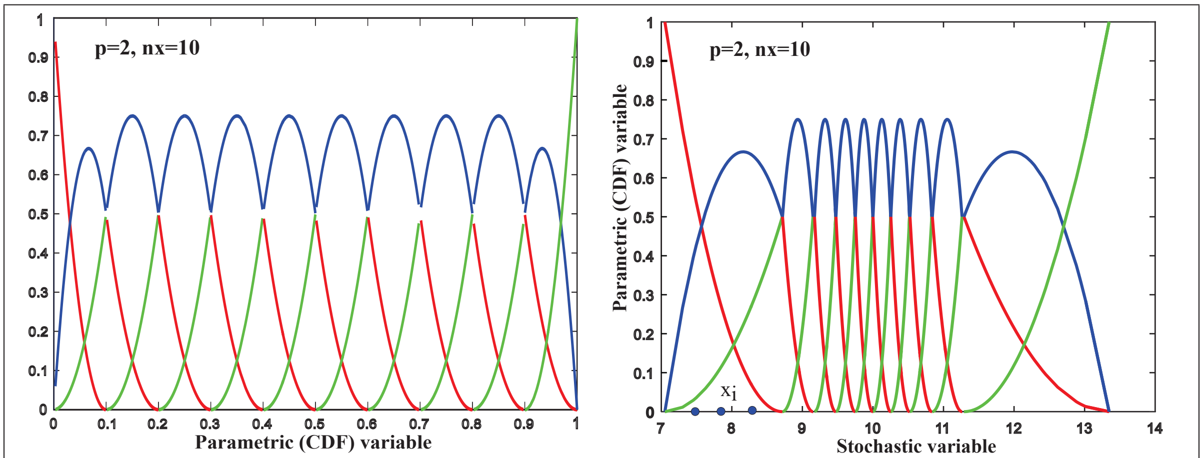
$$\mathbf{\Psi}^e = \begin{bmatrix} \psi_1^e(\xi_1^e) & \cdots & \psi_{p+1}^e(\xi_1^e) \\ \vdots & \ddots & \vdots \\ \psi_1^e(\xi_{n^e}^e) & \cdots & \psi_{p+1}^e(\xi_{n^e}^e) \end{bmatrix} \tag{2.18}$$

Once the collocation points are selected according to a sampling technique, such as uniform random sampling or LHS, within the e^{th} Bézier element $\left(\xi_i^e \in I_e\right)$, their corresponding set of collocation points $\left(x_i^e, i = 1, \ldots, n^e\right)$ on the physical domain are obtained with the inverse cumulative distribution function by :

$$x_i^e = F^{-1}(\xi_i^e) \tag{2.19}$$

The corresponding components Y_i^e of the local output response vector $\mathbf{Y}^e$ are computed by evaluating the deterministic model at these collocation points. To define the $(p+1)$ B-splines functions, $p+1$ independent collocation points are required (see Fig.2.1 for an illustration of the proposed procedure).

As for the aforementioned point collocation method (Pcol), it is also possible to apply the oversampling technique in this approach by selecting $n^e = n_p(p+1)$ collocation points in each Bézier element, in which n_p denotes the oversampling ratio defined in Eq.(2.4). In this case, the



a) $(p+1)$ parametric collocation points (ξ_i^e) b) $(p+1)$ physical collocation points (x_i^e)

Figure 2.1 Sampling procedure using inverse CDF mapping with $(p+1)$ collocation points from the parametric domain (ξ_i^e) to the physical domain (x_i^e) in each e^{th} Bézier element. In this case normal cumulative density function with mean of 10 and standard deviation of 1 is used.

local design (information) matrix and the local output vector response can be computed from the obtained oversized matrix $\boldsymbol{\psi}^e_{n^e\times(p+1)}$ and output response vector $\mathbf{Y}^e_{n^e\times 1}$, respectively. The resulting local design matrix and output response vector can be expressed, respectively, as :

$$\boldsymbol{\Psi}^e = \boldsymbol{\psi}^{e^\top}\boldsymbol{\psi}^e \quad \text{and} \quad \mathbf{Y}^e = \boldsymbol{\psi}^{e^\top}\mathbf{y}^e \tag{2.20}$$

Unlike other multi-element methods where the expansion coefficients are computed separately for each element (Yucel, Bagci & Michielssen, 2013), the proposed approach evaluates the coefficients through a global system of equations resulting from the assembly of the local design matrices and output response vectors over all Bézier elements. The obtained global linear system of equations can be expressed as :

$$\boldsymbol{\Psi}\boldsymbol{\alpha} = \mathbf{Y} \tag{2.21}$$

which can be expanded further as :

$$\begin{bmatrix} \Psi_{11} & \cdots & \Psi_{1M} \\ \vdots & \ddots & \vdots \\ \Psi_{M1} & \cdots & \Psi_{MM} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_M \end{bmatrix} \tag{2.22}$$

where $\mathbf{\Psi} = \underset{e=1}{\overset{nx}{\mathbf{A}}}(\mathbf{\Psi}^e)$ and $\mathbf{Y} = \underset{e=1}{\overset{nx}{\mathbf{A}}}(\mathbf{Y}^e)$ denote the global design matrix and the global output response vector, whose terms are $\Psi_{k,\ell}$ and Y_ℓ, respectively, obtained from the assembly of the local design matrices and response vectors using the assembly operator **A** introduced by Hughes (Hughes, 2012). The assembly procedure makes it possible to use the connectivity array (IEN) that links the global basis functions' indices (k, ℓ) to the local ones (i, j) with respect to the Bézier elements with subscript number (e). For each local subscript number $i \in \{1, \ldots, p+1\}$ and $j \in \{1, \ldots, p+1\}$, the corresponding global numbers (k, ℓ) are computed from $k = IEN(i, e);\ k \in \{1, \ldots, nx+p\}$ and $\ell = IEN(j, e);\ \ell \in \{1, \ldots, nx+p\}$ with $e \in \{1, \ldots, nx\}$. For more implementation details see (Borden *et al.*, 2011 ; Hughes *et al.*, 2005 ; Scott *et al.*, 2011).

The multi-elements aspect of the proposed approach, based on the Bézier extraction operator, provides a simple and efficient random space decomposition strategy. The use of B-spline functions, which are $C^{(p-1)}$ smooth basis functions, have been shown (see (Cottrell *et al.*, 2009)) to be less prone to oscillations than C^0 functions for hyperbolic problems.

2.3.3 Post-processing

Once the coefficients α are computed by solving Eq.(2.22), the local approximation of the output response by the stochastic expansion, given in Eq.(2.17), is performed by extracting the local coefficients $\{\alpha_i^e\}_{i=1}^{p+1}$ from a sequence of global coefficients $\{\alpha_\ell\}_{\ell=1}^{M=nx+p}$ using the same aforementioned IEN connectivity array technique. Calculation of the statistical moments of the output response of the variable of interest becomes straightforward once the surrogate model has been built. Thus, the mean and the standard deviation can be obtained, respectively, as follows :

$$\mu_{\widehat{Y}} = \mathbb{E}\left[\widehat{Y}\right] = \int_{\Xi} \widehat{Y} \varrho(x) dx \tag{2.23}$$

and

$$\sigma^2_{\widehat{Y}} = \mathbb{E}\left[\widehat{Y}^2\right] - \mu^2_{\widehat{Y}} = \left[\int_{\Xi} \left(\widehat{Y}\right)^2 \varrho(x) dx\right] - \mu^2_{\widehat{Y}} \tag{2.24}$$

Using the variable transformation, including the cumulative density function mapping $\xi = F(x)$ and $d\xi = \varrho(x)dx$, the computation of integrals in Eqs.(2.23) and (2.24) can be simplified on the bounded interval $[0, 1]$ corresponding to the cumulative density function image domain (Barth, 2013 ; Xiu, 2010). The expression of the statistical moments can be written for the univariate case as :

$$\mu_{\widehat{Y}} = \sum_{e=1}^{nx} \left[\int_{I_e} \widehat{Y}^e d\xi\right] \tag{2.25}$$

and

$$\sigma^2_{\widehat{Y}} = \sum_{e=1}^{nx} \left[\int_{I_e} \left(\widehat{Y}^e\right)^2 d\xi\right] - \mu^2_{\widehat{Y}} \tag{2.26}$$

The integrals are evaluated using, for instance, the Gauss-Legendre quadrature technique in each Bézier element.

2.3.4 Multivariate case

The aforementioned procedure can be generalized to the multivariate case with m input uncertain parameters $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)^T$, where the output response can be approximated in each of the corresponding multidimensional Bézier elements with Eq.(2.17) using the multivariate B-spline basis functions that can be formed as the tensor product of a sequence of univariate basis functions. The construction procedure can be illustrated, for instance, for the bivariate case in the e^{th} Bézier element $I_e = [\xi_{e_1+p}, \xi_{e_1+p+1}] \times [\xi_{e_2+q}, \xi_{e_2+q+1}]$, with $e_i = 1, \ldots, nx_i$ for $i = 1, 2$, by the following expression (Borden *et al.*, 2011 ; Scott *et al.*, 2011) :

$$\Psi^e_r(\boldsymbol{\xi}) = N^e_{i,p}(\xi_1) N^e_{j,q}(\xi_2) \tag{2.27}$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2)$ denotes the random vector and r denotes the bivariate index defined as $r(i, j) = (p+1)(j-1)+i$ for $i = 1, \ldots, (p+1)$, $j = 1, \ldots, (q+1)$ and $r = 1, \ldots, (p+1)(q+1)$.

The total number of Bézier elements in the case of m input random variables can be computed by $N_{elt} = \prod_{i=1}^{m} nx_i$, where nx_i denotes the number of intervals corresponding to the random variable ξ_i. The sampling procedure consists of selecting $n^e = n_p \prod_{i=1}^{m}(p_i+1)$ collocation points, where p_i denotes the order of the i^{th} random variable in each Bézier element I_e, which can be represented as a hypercube based on the tensor product of the one-dimensional coordinates in each direction (each random variable). The local design matrix given by Eq.(2.18) is then constructed by evaluating the $n_b = \prod_{i=1}^{m}(p_i + 1)$ multivariate local basis functions at these collocation points by $\boldsymbol{\Psi}^e = \left(\Psi_j^e(\boldsymbol{\xi}_i^e)\right)$; $j = 1, \ldots, n_b$, where $\boldsymbol{\xi}_i^e = (\xi_1, \ldots, \xi_m)_i^T$; $i = 1, \ldots, n^e$. The collocation points in the physical domain can be obtained by the cumulative density function mapping given by Eq.(2.19), which can be different from one random variable to another. The assembly procedure remains the same as for the univariate case, with the number of the global coefficients to be computed given by $M = \prod_{i=1}^{m}(nx_i + p_i)$. The evaluation of the statistical moments given by Eqs.(2.25) and (2.26) for the multivariate case can be written as :

$$\mu_{\widehat{Y}} = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} \widehat{Y}^e d\boldsymbol{\xi} \right] \tag{2.28}$$

and

$$\sigma_{\widehat{Y}}^2 = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} \left(\widehat{Y}^e\right)^2 d\boldsymbol{\xi} \right] - \mu_{\widehat{Y}}^2 \tag{2.29}$$

which can be computed using the multidimensional Gauss quadrature formula based on the tensor product technique.

The main steps of the non-intrusive B-spline Bézier elements expansion for uncertainty quantification are illustrated for the bivariate case ($m = 2$) in Algorithm 2.1 :

Algorithme 2.1 Non-intrusive B-Spline Bézier elements method : bivariate case

Input : *p-order, q-order, nx_1, nx_2, n_p, μ_{x_i} and σ_{x_i} (statistical moments of the input uncertain parameters x_i; $i = 1, 2$)*
Output : *$\mu_{\widehat{Y}}$, $\sigma_{\widehat{Y}}$ mean and the standard deviation of the output response.*
Compute the number of Bézier elements : $N_{elt} = nx_1 \times nx_2$, the number of collocation points in each elements : $n^e = n_p(p+1)(q+1)$
for *e=1 :N_{elt}* **do**
    Define interval bounds for each random variable : $\xi_i^e \in \left[\xi_{e_i}, \xi_{e_i+1}\right]$
    Generate n^e samples $\left\{\xi_{i,j}^e\right\}_{j=1}^{n^e} \in \left[\xi_{e_i}, \xi_{e_i+1}\right]$, $i = 1, 2$
    Compute from Eq.(2.16) the corresponding collocation points $\left\{\boldsymbol{\tau}_{i,j}^e\right\}_{j=1}^{n^e}$
    Build the bivariate local basis functions $\Psi_r^e(\boldsymbol{\xi})$ from Eq.(2.27)
    Build the local design matrix $\boldsymbol{\Psi}^e$
    Compute the corresponding physical collocation points $\left\{\mathbf{x}_j\right\}_{j=1}^{n^e}$ from Eq.(2.19)
    Evaluate the deterministic model to get the corresponding output response $\mathbf{Y}^e$
    Assembly $\boldsymbol{\Psi}^e$ and $\mathbf{Y}^e$ into $\boldsymbol{\Psi}$ and $\mathbf{Y}$
end
Compute the global coefficients $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{M}$ from Eq.(2.22)
for *e=1 :N_{elt}* **do**
    Build the surrogate model see Eq.(2.17) by extracting the local coefficients $\boldsymbol{\alpha}^e$ from $\boldsymbol{\alpha}$
    Compute the element contribution to the statistical moments from Eqs.(2.28) and (2.29)
end

The performance of the proposed approach will be examined in the sequel for smooth and discontinuous problems involving a relatively small number of input uncertain parameters ($m = 1, 2$ *and* 3). Its generalization for cases involving a large number of random variables leads to a substantial increase in computational effort, as in the case of polynomial chaos techniques based on Cartesian products. As mentioned earlier, the total number of Bézier elements is $N_{elt} = \prod_{i=1}^{m} nx_i$ and the total number of collocation points in each Bezier element with no oversampling is given by $n^e = \prod_{i=1}^{m}(p_i + 1)$.

The BSBE approach offers several possibilities that should be explored further in order to significantly reduce the computational cost ; for instance, its ability to independently select the polynomial order as well as the number of Bézier elements in each random dimension. To this end, a priori calculations can be achieved by using a minimum required number of collocation points in order to determine, through a sensitivity analysis, a hierarchical structure of the most influential parameters, and to subsequently apply a refinement in the number of Bézier elements (h-type refinement) as well as in the polynomial order (p-type refinement).

2.4 Results and discussion

2.4.1 Bivariate numerical test case

The performance of the proposed method is assessed first by considering a benchmark test where an analytical function output is known as a deterministic model with two input uncertain parameters (x_1, x_2) following a uniform distribution with given values of the mean and standard deviation $\mu_{x_i} = 0.4,\ \sigma_{x_i} = 0.16,\ i = 1, 2$ which is defined as (Hosder *et al.*, 2007) :

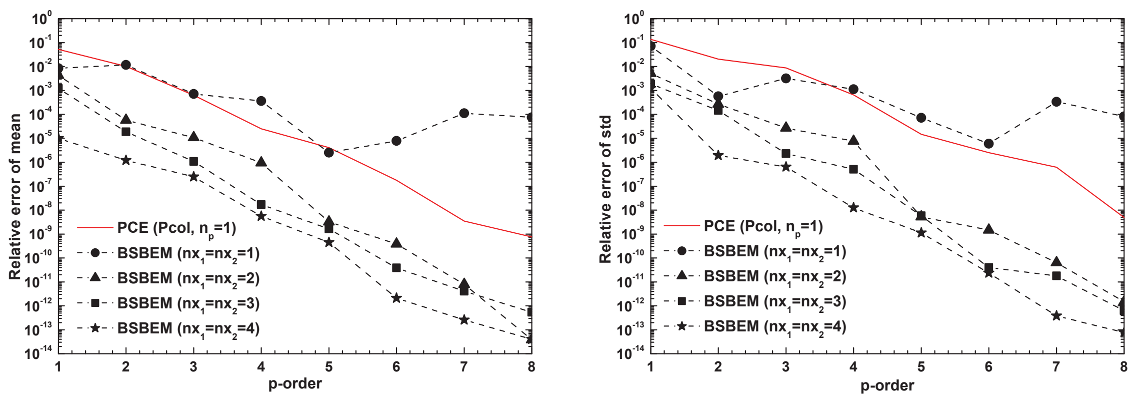
$$Y = f(x_1, x_2) = e^{(x_1 + x_2)} \tag{2.30}$$

The evolution of the relative error with respect to the polynomial order and to the number of Bézier elements for the statistical moments (mean and standard deviation) of the output response is shown in Fig.2.2 and defined as :

$$Error(\Phi_Y) = \frac{|\Phi_Y^{surrogate} - \Phi_Y^{exact}|}{|\Phi_Y^{exact}|} \tag{2.31}$$

where Φ denotes either the mean μ_Y or the standard deviation σ_Y of the output response Y.
As can be seen from Fig.2.2, the increase in the polynomial order leads to a decrease in the relative error for the mean and the standard deviation for both the BSBE and the Pcol methods. In case of one single Bézier element in each dimension ($nx_1 = nx_2 = 1$), one can observe that the point collocation method is showing a better performance than the proposed BSBE approach with less function evaluations, particularly for relatively high polynomial order, as it is clearly illustrated in Fig.2.2 (curve with filled circles) for both statistical moments. However, the

increase of the number of Bézier elements using the BSBE approach contributes to a significant improvement in the convergence of the statistical moments compared to the polynomial chaos point collocation method. This convergence improvement related to the increase of the number of Bézier elements is accompanied by an increase of the total number of collocation points and hence requires more computational effort (more function evaluations).



a) Mean b) Std

Figure 2.2 Relative error of the mean and standard deviation (std) obtained with the BSBE and Pcol methods with respect to polynomial order for different values of Bézier elements (with no oversampling).

To better illustrate the performance of the BSBE expansion, the convergence curves of the statistical moments are compared to the Point collocation (Pcol) and Monte Carlo methods for the same operating conditions (same number of function evaluations) as shown in Fig.2.3. Once the polynomial order is fixed, the number of samples used for each method (Pcol and BSBEM) is adjusted with that of the Monte Carlo method (MC) by varying the number of Bézier elements in both directions (nx_1 and nx_2) for the BSBE method ($N_s = nx_1 \times nx_2 \times (p+1)^2$), and by choosing the oversampling factor n_p in the case of Pcol method, as given by Eq.(2.4). The comparison of the mean and the standard deviation relative error profiles between the three methods clearly shows a large difference in the convergence rates. The Monte Carlo method is known to have a slow convergence rate, on the order of $(\sqrt{N_s})$. It is observed that by increasing the number of samples, the BSBE technique presents a better convergence rate compared to the

Point collocation expansion and the Monte Carlo methods. The effect of the polynomial order (p in the figure) on the convergence rate can also be analysed by examining Fig.2.3. The increase of the polynomial order leads to a significant accuracy improvement for the BSBE expansion compared to the point collocation method.

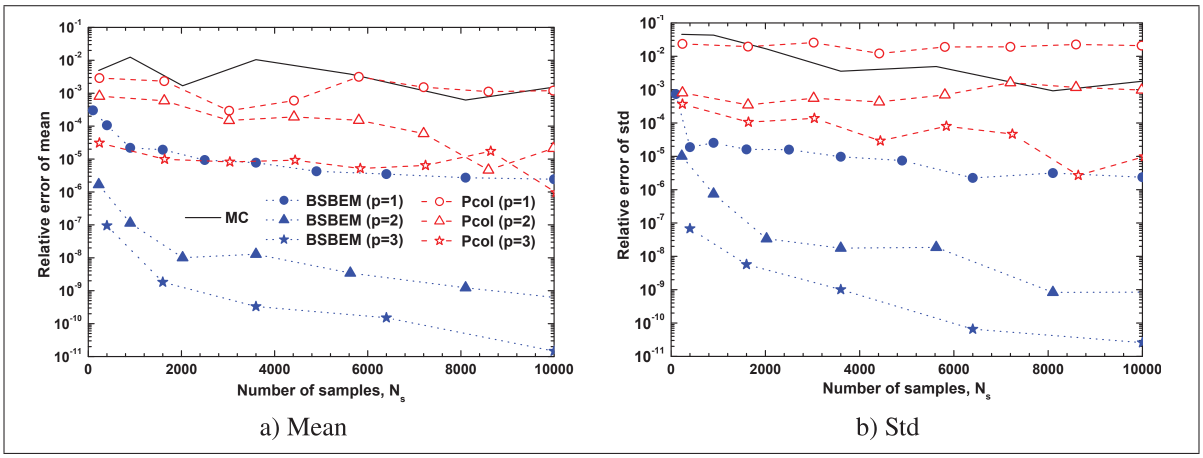
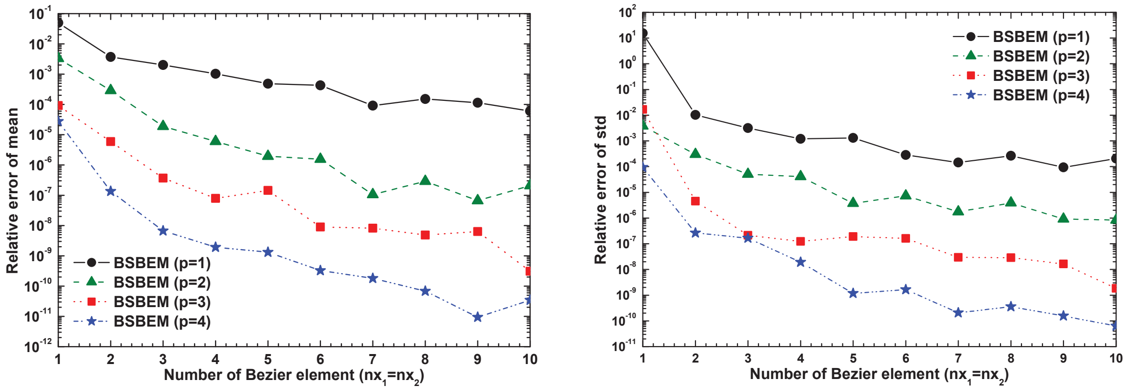


Figure 2.3 Relative error of the mean and the standard deviation (Std) obtained with the BSBE, Pcol and MC methods with respect to the number of samples.

Similarly, the convergence of relative error of the BSBE method with the increase of the number of Bézier elements for different polynomial order is shown in Fig.2.4, and indicates that the convergence of the method is improved considerably by increasing the number of Bézier elements. This improvement is more noticeable for relatively high polynomial order. As mentioned previously, the increase in the number of Bézier elements leads to an increase in the number of collocation points as well as to an optimal distribution of the random variables, which considerably improves the convergence of the statistical moments. This occurs because the coefficients of the BSBE expansion are evaluated accurately with a better use of the stochastic information by an increase in and a more optimal distribution of collocation points.

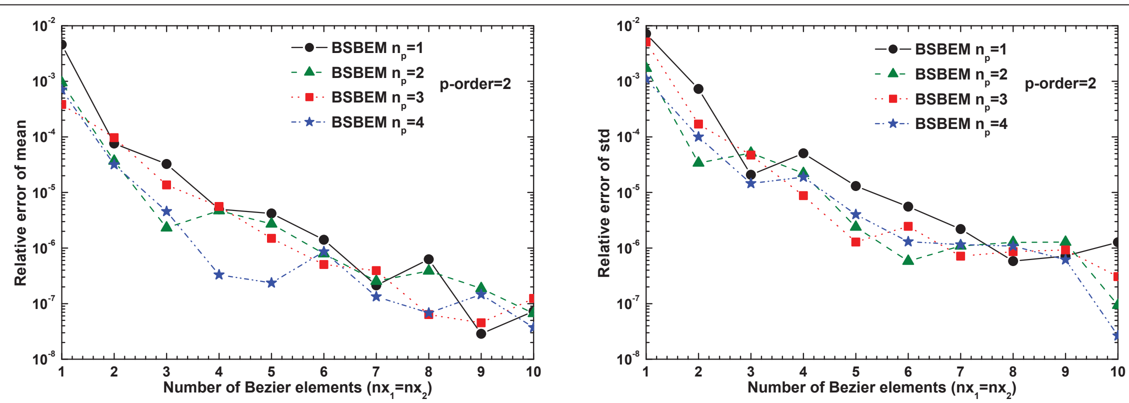
Fig.2.5 shows the convergence of the mean and standard deviation of the BSBE method with respect to the number of Bézier elements for various values of the oversampling ratio. In each



a) Mean b) Std

Figure 2.4 Relative error of the mean and standard deviation (std) obtained with the BSBE method with respect to the number of Bézier elements for different values of polynomial order.

Bézier element, an oversampling is applied by multiplying the minimum required number of collocation points with an oversampling coefficient. This figure shows that the oversampling does not contribute to an improvement of the convergence rate for the statistic moments, whereas the computational effort increases considerably.



a) Mean b) Std

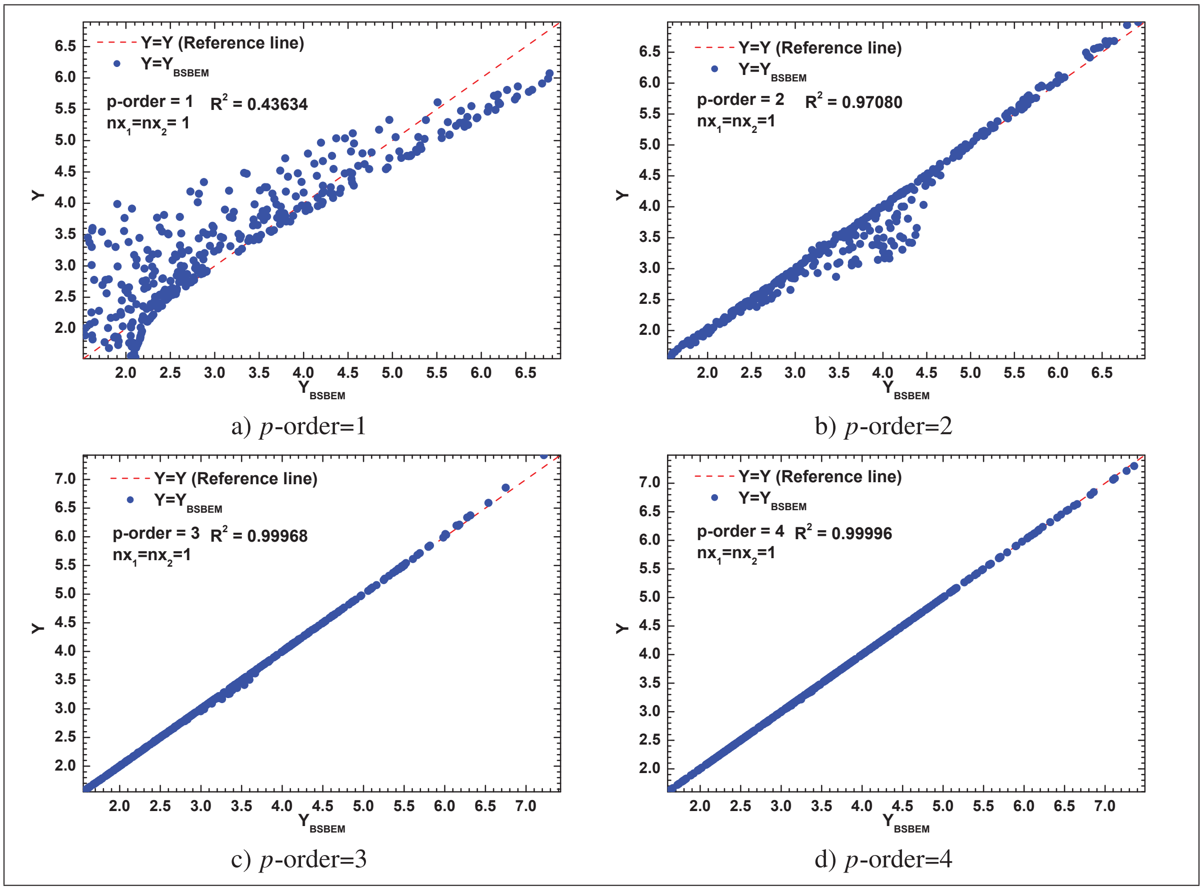
Figure 2.5 Relative error of the mean and standard deviation (std) obtained with the BSBE method with respect to the number of Bézier elements for different oversampling factor values.

In order to verify the accuracy of the established non-intrusive BSBE approach, scatter plots for the results obtained by the exact solution and the BSBE method are displayed in Figs.2.6 and 2.7 for various values of the polynomial order and of the Bézier elements' number, respectively, in which the determination coefficient R^2 and the line $(y = x)$ are also presented for reference. Theses plots depict 400 validation runs that are different than the experimental design points used for the construction of the BSBE surrogate model. Fig.2.6 shows that for low polynomial order values $(p = 1)$, a poor match is observed between the BSBE model and the exact solution, which is also illustrated by a small value of the determination coefficient $(R^2 \approx 0.436)$. The discrepancy between the exact solution and the BSBE model decreases significantly by increasing the polynomial order $(p = 4)$. As we can see, a good match is observed between the two models where all the validation points are superposed on the reference line $(y = x)$, which is also confirmed by the value of the determination coefficient $(R^2 \approx 0.99996)$.

Fig.2.7 shows the effect of the number of the Bézier elements by considering the low value of the polynomial order $(p = 1)$, which is considered as the worst case as stated above. A poor match is obtained between the BSBE model and the analytical solution for low values of the number of the Bézier elements $(nx_1 = nx_2 = 1)$. As we can see, the deviation between the two models decreases with the increase of the number of Bézier elements in the two directions (for each random variable) where an excellent match is observed $(nx_1 = nx_2 = 4)$. The high reliability level of the developed model is clearly illustrated by the good agreement between the validation points and the reference line $(y = x)$ and also by the value of the coefficient of determination $(R^2 \approx 0.99993)$. This improvement can be attributed to the increase of the number of the collocation points and their efficient distribution over the random variables' domain by increasing the number of the Bézier elements.

2.4.2 Ishigami function test case

As a second validation test case, we consider the so-called Ishigami function that is widely used in the literature as a benchmark reference solution for uncertainty quantification. This function, which is known to be highly non-linear and non-monotonic, is defined in its standard form with



a) p-order=1

b) p-order=2

c) p-order=3

d) p-order=4

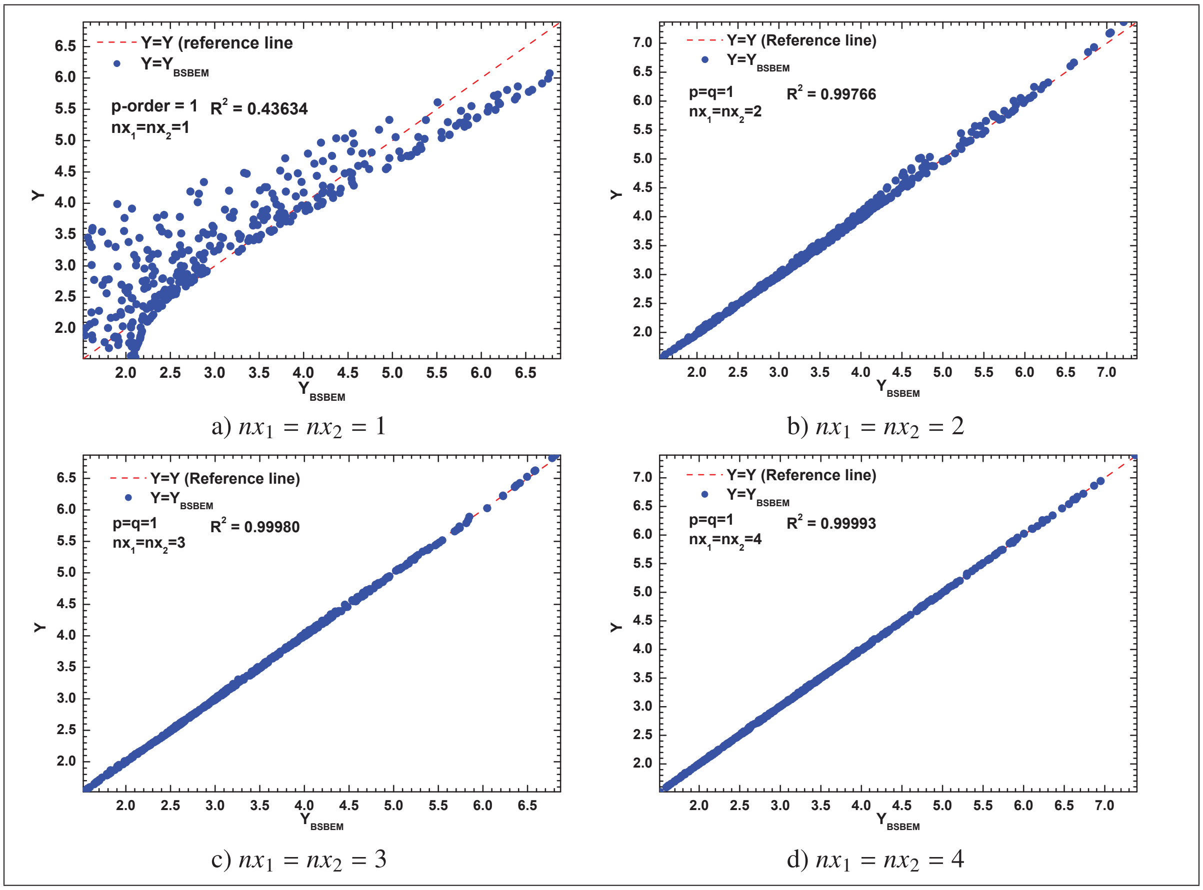
Figure 2.6 Comparison of exact solution results with BSBEM-predicted results on 400 validation points for various polynomial orders.

three random variables as :

$$Y = f(x_1, x_2, x_3) = \sin x_1 + 7\sin^2 x_2 + 0.1 x_3^4 \sin x_1 \tag{2.32}$$

where $x_i (i = 1, \ldots, 3)$ are independent random input variables following a uniform distribution over the domain $[-\pi, \pi]$ with the marginal density probability function $\varrho(x_i) = \frac{1}{2\pi}$.

The performance of the proposed BSBE method is assessed by estimating the mean and the standard deviation of the Ishigami function, which are compared to those of both the stochastic multi-elements approach using piece-wise Lagrange polynomials as multidimensional basis



a) $nx_1 = nx_2 = 1$

b) $nx_1 = nx_2 = 2$

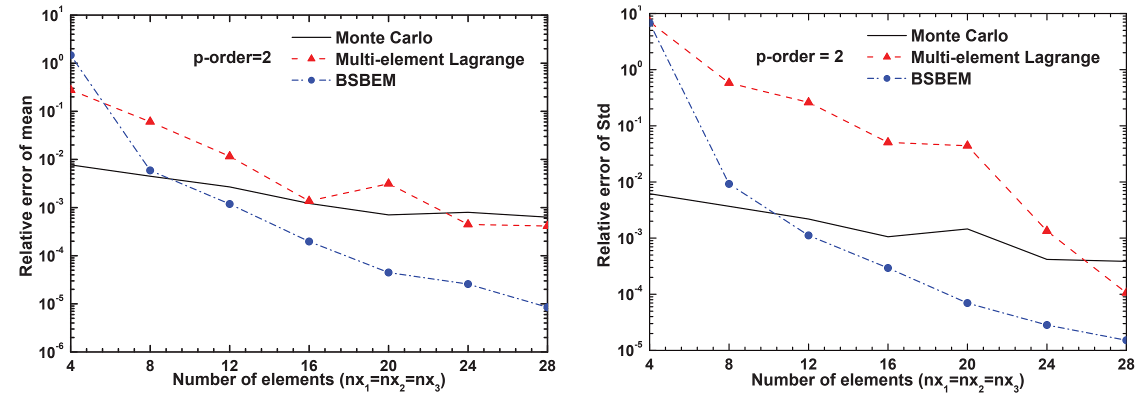
c) $nx_1 = nx_2 = 3$

d) $nx_1 = nx_2 = 4$

Figure 2.7 Comparison of exact solution results with BSBEM-predicted results on 400 validation points for various number of Bézier elements .

functions and the Monte Carlo method by evaluating the relative error of the statistical moments as given by Eq.(2.31). The convergence plots of the mean and the standard deviation of the Ishigami function with respect to the number of elements in each direction (each random variable) are shown in Fig.2.8. The total number of elements is computed by $N_{elt} = \prod_{i=1}^{m} nx_i$, where $m = 3$ denotes the number of random variables. For each value of the number of elements, the corresponding number of samples for which the Monte Carlo results are obtained is given by $N_s = \prod_{i=1}^{m} nx_i(p_i + 1)$, which also represents the calculation effort of the other stochastic expansion approaches.

The results depicted in Fig.2.8 show that the BSBE approach yields better and monotonic convergence profiles of the mean and the standard deviation compared to the Lagrange approach



a) Mean b) Std

Figure 2.8 Relative error of the mean and standard deviation (std) for the Ishigami function, obtained with the BSBE, Lagrange and MC methods with respect to the number of elements in each direction.

and the Monte Carlo method. One can see from these results that the accuracy of the statistical moments' estimates obtained from the BSBE approach increases with respect to the number of Bézier elements. For more comparison details, the numerical values of the mean and the standard deviation of the Ishigami function obtained from the BSBE and Lagrange methods, compared to those of the exact values, are presented in Table 2.1.

Tableau 2.1 Results of the first two statistical moments for the Ishigami function using the BSBE and the Lagrange methods for p-order=2.

nx_i	N_{elt}	N_s	Mean		Std	
			BSBEM	**Lagrange**	**BSBEM**	**Lagrange**
4	64	1728	2.034089	3.223785	10.518115	10.822931
8	512	13824	3.503780	3.560703	3.712677	4.301377
12	1728	46656	3.501187	3.511597	3.719715	3.981993
16	4096	110592	3.500197	3.501374	3.720537	3.771450
20	8000	216000	3.500044	3.496868	3.720761	3.764918
24	13824	373248	3.500025	3.500447	3.720803	3.722173
28	21952	592704	3.500008	3.499587	3.720816	3.720725
Monte Carlo		10^7	3.498761		3.718172	
Exact solution			3.5000		3.7208	

2.4.3 1D dam break flow test cases

2.4.3.1 Description of the test cases

The physical domain of the dam break flow test case consists of a channel with length 100 m and width 2 m. The channel is separated with a vertical gate located at the center ($x = 50$ m) as shown in Fig.2.9. The initial water depth of the upstream reservoir is 10 m. The downstream part of the channel has initial water depths of 1 m and 0 m, representing the wet and the dry beds, respectively. The deterministic model used to describe the dynamics of the free surface flow is given by the shallow water equations which are solved using a finite volume code (Zokagoa & Soulaïmani, 2010,1) :

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{H}(\mathbf{U})}{\partial y} = \mathbf{S}(\mathbf{U}) \tag{2.33}$$

with :

$$\mathbf{U} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \ \mathbf{G}(\mathbf{U}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix}, \ \mathbf{H}(\mathbf{U}) = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}, \ \mathbf{S}(\mathbf{U}) = \begin{bmatrix} 0 \\ gh(S_{0_x} - S_{f_x}) \\ gh(S_{0_y} - S_{f_y}) \end{bmatrix}$$

and :

$$S_{0_x} = -\frac{\partial z}{\partial x}, \ S_{0_y} = -\frac{\partial z}{\partial y}, \ S_{f_x} = \frac{n^2 u\sqrt{u^2+v^2}}{h^{4/3}}, \ S_{f_y} = \frac{n^2 v\sqrt{u^2+v^2}}{h^{4/3}}$$

where u and v denote the depth-averaged velocity components in the x and y directions, respectively. h represents the water depth, which is the difference between the free surface η and the bottom level z as given by $(h = \eta - z)$. n and $\mathbf{S}_f$ denote the Manning roughness coefficient and the friction vector, respectively.

The stochastic analysis was conducted with various input parameters, such as the Manning roughness coefficient and the initial upstream water level, which are considered to be uncertain. Several cases were evaluated to quantify the propagation of the uncertainties of the input parameters through the numerical and analytical models of the shallow water equations that govern the dam break flow.

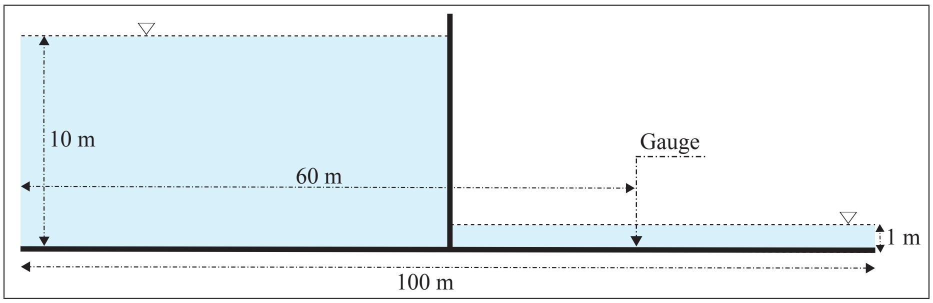


Figure 2.9 Sketch of dam break flow through a wet flat bottom, indicating measurement position.

2.4.3.2 Analytical solutions

Analytical solutions for idealized 1D dam break flows were proposed in the literature for the case of a flat bottom bed with no friction. The first case considered in this study concerns the dam break on an initially partly dry bed (Ritters's solution), and the second case describes the analytical solution of dam break flows on a wet domain (Stoker's solution). The analytical solutions are described in detail in (Delestre, Lucas, Ksinant, Darboux, Laguerre, Vo, James & Cordier, 2013). To demonstrate the effectiveness of BSBEM for physical problems involving strong hyperbolic behaviour, the aforementioned analytical solutions were considered first as deterministic models through which uncertainties are propagated. In this case, only the initial upstream water level is considered as a single uncertain input parameter following a normal distribution with a mean value of 10 m and two values of the coefficient of variation with 5 % and 10 %, which yields to standard variation values of $1m$ and $0.5m$, respectively.

The first analytical solution test case concerns Ritter's solution of dam break flow on a dry bed ($h_{ds} = 0\ m$) with the value of the coefficient of variation of 10 % ($Std = 1\ m$). The results are reported in Fig.2.10 in terms of the 95 % confidence interval ($mean \mp 2std$) (Fig.2.10b) and the standard deviation (Fig.2.10d) profiles of the water depth as a function of the position ($0 \leqslant x \leqslant 100$) at t = 4 s. The statistical moment's profiles of the output quantity of interest obtained by the Monte Carlo method with 10^6 realizations are considered as a reference for comparison. Ritter's solution presents an interesting case with which to describe the hyperbolic behaviour of the expansion wave with a relatively smooth transition of the water depth over the calculation domain. One can observe from these plots that the profiles obtained by the polynomial chaos-based point collocation method (Pcol) present some oscillations, which increase greatly by increasing the polynomial order (from $p = 3$ to $p = 7$) despite the use of an oversampling factor ($n_p = 7$). On the other hand, the results obtained by the proposed BSBEM, with polynomial order value ($p = 3$) and number of Bézier elements ($nx = 15$), present an excellent match with those of the reference Monte Carlo solution as shown by the close up views of the selected zones in the lower bound of the confidence interval (Fig.2.10a) and the standard deviation profiles (Fig.2.10c).

The second analytical test case is devoted to the Stoker's solution of a dam break flow over a wet bed with an initial downstream water depth ($h_{ds} = 1\ m$), which is considered as a deterministic model. The uncertain initial upstream water depth follows a normal distribution with a mean value of 10 m and a coefficient of variation of 5 % ($Std = 0.5\ m$). As for the Ritter's solution, the results are presented in terms of the evolution of the confidence interval and the standard deviation profiles as a function of the channel length at t=4 s. It is well known that such a deterministic solution presents discontinuous outputs with the appearance of a shock in the downstream part of the channel. Fig.2.11 shows the evolution of the statistical moment's profiles over the channel length obtained by BSBEM, which are compared to those of the Pcol expansion and Monte Carlo reference solution with 10^6 realizations. The close up views of the upper bound of the confidence interval (Fig.2.11a) and the standard deviation (Fig.2.11c) profiles in the vicinity of the discontinuity show notable discrepancies in the Pcol results with a strong oscillatory behaviour, particularly for relatively high polynomial order. The results obtained with BSBEM are plotted for ($p = 2, n_p = 2$) and different values of Bézier element numbers. Small discrepancies with slight oscillations in the statistical moments profiles for $nx = 10$ can be

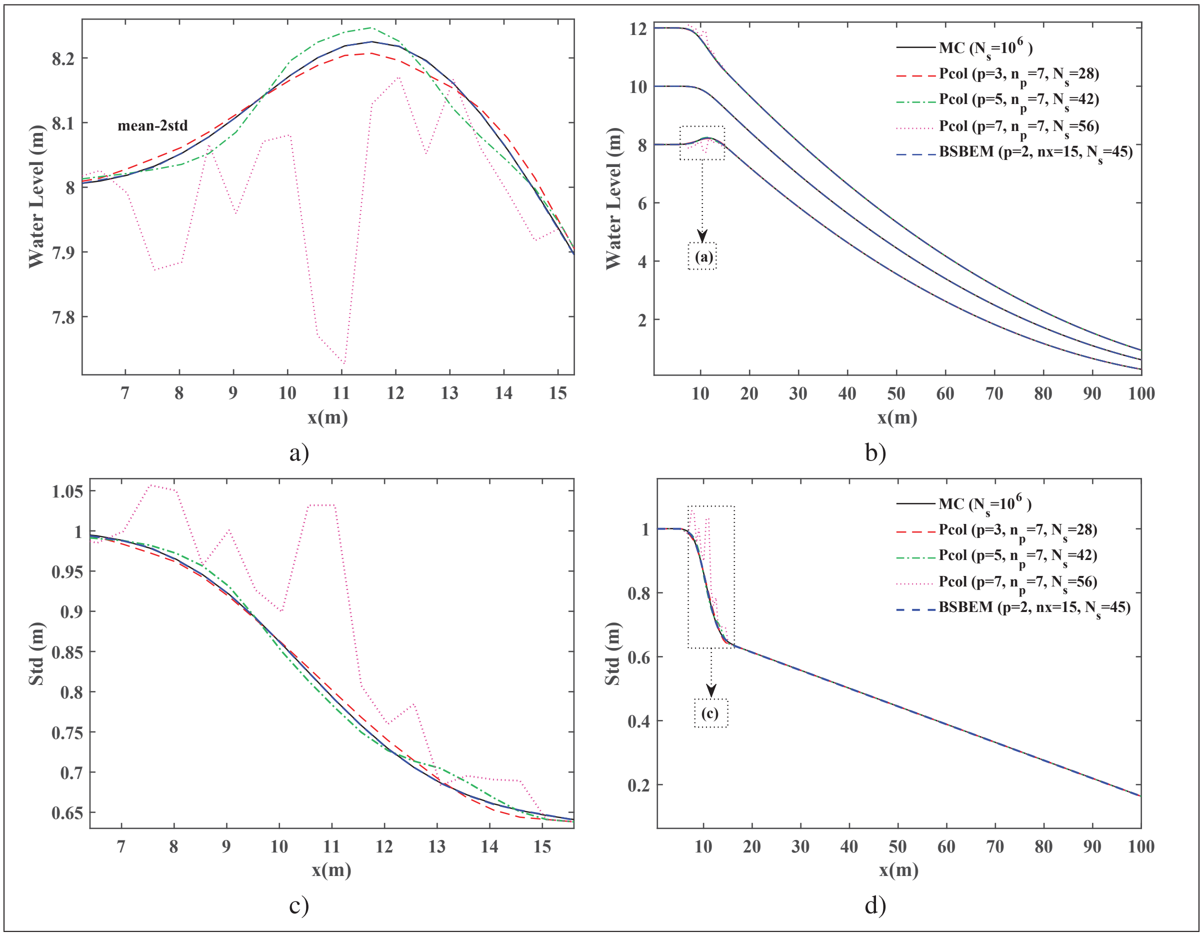


Figure 2.10 Evolution of the 95% confidence interval (b) and standard deviation (d) of water depth over the channel length at t=4 s (Ritter's solution) obtained from various p-order Pcol, MC (with $N_s = 10^6$ realizations) and BSBE methods ($p = 2$, $n_p = 1$, $nx = 15$, $N_s = 45$) : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

observed. These oscillations decrease significantly by increasing the number of Bézier elements ($nx = 25$), where a good match of the profiles is observed in comparison with those of the Monte Carlo reference solution, as shown by a detailed examination of the figures.

2.4.3.3 Numerical solutions

In addition to the aforementioned analytical solutions where the effectiveness of BSBEM has been well assessed for hyperbolic problems at fixed times, we next consider the time dependent problems combined with the output discontinuous solutions. It is known that the

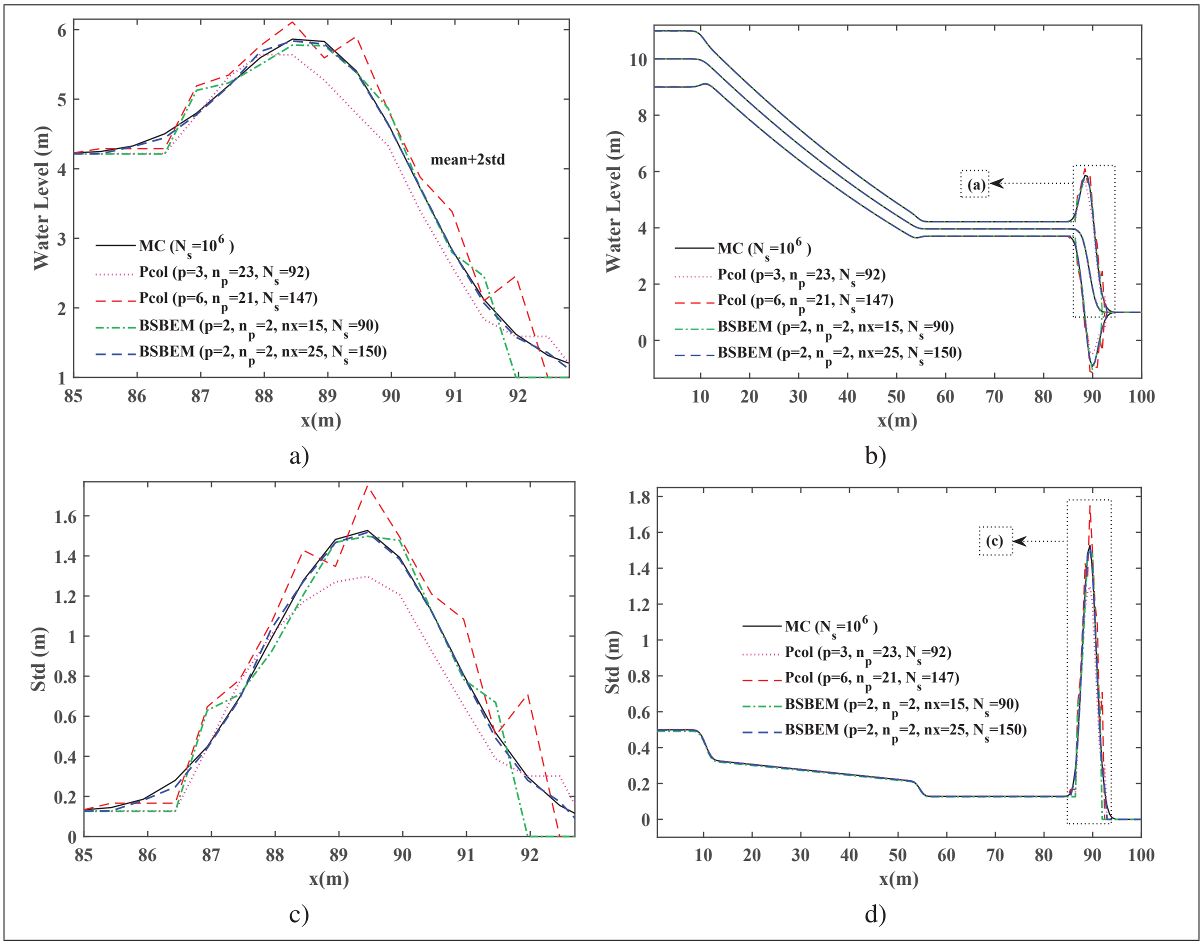


Figure 2.11 Evolution of the 95% confidence interval (b) and standard deviation (d) of water depth over the channel length at t=4 s (Stoker's solution) obtained from various p-order Pcol, MC (with $N_s = 10^6$ realizations) and BSBE methods ($p = 2$, $n_p = 1$, $nx = 15,\ 25$, $N_s = 90,\ 150$) : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

use of polynomial chaos has some limitations in dealing with uncertainties propagation in such cases (Prempraneerach, Hover, Triantafyllou & Karniadakis, 2010 ; Wan & Karniadakis, 2005). The dam break flow is modelled by a numerical deterministic model (an in-house finite volume code) with a wall boundary condition at the end of the channel. A measurement gauge is set at a distance of 60 m from the origin to measure the variations of the water level as a function of time as shown in Fig.2.9.

Case of one uncertain parameter

In this case, as for the analytical solutions cases, only the initial upstream water level is considered, as a single uncertain input parameter following a normal distribution with a mean value of 10 m and a coefficient of variation of 10 % ($Std = 1\ m$). The other input parameters are kept constant with fixed values of the Manning roughness coefficient ($n = 0$) and the initial downstream water level $h_{ds} = 1\ m$. The obtained results are shown in terms of the mean, the standard deviation and the 95 % confidence interval ($mean \mp 2std$) profiles of the water depth as a function of time at the gauging measurement point. The results obtained by the Monte Carlo method with 5000 realizations are considered as reference results with which to assess the accuracy of the BSBE approach. Furthermore, a comparison with polynomial chaos expansions such as the Probabilistic Collocation Method (PCM) and Point collocation method (Pcol) are also conducted at the same operating conditions.

Fig.2.12 shows the mean and the confidence interval (Fig.2.12b) and the standard deviation (Fig.2.12d) profiles of the water depth at the gauging point obtained by the probabilistic collocation method. As mentioned above, the PCM is based on the fact that collocation points in the parametric domain are selected as the roots of the higher-order polynomials, which are ranked decreasingly from the highest to the lowest probability (Tatang *et al.*, 1997). The resulting profiles of the probabilistic collocation approach for a three-polynomial order ($p = 5$, 10 and 20) are compared to the MC reference solution. It is obvious that severe oscillations appear in the statistics profiles (interval bounds and standard deviations) of the water depth around the discontinuity. Only a poor improvement in the smoothness of the profiles is observed by increasing the polynomial order, as shown by the close-up view of the selected zones of the confidence interval (Fig.2.12a) and standard deviation (Fig.2.12c) plots.

The same oscillatory behaviour can be observed with the polynomial chaos-based point collocation method (Pcol) results, as shown by the evolution of the statistical moments' profiles

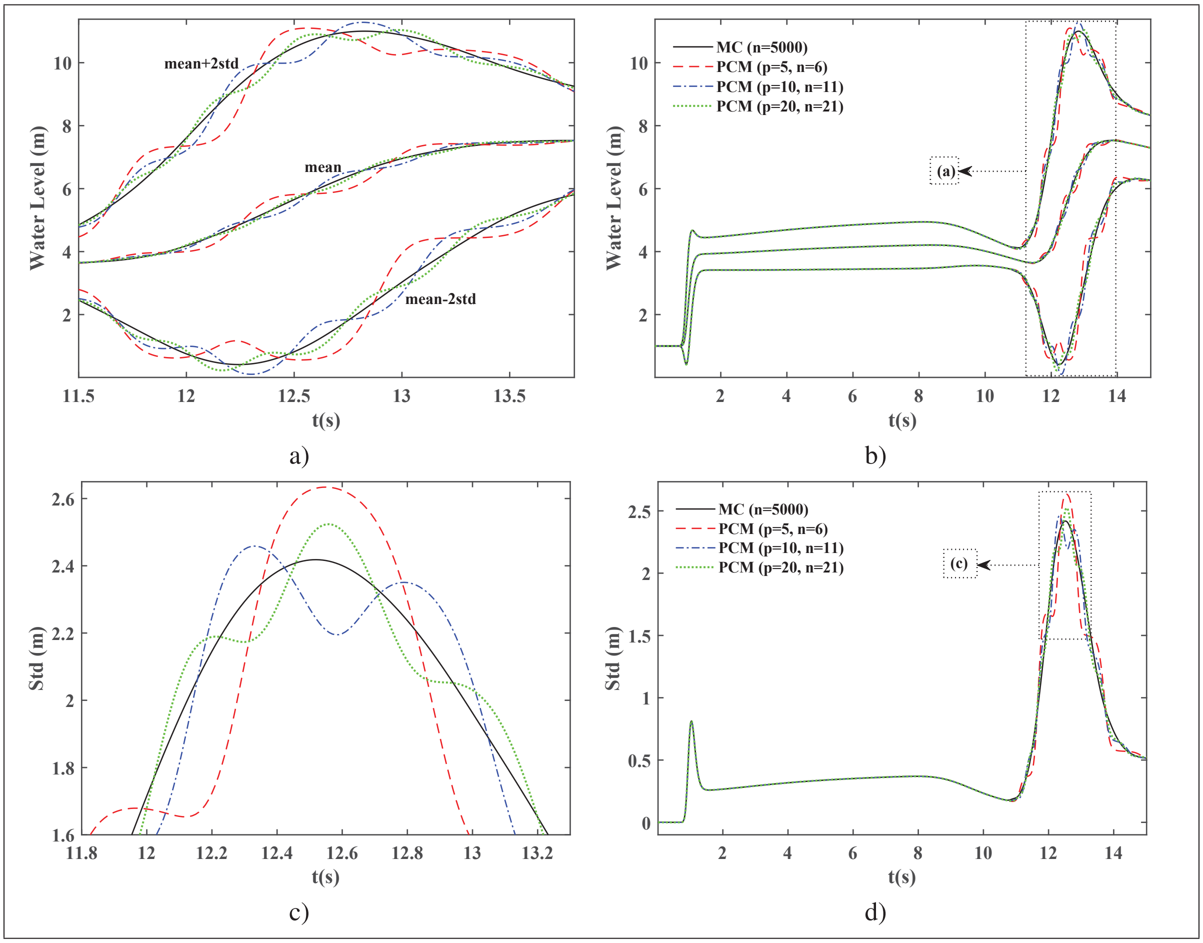


Figure 2.12 Evolution of 95% confidence interval (b) and standard deviation (d) of water depth at the gauging point obtained with various p-order PCM and MC (with N_s = 5000 realizations) solutions : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

plotted in Fig.2.13. In this case, in addition to varying the p-order, the oversampling ratio (n_p), which is considered as a useful means to improve the accuracy of the Pcol method, is also varied in order to increase the number of collocation points (Hosder & Walters, 2010). The results are compared to those obtained by the Monte Carlo method; one can clearly see the apparition of the spurious oscillations on the interval bounds and the standard deviation profiles. These oscillations tend to be intensified by increasing the p-order despite the augmentation of the oversampling ratio. A close-up view of the mean (Fig.2.13a) and the standard deviation (Fig.2.13c) shows that a slight improvement can be observed for small p-order values, whereas the oscillations remain visible, which illustrates a mismatch between the Pcol approach curves and those obtained by the reference MC method.

As illustrated in the aforementioned plots, the use of the orthogonal polynomial basis (polynomial

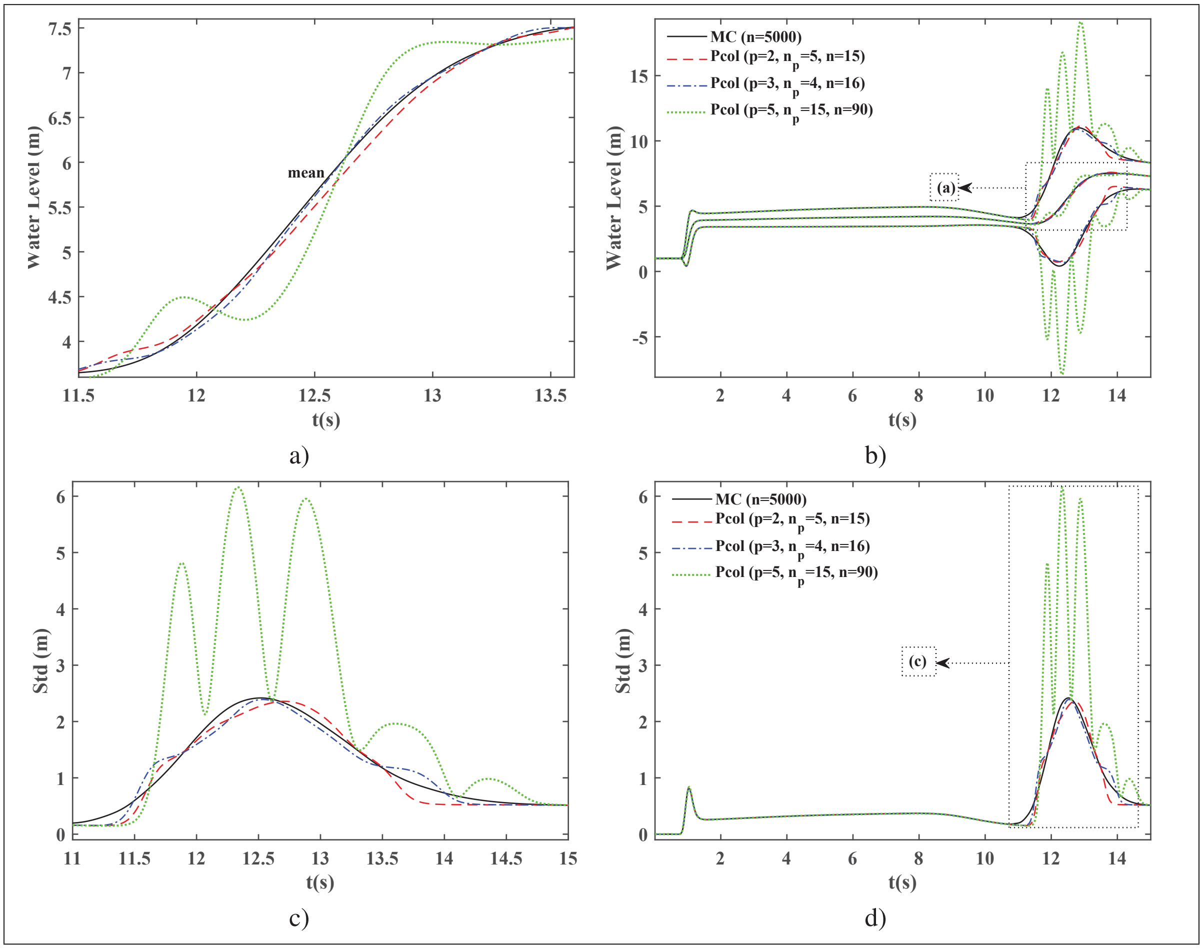


Figure 2.13 Evolution of 95% confidence interval (b) and standard deviation (d) of water depth at the gauging point obtained with various order Pcol and MC methods (with $N_s = 5000$ realizations) : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

chaos) for approximating the stochastic output response of the model presents some limitations with hyperbolic behaviour. To overcome this drawback, the non-intrusive B-Spline Bézier elements-based method (BSBEM) is introduced. Fig.2.14 shows the evolution of the mean and standard deviation profiles over time as obtained by the BSBE method. The results are plotted for p-order value ($p = 2$) and for a various number of Bézier elements ($nx = 10, 15$ and 25).

As can be seen from the close-up views of the figures (Fig.2.14a) and (2.14c), slight oscillations appear on both statistical moments' profiles when 10 Bézier elements are deployed. In contrast to those observed with the polynomial chaos approaches, these oscillations diminish significantly with a higher number of Bézier elements. Indeed, for a relatively high value of the number of Bézier elements ($nx = 25$), the results indicate that the BSBE method can provide an excellent estimation of the statistical moments with a high degree of accuracy. Detailed examination of the curves shows a very good match of the profiles with those obtained by the Monte Carlo reference method.

To further investigate the performance of the BSBE expansion, a comparison of the profiles of the confidence interval and the standard deviation was carried out with different methods (LHS and Pcol) under the same operating conditions (number of calls of the deterministic model), as illustrated in Fig.2.15. The number of realizations (number of samples) is set to ($n_s = 75$) for the LHS method. The same sample size (number of collocation points) is used for the two stochastic expansion methods with a polynomial order of 2 by modifying the oversampling ratio ($n_p = 25$) for the Pcol method, and the number of Bézier elements ($nx = 25$) for the BSBE method. Since the computational efforts of these three approaches remain the same, the comparison will be made on the capacity of these methods to approach the Monte Carlo reference solution. As can be seen from the close-up views of the confidence interval (Fig.2.15a) and standard deviation (Fig.2.15c) profiles, the results obtained by the BSBE approach match very well with those directly produced by the classical Monte Carlo method, in contrast with the Pcol and LHS methods. These results indicate that the non-intrusive BSBE approach can provide a better prediction of the statistical moments with smooth contours, especially in regions surrounding the discontinuity, as clearly shown in Figs.(2.15a) and (2.15c).

Case of two uncertain parameters

The two-dimensional problem has two independent random input parameters where the initial upstream water level has a normal distribution with a mean ($\mu_h = 10\ m$) and a standard deviation

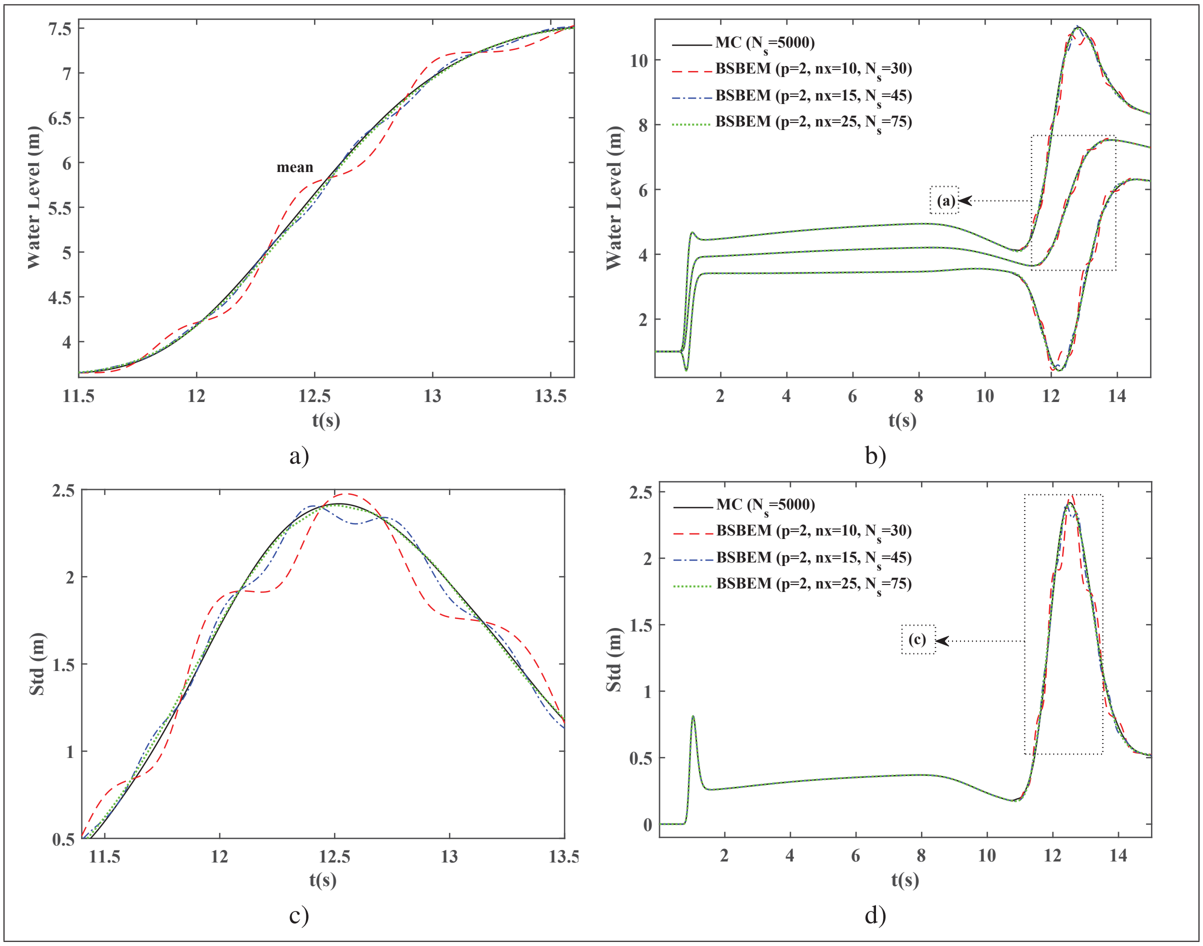


Figure 2.14 Evolution of the 95% confidence interval (b) and standard deviation (d) of water depth at the gauging point obtained with the BSBE method using various Bézier elements and the MC method (with N_s = 5000 realizations) : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

($\sigma_h = 1\ m$), whereas the Manning roughness coefficient has a uniform distribution with a mean ($\mu_n = 0.04$) and a standard deviation ($\sigma_n = 0.004$). The same configuration of the dam break test case, used previously, is again considered in this two-dimensional problem as shown in Fig.2.9. The input initial downstream water level is kept constant with a value of 1 m (wet bed). As in the one-dimensional case, the results are presented in terms of the confidence interval and standard deviation profiles of the water depth as a function of time at the gauging measurement point. The Monte Carlo solution with 10 000 realizations is considered here as a reference solution with which to compare the computational efforts and the output statistical moments' accuracy

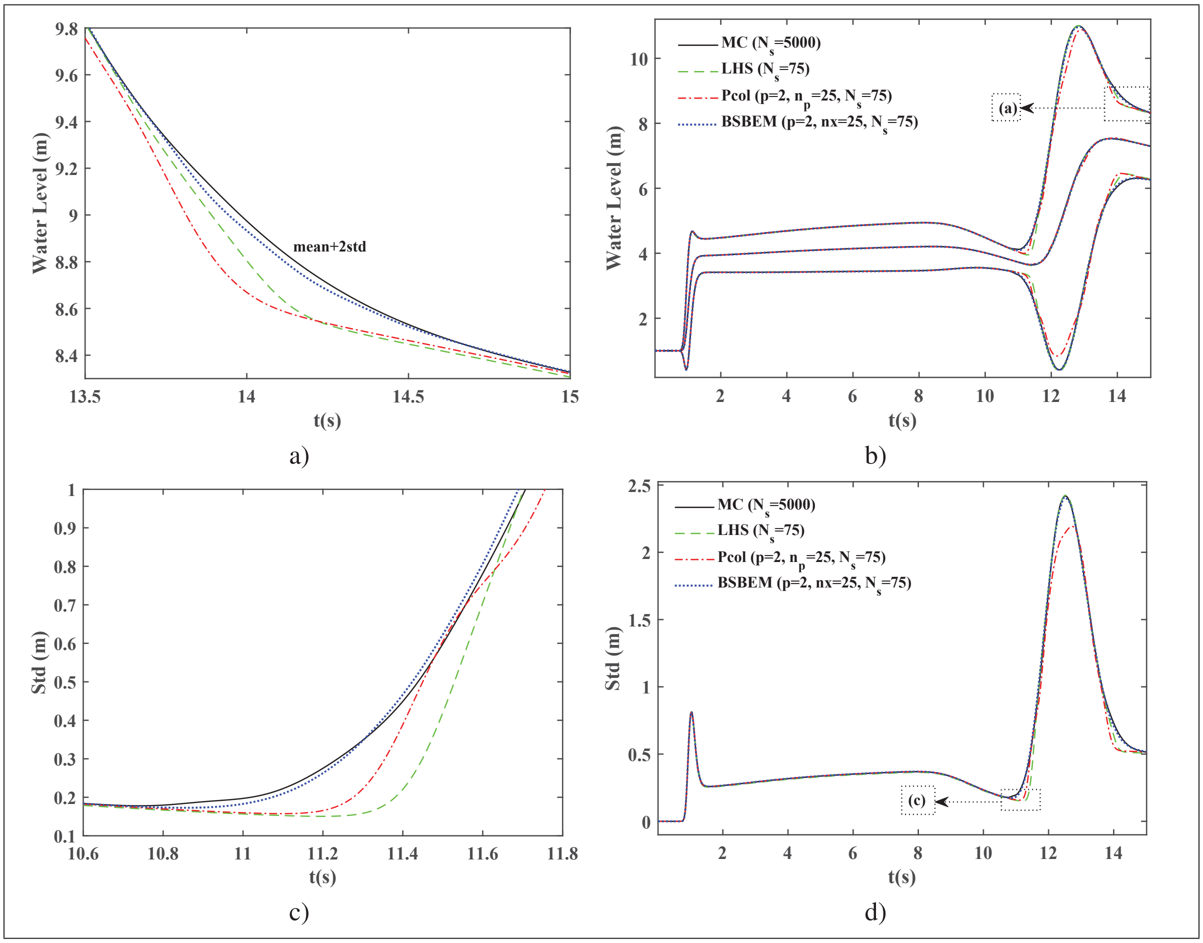


Figure 2.15 Evolution of the 95% confidence interval (b) and standard deviation (d) of water depth at the gauging point obtained from the MC (with $N_s = 5000$ realizations), LHS ($N_s = 75$), Pcol ($p = 2$, $n_p = 25$, $N_s = 75$) and BSBE methods ($p = 2$, $nx = 25$, $N_s = 75$) : (a) and (c) are zooms of the selected zones of (b) and (d), respectively.

between the non-intrusive BSBE approach and the polynomial chaos-based Point collocation method (Pcol).

In the present study, the approximation of the output response of the deterministic model (in this case, the water depth) by the Point collocation method is associated with the orthogonal Hermite basis functions, which are more accurate when the input random variable follows a Gaussian distribution. As mentioned earlier, the Manning roughness coefficient has a uniform distribution that does not belong to the Gaussian random field, and thus it has to be transformed into the

Gaussian random field by an appropriate mapping process, such as the Rosenblatt transformation, which is defined by the following relationship (Xiu, 2010) :

$$\xi_i = F^{-1}_{\mathcal{N}(0,1)}\left[F_x(x_i)\right] \tag{2.34}$$

where F_x is the cumulative density function of the uniform random variable x_i and $F_{\mathcal{N}(0,1)}$ is the marginal cumulative density function of the random variable (ξ_i), which follows a standard normal distribution with a zero mean value and unit standard deviation value.

Figs.2.16 and 2.17 show the comparison of the 95 % confidence interval and the standard deviation of the water depth obtained from both stochastic expansion schemes (BSBE and Pcol) whose results are compared to those of the reference Monte Carlo solution. In this work, and as previously mentioned, the number of Monte Carlo sampling is set at 10 000, which seems to be quite sufficient to achieve an accurate convergence of the statistical moments of the output response. The second polynomial order basis functions are used for both the BSBE and the Pcol expansions, whose number of collocation points is modified by acting, on the one hand, on the oversampling ratio in the case of the Pcol method and, on the other hand, on either the oversampling ratio or the number of Bézier elements in the case of the BSBE approach.

Figs.2.16a and 2.16b show the results obtained with both the BSBE and the Pcol expansions with 18 simulations. As seen from these figures, strong oscillations can be observed in the confidence interval and standard deviation profiles in the regions surrounding the discontinuity, and none of these profiles matches those achieved by the reference Monte Carlo approach, even in areas far from the discontinuity. When the number of collocation points is increased, Figs.(2.16c-2.16d) and (2.17a-2.17b) show that the oscillatory behaviour of the stochastic expansions' responses decreases significantly, although they still deviate from the reference Monte Carlo solution profiles, most precisely in the region surrounding the discontinuity. A notable improvement in the approximation of the output response of the water level, at the gauging point is obtained from the non-intrusive BSBE approach when the number of Bézier elements for both input random parameters is increased ($nx_1 = nx_2 = 5$), as shown on the statistical moments' profiles presented in Figs.2.17c and 2.17d. As can be seen from these figures, the confidence interval and the standard deviation profiles from the BSBE approach are in good agreement with those performed by the Monte Carlo method, unlike the Pcol expansion results where there are still deviations from the reference solution. This deviation becomes larger near the extrema of the standard deviation profile (discontinuity), as it is clearly visible on Fig.2.17d, which gives a good indication of the stability of the BSBE approach.

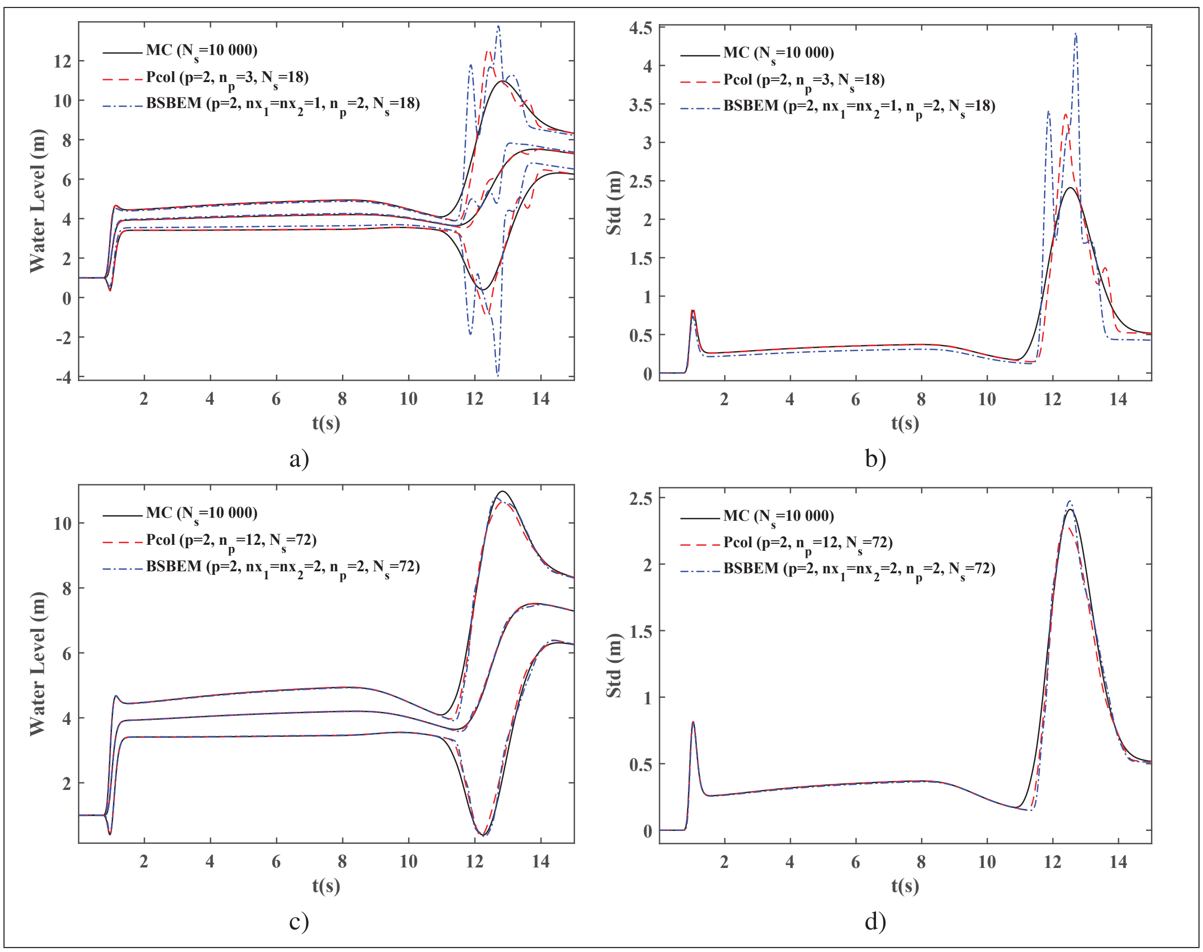


Figure 2.16 Comparison of confidence interval (95% CI) and standard deviation (Std) of the water depth at the gauging point, obtained from the MC (with $N_s = 10000$ realizations), Pcol and BSBE methods : (a) and (b) with $N_s = 18$ realizations ; (c) and (d) with $N_s = 72$ realizations.

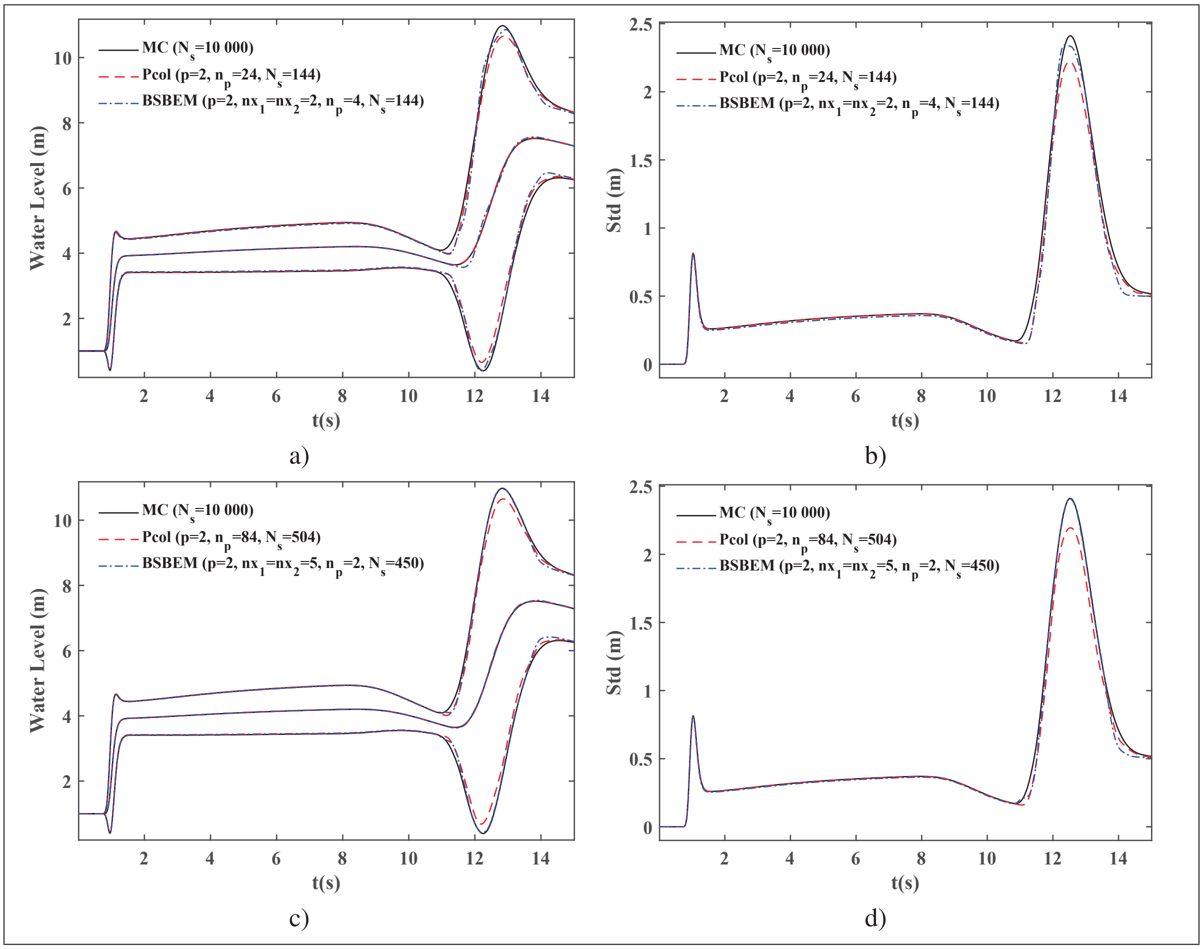


Figure 2.17 Comparison of the confidence interval (95% CI) and standard deviation (Std) of water depth at the gauging point, obtained from the MC (with N_s = 10000 realizations), Pcol and BSBE methods : (a) and (b) with N_s = 144 realizations ; (c) and (d) with N_s = 504 and N_s = 450 realizations for the Pcol and the BSBEM, respectively.

2.4.4 Case of dam break flow over a real topography

As a second numerical experiment, an actual study site (the Bordeaux breakwater) was considered. It covers an area of the Des prairies river, north of Montreal (Quebec) that is approximately 720 m long and 400 m wide. The topography and the mesh data were used in a deterministic in-house finite volume solver for shallow water equations (Zokagoa & Soulaïmani, 2010). In this example, we consider that the above deterministic model simulates a sudden dam break flow located inside the domain. The two uncertain input parameters considered are the upstream water level and the Manning roughness coefficient. These parameters are generated according to their prescribed probability density functions (PDF). The upstream water level η takes its values from a normal distribution with a mean ($\mu_\eta = 27\ m$), and the Manning roughness coefficient from a uniform distribution with a mean ($\mu_n = 0.04$). Various values of the coefficient of variation (c_v = 5 and 10%) are considered for each of the two uncertain parameters, thus allowing a wide range of variability to be covered. The water free surface downstream of the dam is considered to be initially constant with a value ($\eta_d = 17\ m$).

The results are presented and discussed in terms of the statistical moment's evolution of the water level at fixed gauging measurement points as a function of time, and also in terms of the contours of the standard deviation and the confidence interval over the whole domain. As reported in Fig.2.18, three measurement points (Gauge 1, Gauge 2 and Gauge 3) were installed at appropriate positions in order to quantify how the uncertainties related to the variation of the water elevation change with time.

Fig.2.19 shows the evolution of the confidence interval and the standard deviation of the water elevation as a function of time at the three gauging measurement points. As mentioned above, the number of the collocation points is adjusted to be approximately the same for both stochastic methods by varying the oversampling ratio and the number of Bézier elements. For a relatively small value of the number of collocation points ($N_s \approx 50$), achieved for a polynomial basis order of 3, one can see that the match between the statistical moments' profiles provided by the Pcol and BSBE methods compared to those from the 5000 realizations of the LHS simulation is quite poor. This deviation is more pronounced for the results obtained at gauging points 2 and 3, where spurious oscillations in the statistical moments' profiles can clearly be observed, as shown in Figs.(2.19c-2.19f). It must be emphasized that in this case the results of the BSBE expansion are performed with a single Bézier element in each direction (each random variable) and the total number of collocation points is increased through the oversampling ratio, which does not allow

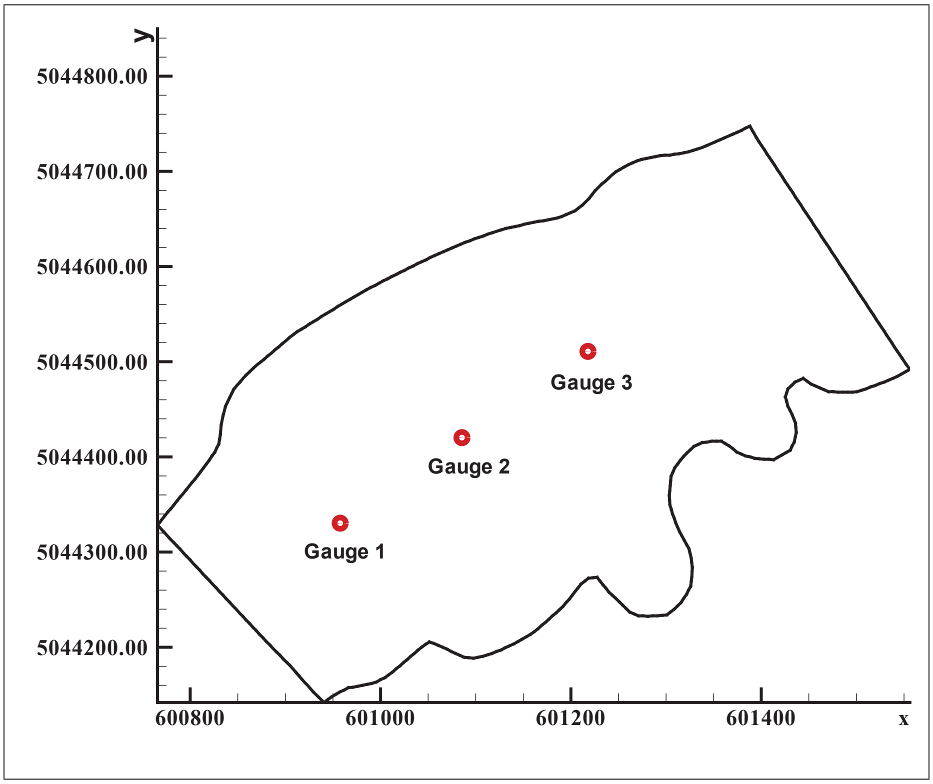
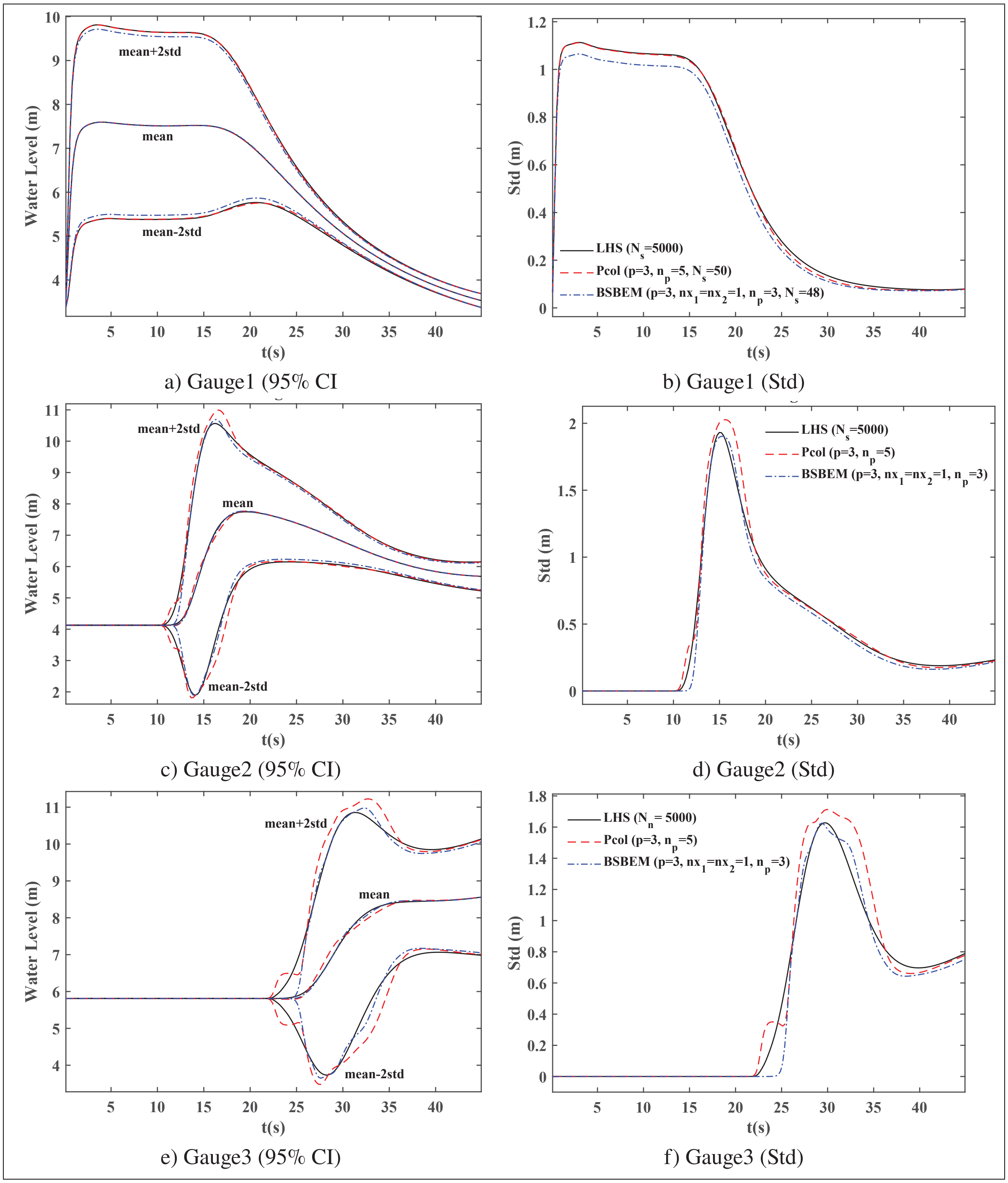


Figure 2.18 Sketch of the study area of the Bordeaux stockade and the gauging measurement points' positions.

the stochastic behaviour of the output response of the variable of interest to be well-reproduced.

To improve the ability of the BSBE method to better capture stochastic information, the number of collocation points is increased by varying the number of Bézier elements in the same way in both directions ($nx_1 = nx_2 = 7$), for a total of ($N_s = 441$) collocation points . The value of the oversampling ratio considered for the point collocation method is ($n_p = 74$), which yields a total number of collocation points relatively close to that of the BSBE method ($N_s = 444$). As in the former case, the evolution of the statistical moments' profiles as a function of time obtained from the BSBE and Pcol approaches are compared with those from 5000 realization of the LHS method, as shown in Fig.2.20. These results show that an obvious improvement is obtained for the confidence interval bounds and standard deviation profiles ; the oscillations have been greatly reduced, thus allowing to obtain smooth curves. Indeed, the increase in the



a) Gauge1 (95% CI b) Gauge1 (Std)

c) Gauge2 (95% CI) d) Gauge2 (Std)

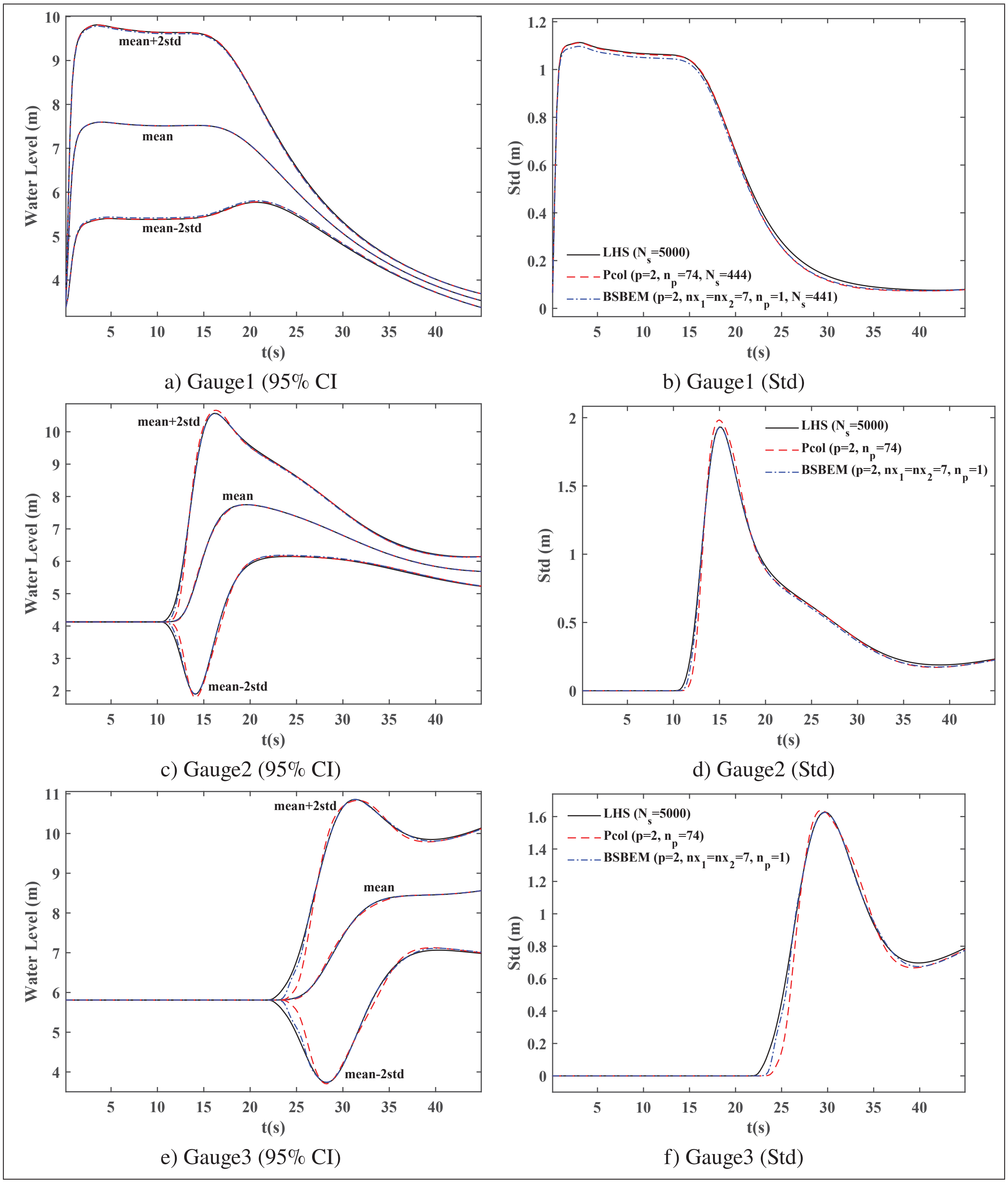
e) Gauge3 (95% CI) f) Gauge3 (Std)

Figure 2.19 Comparison of confidence interval and standard deviation of water depth at gauging points obtained from the LHS (with 5000 realizations), Pcol and BSBE methods (with 50 realizations).

number of Bézier elements of the BSBE expansion improves the prediction of the statistical moments of the output water level, which agree quite well with those produced by the LHS method.

In order to better understand how the propagated uncertainties are spatially distributed over the computational domain, the mean and the standard deviation are computed for each grid cell. Fig.2.21 displays the standard deviation contours of the output water elevation computed from the BSBE approach and those obtained from the reference LHS method for 15 s and 30 s. It can be observed from these figures that the standard deviation contours provided by the BSBE approach, with 225 realizations, are in good agreement with those obtained from 5000 realizations of the LHS method, and for both considered time simulations. The propagation of the wave front is clearly visible on the contour plots, where the standard deviation of the water depth reaches a maximum of 2 m in the areas surrounding the wave front over almost the entire domain width. For a simulation time of 30 s, it is clear that the inundation wave progresses downstream of the domain and where spatial distribution of the standard deviation of the water depth is much more widespread. Furthermore, one can also observe that the maximum values of the standard deviation (2.2 m) are located in areas near both the left and the right banks, as shown in Fig.2.21c and 2.21d. The above results from both approaches, which are similar, yield pertinent information on the areas where there is a potentially large uncertainty of the water depth.

For further details, a comparison of the obtained results is performed by superimposing the contours of both methods on the same plot for simulation times of 15 and 30 s. It is clear from Fig.2.22 that the results from the BSBE expansion are in good agreement with those from the 5000 realizations of the LHS simulation. These plots clearly show the distribution of the standard deviation contours of the water depth over the whole domain and its evolution as a function of time. The values displayed on the contours allow to precisely locate the areas with high variability stemming from the uncertainties of the input parameters. It is clear from these plots that the non-intrusive BSBE expansion yields satisfactory and accurate predictions of the statistical moments of the output response, even for such complex flow cases, and with



a) Gauge1 (95% CI

b) Gauge1 (Std)

c) Gauge2 (95% CI)

d) Gauge2 (Std)

e) Gauge3 (95% CI)

f) Gauge3 (Std)

Figure 2.20 Comparison of confidence interval and standard deviation of water depth at gauging points obtained from the LHS (with 5000 realizations), Pcol and BSBE methods (with 444 realizations).

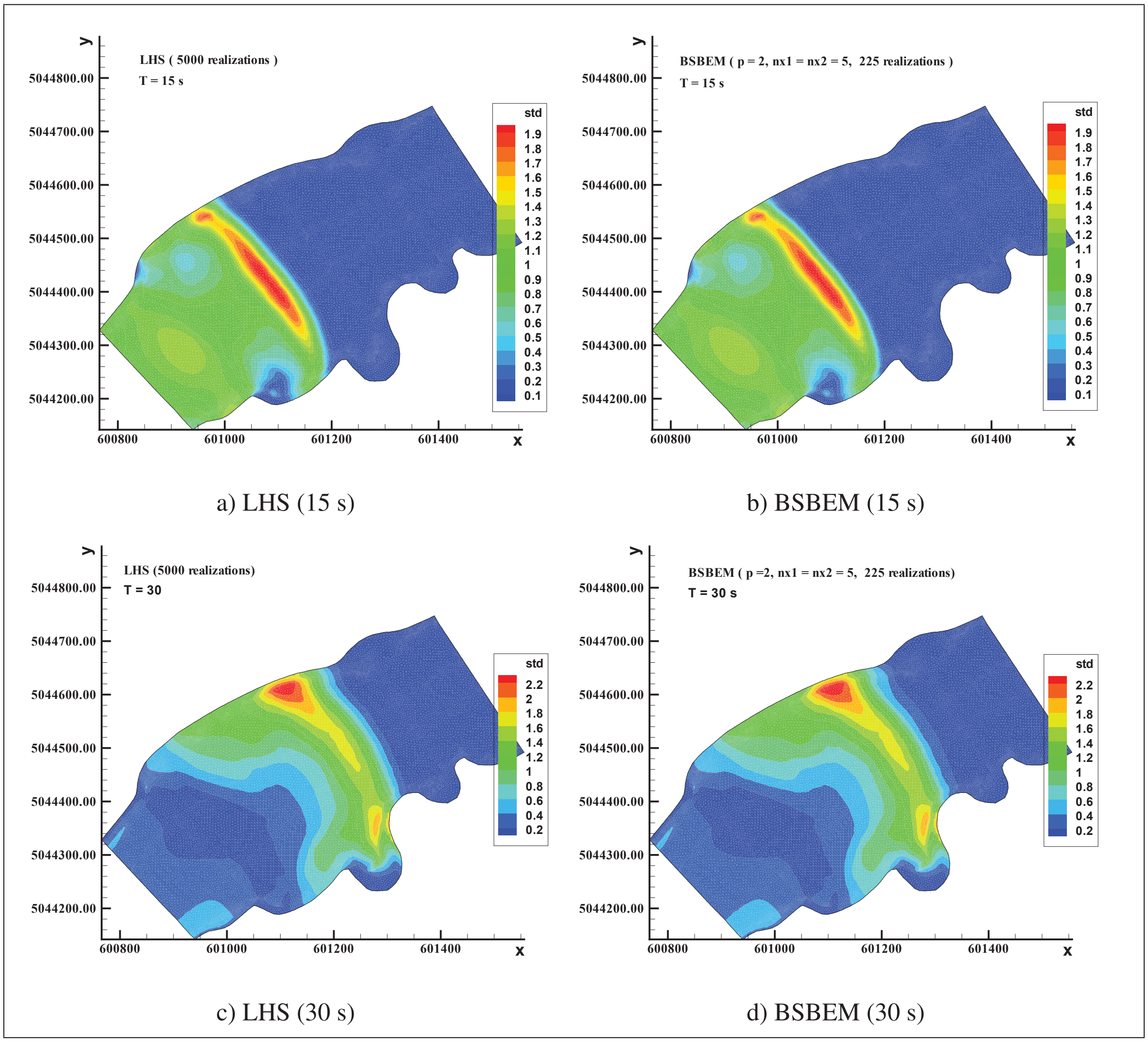
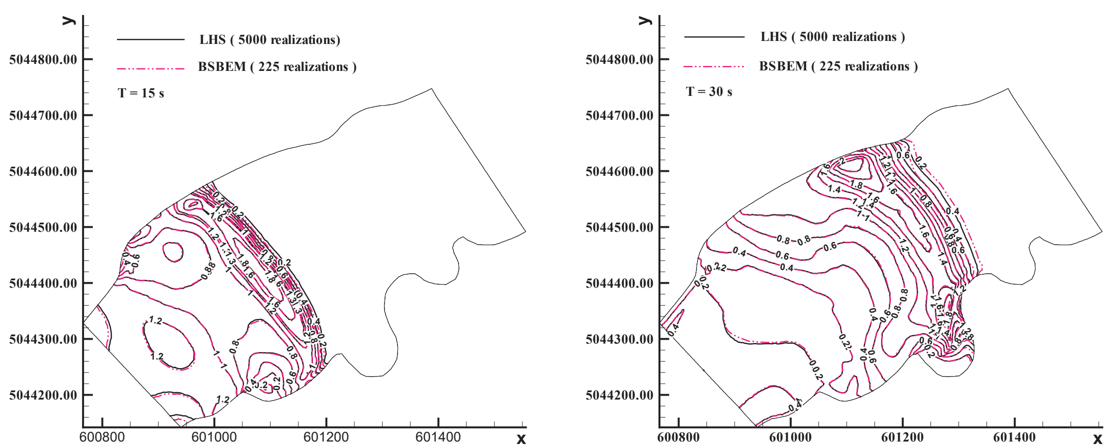


Figure 2.21 Spatial distribution of the standard deviation, comparison between the LHS and the BSBEM results : (a) and (b) for 15 s ; (c) and (d) for 30 s.

much-reduced computational efforts compared to standard sampling methods (Monte Carlo and LHS).

Fig.2.23 depicts the effect of increasing the coefficient of variation (cv) of the input parameters on the uncertainty of the model output response. Indeed, two values of the coefficient of variation ($cv = 5\%$ and $cv = 10\%$) are considered for both the initial upstream water level and the



a) 15 s

b) 30 s

Figure 2.22 Comparison of standard deviation contours between the BSBEM (225 realizations) and the LHS (5000 realizations) : (a) for 15 s and (b) for 30 s.

Manning roughness coefficient. The spatial distributions of the standard deviation of the output water depth through the entire calculation domain for a simulation time of 30 s, corresponding to the aforementioned two values of cv, are shown in Figs.2.23a and 2.23b, respectively. One can see from these figures that the increase of the cv leads to a remarkable change in the standard deviation contours' distributions, where a substantial widening of the area with high variability surrounding the flood wave front can be observed. The maximum value of the standard deviation, which is reached at the areas close to the left and the right banks of the flood wave front, increases from 1.8 m to 2.2 m, with the increase of cv from 5 to 10 %, respectively.

In order to illustrate the ability of the BSBE approach to provide accurate inundation maps, the confidence interval bounds' contours of the output water depth are shown in Fig.2.23c. It must be emphasized that the deterministic model evaluates the output water level at the center of each triangular element over the whole calculation domain, where a post processing is performed to assess the statistical moments of the variable of interest. Once the confidence interval bounds are evaluated and projected over the calculation domain, one can easily determine the spatial variability range surrounding the selected water level. In this particular case, we are focusing on

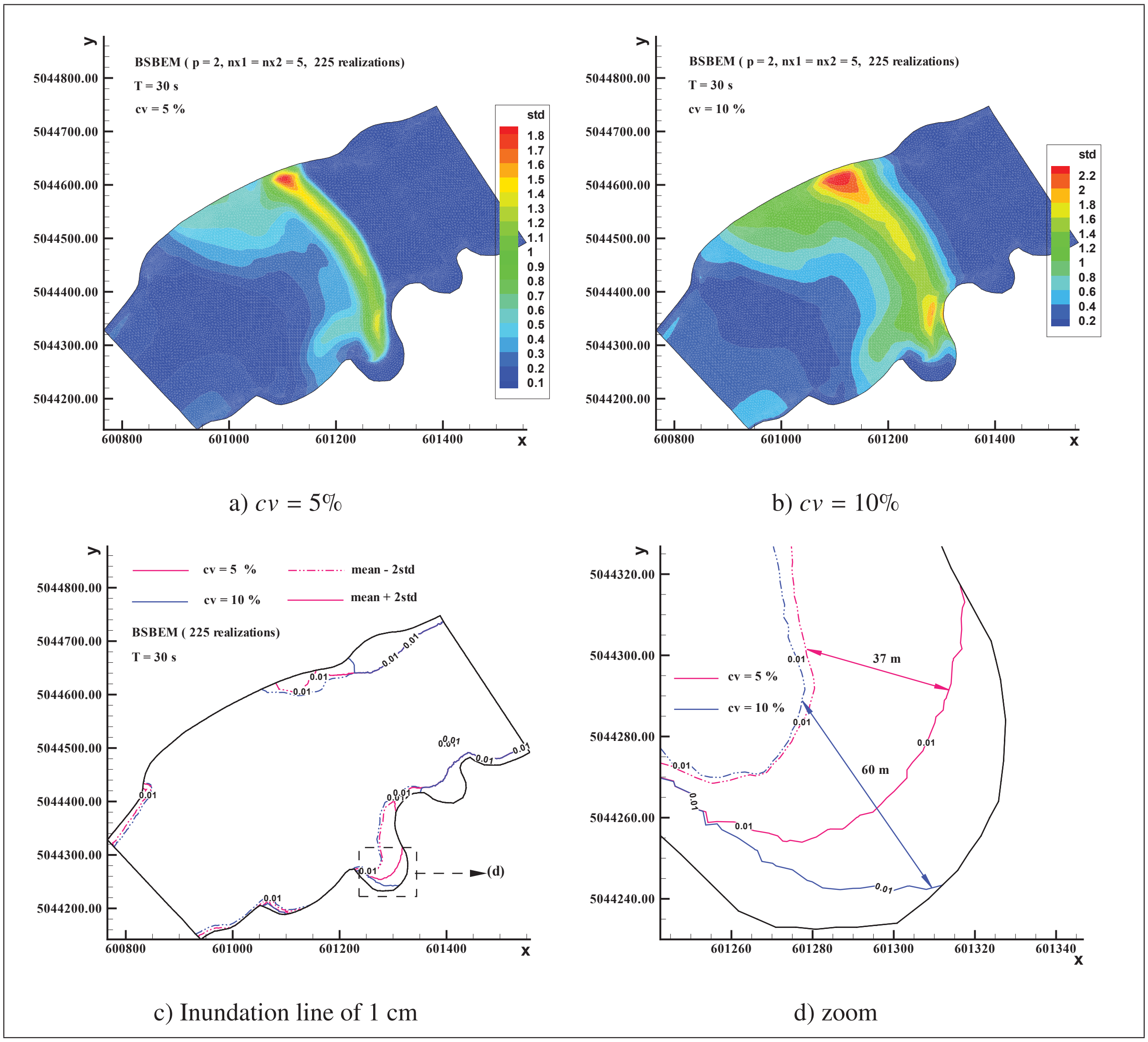


Figure 2.23 Spatial distribution of the standard deviation from the BSBE method for two values of the coefficient of variation : (a) $cv = 5\%$, (b) $cv = 10\%$ and comparison of the confidence interval bounds' contours of the inundation line obtained for both coefficients of variation : (d) is the close-up view of the selected area in (c).

the uncertainty of the inundation line, which represents, by definition, the borderline between the wet and dry zones, and whose selection criteria is defined by the threshold water level of 1 cm. Fig.2.23c illustrates the effect of the increase of the cv of the uncertain parameters on the inundation line for a simulation time of 30 s. The close-up view of the selected area shows that an increase in the coefficient of variation from 5% to 10% leads to a widening of the variability range of the inundation line from approximately 37 m to 60 m, respectively, as depicted in Fig.2.23d.

2.5 Conclusion

This paper contributes to the uncertainty propagation analysis for non-linear hyperbolic problems. The proposed method belongs to the non-intrusive multi-element stochastic expansions that have been developed to overcome the limitations of classical approaches in the presence of discontinuities. The BSBEM combines the smoothness (C^{p-1} continuity) of the B-spline functions with the multi-element structure of the Bézier extraction operator. The expansion is combined with sampling methods to establish a stochastic surrogate model. The calculation of the expansion coefficients was performed by solving a global system of equations obtained from an assembly of local design matrices and response vectors. The procedure's performance was tested on two highly non-linear and non-monotonic analytical benchmark functions with two and three random variables. The BSBEM satisfactorily yields accurate approximations of the statistical moments of the output response compared to the polynomial chaos-based point collocation expansion and the Monte Carlo method at similar calculation efforts.

The introduced BSBE approach was applied to dam break flows governed by highly non-linear shallow water equations where strong discontinuity occurs in the physical field surrounding the front inundation wave. Two applications with a dam break test case and a dam break over a real topography were considered to quantify the uncertainty propagation through the aforementioned deterministic model. The obtained results showed that the use of polynomial chaos expansions (PCM and Pcol) provided strong oscillatory approximations of the statistical moments (mean and standard deviation) of the output response of the water depth at the gauging measurement points. However, it is clear that the use of the BSBE expansion provided smooth statistical moments profiles, thus allowing to greatly reduce the oscillatory behaviour of the output response. This improvement is achieved by reducing the polynomial order of the basis function and increasing the number of Bézier elements in each direction, leading to an optimal distribution of the collocation points over the definition domain of each random variable. Furthermore, the obtained results demonstrated the ability of the BSBEM to approximate the output response in the case of dam break flow over a real topography, where confidence intervals of inundations maps were obtained with high accuracy, comparable to those of 5000 LHS realizations but with much lower computational cost.

Thus, the proposed non-intrusive BSBE method promises to be an efficient tool for uncertainty quantification, particularly for cases with discontinuous physical fields. It must be emphasized that the aforementioned method could also be employed for building surrogate models for optimization purposes. Several features of the BSBEM could be explored to improve its efficiency, for instance, modifying the anisotropy of the polynomial order of the basis function and the number of Bézier elements in each direction, which could greatly reduce the computational cost, especially when combined with an efficient sampling method.

CHAPITRE 3

UNCERTAINTY PROPAGATION OF DAM BREAK FLOW USING THE STOCHASTIC NON-INTRUSIVE B-SPLINES BÉZIER ELEMENTS-BASED METHOD

Azzedine Abdedou[1], Azzeddine Soulaïmani[1], Georges Williams Tchamen[2]

[1] Département de Génie Mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3
[2] Hydro-Québec, 75 Boul. René-Lévesque Ouest, Montréal (Qc), Canada H2Z 1A4

Article publié dans la revue :
Journal of Hydrology, Volume 590, November 2020,
https ://doi.org/10.1016/j.jhydrol.2020.125342

Abstract

The non-intrusive B-spline Bézier elements method (BSBEM) is used to perform uncertainty propagation for complex dam break flow models. The predictive efficiency of the BSBEM is first compared to the polynomial chaos expansion (PCE) and Monte Carlo (MC) methods using Stoker's analytical solution for an idealized dam break flow with discontinuous outputs. The computational efficiency of the proposed methodology is then illustrated by analyzing the uncertainty propagation of a hypothetical failure of an actual dam (Batiscan River) located within the province of Québec (Canada). Three parameters, the input discharge, the Strickler roughness coefficient of the main channel and the average breach width are considered as input random variables from which uncertainties may occur and propagate through the 1D hydraulic numerical model (Mascaret). The obtained results reveal the ability of BSBEM to efficiently predict the statistical moments and probability distributions of the output quantities of interest represented in terms of the downstream discharge, water level and front wave arrival time at different locations of the studied reach, thus yielding useful information for flood risk management with lower computational cost.

keywords

Uncertainty propagation, Polynomial chaos, B-Splines Bézier element method, Dam break flows

3.1 Introduction

Floods triggered by dams breaking are considered among the most devastating type of floods due to their instantaneous and unpredictable nature leading to costly material damages and loss of human lives, as evidenced by numerous accidents (Dai *et al.*, 2005 ; Lempérière, 2017). Considerable efforts have been dedicated over the last decades to develop numerical models to simulate the spatial and temporal variations of the maximum water level and arrival time of a dam break shock wave (Aureli, Maranzoni & Mignosa, 2014 ; Chang, Kao, Chang & Hsu, 2011 ; Erpicum *et al.*, 2010 ; Haltas *et al.*, 2016 ; Marsooli & Wu, 2014 ; Seyedashraf & Akhtari, 2017 ; Zhang *et al.*, 2014). The availability of these hydraulic parameters allows for fair hazard assessment and provides decision-making tools with which to establish emergency action plans for flood-prone areas.

The predictions stemming from these numerical models are considered as deterministic, since they are obtained with fixed values of the input parameters and boundary conditions such as upstream discharge, bathymetry, roughness coefficient, and average breach width. The nominal values of these parameters are estimated by means of in situ measurements and observations, as well as by empirical models, as is the case with the breach width, where several correlations have been proposed to predict breach characteristics (Ahmadisharaf *et al.*, 2016 ; Froehlich, 2008 ; Sammen *et al.*, 2017). It should be stressed that major uncertainties (aleatoric and epistimic) may originate from these estimation processes due to a lack of knowledge, measurement errors, and the random nature of input parameters (Butt *et al.*, 2013 ; Kim & Sanders, 2016), which can significantly influence the deterministic inundation mapping and related planning actions. It is thus crucial to take into account the propagation of uncertainties from the input parameters via numerical models in order to quantify the variability to the output response of quantities of interest (Qoi) (Wu & Hsu, 2018), thereby allowing the establishment of probabilistic flood inundation mapping that can provide credible and realistic predictions (Bharath & Elshorbagy, 2018 ; Tsai *et al.*, 2019).

Sampling methods are some of the most extensively used approaches for uncertainty propagation analysis. The well-known Monte Carlo (MC) method relies on generating a random sampling of each input parameter accordingly to its probability distribution, running the deterministic model and then computing the statistical moments of quantities of interest from the set of the output responses. Despite the attractiveness of the MC method due to its simplicity, it requires a high sample size for an accurate estimation of statistics, which can be a major drawback for models demanding large computational consumption (Ballio & Guadagnini, 2004). A variety of sampling approaches, such as Latin Hypercube Sampling (LHS) (McKay *et al.*, 1979 ; Shields & Zhang, 2016), which is based on stratified sampling, and low-discrepancy sampling (Hammersley & Handscomb, 1964 ; Sobol & Levitan, 1999), have been developed to alleviate the cumbersome computational cost. These approaches improve the sampling strategy and thus reduce the number of sample points by ensuring they are efficiently widespread over the random space.

In addition to the aforementioned sampling techniques, Polynomial chaos expansion (PCE) presents a practical alternative for uncertainty analysis. Based on the spectral representation of the output response as a polynomial expansion of the input variables (Ghanem & Spanos, 1991), PCE selects the orthogonal basis functions in relation to the probability distribution functions of the input random parameters to ensure the best efficacy of the expansion (Xiu & Karniadakis, 2002). Non-intrusive polynomial chaos expansions (NIPCE) have been increasingly applied to uncertainty analysis by many researchers. These methods, unlike the so-called intrusive techniques (Mathelin, Hussaini & Zang, 2005 ; Soize & Ghanem, 2004), treat the deterministic solver as a black box, which renders them more useful since the source codes of most commercial software systems are not available. In addition to the projection technique, which is based on a multidimensional numerical quadrature (Avdonin, Jaensch, Silva, Češnovar & Polifke, 2018 ; Reagana, Najm, Ghanem & Knio, 2003), the so-called regression technique has been successfully employed to compute the expansion's coefficients non-intrusively using the least square minimization procedure (Berveiller *et al.*, 2006). Other approaches originate from the NIPCE, based on the selection of the sampling points in the input domain, including : the probabilistic collocation method (PCM) (Li & Zhang, 2007 ; Tatang *et al.*, 1997), the point collocation method (Pcol) (Hosder & Walters, 2010 ; Hosder *et al.*, 2007) and more recently, a non-intrusive approach that requires fewer collocation points, polynomial chaos decomposition with differentiation (PCDD) (Thapa *et al.*, 2018 ; Thapa, Mulani & Walters, 2019).

The use of NIPC expansions to estimate the statistics of quantities of interest presenting smooth and continuous variations as a function of the input parameters yields accurate values. However, for problems with dominate hyperbolic behaviour, or with discontinuities in the random or physical space, a global polynomial basis that spans the entire random space is used, which generates spurious oscillations in the approximated output response. These spurious oscillations, known as Gibbs phenomena (Bijl *et al.*, 2013), produce large deviations from the exact solution with highly non-physical values (negative water-level values) and less-smooth profiles (Debusschere, Najm, Pébay, Knio, Ghanem & Le Maıtre, 2004). A panoply of methods have been introduced to overcome the limited utility of standard PCEs. Some of these suggest discretizing the input random space into multiple non-overlapping subspaces where either polynomial chaos basis functions (Chouvion & Sarrouy, 2016 ; Foo *et al.*, 2008 ; Wan & Karniadakis, 2005) or other types of basis functions such as multi-wavelet functions (Le Maître *et al.*, 2004a ; Pettersson *et al.*, 2013) can be used to locally approximate the output response (Jakeman, Narayan & Xiu, 2013 ; van Halder, Sanderse & Koren, 2018).

In this context, the B-Splines basis functions present a promising tool to mitigate the oscillatory behaviour of the output response with regard to the mathematical properties they procure. Indeed, Ditkowski *et al.* (2018) recently presented a detailed comparison that addressed the attractive features of B-splines in contrast to some of the limitations of PC basis functions in approximating the statistical moments and more particularly the probability distribution of quantities of interest that do not smoothly depend on their inputs. The application of B-splines basis functions in the field of uncertainty analysis has gradually gained in interest since the advent of Isogeometric analysis concepts (IGA) (Hughes *et al.*, 2005), as indicated by a growing number of studies (Beck *et al.*, 2019 ; Dertimanis, Spiridonakos & Chatzi, 2018 ; Li, Wu, Gao & Song, 2019). Recently, a non-intrusive B-Splines Bézier Elements based Method (BSBEM) was introduced (Abdedou & Soulaïmani, 2019) for uncertainty propagation analysis for hyperbolic problems. Inspired by the IGA approach, this method relies on approximating the stochastic output response using a B-splines polynomial expansion (over the Bézier elements) on the parametric domain, which corresponds to the cumulative distribution function.

The objective of this study is to apply the aforementioned BSBEM to perform uncertainty propagation analysis and surrogate model construction in the case of the hypothetical failure of an actual dam located within the province of Québec (Canada). Realistic terrain data and model parameters are available for this dam, such as the bathymetry, roughness coefficients, initial water level, average breach width, and upstream discharge. The variability of the output responses, such as the downstream water level and the discharge as a function of time at various locations of the reach under study, are evaluated by the proposed expansion and compared to the values from the reference Monte Carlo method and the polynomial chaos-based point collocation approach. Furthermore, this paper complements the previously published one (Abdedou & Soulaïmani, 2019) and pushes the BSBEM method towards its potential in realistic applications for engineering interests, by investigating its abilities to accurately estimate, beyond the statistical moments, the whole probability density functions (pdfs) of the output quantities. The predictions of the pdfs' aspects are particularly worth investigating because, as reported in the literature, even though the classical PCE approaches may be accurate enough to predict the statistical moments, they may fail to accurately estimate the pdfs. This issue is therefore deeply investigated in the present work.

The rest of the paper is organized as follows : Section 3.2 introduces the concepts of the non-intrusive PCE and BSBE methods. Section 4.3 presents a comparative evaluation of the performance of both the BSBEM and Pcol methods, applied to an analytical solution of an idealized dam break test case. The results of a hypothetical rupture of a real dam are then presented in terms of statistical moments and probability distributions. Finally, a summary and concluding remarks are presented in section 4.4.

3.2 Uncertainty propagation methodology

The general framework for uncertainty analysis relies on several steps (De Rocquigny *et al.*, 2008 ; Sudret, 2007) : *i*) *quantifying the sources of uncertainty* which involves the estimation of uncertainty of input parameters (aleatoric and epistimic) *ii*) *propagating the uncertainty* through the model which transforms the measure of uncertainty in the input into the measure of uncertainty in the outputs *iii*) *post-processing the results* by computing the statistics of the outputs and sensitivity indices. As mentioned above, the present study is focused solely on uncertainty propagation. In addition to the classical Monte Carlo (MC) method, two non-intrusive stochastic expansion approaches, the point collocation method (Pcol) and the B-Spline Bézier elements method (BSBEM), are used here to deal with uncertainty propagation in physical problems presenting hyperbolic behaviour. Their fundamental concepts will be summarily addressed in the sequel.

3.2.1 Non-intrusive point collocation method

Let $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_m\} \in \mathbb{R}^m$ be a set of m independent uncertain input parameters defined on the probability space $(\Theta, \varSigma, P)$, where Θ is the event space, P the probability measure and $\varSigma$ is the σ-algebra on Θ. The statistical behaviour of $\eta_i(\theta)$, with its support denoted by Ξ, is characterized by a probability density function $\varrho_i(\eta_i)$. In case of the correlated input parameters, an appropriate mathematical transformation such as the Karhunen-Loève expansion (Li & Zhang, 2007) can be used to remove the correlation prior to the construction of the basis functions (Oladyshkin & Nowak, 2012). Let us consider a deterministic relationship $y = \mathcal{M}(\boldsymbol{\eta})$ that links a set of m independent uncertain input parameters to the output response variable y ; where $\mathcal{M}$ represents the model that may be considered as black box. The model's output response can be expanded onto multivariate orthogonal basis functions as follows (Ghanem & Spanos, 1991) :

$$\widehat{y} = \sum_{j=1}^{M} \alpha_j \Psi_j(\boldsymbol{\eta}) = \widehat{\mathcal{M}}(\boldsymbol{\eta}) \tag{3.1}$$

where $\Psi_j(\boldsymbol{\eta})$ denotes the multivariate basis functions that are orthogonal to the joint probability density function $\varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta}) = \prod_{i=1}^{m} \varrho_i(\eta_i)$. $\widehat{\mathcal{M}}$ denotes the surrogate model of $\mathcal{M}$ and $\widehat{y}$ an approximation of y. α_j are the unknown coefficients to be computed, and which are space and time-dependent ($\alpha_j(x,t)$), representing the deterministic part of the expansion. M is the total number of expansion terms defined as a function of the polynomial chaos order p and the number of the input random parameters m :

$$M = \frac{(p+m)!}{p!m!} \tag{3.2}$$

The one-dimensional basis functions $\psi_k(\eta_i)$, $\{k = 1, \ldots, p;\ i = 1, \ldots, m\}$, from which the multivariate basis functions $\Psi_j(\boldsymbol{\eta})$, $\{j = 1, \ldots, M\}$ are constructed, must be meticulously selected with respect to the probability density function type. Indeed, there is a correspondence between the probability distribution of the input parameter and which appropriate basis functions must be selected in order to ensure the best convergence and accurately represent the output distributions. The most commonly-used orthogonal polynomials with their corresponding probability distributions are listed in Table 3.1 (Xiu & Karniadakis, 2002).

The multivariate basis functions Ψ_i given in Eq.(3.1) are obtained from the product of univariate polynomials (Sudret, 2008 ; Xiu, 2010) :

$$\Psi_j(\boldsymbol{\eta}) = \psi_{p_1^j}(\eta_1) \times \ldots \times \psi_{p_m^j}(\eta_m) \tag{3.3}$$

Tableau 3.1 Continuous probability distributions and associated orthogonal polynomials.

Distribution	**Polynomial**	**Support Range**
Gaussian	Hermite	$[-\infty, +\infty]$
Uniform	Legendre	$[-1, 1]$
Exponential	Laguerre	$[0, +\infty]$
Beta	Jacobi	$[a, b]$

where the subscript p_i^j incorporates the multi-indices notation $\mathbf{q} \in \mathbb{N}^m$ so that the sum of all their possible combinations $\left(|\mathbf{q}| = \sum_{i=1}^{m} p_i^j\right)$ is equal to the total degree of each multivariate polynomial $\Psi_{\mathbf{q}}(\boldsymbol{\eta}) \equiv \Psi_j(\boldsymbol{\eta})$. More implementation details about the multi-indices and the construction of the multivariate basis functions are given in Avdonin *et al.* (2018) ; Sudret (2014) and Miller *et al.* (2018).

The unknown coefficients $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_M\}^T$ of the approximation in Eq.(3.1) can be obtained using the regression (also called the least square minimization) approach, one of the non-intrusive techniques. This method starts with the construction of the experimental design $\mathfrak{D} = \left\{\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N_s)}\right\}^T$, which is defined as a sample set of N_s collocation points generated according to the probability density function of each input parameter $\{\eta_1, \ldots, \eta_m\}$. For each row of the experimental design, the deterministic model is run to obtain the corresponding results which are gathered in an output response vector $\mathbf{y} = \left\{y^{(1)} = \mathcal{M}(\boldsymbol{\eta}^{(1)}), \ldots, y^{(N_s)} = \mathcal{M}(\boldsymbol{\eta}^{(N_s)})\right\}^T$. Evaluating the obtained multivariate basis functions onto these sampling points within the experimental design allows the construction of a design matrix whose terms are defined as : $\boldsymbol{\Psi} = \left\{\Psi_{ij} = \Psi_j(\boldsymbol{\eta}^{(i)}),\ i = 1, \ldots, N_s;\ j = 1, \ldots, M\right\}$. The unknown coefficients can thus be obtained by solving the normal equation :

$$\boldsymbol{\alpha} = \left(\boldsymbol{\Psi}^T \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\Psi}^T \mathbf{y} \tag{3.4}$$

The number of collocation points in the experimental design has to be larger than the number of expansion terms ($N_s \geq M$) in order to avoid a potential ill conditioning of the design matrix (Blatman, 2009). The use of the oversampling concept, known as the Point collocation approach (Pcol), which considers the total number of sampling points as a multiple of the expansion term number $N_s = n_p M$, with $n_p \in \mathbb{N}$ the oversampling ratio, allows a better estimation of the statistical moments, particularly when it is associated with an efficient sampling method (such as the LHS, Halton and Harmemesly methods)(Hosder & Walters, 2010 ; Hosder *et al.*, 2007).

After obtaining the expansion's coefficients, the statistical moments can be estimated easily due to the orthogonality property of the basis functions. Thus, the mean and standard deviation of the output response surface can be expressed as :

$$\mu = \mathbb{E}\left[\widehat{Y}\right] = \int_{\Xi} \left[\sum_{j=1}^{M} \alpha_j \Psi_j(\boldsymbol{\eta})\right] \varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta}) d\boldsymbol{\eta} = \alpha_1 \tag{3.5}$$

and

$$\sigma = \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}\left[\widehat{Y}\right]\right)^2\right]^{1/2} = \left[\sum_{j=2}^{M} \alpha_j^2 \langle \Psi_j^2(\boldsymbol{\eta})\rangle\right]^{1/2} \tag{3.6}$$

respectively, where $\langle \Psi_j^2(\boldsymbol{\eta})\rangle$ is evaluated using the inner product notion whose expression reads : $\langle \Psi_i, \Psi_j \rangle = \int_{\Xi} \Psi_i(\boldsymbol{\eta})\, \Psi_j(\boldsymbol{\eta})\, \varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$, and where $\varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta})$ denotes the joint probability function as mentioned above. The aforementioned fundamental concepts of the Pcol method are summarized in a graphical framework presented in Fig.3.1.

3.2.2 Non-intrusive B-Splines Bézier elements method

The B-splines Bézier element method (BSBEM) can be categorized as a non-intrusive stochastic multi-elements approach that consists of local approximations of the output response of quantities of interest using piecewise B-splines basis functions over each element named a Bézier element Borden *et al.* (2011). The B-splines basis functions present interesting mathematical features ; they offer : C^{p-1} continuity, local interpolation support, a strong convex hull property and a

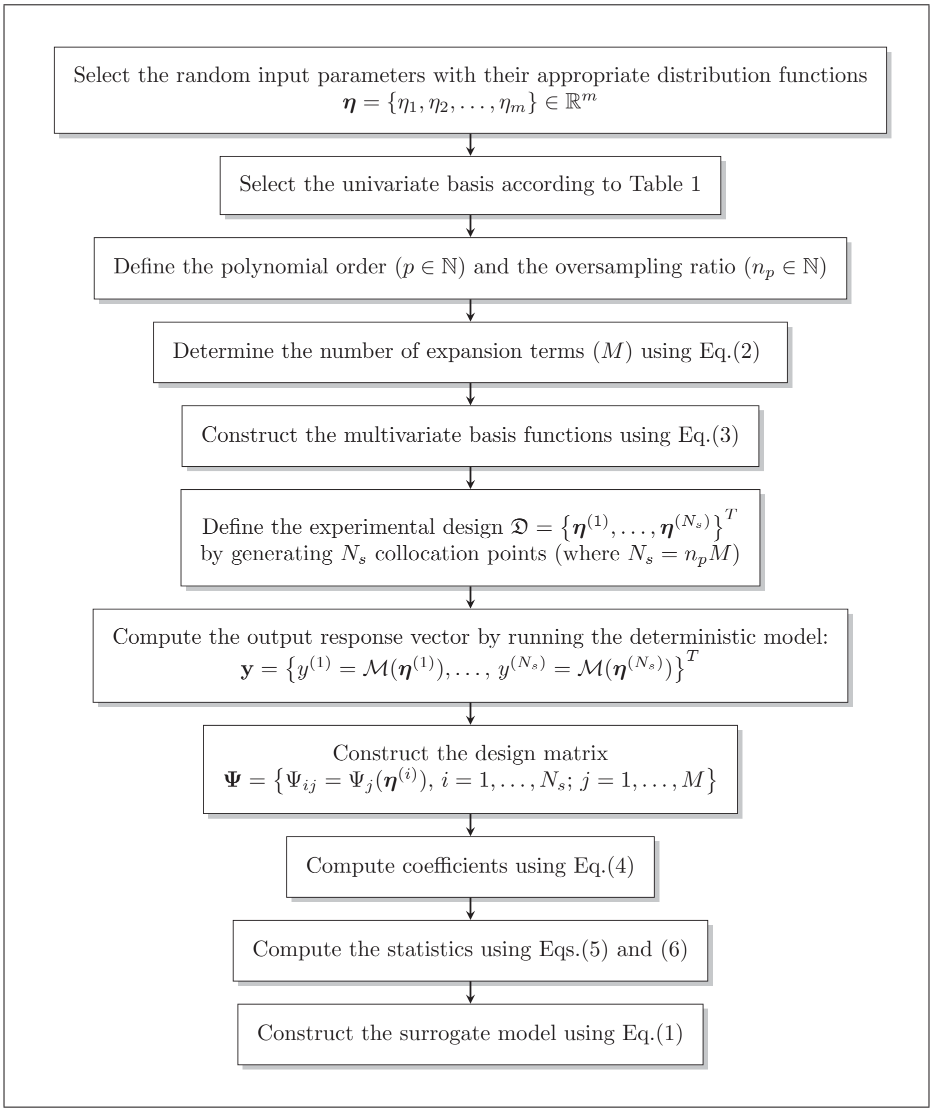


Figure 3.1 A general flowchart of the Pcol-based method.

variation diminishing property. These attractive characteristics combined with the multi-element structure of the Bézier extraction operator render them a promising tool for uncertainty analysis, even for problems with non-smooth or discontinuous outputs (Piegl & Tiller, 1997 ; Rogers, 2001). A schematic representation of the univariate B-splines basis functions (12 global splines) of order 2 over the parametric domain, their traces on the elements are shown in color in Fig.3.2. For every element, their local traces amount into three functions, in order, blue, red and green. The basic framework of this method and its implementation algorithm are addressed in detail in

Abdedou & Soulaïmani (2019).

Let us consider $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_m\} \in \Gamma = [0,\ 1]^m$ as a set of m independent parametric

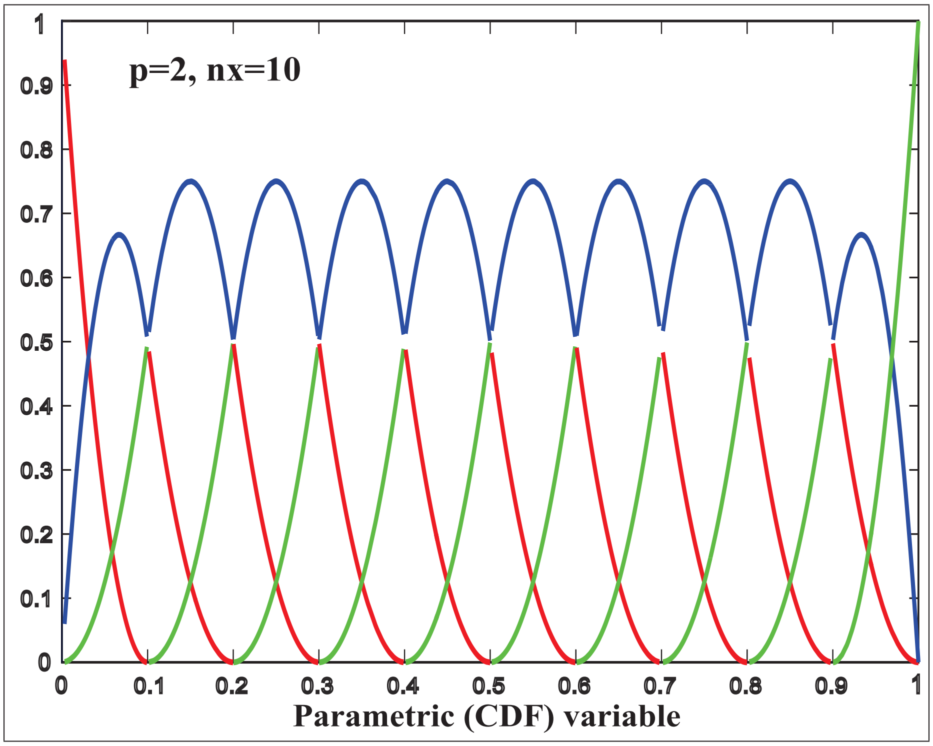


Figure 3.2 Schematic representation of the univariate $(p+1)$ B-splines basis functions (with order $p = 2$) over the parametric domain and the corresponding Bézier elements ($nx = 10$). For every element, their local traces amount into three functions, in order, blue, red and green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

variables defined in the parametric domain and linked to the set of input random parameters $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_m\} \in \mathbb{R}^m$ by the marginal cumulative density function (CDF) whose image domain is $[0,\ 1]$, which corresponds to the image domain of the B-splines basis functions. This allows for $\xi_i(\eta_i) = F_i(\eta_i) = \int_{-\infty}^{\eta_i} \varrho_i(\eta')d\eta'$, which implies that $d\xi_i = \varrho_i(\eta_i)d\eta_i$. The basic principle of the method is that it decomposes the definition domain $[0,\ 1]$ of each variable ξ_i into nx_i non-overlapping intervals whose interfaces are defined by the coordinates $\left(\xi_{i,0}, \ldots, \xi_{i,nx_i}\right)$, with $\xi_{i,0} = 0$ and $\xi_{i,nx_i} = 1$ representing the lower and upper bounds, respectively, of the definition domain. The Bézier element I_e (multidimensional element defined as hypercubes) for the entire parametric domain Γ is defined as a tensor product of univariate intervals $I_e = \left(\xi_{1,e_1}, \xi_{1,e_1+1}\right) \times \ldots \times \left(\xi_{m,e_m}, \xi_{m,e_m+1}\right)$, with $0 \leq e_i < nx_i$ and $i = 1, \ldots, m$. The total number of Bézier elements in the entire parametric domain is given by $N_{elt} = \prod_{i=1}^{m} nx_i$, which increases considerably for high dimensional problems (which have a high number of input random parameters, i.e. the curse of dimensionality problem). The model's output response $y = \mathcal{M}(\boldsymbol{\eta})$ is approximated on each Bézier element I_e as the sum of the piecewise basis functions of the set of parameters $\boldsymbol{\xi}$:

$$\widehat{y}^e = \sum_{j=1}^{nb} \alpha_j^e \Psi_j^e(\boldsymbol{\xi}) \tag{3.7}$$

where α_j^e and Ψ_j^e denote the local coefficients and the local multivariate piecewise B-splines basis functions over the e^{th} Bézier element I_e, respectively. The number of local multivariate basis functions (i.e., the number of local expansion terms) is given by $nb = \prod_{i=1}^{m}(p_i + 1)$, where p_i is the polynomial order corresponding to the random parameter ξ_i. nb grows significantly with increasing dimensionality, unlike in the Pcol method where the number of basis functions is obtained by total-order truncation, as given in Eq.(3.2). The global approximation of the output response can be built by patching the local approximations on each Bézier element I_e :

$$\widehat{y} = \sum_{e=1}^{N_{elt}} \widehat{y}^e(\boldsymbol{\xi}) \lambda_e(\boldsymbol{\xi}) \tag{3.8}$$

where $\lambda_e(\boldsymbol{\xi})$ denotes a standard indicator function satisfying $\lambda_e(\boldsymbol{\xi}) = 1$ if $\boldsymbol{\xi} \in I_e$ and $\lambda_e(\boldsymbol{\xi}) = 0$ otherwise (Jakeman *et al.*, 2013 ; van Halder *et al.*, 2018).

The construction of the multivariate basis functions in each element I_e can be realized by the tensor product of a set of local univariate polynomials as follows :

$$\Psi_j^e(\boldsymbol{\xi}) = \psi_1^e(\xi_1) \times \ldots \times \psi_m^e(\xi_m), \;\; j = 1, \ldots, nb \tag{3.9}$$

where $\psi_i^e(\xi_i)$ corresponds to the univariate B-splines basis functions that ensure C^{p_i-1} continuity, so that $\psi_i^e(\xi_i) \equiv N_{i,p_i}^e(\xi_i)$, $i = 1, \ldots, m$. These basis functions can be obtained with an appropriate mapping from a set of local Bernstein's functions by introducing the so-called local linear Bézier extraction operator $\mathbf{C}^e$, so that $\mathbf{N}^e(\xi) = \mathbf{C}^e\mathbf{B}^e(\xi)$ (Borden *et al.*, 2011), where $\mathbf{B}^e$ denotes the univariate Bernstein functions of degree p on the e^{th} Bézier element, whose explicit recursive form on the parametric domain $[0, 1]$ is given by (Buffa & Sangalli, 2016) :

$$B_{k,p}(\xi) = (1-\xi)\, B_{k,p-1}(\xi) + (1+\xi)\, B_{k-1,p-1}(\xi) \tag{3.10}$$

with $B_{k,0}(\xi) = 1$ and $B_{k,p}(\xi) = 0$ if $k < 1$ or $k > p + 1$. Bernstein functions can also be defined over the interval $\tau \in [-1, 1]$ as follows (Borden *et al.*, 2011) :

$$B_{k,p}(\tau) = \frac{1}{2}(1-\tau)\, B_{k,p-1}(\tau) + \frac{1}{2}(1+\tau)\, B_{k-1,p-1}(\tau) \tag{3.11}$$

This recursive expression is more useful for numerical integration by quadratures. In this case, for a given value of the univariate parameter $\xi_i \in \left[\xi_{i,e_i}, \xi_{i,e_i+1}\right]$, $0 \leq e_i < nx_i$ and $i = 1, \ldots, m$, its corresponding parametric value $\tau_i \in [-1, 1]$ can be computed by adopting the following change of variable :

$$\tau_i = \frac{2\xi_i - (\xi_{i,e_i} + \xi_{i,e_i+1})}{\xi_{i,e_i+1} - \xi_{i,e_i}} \tag{3.12}$$

The sampling procedure in connection with the above polynomial approximations requires the construction of local experimental design $\mathfrak{D}^e$ in each element I_e. Thus, for each i^{th} random variable ξ_i in the parametric random vector $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_m\}$, a set of $n_s^e \geq nb$ collocation points are selected randomly in the e_i^{th} corresponding interval $\xi_i^{e\,(s)} \in \left[\xi_{i,e_i}, \xi_{i,e_i+1}\right]$, $s = 1, \ldots, n_s^e$. Thus, a local experimental design $\mathfrak{D}^e = \left\{\boldsymbol{\xi}^{e\,(1)}, \ldots, \boldsymbol{\xi}^{e\,(s)}, \ldots, \boldsymbol{\xi}^{e\,(n_s^e)}\right\}^T$ with $\boldsymbol{\xi}^{e\,(s)} = \left\{\xi_1^{e\,(s)}, \ldots, \xi_m^{e\,(s)}\right\}$ is constructed. As for the aforementioned Pcol expansion, a local over-sampling technique is also possible in each element by introducing an oversampling ratio $n_s^e = n_p nb$. The local multivariate basis functions given in Eq.(4.7) are evaluated on each row of the local experimental design $\boldsymbol{\psi}^e_{n_s^e \times nb} = \left\{\Psi^e_{i,j} = \Psi^e_j(\boldsymbol{\xi}^{e\,(i)}), i = 1, \ldots, n_s^e,\ j = 1, \ldots, nb\right\}$. The components of the local experimental design are also used to run the deterministic model to obtain the local output response vector $\mathbf{y}^{e}_{n^e_s\times 1} = \left\{y^{e\,(1)} = \mathcal{M}(\boldsymbol{\eta}^{e\,(1)}), \ldots,\ y^{e\,(s)} = \mathcal{M}(\boldsymbol{\eta}^{e\,(s)}), \ldots,\ y^{e\,(n^e_s)} = \mathcal{M}(\boldsymbol{\eta}^{e\,(n^e_s)})\right\}^T$, where $\boldsymbol{\eta}^{e\,(s)} = \left\{\eta^{e\,(s)}_1, \ldots, \eta^{e\,(s)}_m\right\}$ represents the s^{th} component of the corresponding local experimental design in the random physical domain obtained by the inverse CDF mapping :

$$\eta^{e\,(s)}_i = F_i^{-1}(\xi^{e\,(s)}_i),\ s = 1, \ldots, n^e_s,\ i = 1, \ldots, m. \tag{3.13}$$

The resulting local design matrix and the output response vector are expressed as :

$$\boldsymbol{\Psi}^{e}_{nb\times nb} = \boldsymbol{\psi}^{e^{\mathsf{T}}}\boldsymbol{\psi}^{e} \quad \text{and} \quad \mathbf{Y}^{e}_{nb\times 1} = \boldsymbol{\psi}^{e^{\mathsf{T}}}\mathbf{y}^{e} \tag{3.14}$$

A global design matrix $\boldsymbol{\Psi}_{M\times M}$ and global output response vector $\mathbf{Y}_{M\times 1}$, where $M = \prod_{i=1}^{m} (nx_i + p_i)$, are constructed from the assembly of local design matrices and local output response vectors using the assembly operator $\boldsymbol{\Psi} = \underset{e=1}{\overset{N_{elt}}{\mathbf{A}}}(\boldsymbol{\Psi}^e)$ and $\mathbf{Y} = \underset{e=1}{\overset{N_{elt}}{\mathbf{A}}}(\mathbf{Y}^e)$ introduced by Hughes (2012), where the subscripts of the global design matrix components $\Psi_{k,\ell}$ are linked to those of the local design matrices $\Psi^{e}_{r,s}$ by the connectivity array (IEN) with respect to the e^{th} Bézier element, expressed as : $k = IEN(r, e)$ and $\ell = IEN(s, e)$ with $k, \ell = 1, \ldots, M$ and $r, s = 1, \ldots, nb$ (Cottrell *et al.*, 2009 ; Scott *et al.*, 2011). The unknown coefficients can be computed by solving the resulting global system of equations :

$$\boldsymbol{\alpha} = \boldsymbol{\Psi}^{-1}\mathbf{Y} \tag{3.15}$$

Once the coefficients are computed, the surrogate model given in Eq.(3.7) is realized by extracting the local coefficients $\boldsymbol{\alpha}^e$ from the set of global coefficients $\boldsymbol{\alpha}$, found with Eq.(4.16), using the same IEN connectivity array mentioned above. The constructed substitution model can then evaluate the statistical moments of the output quantities of interest, whose mean and variance are expressed, respectively, as follows :

$$\mu_{\widehat{Y}} = \mathbb{E}\left[\widehat{Y}\right] = \int_{\Xi} \widehat{Y}\varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta})d\boldsymbol{\eta} \tag{3.16}$$

and

$$\sigma^2_{\widehat{Y}} = \mathbb{E}\left[\widehat{Y}^2\right] - \mu^2_{\widehat{Y}} = \left[\int_{\Xi} \left(\widehat{Y}\right)^2 \varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta}) d\boldsymbol{\eta}\right] - \mu^2_{\widehat{Y}} \tag{3.17}$$

The evaluation of these integrals can be greatly simplified by considering the cumulative density function mapping, as detailed above, $\xi_i = F_i(\eta_i)$ and $d\xi_i = \varrho_i(\eta_i)d\eta_i$. This change of variable allows the integrals to be evaluated over the CDF image domain $[0, 1]$ that corresponds to the parametric domain of the basis functions. Thus, the mean and the variance given in Eqs.(3.16) and (3.17) can be expressed, respectively, as :

$$\mu_{\widehat{Y}} = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} \widehat{Y}^e d\boldsymbol{\xi}\right] \tag{3.18}$$

and

$$\sigma^2_{\widehat{Y}} = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} \left(\widehat{Y}^e\right)^2 d\boldsymbol{\xi}\right] - \mu^2_{\widehat{Y}} \tag{3.19}$$

Based on the detailed steps presented above, a general flowchart for the implementation of the proposed BSBEM is shown in Fig.3.3. The efficiency and accuracy of the aforementioned non-intrusive stochastic expansions will be compared in two 1D hydraulic tests.

3.3 Results and discussion

In the following subsections, numerical results for an idealized univariate test case (Stoker's analytical solution) and a hypothetical break of an actual dam are presented and discussed in terms of statistical moments and probability density functions profiles of output quantities of interest.

3.3.1 Univariate dam break flow test case (analytical solution)

In this benchmark test case, the physical domain is a 100 m flat frictionless channel. The dam, represented by a vertical gate, is located in the middle of the channel ($x_0 = 50\ m$). The initial conditions are defined by unequal water levels on both sides of the dam, as shown in Fig.3.4. The downstream water depth is kept constant at a deterministic value $h_{ds} = 1\ m$, while the

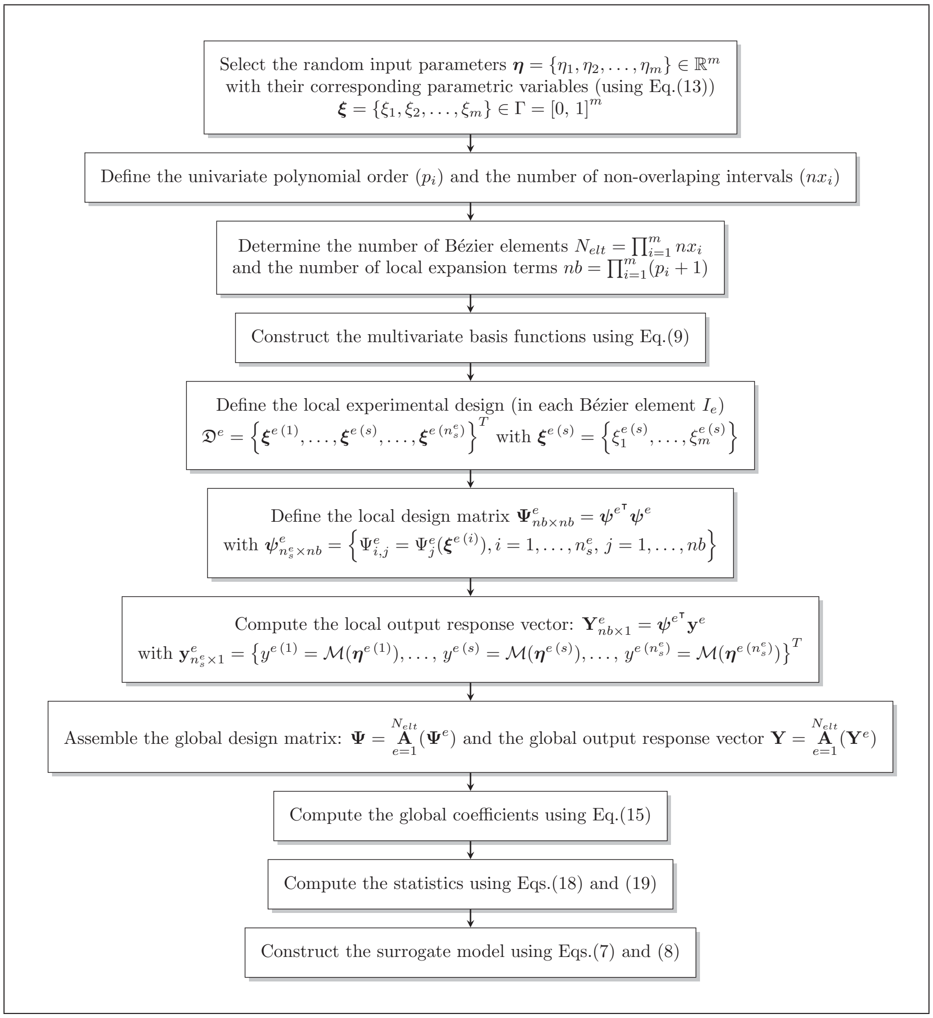


Figure 3.3 A general flowchart of the BSBEM.

upstream water depth (h_{up}) is considered as a noisy input parameter with various values in the random space. Stoker's solution (Stoker, 1957), an analytical solution of the aforementioned idealized problem, is obtained by solving the following polynomial equations (Delestre *et al.*, 2013 ; Seyedashraf & Akhtari, 2017 ; Seyedashraf, Mehrabi & Akhtari, 2018) :

$$h(x,t) = \begin{cases} h_{up} \\ \frac{4}{9g}(\sqrt{gh_{up}} - \frac{x}{2t})^2 \\ \frac{c_m^2}{g} \\ h_{ds} \end{cases} \quad , \quad u(x,t) = \begin{cases} 0 \quad m/s & \text{if } x \leq x_A(t) \\ \frac{2}{3}(\frac{x}{t} + \sqrt{gh_{up}}) & \text{if } x_A(t) \leq x \leq x_B(t) \\ 2(\sqrt{gh_{up}} - c_m) & \text{if } x_B(t) \leq x \leq x_C(t) \\ 0 \quad m/s & \text{if } x_C(t) \leq x \end{cases} \tag{3.20}$$

where x denotes the axial position, $x_A(t) = x_0 - t\sqrt{gh_{up}}$, $x_B(t) = x_0 + t(\sqrt{gh_{up}} - 3c_m)$ and $x_C = x_0 + t\frac{2c_m^2(\sqrt{gh_{up}} - c_m)}{c_m^2 - gh_{ds}}$, and where $c_m = \sqrt{gh_m}$ denotes the selected solution of $-8gh_{ds}c_m^2(gh_{up} - c_m^2)^2 + (c_m^2 + gh_{ds})(c_m^2 - gh_{ds})^2 = 0$, that fulfil both the mathematical and the physical conditions.

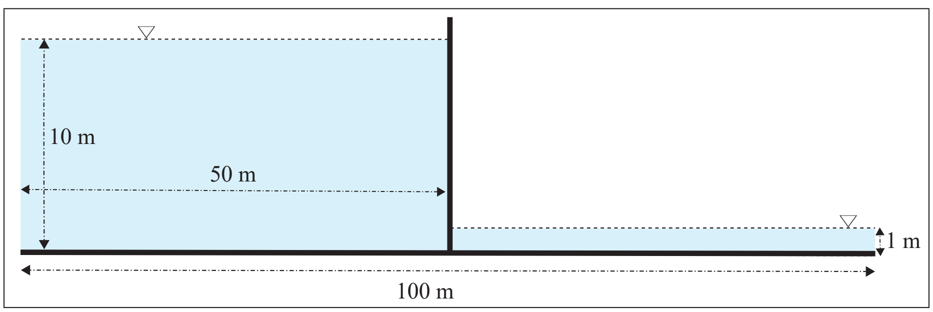


Figure 3.4 Sketch of initial conditions of dam break flow through a wet flat bottom.

The test case described above depicts an interesting case study for uncertainty propagation analysis and surrogate model building because it presents a low computational cost (analytical solution) combined with a high degree of difficulty due to the presence of a discontinuity in the output quantities of interest. As mentioned earlier, such problems have been addressed in the literature and are known to present a serious challenge for polynomial chaos methods for estimation of the stochastic behaviour of output responses in the vicinity of discontinuities (Pettersson, Iaccarino & Nordström, 2014). To further investigate this aspect, the analytical solution (Stoker's solution) is considered as a deterministic model through which uncertainties are propagated. Thus, the initial upstream water level is considered as a random parameter following a normal distribution centred on the nominal value $h_{up} = 10\,m$ with a standard deviation value of $0.5\,m$ which represents a coefficient of variation of 5%.

It has been shown that using a number of sampling points that is twice the minimum required number represents a fairly optimal value of the oversampling ratio ($n_p = 2$), which when used with a relatively high polynomial order of basis functions, leads to a substantial improvement in the accuracy of the point collocation method for smooth and continuous problems (Hosder & Walters, 2010; Hosder *et al.*, 2007). In the sequel, we address the effect of the oversampling ratio and its interaction with the polynomial order on the accuracy of the output response, especially for cases presenting hyperbolic behaviour with a discontinuity in the physical domain. It should be noted that a part of this aspect has been briefly outlined in Abdedou & Soulaïmani (2019).

The evolution of the statistics profiles (mean and standard deviation) of the water depth as a function of the channel length at simulation time $t = 4\,s$, obtained from the Pcol method using various polynomial order values and an oversampling ratio of two, are depicted in Fig.3.5 and compared to that of 10^6 Monte Carlo simulations. The distributions of the mean of the water depth (Fig.3.5a) obtained from the Pcol approach show a smooth profile with a good agreement with those obtained from the MC method over the whole domain, except in the regions near the discontinuity where there is a strong oscillatory behaviour. Indeed, by keeping an oversampling ratio of two, any increase in the polynomial order (from $p = 2$ to $p = 5$) leads to an increase in the intensity of the oscillations. This may include unphysical negative values for both water level and velocity, as can be seen from the zooms of the selected areas (discontinuity region). The described behaviour is more remarkable on the standard deviation distribution, where it seems to be much more sensitive to the increase of the polynomial order. The Pcol results clearly change significantly when going from $p = 2$ to $p = 5$ where strong oscillations, in terms of amplitude and frequencies, appear in the discontinuity region. This leads to multiple values of the maximum standard deviation, which could be observed in terms of local peak values obtained for $p = 5$ in the region surrounding the discontinuity, causing an unrealistic over-prediction of the standard deviation as shown in the close up view of the selected zones of the water depth profiles (Fig.3.5b). Similar error structures have been observed in the velocity profiles, and because of space limitation, their corresponding plots are not shown.

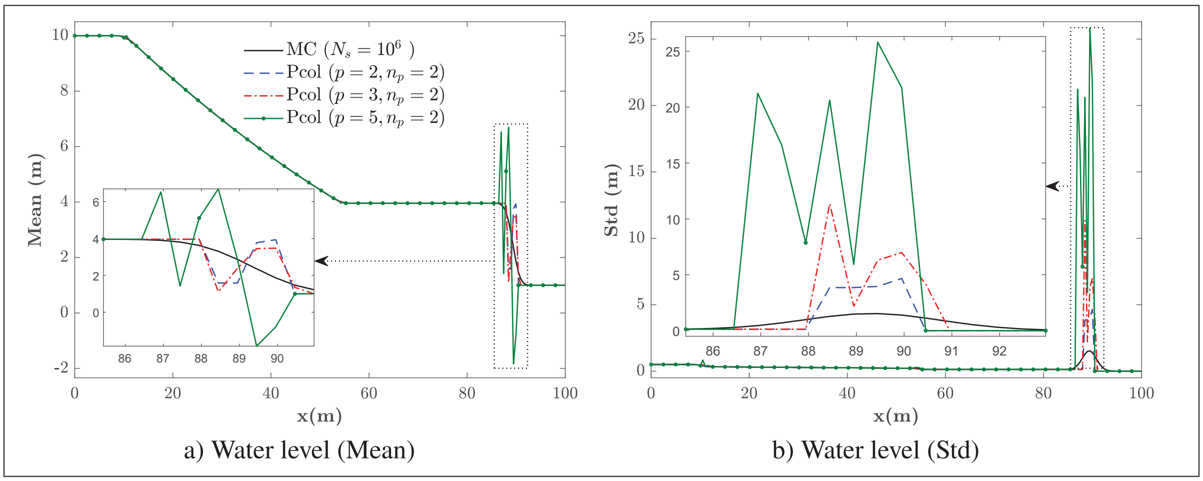


Figure 3.5 Distribution of mean and standard deviation of the water level (a and b) and along the channel length at time $t = 4\ s$, obtained from the Pcol (with $p = \{2,\ 3,\ 5\}$) and MC (with $N_s = 10^6$ realizations) methods.

To further investigate the effect of the interaction between the oversampling ratio and the polynomial order on the statistical moments' prediction of discontinuous output responses, the convergence of the relative least square mean error (L^2-norm), given by Eq.(3.21), is assessed for a wide range of the aforementioned parameter values :

$$L^2_{\Phi} = \sqrt{\frac{\sum_{k=1}^{N_e} (\Phi_{k,surrogate} - \Phi_{k,MC})^2}{\sum_{k=1}^{N_e} \Phi^2_{k,MC}}} \tag{3.21}$$

where Φ_k stands for the mean or the standard deviation of the output quantity of interest (in this case water depth) at the k^{th} axial position. This error is evaluated by comparing the statistical moments obtained by surrogate models (Pcol and BSBEM) with those from the Monte Carlo reference solution (with 10^6 realizations) over the total number (N_e) of calculation points (axial positions).

Fig.3.6 shows the variation of the L^2 error of the statistical moments of the water level as a function of the polynomial order for various oversampling ratios. One can observe from Fig.3.6a that the error in the mean increases greatly with the increase of the polynomial order, thus providing high values of relative errors, especially for low oversampling ratios (ex. $n_p = 2$). In the case of hyperbolic discontinuous problems, the PCE basis functions show spurious oscillations in the vicinity of the discontinuity, as shown in Fig.3.5. Even though theses oscillations are localized in a small region surrounding the shock, their intensity is so strong (and increases with the polynomial order) that the statistic profiles present as unrealistic and even non-physical values (Fig.3.5b), and so they impact the global tendency of the global L^2 error (this is similar to the well-known Gibbs phenomenon in the approximation theory). It can also be seen that as the oversampling ratio is increased (from $n_p = 2$ to $n_p = 10$), the intensity of the error generally diminishes for low chaos orders ($p \leqslant 5$), and is less remarkable for high chaos orders. The same behaviour can be observed in Fig.3.6b, which shows the error in the standard deviation with respect to the polynomial chaos order. This figure reveals that as the polynomial order increases, the plots display rapid divergence (almost exponential) where the percentage error reaches significant values ($\approx 10^4$) at a higher polynomial order, even more distinctly when a low oversampling ratio ($n_p = 2$) is used. Thus, by considering an oversampling ratio of two ($n_p = 2$), what is found to be an optimal value for smooth continuous solutions, as mentioned above, does not yield a suitable convergence with the increase of the polynomial orders for cases with discontinuous output response. In contrast, by choosing a relatively high value of oversampling ratio ($n_p = 10$), the error decreases slightly, particularly for low values of the polynomial order ($p \leqslant 5$).

To assess the ability of the BSBEM to accurately reproduce the Monte Carlo reference solution, the variation of the statistical moments' relative errors in the water depth as a function of the Bézier elements (nx) for a fixed polynomial order ($p = 2$) and different oversampling ratios are

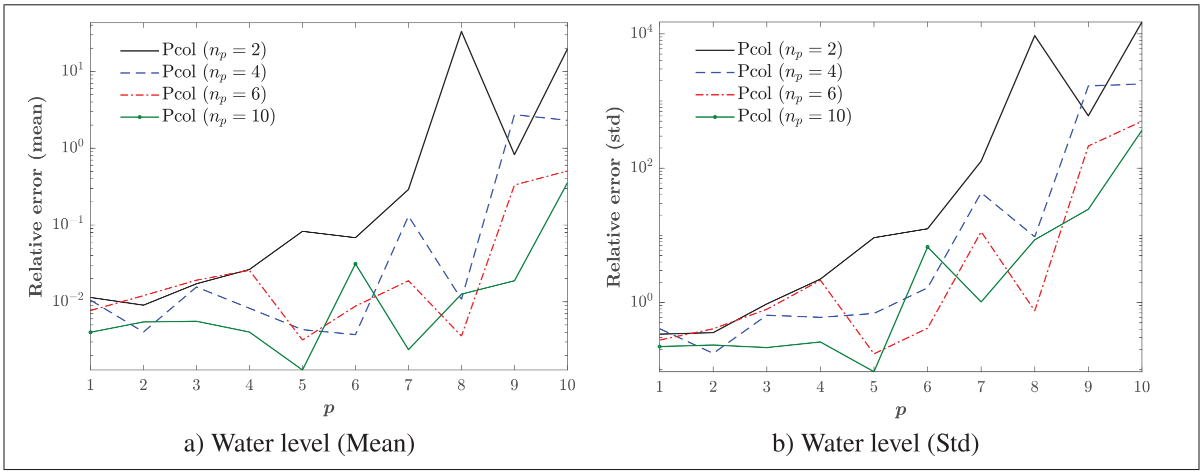


Figure 3.6 Error analysis of water level with increasing polynomial orders of a Pcol expansion for different values of oversampling ratio. (a) : error in mean ; (b) : error in standard deviation.

presented in Fig.3.7. These variations indicate that the relative errors of the BSBEM results, which are estimated with respect to reference Monte Carlo statistics, decrease monotonically when the number of Bézier elements are gradually increased from $nx = 1$ to $nx = 10^2$. It is worth noting that the BSBEM approach yields reasonable error values ($\approx 10^{-3}$ for the mean and $\approx 10^{-1}$ for the standard deviation), even for low Bézier element values, which are considerably lower than those from the Pcol expansion obtained with low oversampling ratios. Fig.3.7 also shows that when the oversampling ratio is increased locally (in each Bézier element) from $n_p = 1$ to $n_p = 4$, the smoothness of the convergence is slightly improved and the relative error values are reduced for both the mean and the standard deviation.

For a better interpretation of the relative errors presented above, the distributions of the mean and standard deviation of the water level along the channel length obtained from the BSBEM for different Bézier element numbers are shown in Fig.3.8. One can observe that errors due to the coarse B-spline approximation appear in the mean distribution profile obtained for $nx = 1$ in the discontinuity region (Fig.3.8a). As nx increases, the mean distribution gets smoother and becomes completely superimposed to that of the Monte Carlo reference solution. The

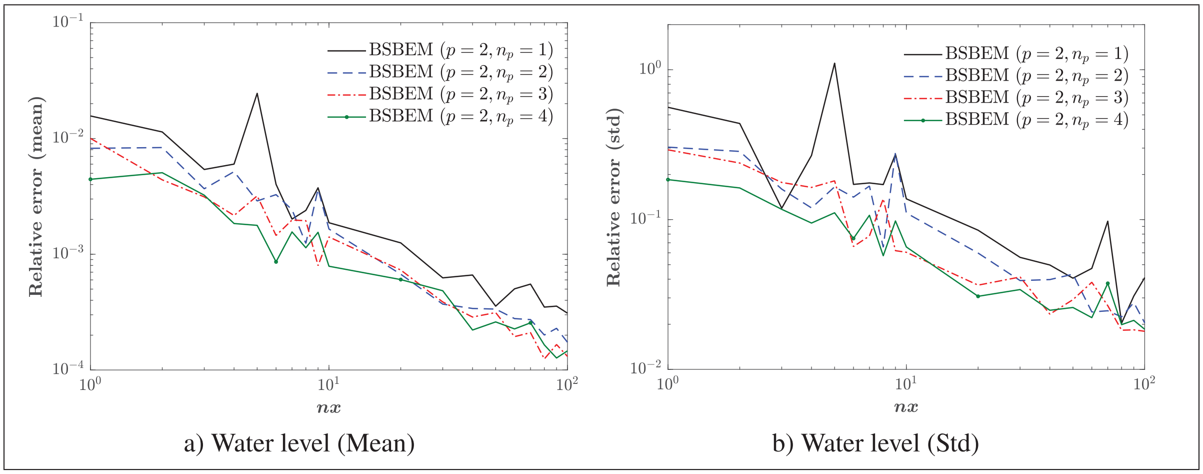


Figure 3.7 Water level error analysis when the number of Bézier elements of a BSBEM expansion are increased for different values of oversampling ratio. (a) : error in mean; (b) : error in standard deviation.

same oscillatory behaviour is also observed in the standard deviation profile, with more visible discrepancies in the discontinuity region where a high peak value can be seen for ($nx = 1$), thus leading to an over-prediction of the standard deviation at that location. As for the mean distribution, this effect is greatly reduced by increasing the number of Bézier elements, thereby allowing an accurate reproduction of the standard deviation distribution of the water level as shown in Fig.3.8b. The combination of the local over-sampling and relatively high number of Bézier elements shows a remarkable improvement in the prediction of the mean and standard deviation of the water depth, where a very good agreement can be observed between the BSBEM and Monte Carlo distributions for $nx = 100$ and $n_p = 4$ (Figs.3.8c and 3.8d). It should be pointed out that the observed errors are well-controlled using the BSBEM across the whole range of the tested configurations, unlike the Pcol approach which yields spurious oscillations and fails to converge.

Fig.3.9 presents a comparison of the mean and standard deviation distributions of the water level along the channel length obtained from the BSBEM and the Pcol using almost the same operating conditions at simulation time ($t = 4\ s$). It is clear that mean distributions are more

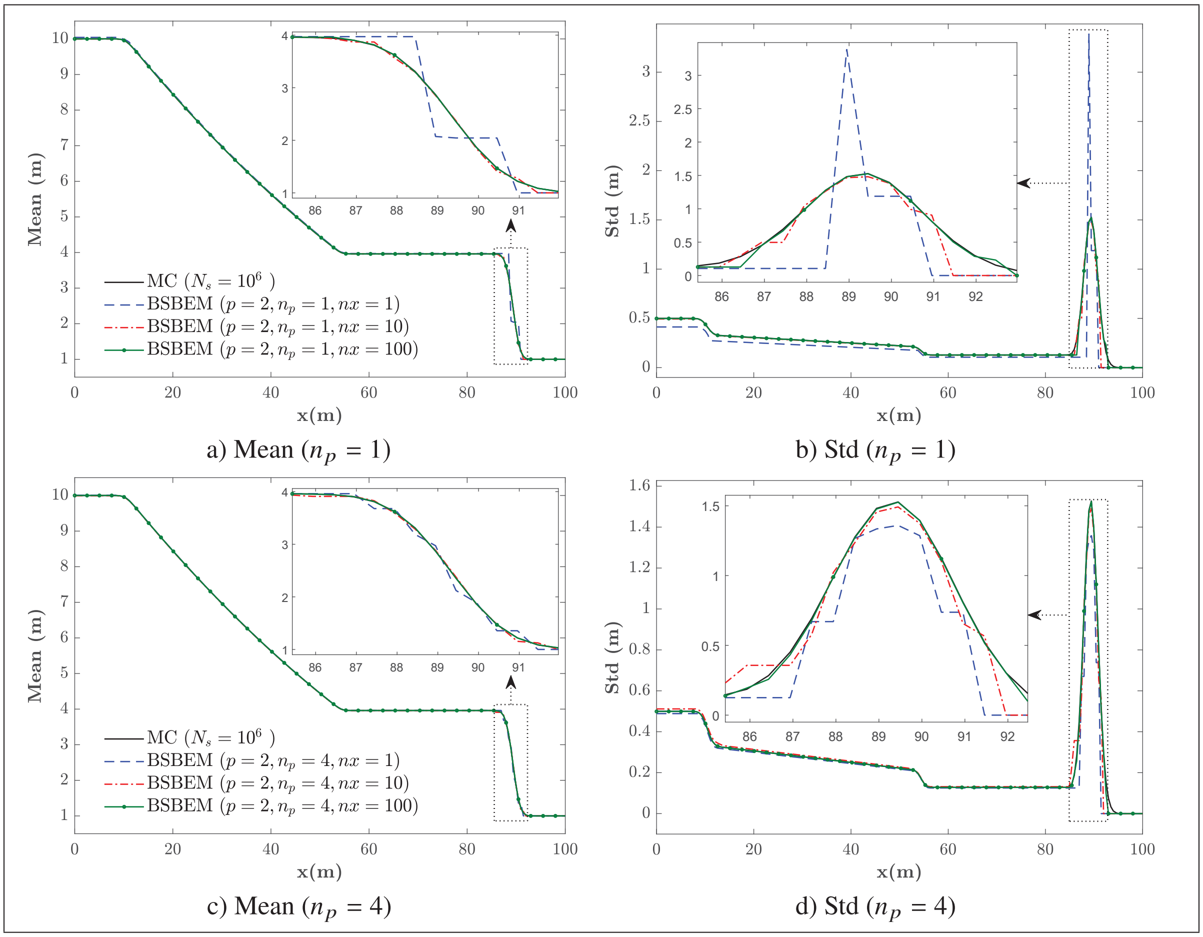
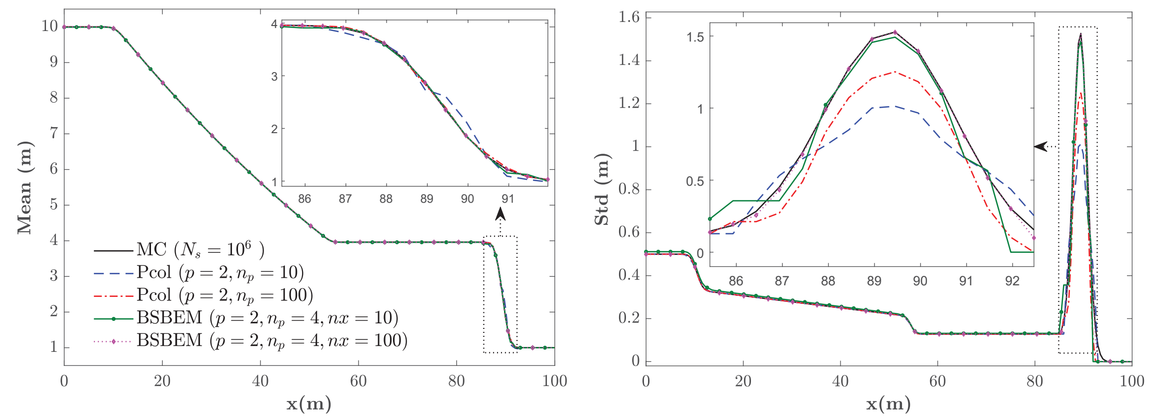


Figure 3.8 Distribution of the mean and standard deviation of a water level along the channel length at time $t = 4\ s$, obtained from the BSBEM (with $nx = \{1,\ 10,\ 100\}$) and MC (with $N_s = 10^6$ realizations) methods. (a) and (b) : $n_p = 1$; (c) and (d) : $n_p = 4$.

accurately reproduced by the BSBEM and Pcol methods than the standard deviation distributions. Indeed, by increasing the oversampling ratio, the profiles obtained from the Pcol expansion show a smooth distribution, but still cannot adequately match the reference MC distribution near the discontinuity where the peak value is under-predicted. In contrast, the distribution obtained from the BSBEM with $nx = 100$ illustrates an excellent superimposition with that of the MC solution over the whole domain study. It should be emphasized that the computational effort of the BSBEM in this univariate case, without local oversampling is almost equal to that of the Pcol method. However, when a local oversampling procedure is applied to each Bézier element, the computational effort required for the BSBEM is higher than that of the Pcol approach for the same polynomial order.



a) Water level (Mean) b) Water level (Std)

Figure 3.9 Distribution of the mean and standard deviation of the water level (a and b) along the channel length at time $t = 4\ s$, obtained from the Pcol (with $n_p = \{10,\ 100\}$), BSBEM (with $nx = \{10,\ 100\}$) and MC (with $N_s = 10^6$ realizations) methods.

3.3.2 Application to a hypothetical break of an actual dam

3.3.2.1 Description of the study case

The aforementioned modelling approaches have been applied for uncertainty propagation analysis regarding a hypothetical break of a 21 m height actual dam located on the Batiscan River (province of Québec), which takes its source in the Laurentides and discharging south into the Saint-Laurent River. The hydraulic behaviour was simulated by using the 1D Mascaret software (Goutal & Maurel, 2002) which is widely adopted to solve the one-dimensional form of the shallow water equations describing a dam-break wave. However, it should be noted that a growing number of 2D and 3D computational dam break models are being used (Ferrari *et al.*, 2009 ; Marsooli & Wu, 2014 ; Ozmen-Cagatay & Kocaman, 2011). The site under study, shown in Fig.3.10, is represented by a series of 74 cross sections describing the topography and the bathymetry of the terrain.

Figure 3.10 Map of the study area on the Batiscan River. Circle represents the location of the dam. Positions of cross-sections where results are depicted and discussed are shown on the map. (Source : Google Map)

The hydraulic parameters considered in building the deterministic model in Mascaret correspond to 1000-year flood episodes. The upstream condition is a fixed discharge during the simulation period, however, its value is uncertain. It is randomly selected following a normal distribution with given statistical moments (a mean value of 1 343 m^3/s, which corresponds to a 1000-year flood event, and a coefficient of variation of 5%). Thus, this variability of the upstream discharge, combined with those from other considered input random parameters, are propagated through the numerical model (Mascaret) to obtain the random output responses at selected cross-sections of interest. The corresponding water level of a value of 64.86 m represents the water level at the crest of the dam, at which the breach is initiated. The downstream boundary condition is described by a limnigram with a value of 4.2 m corresponding to the Saint-Laurent River's water depth under high tide conditions. The roughness parameters, which are represented in this study in terms of the Strickler coefficient (Ks), were equally distributed throughout the entire stretch and lumped into two values, one of the main channel and one of the floodplain, with their corresponding values of 30 and 20, respectively. Other parameters, such as breach characteristics, are considered as key parameters due to their great contribution to estimating the peak discharge during the breaking process. In the absence of observation evidence, various empirical models, based on several assumptions, have been proposed to estimate these characteristics. These estimations of the nominal values are prone to high uncertainty, which significantly impact the reliability of the dam-break analysis according to the results from some relevant investigations (Sammen *et al.*, 2017 ; Vorogushyn, Apel & Merz, 2011). In the current study, the dam's breach height is fixed at 18.13 m and its average width is considered equal to four times its height ($B_{av} = 72.52\ m$).

The input parameters described above may be subjected to high variability in their nominal values, which could induce several uncertainties in the numerical model and thus affect the reliability of the predictions of the output quantities of interest. In the present work, the upstream discharge, the roughness coefficient of the main channel and the average width of the breach have been identified as independent random variables with their a priori probability density functions (pdf) from which uncertainties are propagated throughout the one-dimensional Mascaret hydraulic model. The upstream discharge is described by a truncated normal distribution with a mean of 1344 m^3/s (which corresponds to the safety flood event discharge) and a coefficient of variation of 5 % ($std \approx 67.2\ m^3/s$). The pdf was truncated by a factor of $\pm\ 3std$ to avoid generating values of the input discharge outside the bounds of the rating curve of the dam's evacuator. The roughness parameter (Strickler coefficient) of the main channel is considered to be uniformly distributed within its plausible range, with a mean value of 30 and coefficient of variation of 5 % ($std \approx 1.5$). The average width of the breach is considered to follow a normal distribution with a mean of 72.52 m (which corresponds to the widely used value of 4 times the breach height), while the standard deviation is set to 3.6 m (coefficient of variation of 5 %). The other parameters, i.e., the roughness coefficient of the floodplain and the downstream water level at the junction with the Saint-Laurent River, are treated as deterministic parameters. The statistical characteristics of the aforementioned uncertain input parameters are summarized in Table 3.2.

Tableau 3.2 Statistical characteristics of the input uncertain parameters.

Parameter	Mean	Standard deviation	Cov (%)	Distribution
$Q_{up}\ \left[m^3/s\right]$	1 344	67.2	5	Normal
$Ks\ \left[s/m^{\frac{1}{3}}\right]$	30	7.5	5	Uniform
$B_{av}\ [m]$	72.52	3.6	5	Normal

Stochastic processing requires the generation of samples for random variables according to their probability density functions for a given sample size (N_s). A single value of each random parameter was selected to run the deterministic model, for which an in-house MATLAB code was developed to assign the selected values of the input discharge, the Strickler coefficient and the average width of the breach to their appropriate locations in the pre-process files used by the Mascaret software. This deterministic model was run for each combination of the aforementioned input parameters and the obtained output responses of quantities of interest were extracted from the Mascaret output files by the routines developed for statistical post-processing and surrogate modelling. Special treatment is required to take into account the variability of the average width of the breach in the deterministic model call. This breach is represented in Mascaret by an idealized shape defined by two vertical sides whose gap delineates the breach width, as shown in Fig.3.11, which illustrates a typical terrain profile at the location of the dam with a schematic representation of the breach. Therefore, once the value of the breach width is selected within the generated samples, the positions of both the left and right sides (points 1-2 and 3-4, respectively, in Fig.3.11) are updated accordingly using Eq.(3.22) :

$$\begin{cases} x_{1,2}^{(i)} = x_0 - \frac{B_{av_i}}{2} \\ \\ x_{3,4}^{(i)} = x_0 + \frac{B_{av_i}}{2} \end{cases} \tag{3.22}$$

where x_0 represents the position of the breach centre fixed with the deterministic value and B_{av_i} denotes the selected value of the average breach width from the generated sample in the random space. After this updating, these positions are then replaced in the input geometric file used by Mascaret to run the deterministic model.

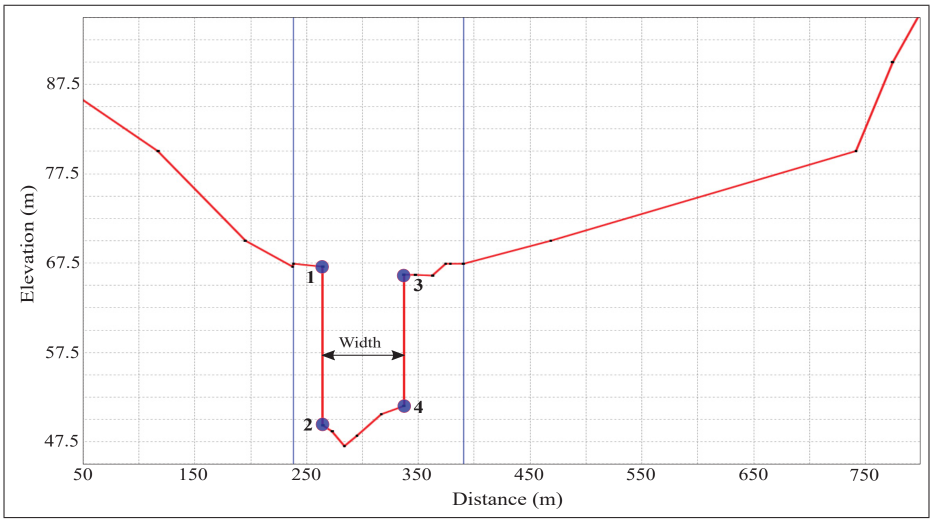


Figure 3.11 Typical cross-sectional profile at the dam location with a schematic representation of the breach. Points 1-2 and 3-4 represent the left and right vertical sides of the breach, respectively. The blue vertical lines indicate the left and right banks.

3.3.2.2 Statistical moments predictions

The results are presented here mainly in terms of statistical moment profiles (mean and standard deviation) of the downstream discharge and water level at selected cross-sections. In the absence of the analytical solution of such complex situations, the Monte Carlo method was considered as a benchmark reference solution to which other stochastic expansion approaches (BSBEM and Pcol) were compared to evaluate their accuracy. It is well known that the convergence of the Monte Carlo method strongly depends on the sample size, whose convergence rate is on the order of $\frac{1}{\sqrt{N_s}}$, which requires a relatively high number of realizations for suitable accuracy. Thus, the required optimal number of samples has to be defined by conducting a convergence study of the standard deviation of the output quantities of interest (water level and downstream discharge) as a function of the number of realizations.

Fig.3.12 shows the convergence of the standard deviation of the water level and discharge at specific locations downstream of the dam (x = 8 168 m and 14 555 m) for two simulation times (t = 1.5 h and 3 h) versus the number of MC realizations. The total number of 50 000 MC realizations is retained in this study to generate results that can be considered as reference solutions to use in assessing the reliability of other approaches. The runtime required to achieve 50 000 MC realizations using a personal computer (Intel (R) Xeon(R) CPU E3-125 v6 @3.30 GHZ) is approximately 359 hours.

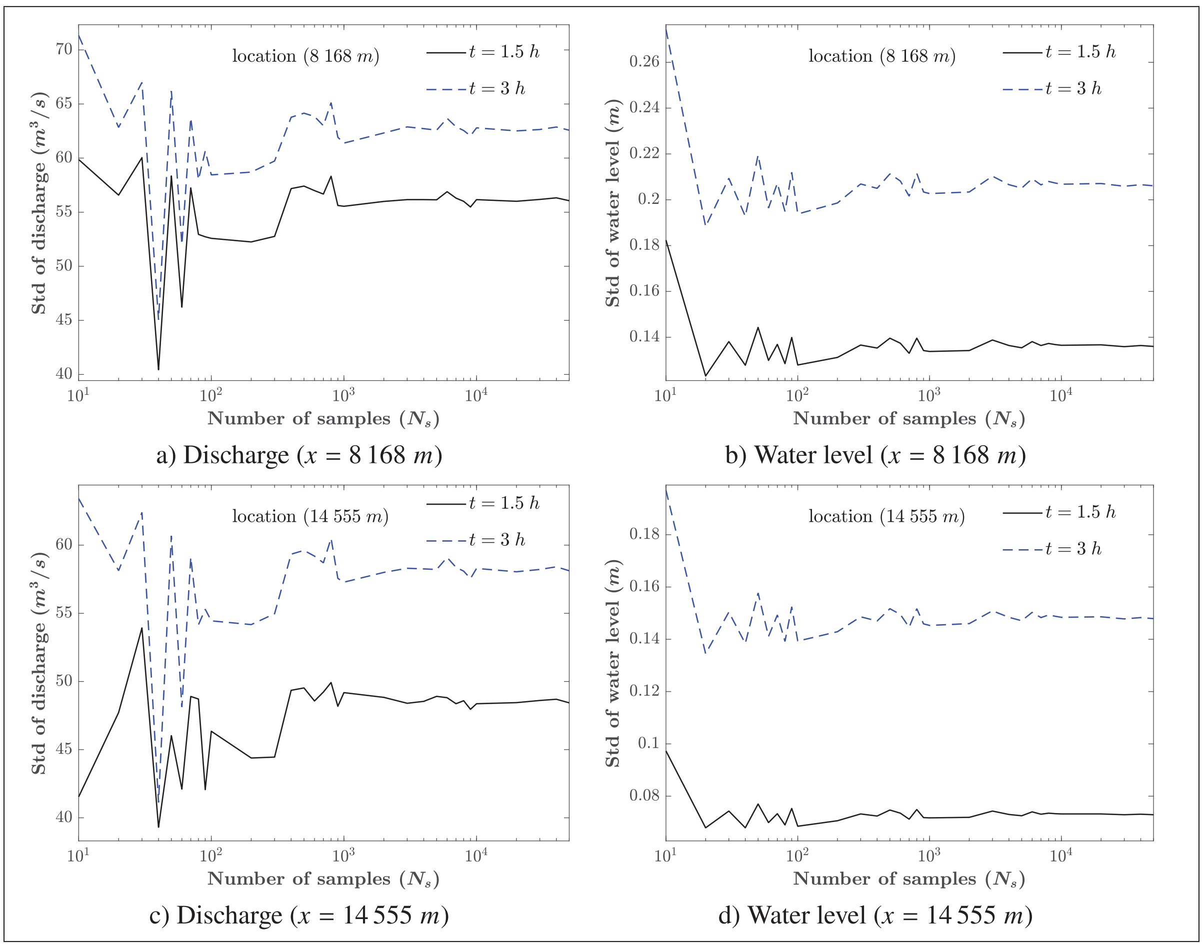
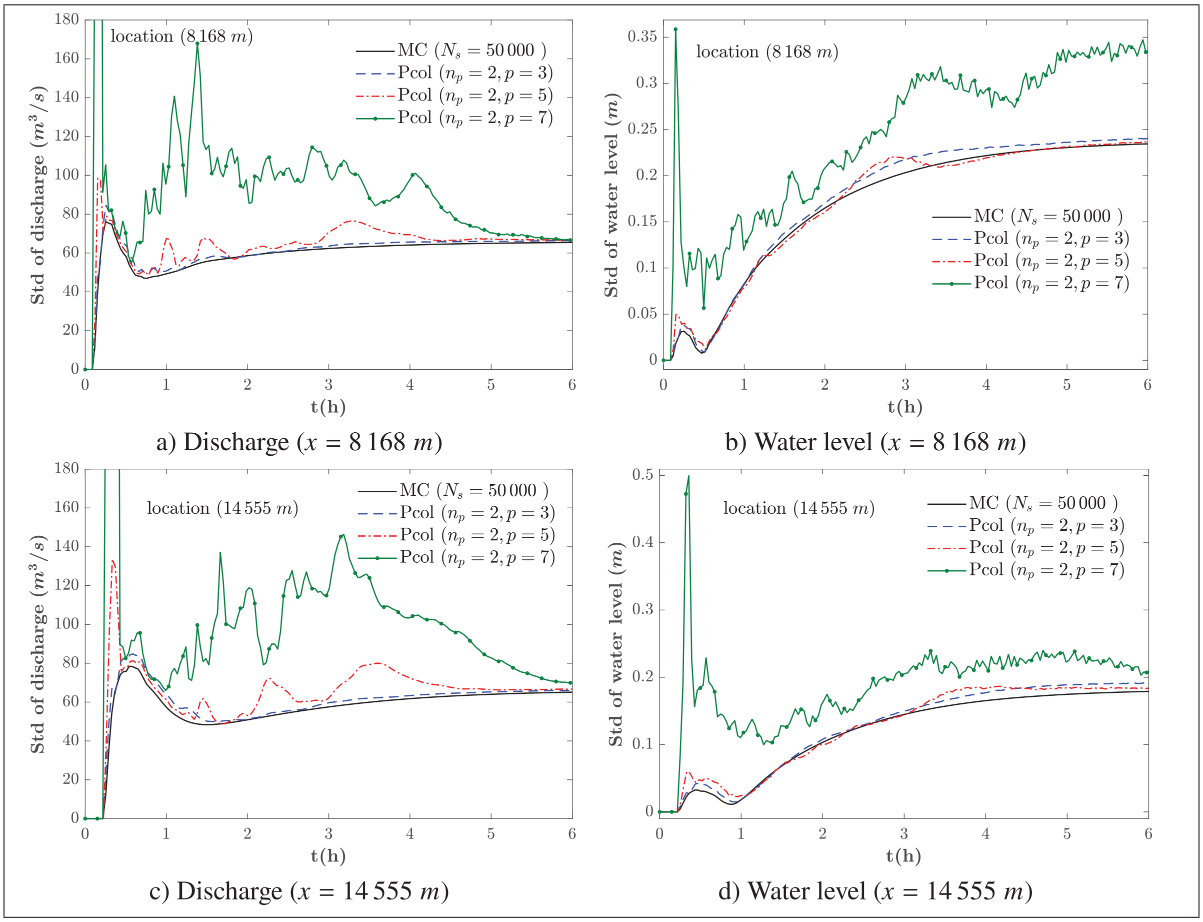


Figure 3.12 Convergence of the sample standard deviation of discharge (left) and water level (right) as functions of the number of Monte Carlo realizations (N_s) at times t = 1.5 h and t = 3 h and different locations : (a and b) : x = 8 168 m, (c and d) : x = 14 555 m.

Fig.3.13 displays the variation of the standard deviations of the discharge and of the water level as a function of time at two locations ($x = 8\,168$ m and $14\,555$ m) obtained from Pcol expansion with various polynomial orders. The results are presented for a value of oversampling ratio that is twice the minimum required number of collocation points ($n_p = 2$), which is highly recommended for smooth continuous problems (Hosder & Walters, 2010; Hosder *et al.*, 2007). As can be seen from these figures, slight oscillations appear near the front wave when a relatively low value of polynomial order ($p = 3$) is used. The intensity of these oscillations increases greatly as the polynomial order is increased from $p = 3$ to $p = 7$, where a strong oscillatory behaviour is observed with high oscillations amplitude, leading to an overestimation and non-physical value of the statistical moment for both downstream discharge (Figs.3.13a and 3.13c) and water level (Figs.3.13b and 3.13d). It should be emphasized that for smooth continuous problems, once the oversampling ratio is fixed to $n_p = 2$, any increase in the polynomial order leads to a substantial improvement in the accuracy of the PCE approximation, conversely to the present case where the polynomial chaos expansion cannot effectively describe the high hyperbolic behaviour of the flooding wave. Indeed, as shown in Fig.3.13, any increase in the polynomial order leads to the appearance of spurious oscillations with non-physical values of the output responses of quantities of interest.

To illustrate the ability of the BSBEM to accurately approximate the statistical moments of the output quantities, Fig.3.14 shows the evolution of the standard deviation profiles of the downstream discharge and water level at selected locations ($x = 8\,168$ m and $x = 14\,555$ m) as a function of time. The results obtained from the BSBEM with a fixed polynomial order of the basis functions ($p_i = 2, i \in \{1, 2, 3\}$) and various numbers of Bézier elements are compared to those from the 50 000-realization Monte Carlo reference solution. These plots reveal that even for a small value of the number of Bézier elements ($nx_i = 2, i \in \{1, 2, 3\}$), the results show a smooth profile, even in regions surrounding the front wave. The increase in the number of Bézier elements from $nx_i = 2$ to 5 leads to a real improvement in the BSBEM's accuracy, to where an excellent agreement is observed with those from the 50 000 MC realizations. This improvement



a) Discharge ($x = 8\,168\ m$)

b) Water level ($x = 8\,168\ m$)

c) Discharge ($x = 14\,555\ m$)

d) Water level ($x = 14\,555\ m$)

Figure 3.13 Standard deviations of the discharge (left) and water level (right) as functions of time obtained from Pcol with various polynomial orders ($n_p = 2$, $p = 3,\ 5,\ 7$), and by the Monte Carlo method (with 50 000 realizations) : (a and b) : $x = 8\,168\ m$, (c and d) : $x = 14\,555\ m$.

is the result of an appropriate split of the random space of each variable into nx_i sub-domains from which the minimum required number of collocation points ($N_s^e = \prod_{i=1}^{3}(p_i+1)$) are selected. This procedure provides an efficient distribution of the collocation points within the random space, thereby considerably improving the accuracy of the approximation.

The abilities of the Pcol and BSBEM approaches to fit the statistical moments obtained by the reference Monte Carlo method are evaluated together in Fig.3.15, which shows the confidence interval profiles (CI), quantified in terms of the mean plus and minus two standard deviations

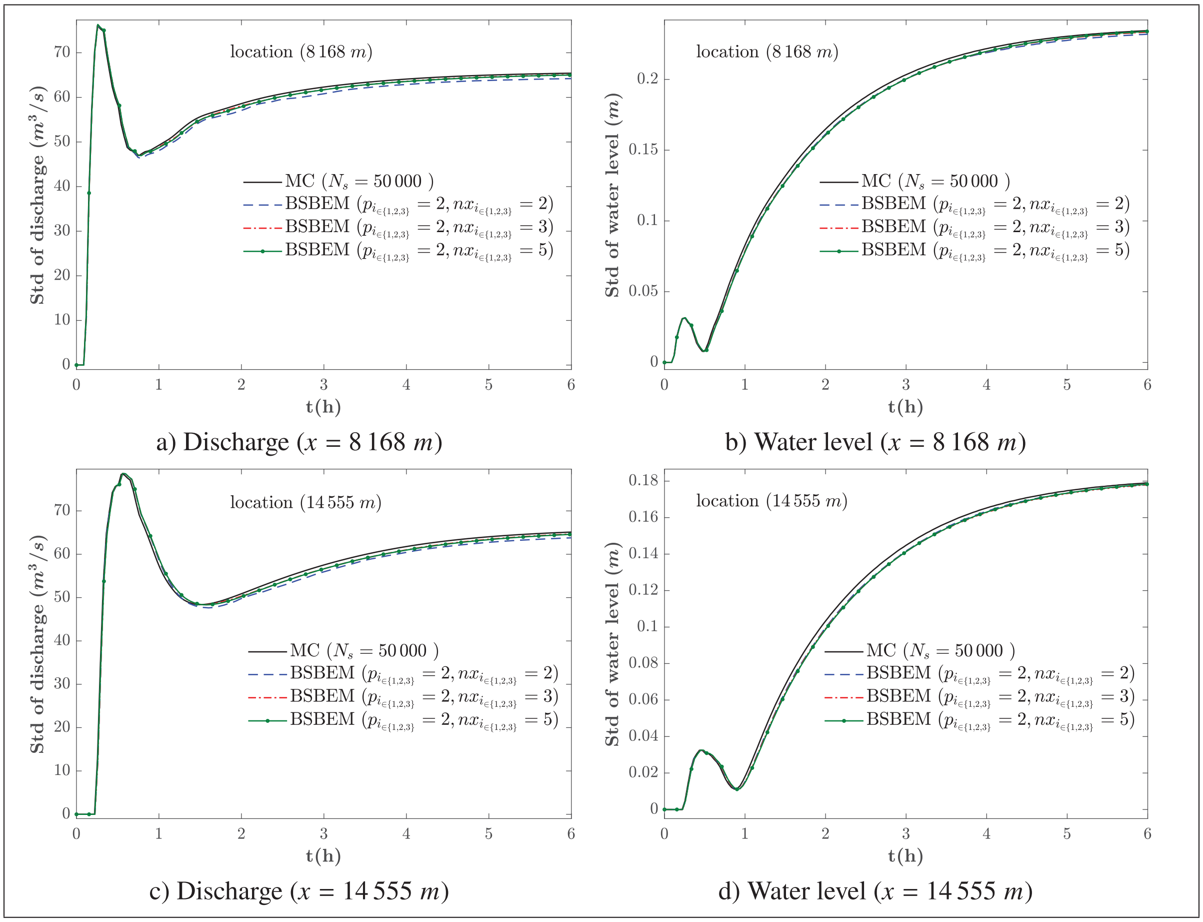
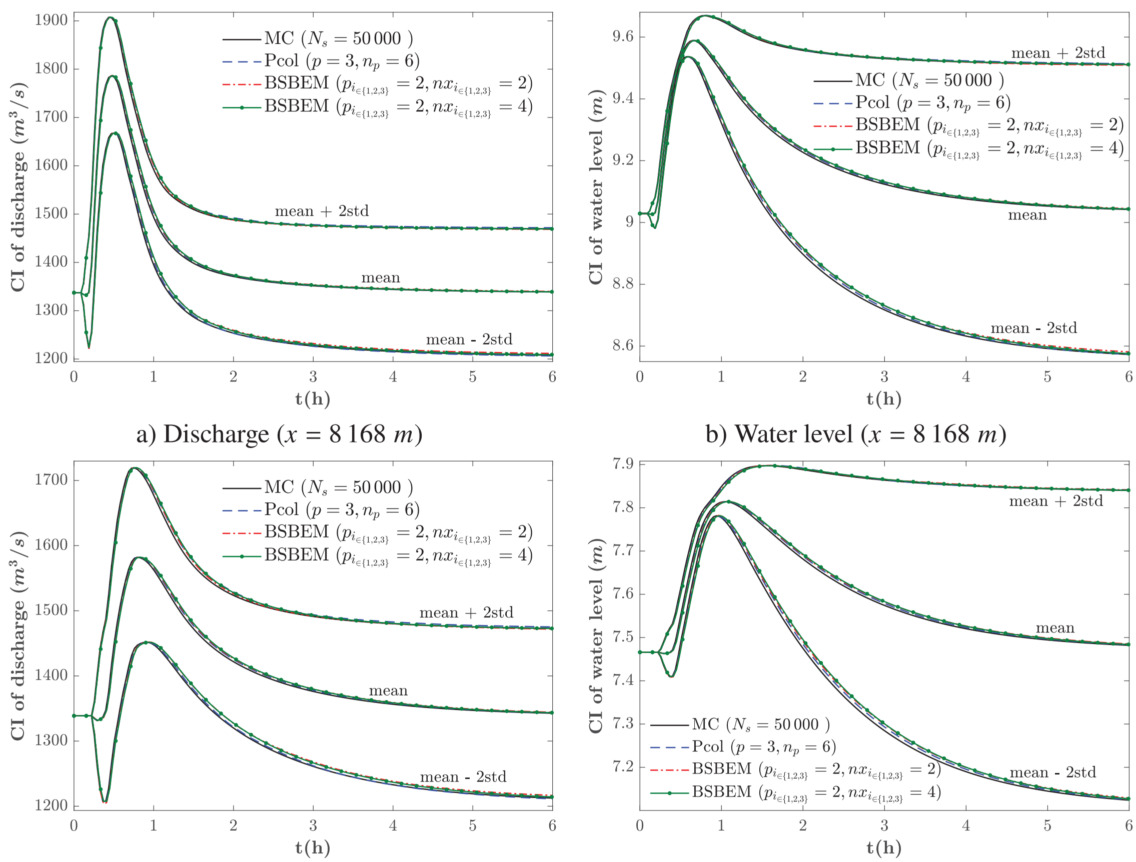


Figure 3.14 Standard deviations of discharge (left) and water level (right) as a function of time obtained from the BSBEM with various Bézier element numbers ($p_i = 2$, $nx_i = 2,\ 3,\ 5$) and from the Monte Carlo method (with 50 000 realizations) : (a and b) : $x = 8\,168\ m$, (c and d) : $x = 14\,555\ m$.

of downstream discharge and water level as a function of time at selected cross-sections ($x = 8\,168\ m$ and $14\,555\ m$). These plots clearly indicate that the statistical moment profiles obtained with the Pcol expansion (with $p = 3$ and $n_p = 6$) and with the BSBEM approach (with $p_i = 2$ and $nx_i = 4$, $i \in \{1,\ 2,\ 3\}$) present an excellent agreement with those from the Monte Carlo method with 50 000 realizations. The propagation of the inundation wave is outlined by the evolution in time for the peak value of discharge and water level, which progresses downstream from one cross section to another. This progression is accompanied by a high variability range around the mean, as illustrated by the confidence interval delineated by upper and lower bounds, which transmit the effect of uncertainty propagation stemming from the input parameters into the output responses of the quantities of interest. The confidence interval of the water level may be used to establish the probability inundation map. By calculating the confidence interval of the uncertain free surface from the uncertain water level and the terrain profile, and projecting its lower and upper bounds, at a selected time, onto the topography of the terrain, it is possible to provide an estimated range of the inundated banks of the selected cross section. This probabilistic inundation mapping is especially worthwhile, as it considers the variability that may be introduced by the input parameters.



a) Discharge (x = 8 168 m)

b) Water level (x = 8 168 m)

c) Discharge (x = 14 555 m)

d) Water level (x = 14 555 m)

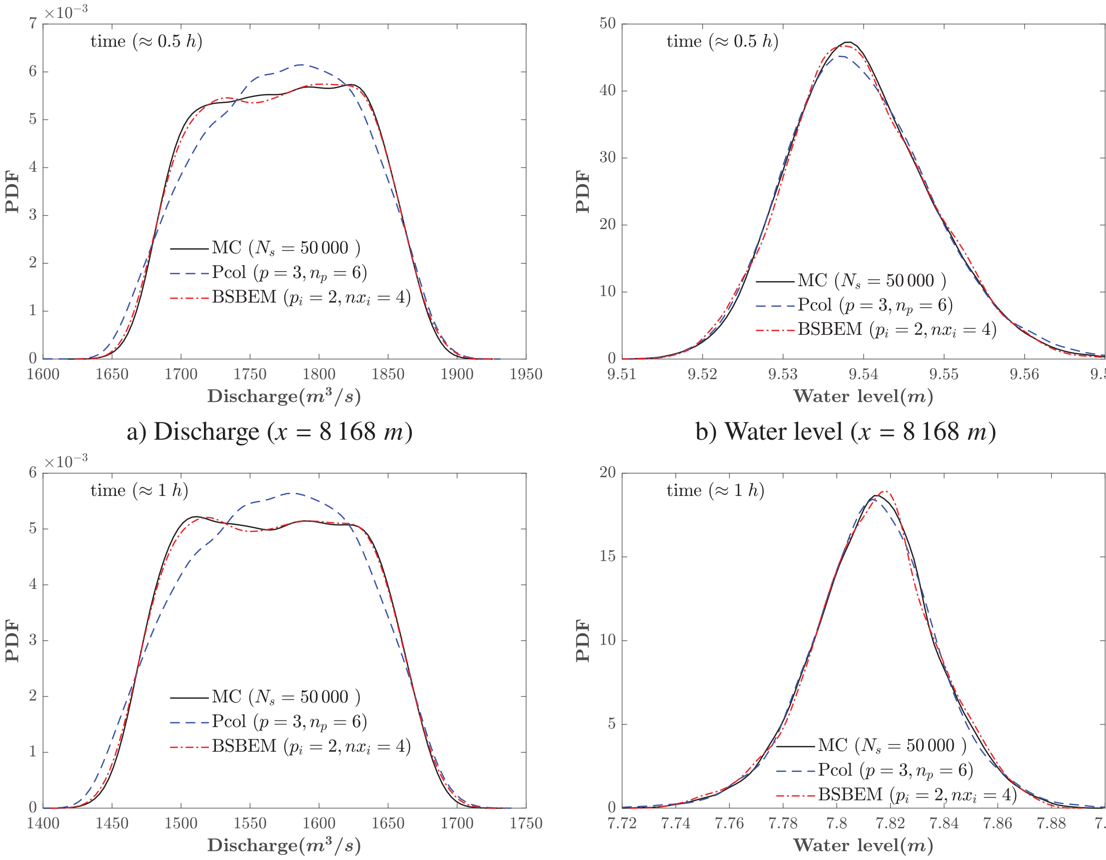
Figure 3.15 Evolution of the confidence interval of discharge (left) and of the water level (right) as a function of time, obtained from the BSBEM (with $p_i = 2$, $nx_i = 2,\ 4$), Pcol (with $p = 3$, $n_p = 6$) and Monte Carlo (with 50 000 realizations) methods : (a and b) : $x = 8\,168\ m$, (c and d) : $x = 14\,555\ m$.

3.3.2.3 Estimated probability density function

In this subsection, probability density function of the downstream discharge, water level and front wave arrival time, at selected cross-sections, obtained from Pcol and BSBEM are presented and discussed in comparison with those from the Monte Carlo and LHS sampling methods.

Fig.3.16 shows the probability density function of the downstream discharge and water level at selected cross-sections ($x = 8\,168\ m$) and ($x = 14\,555\ m$) at simulation times $t = 0.5\ h$ and $t = 1\ h$, respectively, which closely correspond to the time of the peak discharge and water level describing the inundation wave. The Monte Carlo sampling method is applied by using the constructed surrogate model from Pcol (with $p = 3,\ n_p = 6$) and BEBEM (with $p_i = 2$, $nx_i = 4$) to estimate Pdfs of the output quantities of interest. The results are then compared with those computed from the 50 000 Monte Carlo realizations using the direct deterministic model. The distributions of the downstream discharge show that the pdf resulting from the BSBEM is in excellent agreement with the pdf obtained from direct MC simulations, unlike the Pcol expansion which fails to approximate PDFs, as shown in Figs.3.16a and 3.16c. The deviation observed in the discharge distribution is less visible in the probability distribution of the water level pdfs of the surrogate models, and so they compare well with the reference Monte Carlo pdf, whose shape is almost Gaussian as illustrated in Figs.3.16b and 3.16d. Thus, the obtained results clearly show the strong abilities of the BSBEM to closely approximate a wider range of PDFs than the Pcol expansion.

To further illustrate the BSBEM's ability to well estimate the probability density function of the outputs, in addition to the statistical moments, the probability distributions of the arrival time of the front wave at different selected locations over the reach are depicted in Fig.3.17. It is important to note that in this work the arrival time of the front wave is defined as the time recorded by the deterministic solver (Mascaret) when the water level at a given downstream cross-section is raised by 30 cm compared to the initial water level. Fig.3.17 shows that the



a) Discharge (x = 8 168 m) b) Water level (x = 8 168 m)
c) Discharge (x = 14 555 m) d) Water level (x = 14 555 m)

Figure 3.16 Probability distribution of the discharge (left) and water level (right) estimated with Pcol ($p = 3,\ n_p = 6$) and BSBEM ($p_i = 2,\ nx_i = 4$) surrogate models and compared to the Monte Carlo sampling estimation (with $N_s = 50\,000$ realizations) : (a and b) at location ($x = 8\,168\ m$) and time ($t = 0.5\ h$) ; (a and b) at location ($x = 14\,555\ m$) and time ($t = 1\ h$).

pdf resulting from the BSBEM surrogate is almost superimposed with that resulting from the reference Monte Carlo method, unlike the distributions obtained using the Pcol surrogate model where significant deviations can be observed. The uncertainty accompanying the wave arrival time increases downstream of the dam, as illustrated by the wider range of the probability density function, which does not follow a normal distribution.

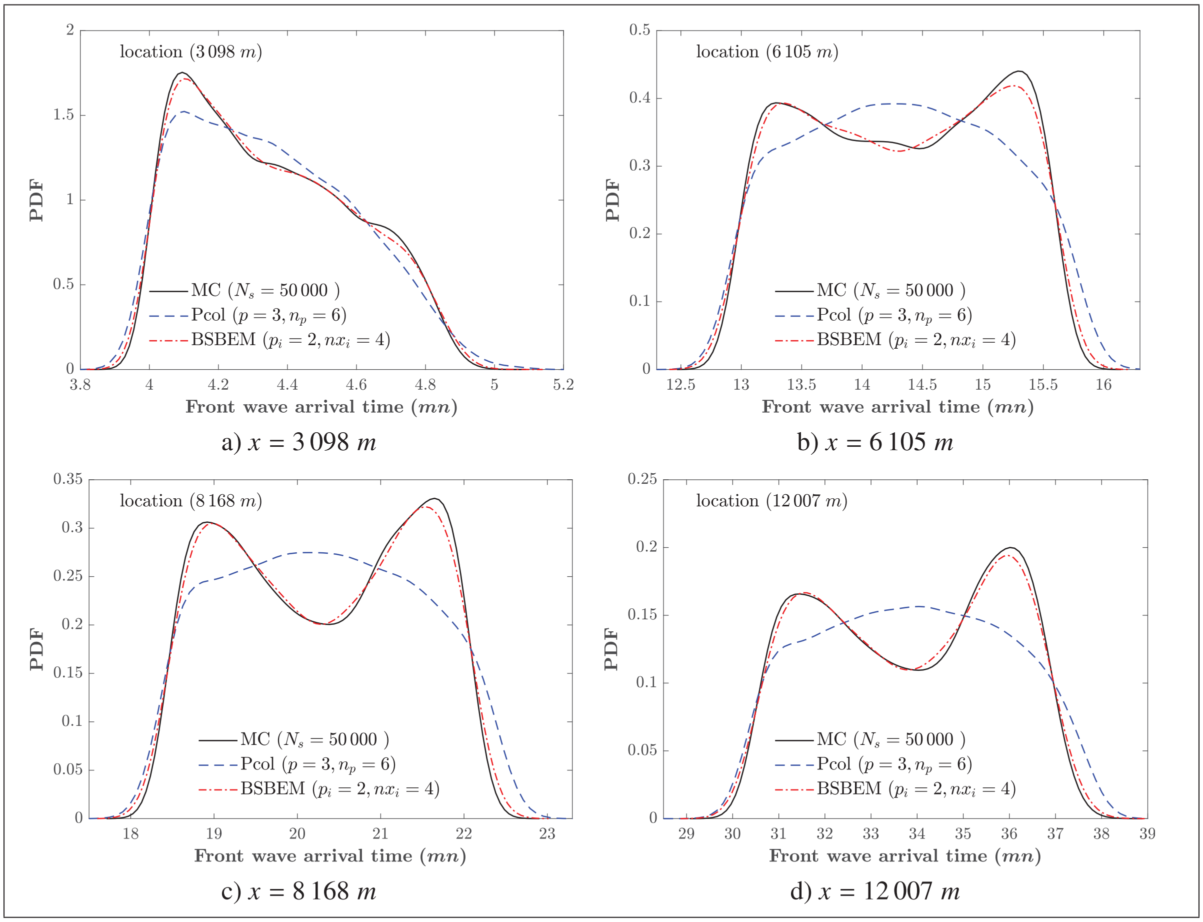
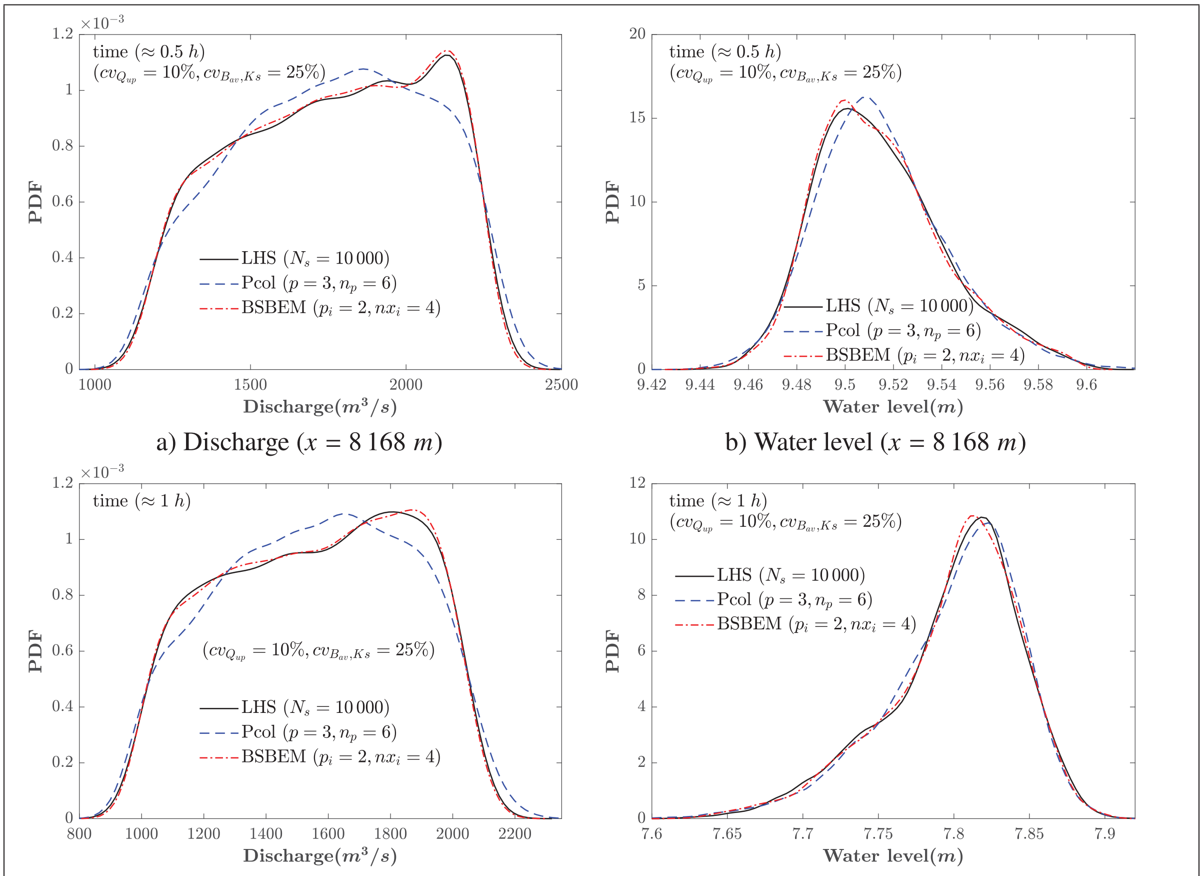


Figure 3.17 Probability distribution of the front wave arrival time estimated with Pcol ($p = 3,\ n_p = 6$) and BSBEM ($p_i = 2,\ nx_i = 4$) surrogate models and compared to the Monte Carlo sampling estimation (with $N_s = 50\,000$ realizations) : (a) $x = 3\,098\ m$; (b) $x = 6\,105\ m$, (c) $x = 8\,168\ m$ and (d) $x = 12\,007\ m$.

The pdfs shown in Figs.3.16 and 3.17 are obtained for a low variability in the input parameters with coefficients of variation of $cv = 5\,\%$. In order to show the BSBEM's ability to accurately reproduce the pdfs of the outputs with a relatively high variability of the inputs, the effect of increasing the coefficient of variation of the input parameters is depicted in Figs.3.18 and 3.19. Indeed, the upstream discharge values are generated by a normal distribution $\mathcal{N}(1\,344, 134.4)$ with a coefficient of variation $cv_{Q_{up}} = 10\,\%$, truncated at 950 and 1745 m^3/s. The average breach width is described by a normal distribution with a coefficient of variation of $cv_{B_{av}} = 25\,\%$ which yields $\mathcal{N}(72.52, 18.13)$ truncated at 36 and 109 m, which approximately corresponds to $2H$ and $6H$, respectively, where H represents the dam's breach height (Wahl, 2004). The variability of the Strickler coefficient of the river's main channel is represented by a uniform distribution $\mathcal{U}[17, 43]$ with a coefficient of variation $cv_{Ks} = 25\,\%$ (Johnson, 1996).

Fig.3.18 shows the probability distribution functions of the downstream discharge and water level at two selected cross-sections and simulation times, where a good agreement can be observed between the pdfs obtained from BSBEM surrogate model and those computed from the 10 000 LHS realizations (considered as a reference solution), unlike the Pcol surrogate model, where deviations can be detected, particularly in the output discharge distributions at both selected locations. The effect of increasing the variability of the input parameters is clearly visible in the uncertainty of the output discharge which is traduced by a high variability range. Indeed, as an example, we can clearly see the increase in the variability range of the output discharge at $x = 14\,555\,m$, from $\approx$ $[1420, 1730]$ m^3/s for a coefficient of variation of 5 %, to considerably widen to $\approx$ $[800, 2300]$ m^3/s for $cv_{Q_{up}} = 10\,\%$ and $cv_{B_{av},Ks} = 25\,\%$, as illustrated by Figs.3.16c and 3.18c, respectively.

The same behaviour is observed in the probability distributions of the front wave arrival time at selected locations over the studied area shown in Fig.3.19. Overall, an excellent agreement can be clearly observed between the pdfs from the BSBEM surrogate model and those from the LHS realizations, where the profiles are almost superimposed, with the exception of the results obtained at the location $x = 3\,098$ m, where some slight discrepancies can be noted between both distributions (BSBEM and LHS). On the other hand, one can clearly note that the Pcol surrogate model fails to reproduce the shape of the front wave arrival time distributions over the whole selected locations. The uncertainty related to the arrival time of the front wave, which increases as the wave evolves through the studied reach, is greatly affected by the increase in the uncertainty of the input parameters, as illustrated by the variability range of the distribution functions of the front wave arrival time, shown in Figs.3.17d and 3.19d, which correspond, respectively, to the low and high coefficient of variation of the input parameters.



a) Discharge ($x = 8\,168\ m$) b) Water level ($x = 8\,168\ m$)

c) Discharge ($x = 14\,555\ m$) d) Water level ($x = 14\,555\ m$)

Figure 3.18 Probability distribution of the discharge (left) and water level (right) estimated with Pcol ($p = 3,\ n_p = 6$) and BSBEM ($p_i = 2,\ nx_i = 4$) surrogate models and compared to the LHS estimation (with $N_s = 10\,000$ realizations), obtained with a relatively high variability of the input parameters ($cv_{Q_{up}} = 10\%,\ cv_{B_{av},Ks} = 25\%$) : (a and b) at location ($x = 8\,168\ m$) and time ($t = 0.5\ h$) ; (c and d) at location ($x = 14\,555\ m$) and time ($t = 1\ h$).

3.3.2.4 Computational efficiency

This subsection provides a computational efficiency analysis of the BSBEM approach in comparison to the Pcol and the LHS methods. It should be emphasized that all programs involved in this paper were implemented serially and the computations were performed on a personal computer : Intel (R) Xeon(R) CPU E3-125 v6 @3.30 GHz and 16 GB memory.

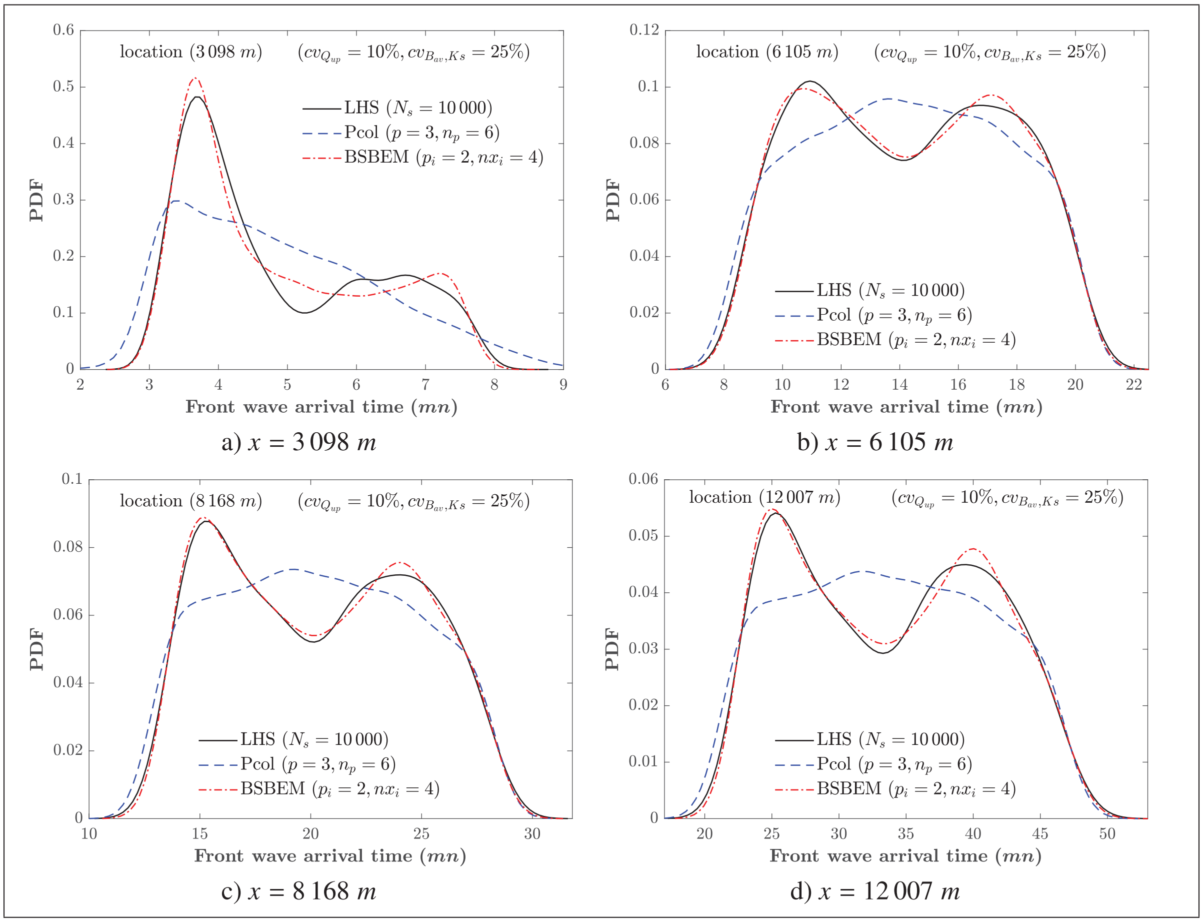


Figure 3.19 Probability distribution of the front wave arrival time estimated with Pcol ($p = 3,\ n_p = 6$) and BSBEM ($p_i = 2,\ nx_i = 4$) surrogate models and compared to the LHS estimation (with $N_s = 10\,000$ realizations), obtained with a relatively high variability of the input parameters ($cv_{Q_{up}} = 10\%,\ cv_{B_{av},Ks} = 25\%$) : (a) $x = 3\,098\ m$; (b) $x = 6\,105\ m$, (c) $x = 8\,168\ m$ and (d) $x = 12\,007\ m$.

The computational efficiency can be estimated, among other ways, by comparing the number of numerical solver runs. Thus, for instance, a BSBEM with $p_i = 2$, $nx_i = 2$, and $nx_i = 4$ requires 216 and 1 827 runs, respectively , which makes the BSBEM 231 and 30 times, respectively, faster than the Monte Carlo sampling method (with 50 000 runs). The Pcol approach with $p = 2$ an $n_p = 6$ requires 84 model evaluations, which makes it 595 times faster than the Monte Carlo method. Another aspect of the comparison concerns the required cost to approximate the pdfs of the outputs. Thus, the computed pdf with the direct Monte Carlo method requires a high computational cost (359 hours to achieve 50 000 realizations), while for the Pcol and

BSBEM approaches, the computational time is on the order of minutes (13 minutes for BS-BEM and 14 minutes for Pcol to perform 50 000 realizations with the generated surrogate models).

It is worth mentioning that the gain in the computational effort with BSBEM is expected to be more substantial for cases involving 2D and 3D configurations with a computational domain consisting of millions of nodes in some applications, where the classical sampling techniques (MC, LHS . . .) may require a huge computational cost.

3.4 Conclusion

This study describes the implementation of the non-intrusive B-splines Bézier element method (BSBEM) for uncertainty propagation analysis for a hypothetical rupture of a real dam. The method, which belongs to the multi-element approaches, is based on splitting the random space of the input parameters, which corresponds to the parametric domain, into subspaces-called Bézier elements, within which the stochastic output response is approximated by local B-spline basis functions. The coefficients allowing this local approximation are calculated non-intrusively by the regression technique.

Stoker's analytical solution of the downstream water level and velocity for an idealized dam break flow was first considered as a test case to assess the BSBEM's performance compared to that of the polynomial chaos-based Pcol expansion. The variation of L^2 error norms of mean and standard deviation for both methods has clearly revealed the limitation of PC expansion, particularly for relatively high polynomial order, to accurately approximate the statistics of problems where discontinuity occurs in the output quantities, where a strong divergence in the L^2 error is observed. Our assessment also showed that using an oversampling ratio beyond the recommended value for smooth problems, or twice the number of expansion terms, combined with a lower polynomial order, contributes to mitigating the divergence rate as much as possible. In contrast, the BSBEM showed a better and monotonic convergence rate of the mean and standard deviation errors as the number of Bézier elements is increased, producing smooth profiles of the water level and velocity.

The introduced method was then applied to analyze the propagation of uncertainties over numerical models describing the flood wave resulting from a hypothetical break of an actual dam located in the province of Québec. The variability of the input parameters, namely the upstream discharge, the Strickler roughness coefficient of the main channel, and the average breach width, was defined in accordance with their corresponding probability distributions around the deterministic nominal values. The generated uncertainties were then propagated through the one-dimensional hydrodynamic deterministic model and their impact on the output response quantified in terms of the statistical moments and probability density functions of the downstream water level and discharge as a function of time at different locations of the considered reach. The results reveal that the BSBEM provides smooth profiles of the standard deviation and confidence interval as a function of time, accurately describing the variability accompanying the flood wave propagation at different locations. Furthermore, the predictions obtained using the BSBEM were found to be in excellent agreement with those from the 50 000 Monte Carlo and 10 000 LHS realizations, unlike the Pcol expansion that estimated spurious oscillations. This oscillatory behaviour was slightly mitigated by using a low polynomial order and a relatively high oversampling ratio. Another meaningful result stemming from this study is the remarkable efficiency of the BSBEM to successfully approximate the probability distribution of the outputs, unlike the Pcol approach which cannot accurately predict these pdfs, especially for the front wave arrival time.

This study's main findings indicate that the BSBEM presents an efficient tool for dealing with uncertainty propagation through complex dam break flow models, thereby making it possible to establish probabilistic inundations maps for flood hazard predictions that takes into account uncertainty stemming from input parameters. The applicability of the presented approach can be successfully extended to other complex system with hyperbolic, non-smooth or even discontinuous behaviour. For high-dimensional problems, the total number of collocation points is greatly affected by the number of Bézier elements and the number of random variables, which could present a drawback in terms of computational cost (curse of dimensionality). More research effort is still required to solve the problem of curse of dimensionality. In the BSBEM, the curse of dimensionality is inherited from the use of a simple tensor product of a one-dimensional basis to build the multivariate basis. This aspect highlights an area of potential research improvements. One idea that could be pursued to overcome (at least partially) the problem of curse of dimensionality, is the use of local refinements while still using the tensor product. For instance, if more accuracy is required for the tails of the CDF, the number of elements could be increased locally. Therefore, patches will constitute the mesh, and nx_i will be defined for each patch. The use of unstructured splines, as in U-splines (Thomas, Engvall, Schmidt, Tew & Scott, 2018) or T-splines (Sangalli, Takacs & Vázquez, 2016 ; Scott *et al.*, 2011 ; Sederberg, Zheng, Bakenov & Nasri, 2003) is another idea to be investigated (though not trivial for a dimension greater than 3). Nevertheless, once the surrogate model based on BSBEM has been built, the computational effort, which is on the order of minutes, remains insignificant compared to that of the standard Monte Carlo sampling method.

CHAPITRE 4

A NON-INTRUSIVE REDUCED-ORDER MODELING METHOD FOR UNCERTAINTY PROPAGATION OF TIME-DEPENDENT PROBLEMS USING A B-SPLINES BÉZIER ELEMENTS-BASED METHOD AND PROPER ORTHOGONAL DECOMPOSITIONS : APPLICATION TO DAM-BREAK FLOWS

Azzedine Abdedou, Azzeddine Soulaïmani

Department of Mechanical Engineering, Ecole de Technologie Supérieure, 1100 Notre-Dame W., Montreal (QC), Canada H3C 1K3

Article soumis à la revue « Journal of Computational Physics » le 11 Décembre 2020.

Abstract

A proper orthogonal decomposition-based B-splines Bézier elements method (POD-BSBEM) is proposed as a non-intrusive reduced-order model for uncertainty propagation analysis for stochastic time-dependent problems. The method uses a two-step proper orthogonal decomposition (POD) technique to extract the reduced basis from a collection of high-fidelity solutions called snapshots. A third POD level is then applied on the data of the projection coefficients associated with the reduced basis to separate the time-dependent modes from the stochastic parametrized coefficients. These are approximated in the stochastic parameter space using B-splines basis functions defined in the corresponding Bézier element. The accuracy and the efficiency of the proposed method are assessed using benchmark steady-state and time-dependent problems and compared to the reduced-order model-based artificial neural network (POD-ANN) and to the full-order model-based polynomial chaos expansion (Full-PCE). The POD-BSBEM is then applied to analyze the uncertainty propagation through a flood wave flow stemming from a hypothetical dam-break in a river with a complex bathymetry. The results confirm the ability of the POD-BSBEM to accurately predict the statistical moments of the output quantities of interest with a substantial speed-up for both offline and online stages compared to other techniques.

keywords

Uncertainty propagation, Proper orthogonal decomposition, B-Splines Bézier elements method, Dam-break flows

4.1 Introduction

The recent developments in computational power have led to great improvements in numerical modeling, leading to the solutions of more and more complex physical problems in several engineering fields. However, when these problems are related to large scale configurations and involve mathematical models expressed as time-dependent and parametrized governing equations whose discretization may be significantly demanding in terms of the computational efforts required (Bērziņš, Helmig, Key & Elgeti, 2020 ; Sanderse, 2020). This situation becomes all the more demanding when the input parameters are subject to uncertainties that propagate through numerical solvers (Kalinina, Spada, Vetsch, Marelli, Whealton, Burgherr & Sudret, 2020), and which must be considered, further increasing the requirements both in terms of CPU time and memory demands. To address this issue, considerable efforts have been dedicated over the last decades to develop reduced-order modeling methods (ROM) to approximate the original full-order model based on one with a significantly reduced dimensionality (Georgaka, Stabile, Star, Rozza & Bluck, 2020 ; Zokagoa & Soulaïmani, 2018).

One of the widely-used methods to construct such a reduced-order model is proper orthogonal decomposition (POD) (Chatterjee, 2000 ; Sirovich, 1987), which adopts the singular value decomposition technique (SVD) to extract the reduced basis from a collection of high-fidelity solutions, called a snapshot matrix. Two broad categories can be considered to compute the associated coefficients of a reduced basis, intrusive and non-intrusive techniques. The intrusive methods combine the POD with Galerkin's projection and require a modification of the source code that solves the resulting discretized governing equations (Couplet *et al.*, 2005 ; Fang, Pain, Navon, Elsheikh, Du & Xiao, 2013 ; Zokagoa & Soulaïmani, 2012). Despite the improvements that have been introduced (Ballarin, Manzoni, Quarteroni & Rozza, 2015 ; Willcox, 2006), intrusive methods may require cumbersome modifications in complex codes, and the inaccessibility of these codes in most commercial software poses another difficulty. As a valuable alternative to overcome these issues, non-intrusive techniques have become increasingly attractive due to their flexibility, given that they treat the source code describing the physical model as a black box. A growing number of studies based on regression interpolations have been proposed to compute the associated coefficients of the reduced basis, including Gaussian process regression (GPR) (Guo & Hesthaven, 2018), Radial basis functions interpolation (RBF)(Walton, Hassan & Morgan, 2013 ; Xiao, Yang, Fang, Xiang, Pain & Navon, 2016a), Kriging interpolation (Xiao, Breitkopf, Coelho, Knopf-Lenoir, Sidorkiewicz & Villon, 2010), and Smolyak sparse grid interpolation (Xiao, Fang, Buchan, Pain, Navon & Muggeridge, 2015a). Recently, a panoply of configurations dealing with regression techniques based on artificial neural networks (ANN) has been introduced in the ROM construction process to accurately approximate unknown coefficients (Hesthaven & Ubbiali, 2018 ; Zhiwei, Chen, Zheng, Junqiang, Zheng, Qiang & Qiujun, 2020). Other non-intrusive approaches have been developed to deal with parametrized time-dependent problems that require specific treatments (Guo & Hesthaven, 2019 ; Swischuk, Mainini, Peherstorfer & Willcox, 2019). Some of them use the two-level POD technique to extract the reduced basis from a large snapshot matrix (Wang *et al.*, 2019), more particularly when the associated parametric values are generated by sampling in the random parametric domain (Jacquier *et al.*, 2021).

The effect of the variability stemming from the input parameters has to be quantified in the output quantities of interest through the uncertainty propagation analysis framework. The sampling approaches, i.e. Monte Carlo (MC) and Latin Hypercube Sampling (LHS) (McKay *et al.*, 1979) are the most popular and are known to require high sample sizes to reach an accurate estimation of the statistics. Being able to combine them with the above-mentioned reduced-order models renders them a useful tool for uncertainty propagation analysis, and with lower cost. In addition to the sampling methods mentioned above, polynomial chaos expansions (PCEs)

(Ghanem & Spanos, 1991 ; Sudret, 2014) are known to present alternative techniques to deal with uncertainty propagation analysis with a straightforward method to approximate the statistics. They are mainly associated with POD to propose a non-intrusive stochastic reduced-order model for polynomial chaos representation (Hijazi *et al.*, 2020 ; Raisee *et al.*, 2015,1). Recently, the so-called POD-PCE methods (El Moçayd *et al.*, 2020 ; Sun *et al.*, 2019), which are based on an offline-online paradigm, have been introduced as a stochastic reduced-order model, in which the projection data of the associated coefficient of the reduced basis is expressed as a stochastic expansion with respect to the orthogonal polynomials of the random input parameters. Despite the multiple advantages they offer, these approaches may present some limitations to accurately reproduce the stochastic outputs that present strong hyperbolic behavior or even discontinuity where spurious oscillations may result, as is the case for dam-break flows.

This work presents a non-intrusive reduced order model for stochastic time-dependent problems. The method, referred to as the POD-BSBEM, combines the advantages of the POD and the B-splines Bézier element method (BSBEM) (Abdedou & Soulaïmani, 2019 ; Abdedou *et al.*, 2020) to extract the reduced basis and to build a stochastic expansion over the random space, respectively. A two-level POD is applied to the snapshot matrix obtained by the collection of time-trajectory high-fidelity solutions for each given value of the random parametric vector. A third POD level is then performed for each associated coefficient of the reduced basis to separate the time-dependent modes from the stochastic parameterized weights. A series of appropriate transformations are then applied on the resulting projection data, considered as outputs, with which a local stochastic expansion is constructed based on local B-splines basis functions within each Bézier element constituting the parametric domain. The use of such local interpolations with the multi-element aspect offers a promising tool to accurately estimate the statistics of the output quantities of interest, particularly those with a strong hyperbolic or even discontinuous behavior.

The paper is organized as follows. Section 2 presents the fundamental concepts of the POD (subsection 4.2.1) and the mathematical framework of the proposed POD-BSBEM (subsection 4.2.2). Section 4.3 evaluates the performance and accuracy of the proposed method through two benchmark numerical test cases for the stochastic steady-state Ackley function (subsection 4.3.1), the time-dependent Burgers' equation (subsection 4.3.2), and for its application to a hypothetical dam-break with a real terrain database (subsection 4.3.3). Finally, a summary and concluding remarks are presented in Section 4.4.

4.2 Methodology

This section presents the main aspects of the reduced-order model formulation based on the proper orthogonal decomposition concept combined with the non-intrusive B-splines Bézier element method (POD-BSBEM) for both steady-state and time-dependent problems with uncertain input parameters.

Consider a parametrized time-dependent relationship $u = \mathcal{M}(\mathbf{x}, \boldsymbol{\eta}, t)$ that links an input of uncertain parameters to the output response, where $\mathcal{M}$ stands for the computational model that may be considered as a black box, $\mathbf{x} \in \Omega \subset \mathbb{R}^q$, $q = 1$, 2 or 3 is the space domain composed of N_e computational elements, and $t \in \mathcal{T} = [0,\ T]$ represents the temporal domain which is decomposed into N_t time steps. $\boldsymbol{\eta} \in \mathcal{P} \subset \mathbb{R}^m$ denotes the parametric domain that contains a random vector $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_m\}$, with m independent input uncertain parameters described by their probability density functions $\varrho_i(\eta_i)$ on the probability space (Θ, Σ, P), where Θ is the event space, P the probability measure and Σ is the σ-algebra on Θ.

4.2.1 Proper Orthogonal Decomposition

The main framework of reduced-order models is based on the approximation of the aforementioned high-fidelity solution of parameterized time-dependent problems by a combination of modes in the reduced-order basis stemming from the proper orthogonal decomposition (POD) as follows :

$$\mathbf{u}(\mathbf{x},\boldsymbol{\eta},t) \approx \widehat{\mathbf{u}}(\mathbf{x},\boldsymbol{\eta},t) = \sum_{\ell=1}^{L} \beta_\ell(\boldsymbol{\eta},t)\Phi_\ell(\mathbf{x}) \tag{4.1}$$

where $\Phi_\ell(\mathbf{x})$ $(\ell = 1,\ldots,L)$ represents the orthonormal modes that define the low dimensional basis in which L is the total number of modes. $\beta_\ell(\boldsymbol{\eta},t)$ denotes the coefficients of the ℓ^{th} mode that are dependent upon time and the random parameter vector $\boldsymbol{\eta}$. As outlined above, the extraction of the low dimensional basis is performed with the proper orthogonal decomposition procedure (POD) (Hesthaven & Ubbiali, 2018 ; Liang, Lee, Lim, Lin, Lee & Wu, 2002) from a snapshot matrix $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{N_e \times N_s N_t}$, a collection of high-fidelity solutions vectors, called snapshots (Sirovich, 1987), generated by the numerical solver $\left\{\mathbf{u}(\boldsymbol{\eta}^{(s)},t_j) \in \mathbb{R}^{N_e\times 1};\ j = 1,\ldots,N_t;\ s = 1,\ldots,N_s\right\}$, with N_s denotes the sample size. In the present work, the aforementioned global snapshot matrix $\boldsymbol{\mathcal{U}}$ is organized by concatenating local snapshot matrices $\mathcal{U}_s \in \mathbb{R}^{N_e\times N_t}$ that collect the high fidelity solutions for a given value of the random parametric vector $\boldsymbol{\eta}^{(s)},\ s = 1,\ldots,N_s$ over all of the time-steps $t = t_1,\ldots,t_{N_t}$. It is expressed as (Guo & Hesthaven, 2019) :

$$\mathcal{U}_s = \left[\mathbf{u}(\boldsymbol{\eta}^{(s)},t_1) \mid \ldots \mid \mathbf{u}(\boldsymbol{\eta}^{(s)},t_{N_t})\right] \in \mathbb{R}^{N_e\times N_t} \tag{4.2}$$

The global snapshot matrix assembling all the snapshots is defined as :

$$\boldsymbol{\mathcal{U}} = \left[\ \mathcal{U}_1 \mid \ldots \mid \mathcal{U}_s \mid \ldots \mid \mathcal{U}_{N_s}\right] \in \mathbb{R}^{N_e\times N_s N_t} \tag{4.3}$$

The reduced basis can be recovered with POD using the singular value decomposition (SVD) of the snapshot matrix $\boldsymbol{\mathcal{U}}$ (Wang *et al.*, 2019) :

$$\boldsymbol{\mathcal{U}} = \mathbb{W}\begin{bmatrix}\mathbb{D} & 0\\ 0 & 0\end{bmatrix}\mathbb{V}^T \tag{4.4}$$

where $\mathbb{W} = \left[\ \mathcal{W}_1 \mid \ldots \mid \mathcal{W}_{N_e}\right] \in \mathbb{R}^{N_e\times N_e}$ and $\mathbb{V} = \left[\ \mathcal{V}_1 \mid \ldots \mid \mathcal{V}_{N_s N_t}\right] \in \mathbb{R}^{N_s N_t\times N_s N_t}$ are orthogonal matrices, and $\mathbb{D} = diag\,(\sigma_1,\ldots,\sigma_{N_r})$ is the diagonal matrix containing the N_r singular values $\sigma_1 \geq \sigma_1 \geq \ldots \geq \sigma_{N_r} > 0$, in which $N_r \leq min(N_e, N_s N_t)$ represents the number of non-zero singular values. The low-rank POD basis can be constructed by selecting the L-most dominant singular values, i.e., those that capture the essential of the cumulative energy, by adopting the following criterion (Sun *et al.*, 2019) :

$$\frac{\sum_{\ell=1}^{L} \sigma_\ell^2}{\sum_{\ell=1}^{N_r} \sigma_\ell^2} > 1 - \epsilon \tag{4.5}$$

where ϵ is a user-defined hyperparameter representing the relative error tolerance. Thus, the POD mode vector $\Phi_\ell \in \mathbb{R}^{N_e \times 1}$ can be obtained by combining the snapshot matrix $\boldsymbol{\mathcal{U}}$ and the ℓ^{th} column of $\mathbb{V}$, $\mathcal{V}_\ell \in \mathbb{R}^{N_s N_t \times 1}$:

$$\Phi_\ell = \frac{\boldsymbol{\mathcal{U}} \mathcal{V}_\ell}{\sigma_\ell} \tag{4.6}$$

The orthonormal reduced POD basis obtained above $\boldsymbol{\Phi} = [\Phi_1 \mid \ldots \mid \Phi_L] \in \mathbb{R}^{N_e \times L} = POD\,(\boldsymbol{\mathcal{U}}, \epsilon)$ provides an optimal low-dimensional approximation of the snapshot data.

The POD construction procedure described above (Eqs.(4.4), (4.5) and (4.6)) for parameterized time-dependent problems with large computational domains and many time-steps will likely be cumbersome and thus incur significant computational costs and require too much memory to perform the SVD of the collected snapshot matrix. To overcome this problem, the so-called two-step POD algorithm (Wang *et al.*, 2019) is adopted by performing the first-level POD on the time-trajectory matrix $\mathcal{U}_s \in \mathbb{R}^{N_e \times N_t}$ given by Eq.(4.2), corresponding to the s^{th} value of the random parametric vector sampling set $\boldsymbol{\eta}^{(s)}$, i.e., $\mathbb{T}_s = \left[T_1 \mid \ldots \mid T_{L_s}\right] \in \mathbb{R}^{N_e \times L_s} = POD\,(\mathcal{U}_s, \epsilon_t)$. The compressed time-trajectory matrices thereby obtained for the whole parametric sampling set components ($s = 1, \ldots, N_s$) are concatenated, and then a second POD level is applied on the resulting matrix to extract the final POD basis, expressed as $\boldsymbol{\Phi} = [\Phi_1 \mid \ldots \mid \Phi_L] \in \mathbb{R}^{N_e \times L} = POD\left(\left[\mathbb{T}_1 \mid \ldots \mid \mathbb{T}_{N_s}\right] \in \mathbb{R}^{N_e \times \sum_{s=1}^{N_s} Ls}, \epsilon_s\right)$, where ϵ_t and ϵ_s represent the relative error tolerance for time and random parametric spaces, respectively. Detailed algorithms with illustrating schematics of the two-step POD procedure can be found in (Jacquier *et al.*, 2021 ; Wang *et al.*, 2019).

4.2.2 Regression-based non-intrusive B-Splines Bézier elements method

Once the construction of the reduced POD basis has been performed, the corresponding parametrized time-dependent POD coefficients involved in Eq.(4.1) must be computed. The B-splines Bézier element method is combined with the POD model reduction technique to propose a novel stochastic non-intrusive reduced order-based regression approach (POD-BSBEM) to accurately estimate the statistics of the output response of quantities of interest with a low computational cost. This approach belongs to the multi-element techniques that use local piecewise bases expressed as a function of input random parameters. It is worth noting that the basic framework of the BSBEM in its full-order version has been presented in detail in (Abdedou & Soulaïmani, 2019 ; Abdedou *et al.*, 2020). Thus, only a summary description of its fundamental aspects is outlined in the following.

Let us consider $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_m\} \in \Gamma = [0,\ 1]^m$ as an ensemble of m independent parametric variables defined in the parametric domain and linked to their corresponding sets of input random parameters in the physical domain $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_m\} \in \mathbb{R}^m$ by the marginal cumulative density function (CDF) whose image domain corresponds to that of the B-splines basis functions, i.e., $[0,\ 1]$. This allows $\xi_i(\eta_i) = F_i(\eta_i) = \int_{-\infty}^{\eta_i} \varrho_i(\eta')d\eta'$, which implies that $d\xi_i = \varrho_i(\eta_i)d\eta_i$. The parametric domain of each variable $\xi_i \in [0,\ 1]$ is decomposed into nx_i non-overlapping segments. The tensor product of these univariate intervals forms a multidimensional element defined as a Bézier element $I_e \subset \Gamma = [0,\ 1]^m$ whose total number in the entire parametric domain is given by $N_{elt} = \prod_{i=1}^{m} nx_i$. A set of $n_b = \prod_{i=1}^{m}(p_i + 1)$ multivariate local piecewise basis functions is constructed within each Bézier element using a tensor product of the univariate basis, expressed as

$$\Psi_j^e(\boldsymbol{\xi}) = \psi_1^e(\xi_1) \times \ldots \times \psi_m^e(\xi_m),\ \ j = 1, \ldots, nb \tag{4.7}$$

where e stands for the e^{th} Bézier element I_e and p_i denotes the polynomial order of the univariate B-splines basis function $\psi_i(\xi_i)$. Details of the construction of the B-splines basis and their extraction from the local Bernstein's functions can be found in (Abdedou & Soulaïmani, 2019 ; Abdedou *et al.*, 2020 ; Borden *et al.*, 2011). In each Bézier element, a set of $n_s^e \geq n_b$ collocation points are selected $\boldsymbol{\xi}^{e(r)} \in I_e;\ r = 1, \ldots, n_s^e;\ e = 1, \ldots, N_{elt}$ and used to construct the snapshot matrix by running the numerical solver with their corresponding values in the physical domain $\left\{\mathbf{u}(\boldsymbol{\eta}^{e(r)}, t_j) \in \mathbb{R}^{N_e \times 1};\ j = 1, \ldots, N_t;\ r = 1, \ldots, n_s^e\right\}$, obtained by an appropriate inverse CDF mapping :

$$\eta_i^{e\,(r)} = F_i^{-1}(\xi_i^{e\,(r)}),\ r = 1, \ldots, n_s^e,\ i = 1, \ldots, m \tag{4.8}$$

where F_i^{-1} denotes the cumulative density function that links the random parametric variable ξ_i to its corresponding random physical variable η_i. The total number of sampling points over the entire parametric domain is $N_s = n_s^e N_{elt}$.

4.2.2.1 A third level proper orthogonal decomposition

The unknown parametric time-dependent coefficients associated with the orthonormal column bases of $\mathbf{\Phi}$, involved in Eq.(4.1), can be expressed as :

$$\boldsymbol{\mathcal{B}}(\boldsymbol{\eta}, t) = \mathbf{\Phi}^T \boldsymbol{\mathcal{U}} = \left[\begin{array}{ccc|c|ccc} \beta_{11}^1 & \cdots & \beta_{1N_t}^1 & \cdots & \beta_{11}^{N_s} & \cdots & \beta_{1N_t}^{N_s} \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ \beta_{\ell 1}^1 & \cdots & \beta_{\ell N_t}^1 & \cdots & \beta_{\ell 1}^{N_s} & \cdots & \beta_{\ell N_t}^{N_s} \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ \beta_{L1}^1 & \cdots & \beta_{LN_t}^1 & \cdots & \beta_{L1}^{N_s} & \cdots & \beta_{LN_t}^{N_s} \end{array}\right] \in \mathbb{R}^{L \times N_s N_t} \tag{4.9}$$

It is worth noting that the procedure adopted to compute these coefficients is detailed in the following paragraphs only for the ℓ^{th} projection coefficient, i.e., $\boldsymbol{\beta}_\ell(\boldsymbol{\eta}, t) \in \mathbb{R}^{1 \times N_s N_t}$, representing the ℓ^{th} row vector reported in Eq.(4.9), and will be generalized for the whole L modes. Thus, this ℓ^{th} row vector can be written in a matrix form, $\boldsymbol{\beta}_\ell \in \mathbb{R}^{1 \times N_s N_t} \equiv \boldsymbol{\beta}^\ell \in \mathbb{R}^{N_t \times N_s}$, with N_t rows and N_s columns, representing the time and the parameter locations, respectively, as (Guo & Hesthaven,

2019) :

$$\boldsymbol{\beta}^{\ell}(t,\boldsymbol{\eta}) = \begin{bmatrix} \beta^{\ell}_{11} & \cdots & \beta^{\ell}_{1N_s} \\ \vdots & \ddots & \vdots \\ \beta^{\ell}_{N_t1} & \cdots & \beta^{\ell}_{N_tN_s} \end{bmatrix} \in \mathbb{R}^{N_t \times N_s} \tag{4.10}$$

A third POD level is then applied on $\boldsymbol{\beta}^{\ell}(t,\boldsymbol{\eta})$ to decompose the data into time-dependent modes with their associated parameter-dependent coefficients, expressed as :

$$\beta^{\ell}(t,\boldsymbol{\eta}) \approx \sum_{k=1}^{\mathcal{K}_\ell} \lambda^{\ell}_k(\boldsymbol{\eta})\chi^{\ell}_k(t) \tag{4.11}$$

where $\boldsymbol{X}^{\ell}(t) = \left[\chi^{\ell}_1 \mid \ldots \mid \chi^{\ell}_{\mathcal{K}_\ell}\right] \in \mathbb{R}^{N_t \times \mathcal{K}_\ell} = POD\left(\boldsymbol{\beta}^{\ell}(t,\boldsymbol{\eta}) \in \mathbb{R}^{N_t \times N_s}, \epsilon_s\right)$ and $\boldsymbol{\Lambda}^{\ell}(\boldsymbol{\eta}) \in \mathbb{R}^{\mathcal{K}_\ell \times N_s}$ represent the time-dependent orthogonal basis modes with a truncating rank of $\mathcal{K}_\ell$ and the associated parameter-dependent coefficients, respectively, of the ℓ^{th} mode. These associated coefficients can be expressed in a matrix form obtained by projection from Eq.(4.11) by exploiting the orthonormality feature of the $\boldsymbol{X}^{\ell}$ basis :

$$\boldsymbol{\Lambda}^{\ell} = \boldsymbol{X}^{\ell^T}\boldsymbol{\beta}^{\ell} = \begin{bmatrix} \lambda^{\ell}_{11} & \cdots & \lambda^{\ell}_{1N_s} \\ \vdots & \ddots & \vdots \\ \lambda^{\ell}_{k1} & \cdots & \lambda^{\ell}_{kN_s} \\ \vdots & \ddots & \vdots \\ \lambda^{\ell}_{\mathcal{K}_\ell 1} & \cdots & \lambda^{\ell}_{\mathcal{K}_\ell N_s} \end{bmatrix} \in \mathbb{R}^{\mathcal{K}_\ell \times N_s} \tag{4.12}$$

The procedure used to construct a regression as a function of stochastic basis functions is detailed below. As with the above coefficient calculation, the calculations are restricted, in this case, to the k^{th} row vector, and must be generalized for the $\mathcal{K}_\ell$ associated coefficients given by Eq.(4.12). The collocation points are selected locally within each Bézier element I_e with a minimum required number of n^e_s, as noted earlier. The set of sampling points $N_s = n^e_s N_{elt}$ is obtained by concatenating the local points n^e_s over the complete set of N_{elt} Bézier elements that contains entire parametric domain. Thus, the k^{th} row vector of $\boldsymbol{\Lambda}^{\ell}$ is a concatenation of local vectors corresponding to each Bézier element, i.e., $\boldsymbol{\lambda}^{\ell,e}_k \in \mathbb{R}^{1 \times n^e_s} \subset \boldsymbol{\lambda}^{\ell}_k \in \mathbb{R}^{1 \times N_s}$, where

$e = 1, \ldots, N_{elt}$ denotes the e^{th} Bézier element. These local data $\boldsymbol{\delta}_k^{\ell,e} \in \mathbb{R}^{n_s^e \times 1} = (\boldsymbol{\lambda}_k^{\ell,e})^T$ are considered as outputs with which a regression is constructed in each Bézier element as the sum of the local piecewise basis functions :

$$\delta_k^{\ell,e} = \sum_{j=1}^{nb} \alpha_{j,k}^{\ell,e} \Psi_j^e(\boldsymbol{\xi}) \tag{4.13}$$

where $\alpha_{j,k}^{\ell,e}$ and Ψ_j^e denote the local coefficients and the n_b local multivariate piecewise B-splines basis functions, given by Eq.(4.7), over the e^{th} Bézier element I_e, respectively. These local basis functions are evaluated on each component $\boldsymbol{\xi}^{e(i)}$ of the collocation points set in the e^{th} Bézier element, i.e. $\boldsymbol{\psi}^e_{n_s^e \times nb} = \left\{ \Psi_{i,j}^e = \Psi_j^e(\boldsymbol{\xi}^{e\,(i)}), i = 1, \ldots, n_s^e,\ j = 1, \ldots, nb \right\}$. Thus, the resulting local design matrix and the output response vector are expressed as :

$$\boldsymbol{\Psi}^e = \boldsymbol{\psi}^{e^T} \boldsymbol{\psi}^e \in \mathbb{R}^{nb \times nb} \quad \text{and} \quad \boldsymbol{\Delta}_k^{\ell,e} = \boldsymbol{\psi}^{e^T} \boldsymbol{\delta}_k^{\ell,e} \in \mathbb{R}^{nb \times 1} \tag{4.14}$$

The unknown coefficients are computed by solving a global system of equations instead of the constructed local systems in each Bézier element. This global system of equations is constructed from the assembly of local design matrices and local output vectors through the assembly operator (Hughes, 2012). The global design matrix and output vector are thus given by :

$$\boldsymbol{\Psi} = \mathop{\mathbf{A}}_{e=1}^{N_{elt}} (\boldsymbol{\Psi}^e) \in \mathbb{R}^{M \times M} \quad \text{and} \quad \boldsymbol{\Delta}_k^{\ell} = \mathop{\mathbf{A}}_{e=1}^{N_{elt}} (\boldsymbol{\Delta}_k^{\ell,e}) \in \mathbb{R}^{M \times 1} \tag{4.15}$$

where $M = \prod_{i=1}^{m} (nx_i + p_i)$ denotes the total number of the global coefficients to be computed. The components of the global design matrix $\Psi_{q,f}$ are linked to their corresponding components in the local design matrices $\Psi_{i,j}^e$ by the connectivity array (IEN) with respect to the e^{th} Bézier element : $q = IEN(i, e)$ and $f = IEN(j, e)$, with $q,\ f = 1, \ldots, M$ and $i,\ j = 1, \ldots, nb$ (more implementation details can be found in (Cottrell *et al.*, 2009 ; Scott *et al.*, 2011)). Thus, the coefficients can be computed by solving the following global system of equations :

$$\boldsymbol{\alpha}_k^{\ell} = \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta}_k^{\ell} \in \mathbb{R}^{M \times 1} \tag{4.16}$$

Once these coefficients are computed, the stochastic output response, given by Eq.(4.1), can be locally approximated $\widehat{\mathbf{u}}^e$ in each Bézier element, which represents a part of the random parametric space, using the online stage of the proposed POD-BSBEM. Thus, let us consider a new set of n_s^e collocation points in the e^{th} Bézier element, $\widehat{\boldsymbol{\xi}}^{e^{(i)}}$, $i = 1, \ldots, n_s^e$, with which the local B-splines basis functions $\left\{\Psi_j^e\right\}_{j=1}^{n_b}$ are evaluated to build the so-called local design matrix : $\widehat{\boldsymbol{\psi}}^e_{n_s^e \times nb} = \left\{\Psi_{i,j}^e = \Psi_j^e\left(\widehat{\boldsymbol{\xi}}^{e\,(i)}\right), i = 1, \ldots, n_s^e,\ j = 1, \ldots, nb\right\}$. The local coefficients $\boldsymbol{\alpha}_k^{\ell,e} \in \mathbb{R}^{n_b \times 1}$ are extracted from the set of computed global coefficients, $\boldsymbol{\alpha}_k^{\ell} \in \mathbb{R}^{M \times 1}$, with the same connectivity array procedure mentioned above. The approximated local projection data associated with the $\mathcal{K}_\ell$ time-dependent modes can be expressed as :

$$\widehat{\boldsymbol{\Lambda}}^{\ell,e} = \left[\widehat{\boldsymbol{\psi}}^e \boldsymbol{\alpha}_1^{\ell,e} \mid \ldots \mid \widehat{\boldsymbol{\psi}}^e \boldsymbol{\alpha}_k^{\ell,e} \mid \ldots \mid \widehat{\boldsymbol{\psi}}^e \boldsymbol{\alpha}_{\mathcal{K}_\ell}^{\ell,e}\right]^T \in \mathbb{R}^{\mathcal{K}_\ell \times n_s^e} \tag{4.17}$$

The components of the ℓ^{th} projection coefficient can be approximated locally in the e^{th} Bézier element, by combining the time-dependent basis functions $\boldsymbol{X}^\ell \in \mathbb{R}^{N_t \times \mathcal{K}_\ell}$, as in Eq.(4.11), and the local projection data from Eq.(4.17) : $\widehat{\boldsymbol{\beta}}^{\ell,e} = \boldsymbol{X}^\ell \widehat{\boldsymbol{\Lambda}}^{\ell,e} \in \mathbb{R}^{N_t \times n_s^e}$. The obtained data are then transformed from an $N_t \times n_s^e$ matrix to a $1 \times n_s^e N_t$ row vector, i.e., $\widehat{\boldsymbol{\beta}}_\ell^e \in \mathbb{R}^{1 \times n_s^e N_t} \equiv \widehat{\boldsymbol{\beta}}^{\ell,e} \in \mathbb{R}^{N_t \times n_s^e}$, and so the approximated projection coefficients associated with the L modes reduced basis are obtained as :

$$\widehat{\mathcal{B}}^e = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^e \\ \vdots \\ \widehat{\boldsymbol{\beta}}_\ell^e \\ \vdots \\ \widehat{\boldsymbol{\beta}}_L^e \end{bmatrix} \in \mathbb{R}^{L \times n_s^e N_t} \quad \text{with} \quad \widehat{\boldsymbol{\beta}}_\ell^e = \left[\widehat{\beta}_{\ell 1}^1 \cdots \widehat{\beta}_{\ell N_t}^1 \mid \cdots \mid \widehat{\beta}_{\ell 1}^{n_s^e} \cdots \widehat{\beta}_{\ell N_t}^{n_s^e}\right] \in \mathbb{R}^{1 \times n_s^e N_t} \tag{4.18}$$

Thus, the approximation of the output response can be expressed by the local snapshot matrix constructed within the e^{th} Bézier element :

$$\widehat{\mathcal{U}}^e = \boldsymbol{\Phi} \widehat{\mathcal{B}}^e \in \mathbb{R}^{N_e \times n_s^e N_t} \tag{4.19}$$

The approximation of the snapshot matrix over the whole random parametric domain is given by :

$$\widehat{\boldsymbol{\mathcal{U}}} = \left[\widehat{\boldsymbol{\mathcal{U}}}^{1} \mid \ldots \mid \widehat{\boldsymbol{\mathcal{U}}}^{e} \mid \ldots \mid \widehat{\boldsymbol{\mathcal{U}}}^{N_{elt}}\right] \in \mathbb{R}^{N_e \times N_s N_t} \tag{4.20}$$

The statistical moments can be estimated through the constructed surrogate model of the stochastic output response $\widehat{u}$ as follows :

$$\mu_{\widehat{u}} = \mathbb{E}\left[\widehat{u}\right] = \int_{\Xi} \widehat{u} \varrho_{\boldsymbol{\eta}}(\boldsymbol{\eta}) d\boldsymbol{\eta} \tag{4.21}$$

and

$$\sigma_{\widehat{u}}^2 = \mathbb{E}\left[\widehat{u}^2\right] - \mu_{\widehat{u}}^2 = \left[\int_{\Xi} (\widehat{u})^2 \varrho_{\eta}(\boldsymbol{\eta}) d\boldsymbol{\eta}\right] - \mu_{\widehat{u}}^2 \tag{4.22}$$

The integrals involved in the mean and variance expressions can be simplified by exploiting the correspondence between the CDF image domain and the parametric domain of the basis functions $[0,\ 1]$. As indicated above, this CDF mapping, i.e., $\xi_i = F_i(\eta_i)$ and $d\xi_i = \varrho_i(\eta_i) d\eta_i$, allows a simplified evaluation of the integrals over each Bézier element definition domain. Thus, the statistical moments of the approximated stochastic output response of a quantity of interest over each computational node for a given simulation time are expressed as :

$$\mu_{\widehat{u}} = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} \widehat{u}^e d\boldsymbol{\xi}\right] \tag{4.23}$$

and

$$\sigma_{\widehat{u}}^2 = \sum_{e=1}^{N_{elt}} \left[\int_{I_e} (\widehat{u}^e)^2 \, d\boldsymbol{\xi}\right] - \mu_{\widehat{u}}^2 \tag{4.24}$$

Based on the detailed steps presented above, the implementations of the offline and online stages of the proposed POD-BSBEM are summarized in Algorithm.4.1 and Algorithm.4.2, respectively.

In addition to the sampling approaches such as Monte Carlo (MC) and Latin Hypercube Sampling (LHS), the performance and accuracy of the proposed POD-BSBEM are also compared to the non-intrusive reduced-order model-based artificial neural network (POD-ANN) and the

Algorithme 4.1 Offline stage of the POD-BSBEM for stochastic time-dependent problems

Input : p_i, nx_i, μ_{η_i}, σ_{η_i}, $\varrho_i(\eta_i)$, $n_s^e = \prod_{i=1}^{m}(p_i+1)$, $N_{elt} = \prod_{i=1}^{m} nx_i$, $i = 1, \ldots, m$

Output : $\mathbf{\Phi} \in \mathbb{R}^{N_e \times L}$, $\boldsymbol{\alpha}^{\ell} \in \mathbb{R}^{M \times 1}$, $\boldsymbol{X}^{\ell} \in \mathbb{R}^{N_t \times \mathcal{K}_\ell}$, $\ell = 1, \ldots, L$

Generate the local collocation points : $\boldsymbol{\xi}^{e^{(r)}} \in I_e$, $r = 1, \ldots, n_s^e$; $e = 1, \ldots, N_{elt}$

Compute the high fidelity solutions : $\mathbf{u}(\boldsymbol{\eta}^{e^{(r)}}, t_j)$, $j = 1, \ldots, N_t$ with $\eta_i^{e\,(r)} = F_i^{-1}(\xi_i^{e\,(r)})$

Construct the time-trajectory snapshot matrices $\boldsymbol{\mathcal{U}}_s \in \mathbb{R}^{N_e \times N_t}$ from Eq.(4.2) and the global snapshot matrix $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{N_e \times N_s N_t}$ from Eq.(4.3)

Compress the time-trajectory basis : $\mathbb{T}_s = \left[T_1 \mid \ldots \mid T_{L_s}\right] = POD\,(\boldsymbol{\mathcal{U}}_s, \epsilon_t)$, with $s = 1, \ldots, N_s$

Extract the L reduced basis functions :
$\mathbf{\Phi} = [\Phi_1 \mid \ldots \mid \Phi_L] = POD\left(\left[\mathbb{T}_1 \mid \ldots \mid \mathbb{T}_{N_s}\right], \epsilon_s\right)$

Project the POD-associated coefficients : $\mathcal{B}(\boldsymbol{\eta}, t) = \mathbf{\Phi}^T \boldsymbol{\mathcal{U}} \in \mathbb{R}^{L \times N_s N_t}$ from Eq.(4.9)

for $\ell = 1 : L$ **do**

    Transform the ℓ^{th} row vector to a matrix : $\boldsymbol{\beta}^{\ell} \in \mathbb{R}^{N_t \times N_s} \equiv \boldsymbol{\beta}_{\ell} \in \mathbb{R}^{1 \times N_s N_t}$ from Eq.(4.10)

    Extract the $\mathcal{K}_\ell$ time-dependent basis $\boldsymbol{X}^{\ell}(t) = \left[\chi_1^{\ell} \mid \ldots \mid \chi_{\mathcal{K}_\ell}^{\ell}\right] = POD\left(\boldsymbol{\beta}^{\ell}, \epsilon_s\right)$

    Perform the projection $\mathbf{\Lambda}^{\ell} = \boldsymbol{X}^{\ell^T} \boldsymbol{\beta}^{\ell} \in \mathbb{R}^{\mathcal{K}_\ell \times n_s^e N_{elt}}$ from Eq.(4.12)

    for $k = 1 : \mathcal{K}_\ell$ **do**

        Extract $\boldsymbol{\delta}_k^{\ell} = \{\lambda_{ks};\ s = 1, \ldots, N_s\}^T \in \mathbb{R}^{N_s \times 1}$ with $N_s = n_s^e N_{elt}$

        for $e = 1 : N_{elt}$ **do**

            Extract the local output projection data $\boldsymbol{\delta}_k^{\ell,e} \in \mathbb{R}^{n_s^e \times 1} \subset \boldsymbol{\delta}_k^{\ell} \in \mathbb{R}^{N_s \times 1}$

            Define the local design matrix $\mathbf{\Psi}^e = \boldsymbol{\psi}^{e^T} \boldsymbol{\psi}^e \in \mathbb{R}^{nb \times nb}$ and the local output vector $\mathbf{\Delta}_k^{\ell,e} = \boldsymbol{\psi}^{e^T} \boldsymbol{\delta}_k^{\ell,e} \in \mathbb{R}^{nb \times 1}$ according to Eq.(4.14), with
$\boldsymbol{\psi}^e_{n_s^e \times nb} = \left\{\Psi_{i,j}^e = \Psi_j^e(\boldsymbol{\xi}^{e\,(i)}), i = 1, \ldots, n_s^e,\ j = 1, \ldots, nb\right\}$

            Assemble the global design matrices : $\mathbf{\Psi} = \underset{e=1}{\overset{N_{elt}}{\mathbf{A}}}(\mathbf{\Psi}^e) \in \mathbb{R}^{M \times M}$ and global output vectors : $\mathbf{\Delta}_k^{\ell} = \underset{e=1}{\overset{N_{elt}}{\mathbf{A}}}(\mathbf{\Delta}_k^{\ell,e}) \in \mathbb{R}^{M \times 1}$ according to Eq.(4.15)

        end

        Compute the global coefficients : $\boldsymbol{\alpha}_k^{\ell} = \mathbf{\Psi}^{-1} \mathbf{\Delta}_k^{\ell} \in \mathbb{R}^{M \times 1}$ according to Eq.(4.16)

    end

end

full-order polynomial chaos expansion (Full-PCE). It should be noted that a detailed discussion of the development and implementation steps of these methods is beyond the scope of this work, and the reader is encouraged to consult the corresponding references (Hesthaven & Ubbiali, Algorithme 4.2 Online stage of the POD-BSBEM for stochastic time-dependent problems

Input : p_i, nx_i, $\left\{\widehat{\boldsymbol{\xi}}^{e(i)}\right\}_{i=1}^{n_s^e}$, $\left\{\Psi_j\right\}_{j=1}^{n_b}$, $\boldsymbol{\Phi}$, $\boldsymbol{\alpha}^\ell$, $\boldsymbol{X}^\ell$, $\ell = 1, \ldots, L$

Output : $\widehat{\boldsymbol{\mathcal{U}}}^e$, $\widehat{\boldsymbol{\mathcal{U}}}$, $\mu_{\widehat{u}}$, $\sigma^2_{\widehat{u}}$

for $e = 1 : N_{elt}$ **do**

  Generate a new set of local collocation points : $\widehat{\boldsymbol{\xi}}^{e(i)} \in I_e$, $i = 1, \ldots, n_s^e$

  Construct local design matrix $\widehat{\boldsymbol{\psi}}^e_{n_s^e \times nb} = \left\{\Psi^e_{i,j} = \Psi^e_j\left(\widehat{\boldsymbol{\xi}}^{e\,(i)}\right),\ j = 1, \ldots, nb\right\}_{i=1}^{n_s^e}$

  for $\ell = 1 : L$ **do**

    Select the time-dependent matrix $\boldsymbol{X}^\ell \in \mathbb{R}^{N_t \times \mathcal{K}_\ell}$

    for $k = 1 : \mathcal{K}_\ell$ **do**

      Extract local coefficients from the global ones : $\boldsymbol{\alpha}_k^{\ell,e} \in \mathbb{R}^{n_b \times 1} \subset \boldsymbol{\alpha}_k^\ell \in \mathbb{R}^{M \times 1}$

      Construct the local regression : $\widehat{\boldsymbol{\delta}}_k^{\ell,e} = \widehat{\boldsymbol{\psi}}^e \boldsymbol{\alpha}_k^{\ell,e} \in \mathbb{R}^{n_s^e \times 1}$ following Eq.(4.13)

    end

    Concatenate the approximated output data : $\widehat{\boldsymbol{\Lambda}}^{\ell,e} = \left[\widehat{\boldsymbol{\delta}}_1^{\ell,e} \mid \ldots \mid \widehat{\boldsymbol{\delta}}_{\mathcal{K}_\ell}^{\ell,e}\right]^T \in \mathbb{R}^{\mathcal{K}_\ell \times n_s^e}$

    Compute the data of the ℓ^{th} projection coefficient : $\widehat{\boldsymbol{\beta}}^{\ell,e} = \boldsymbol{X}^\ell \widehat{\boldsymbol{\Lambda}}^{\ell,e} \in \mathbb{R}^{N_t \times n_s^e}$

    Transform the obtained matrix to a row vector : $\widehat{\boldsymbol{\beta}}_\ell^e \in \mathbb{R}^{1 \times n_s^e N_t} \equiv \widehat{\boldsymbol{\beta}}^{\ell,e} \in \mathbb{R}^{N_t \times n_s^e}$

  end

  Concatenate over the L-modes : $\widehat{\boldsymbol{\mathcal{B}}}^e = \left[\widehat{\boldsymbol{\beta}}_1^{e^T} \mid \ldots \mid \widehat{\boldsymbol{\beta}}_\ell^{e^T} \mid \ldots \widehat{\boldsymbol{\beta}}_L^{e^T}\right]^T \in \mathbb{R}^{L \times n_s^e N_t}$

  Compute the local approximated snapshot matrix : $\widehat{\boldsymbol{\mathcal{U}}}^e = \boldsymbol{\Phi}\widehat{\boldsymbol{\mathcal{B}}}^e \in \mathbb{R}^{N_e \times n_s^e N_t}$

  Compute the statistical moments according to Eqs.(4.23) and (4.24)

end

Concatenate over the N_{elt} Bézier elements to approximate the global snapshot matrix :
$\widehat{\boldsymbol{\mathcal{U}}} = \left[\widehat{\boldsymbol{\mathcal{U}}}^1 \mid \ldots \mid \widehat{\boldsymbol{\mathcal{U}}}^e \mid \ldots \mid \widehat{\boldsymbol{\mathcal{U}}}^{N_{elt}}\right] \in \mathbb{R}^{N_e \times N_s N_t}$

2018 ; Jacquier *et al.*, 2021 ; Wang *et al.*, 2019) and (Abdedou *et al.*, 2020 ; Hosder & Walters, 2010 ; Hosder *et al.*, 2007), respectively, where modeling details can be found.

4.3 Results and discussion

This section presents the numerical results for two benchmark test cases, a steady-state Ackley function and a time-dependent Burgers equation, along with the results from a hypothetical dam break with real terrain data. These results are then discussed in terms of statistical moments and probability density function profiles of the quantities of interest, as well as the resulting probabilistic inundation mapping.

4.3.1 Ackley function test case

The so-called Ackley function is considered as a first test function to showcase the performance of the proposed approach. It presents a challenging candidate due to its high irregularity and non-linearity aspects. This function is defined over a 2D space domain $-5 \leq x, y \leq 5$, which is discretized with a mesh of $N_e = 160 \times 160$ nodes (Raisee *et al.*, 2015). The stochastic version of the Ackley function is formulated as follows (Raisee *et al.*, 2015 ; Sun *et al.*, 2019) :

$$
\begin{aligned}
u(x, y, \boldsymbol{\xi}) = -20(1 + 0.1\xi_3)\left(exp\left[-0.2(1 + 0.1\xi_2)\sqrt{0.5(x^2 + y^2)}\right]\right) \\
-exp\left(0.5\left[cos(2\pi(1 + 0.1\xi_1)x + cos(2\pi(1 + 0.1\xi_1)y)\right]\right) + 20 + e
\end{aligned}
\tag{4.25}
$$

where $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3]$ denotes a random vector defined by three uncertain input parameters assumed to be described by uniform distributions over $[-1, 1]^3$ and $e \approx\ 2.718281$.

The results of the non-intrusive reduced order models (POD-BSBEM and POD-ANN) are compared with the solution obtained using 50 000 realizations on a sample generated with the LHS algorithm, considered as a reference solution. Fig.4.1 presents the 3D view and contour plots of the statistics in order to show the high irregular aspect of the Ackley function. In the following, the results concerning the present test case are mainly displayed and discussed in terms of the variation of the mean and the standard deviation profiles as a function of y at $x = 0$ (cross-section chosen arbitrarily).

Fig.4.2 presents the variation of the mean and the standard deviation as a function of y (at cross-section $x = 0$) obtained by the non-intrusive reduced order model POD-BSBEM with polynomial order $p_i = 2$ and various values of the number of Bézier elements ($nx_i = 2, 3$ and 5, with $i \in \{1, 2, 3\}$). The results are displayed for two values of the energy reconstruction

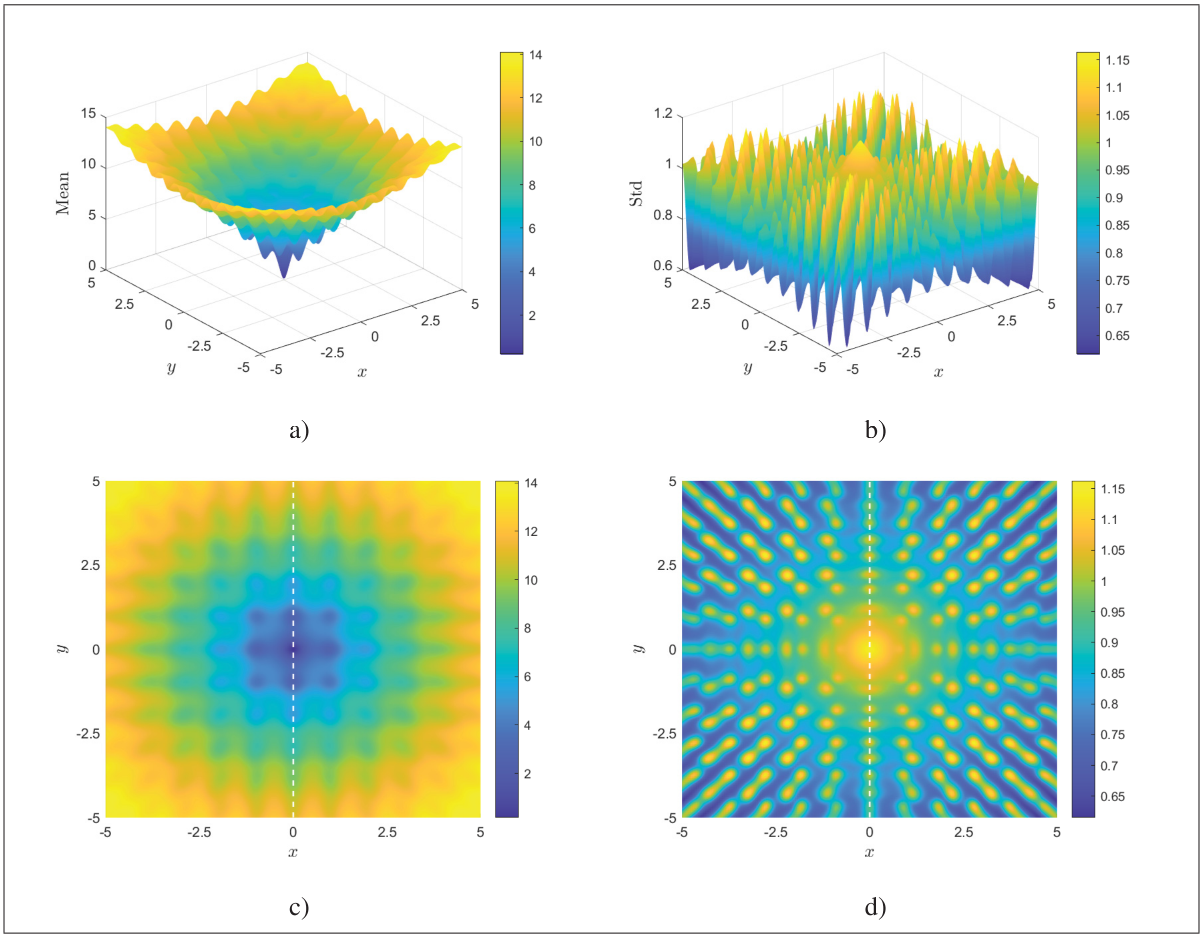
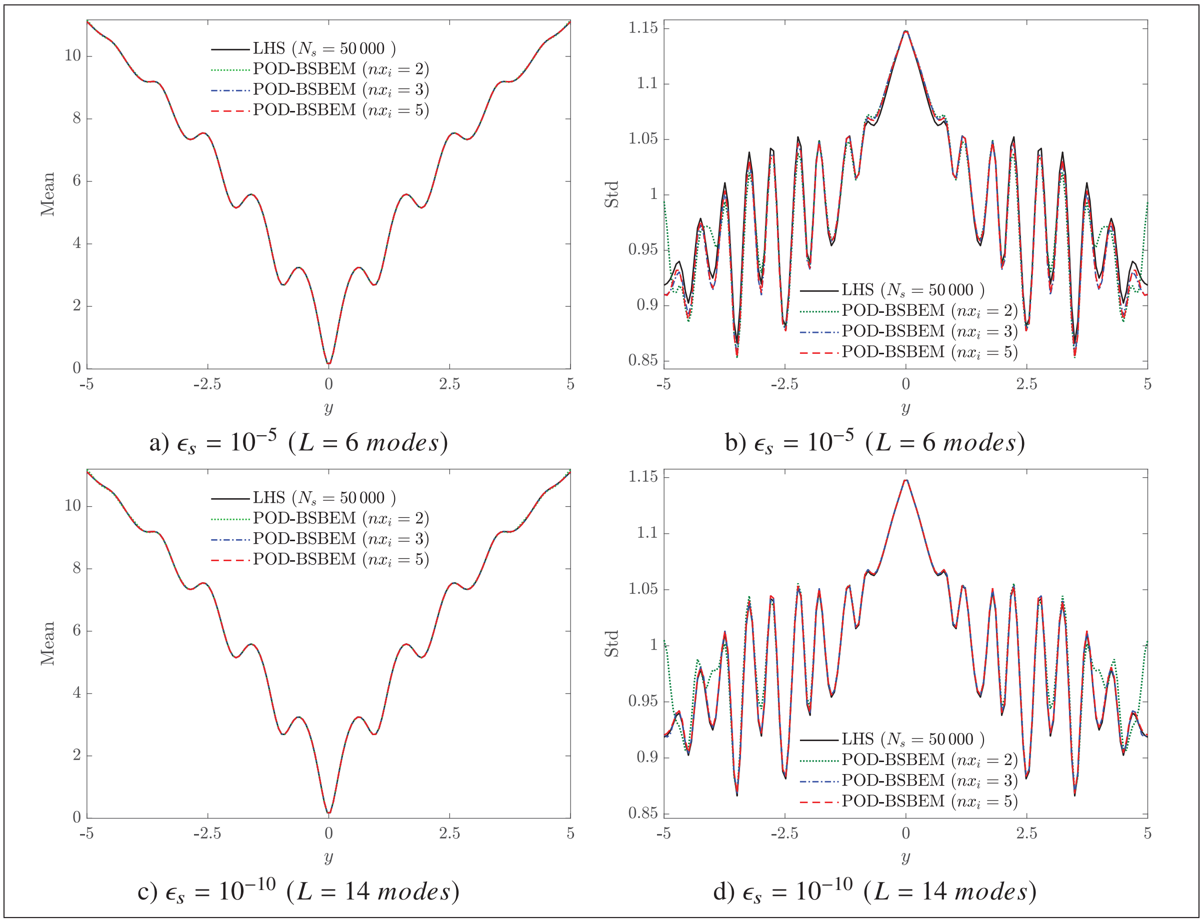


Figure 4.1 Mean and standard deviation fields for the Ackley function obtained from $N_s = 50\,000$ LHS realizations. First row : 3D plots, second row : contour plots. (For interpretation of the references to color in this figure, please refer to the web version of this article).

level, $\epsilon_s = 10^{-5}$ and 10^{-10}, corresponding to the number of modes $L = 6$ and 14, respectively. These plots show that the results obtained by the proposed reduced model (POD-BSBEM) for the mean statistics accurately reproduce the reference LHS solution (with 50 000 realizations), independently of the number of Bézier elements and the number of POD modes. The standard deviation profile reveals a clear deviation between the solution obtained by the POD-BSBEM approach and that from the LHS method for low values of the number of Bézier elements ($nx_i = 2$) and of the number of modes $L = 6$ (Fig.4.2b). This deviation decreases as the number of Bézier elements increases, where a satisfactory matching can be noted with the LHS reference solution, more particularly for $nx_i = 5$ and $L = 14$, as shown in Fig.4.2d.

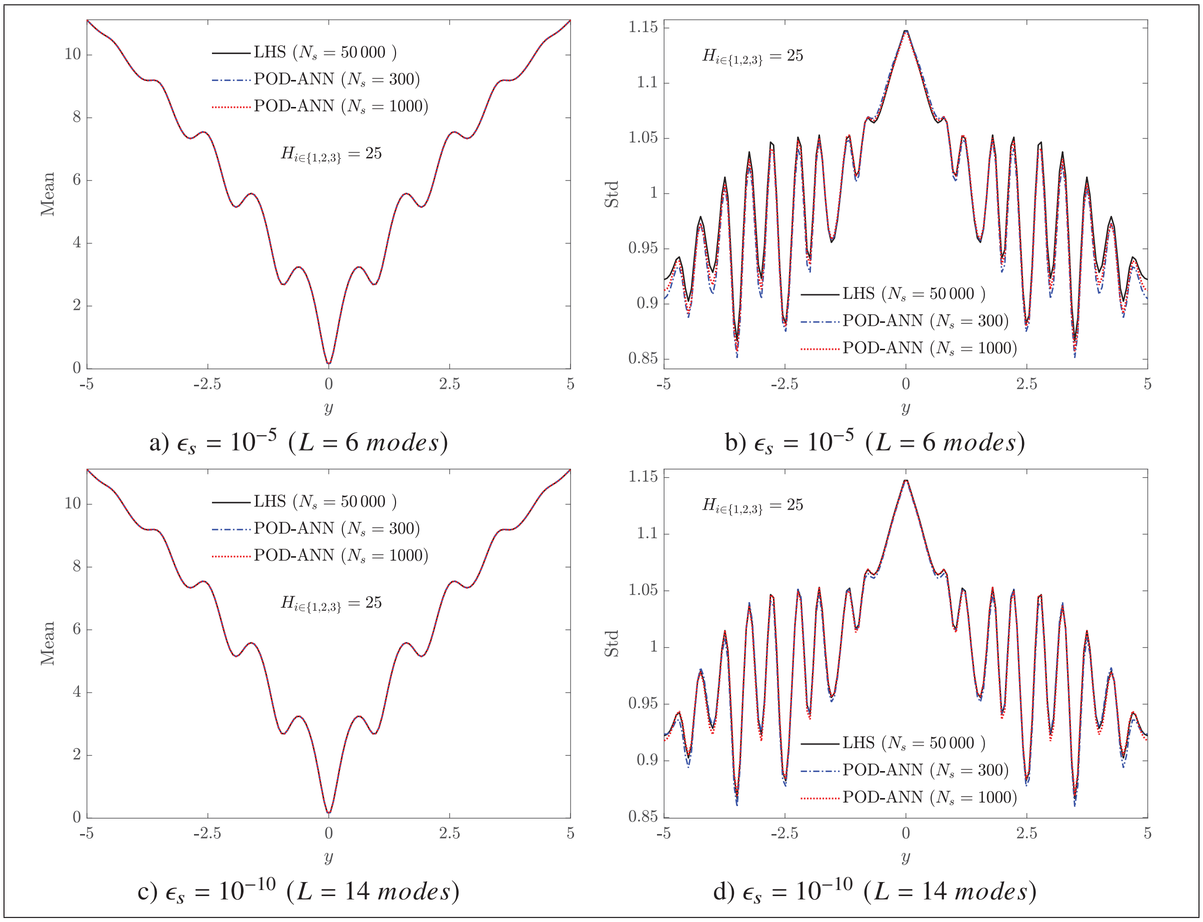


a) $\epsilon_s = 10^{-5}$ ($L = 6$ $modes$)
b) $\epsilon_s = 10^{-5}$ ($L = 6$ $modes$)
c) $\epsilon_s = 10^{-10}$ ($L = 14$ $modes$)
d) $\epsilon_s = 10^{-10}$ ($L = 14$ $modes$)

Figure 4.2 Statistical moments as a function of y (at $x = 0$) obtained using the non-intrusive POD-BSBEM with $p_{i\in\{1,2,3\}} = 2$ and $nx_{i\in\{1,2,3\}} = 2,\ 3,\ 5 : L = 6$ modes (a-b) and $L = 14$ modes (c-d), compared to the LHS method (with $N_s = 50\,000$ realizations).

The variation of the statistical moments of the Ackley function as a function of y at cross-section $x = 0$ obtained using the reduced model based on artificial neural networks (POD-ANN) with three hidden size layers $H_1 = H_2 = H_3 = 25$ is shown in Fig.4.3. It should be mentioned that the neural network topology adopted in the present test case was selected after a series of numerical tests by varying the number of hidden layers and the width of each layer, in which a trade-off was made between the network performance (based on the mean squared error) and the learning time. As mentioned above, the Bayesian regularization of the training algorithm is employed in the supervised learning process to iteratively adjust the regularizing hyperparameter. Two sets of the uncertain input parameters (N_s = 300 and 1000) are randomly selected within the parameter domain using the LHS approach. The generated sampling data are all scaled to the same range using the mean normalization (Wang *et al.*, 2019), i.e., $\widehat{\eta}_i = \frac{\eta_i - \tilde{\eta}}{\eta_{max} - \eta_{min}}$, $i = 1 \ldots, N_s$, where $\tilde{\eta}$, η_{min}, and η_{max} represent the mean, minimum and maximum of the input data, respectively. Their corresponding outputs, which represent the snapshot matrix, are obtained from N_s runs of the high-fidelity solution (Ackley function). The results in Fig.4.3 show that for a relatively low number of eigenmodes ($L = 6$), the standard deviation profiles predicted by POD-ANN present a visible deviation with those of the reference LHS solution for both sizes of the learning sets (Fig.4.3b). When the number of modes increases to $L = 14$ (obtained with an energy construction level of $\epsilon_s = 10^{-10}$), a remarkable improvement in the predicted standard deviation profile can be observed, where the POD-ANN solution agrees well with the LHS reference solution, particularly for $N_s = 1000$, as shown in Fig.4.3d.

To further investigate the ability of the proposed non-intrusive POD-BSBEM to accurately approximate the output stochastic response of quantities of interest, its results are compared with those from the Full-PCE and POD-ANN approaches, and to the results of 50 000 realizations of the LHS (reference) method, as shown in Fig.4.4. These results are presented solely in terms of the variation of the standard deviation profile, as the mean is well reproduced by all four methods and for any value of their hyper parameters, as shown in Figs.4.2-4.3. The results from the Full Polynomial chaos expansion model are provided in order to show the benefits of introducing the reduced order model concept in the estimation of the stochastic output response. The Full-PCE solution, obtained with a polynomial order of $p = 13$ and oversampling ratio $n_p = 2$, as suggested in (Raisee *et al.*, 2015), shows a good agreement with the LHS solution. It should be emphasized that the Full-PCE method is implemented by adopting the classical polynomial chaos decomposition such that the expansion's coefficients are computed by regression for each of the $N_e = 25\,600$ nodes representing the computational mesh considered in

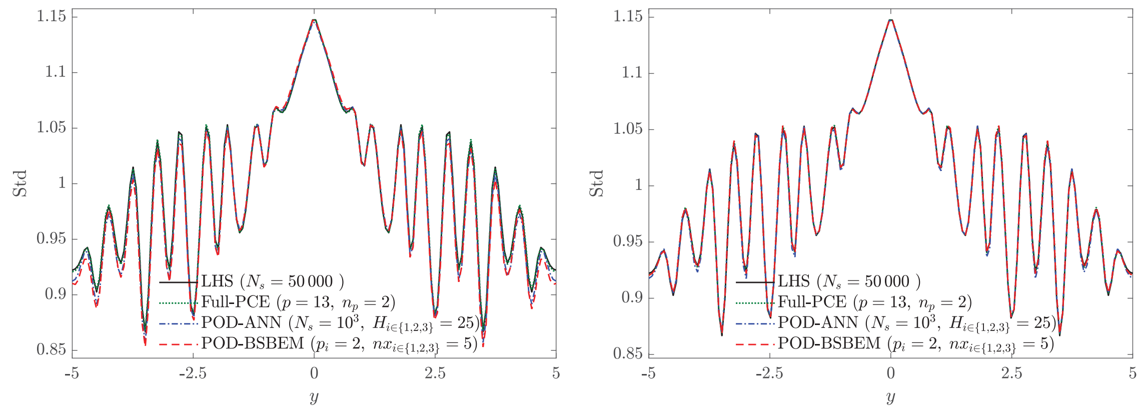


a) $\epsilon_s = 10^{-5}$ $(L = 6\ modes)$ b) $\epsilon_s = 10^{-5}$ $(L = 6\ modes)$

c) $\epsilon_s = 10^{-10}$ $(L = 14\ modes)$ d) $\epsilon_s = 10^{-10}$ $(L = 14\ modes)$

Figure 4.3 Statistical moments as a function of y (at $x = 0$) obtained using the non-intrusive POD-ANN with $H_{i\in\{1,2,3\}} = 25$ and $N_s = 300,\ 1000$: $L = 6$ modes (a-b) and $L = 14$ modes (c-d), compared to the LHS method (with $N_s = 50\,000$ realizations).

the present test case. As illustrated in Fig.4.4a, the results obtained by the reduced-order models POD-BSBEM and POD-ANN, built with $L = 6$ significant modes (corresponding to a threshold $\epsilon_s = 10^{-5}$), show a slight deviation in the std profile compared with those from the Full-PCE and LHS methods. Fig.4.4b shows a significant improvement in the prediction of the std profile obtained by the reduced-order models as the number of dominant modes increases ($L = 14$), particularly with the POD-BSBEM, where a faithful reproduction of the statistical moment can be observed, compared to that of the LHS reference approach, and with much less computational effort.



a) $\epsilon_s = 10^{-5}$ ($L = 6$ *modes*) b) $\epsilon_s = 10^{-10}$ ($L = 14$ *modes*)

Figure 4.4 Comparison of the standard deviation profiles obtained with the Full-PCE, POD-ANN and POD-BSBEM with that of the LHS reference solution (with 50 000 realizations). (a) : $\epsilon_s = 10^{-5}$ ($L = 6$ modes) ; (b) : $\epsilon_s = 10^{-10}$ ($L = 14$ modes).

For a more profound insight, the accuracy of POD-BSBEM in the approximation of output statistics of quantities of interest is investigated by evaluating the relative error in the ℓ^2-norm, expressed as follow :

$$Err^{Surr}_{\ell^2,\Phi} = \sqrt{\frac{\sum_{i=1}^{N_e} (\Phi_{i,Surr} - \Phi_{i,LHS})^2}{\sum_{i=1}^{N_e} (\Phi_{i,LHS})^2}} \tag{4.26}$$

with Φ denotes either the mean or the standard deviation, and $Surr$ stands for Full-PCE, POD-ANN and POD-BSBEM. Errors are computed by evaluating the statistical moments obtained by both full and reduced order surrogate models (Full-PCE, POD-ANN and POD-BSBEM) and comparing them with those of the LHS reference solution over the N_e total number of nodes representing the computational domain. It is obvious from the results reported in Table 4.1 that the decrease in the tolerance ϵ_s (which results in an increase in the POD dimension) leads to a decrease in the ℓ^2-error for both reduced-order models (POD-BSBEM and POD-ANN). This indicates that the proposed POD-BSBEM approximates the mean and standard deviation with the same accuracy as that of the Full-PCE model. However, it is worth mentioning that expansions' construction costs vary significantly between both models, as only 14 decompositions are required for POD-BSBEM (which corresponds to $\epsilon_s = 10^{-10}$) as a function of the random input parameters, where a vector of global coefficients has to be computed for each mode, while the

Full-PCE must build a decomposition for each node of the $N_e = 25\,600$ mesh points representing the entire computational domain. Hence, the proposed reduced-order models offers a very real gain in terms of reducing the computational efforts required for the prediction of the statistical moments without loss of accuracy compared to the Full-PCE, particularly for cases that require a high number of mesh nodes, such as for dam-break flow applications, where the computational domain may contain millions of nodes.

Tableau 4.1 Effect of ϵ_s on the relative error in the ℓ^2-norm for the mean and standard deviation obtained from Full-PCE ($p = 13$, $n_p = 2$), POD-ANN ($H_{i\in\{1,2,3\}} = 25$, $N_s = 10^3$) and POD-BSBEM ($p_{i\in\{1,2,3\}} = 2$, $nx_{i\in\{1,2,3\}} = 5$). Errors are computed with respect to the LHS reference solution (with $N_s = 50\,000$ realizations).

ϵ_s	L	$\mathrm{Err}^{\mathrm{Surr}}_{\ell^2,\mathrm{Mean}}$		$\mathrm{Err}^{\mathrm{Surr}}_{\ell^2,\mathrm{Std}}$	
		POD-ANN	**POD-BSBEM**	**POD-ANN**	**POD-BSBEM**
1E-03	3	9.1591E-05	1.9050E-04	0.0673	0.072232
1E-05	6	6.7062E-05	7.8300E-05	0.009251	0.007650
1E-08	11	5.3319E-05	3.3817E-05	0.004231	0.002541
1E-10	14	1.1563E-05	2.6996E-05	0.003692	0.002540
Full-PCE		1.7572E-05		0.002552	

To gain a better understanding of the proposed non-intrusive reduced-order model performance, the effect of the number of Bézier elements (nx_i) on the ℓ^2-error norm of the statistical moments for the tolerance $\epsilon_s = 10^{-10}$ is investigated. As with the former case, the results, reported in Table 4.2, are compared with those from the POD-ANN and the Full-PCE in terms of computational efficiency. The number of Bézier elements represents a key controlling parameter in the POD-BSBEM concept ; it defines the number of elements within which local basis functions with compact support are used to approximate the stochastic output response. It also contributes to defining the total number of collocation points (number of numerical solver calls) given by : $N_s = N_{elt}.nb$ where $N_{elt} = \prod_{i=1}^{m} nx_i$ and $nb = \prod_{i=1}^{m} (p_i + 1)$, where m denotes the number of random input parameters and p_i is the polynomial order of the basis functions. It is worth noting that comparing the computational costs of both reduced order-models (POD-ANN and POD-BSBEM) is especially interesting given that their accuracies are almost the same. Thus, for each value of nx_i, the aforementioned resulting number of collocation points N_s is considered as a size of the learning set to evaluate the ℓ^2-error in the mean and standard deviation of the POD-ANN. The reported results show that the POD-BSBEM approaches the LHS reference solution with the same accuracy as the Full-PCE as the number of Bézier elements increases, and with much lower decomposition costs. It is also worth noting that predictions from the POD-BSBEM are slightly more accurate than those from the reduced-order model-based POD-ANN, particularly when N_s takes relatively high values.

Tableau 4.2 Effect of the number of Bézier elements nx_i on the relative error in the ℓ^2-norm for the mean and standard deviation obtained from the Full-PCE ($p = 13$, $n_p = 2$), POD-ANN ($H_{i\in\{1,2,3\}} = 25$) and POD-BSBEM ($p_{i\in\{1,2,3\}} = 2$, $nx_{i\in\{1,2,3\}} = 5$) with $\epsilon_s = 10^{-10}$ ($L = 14$). Errors are computed with respect to the LHS reference solution (with $N_s = 50\,000$ realizations).

$\mathbf{nx_{i\in\{1,2,3\}}}$	$\mathbf{N_s}$	$\mathbf{Err}^{\mathbf{Surr}}_{\ell^2,\mathbf{Mean}}$		$\mathbf{Err}^{\mathbf{Surr}}_{\ell^2,\mathbf{Std}}$	
		POD-ANN	**POD-BSBEM**	**POD-ANN**	**POD-BSBEM**
2	216	2.9978E-05	0.001143	0.005805	0.007084
3	729	4.9126E-05	8.1399E-05	0.002548	0.003234
4	1 728	1.8819E-05	4.9126E-05	0.006231	0.002578
5	3 375	3.3262E-05	2.6996E-05	0.004934	0.002540
Full-PCE		1.7572E-05		0.002552	

The computational costs of the presented approaches are summarized in Table 4.3. All the programs involved in this test case were implemented serially and the computations were performed using a personal computer : Intel (R) Xeon (R) CPU E3-125 v6 @3.30 GHz with 16 GB of memory. It is worth mentioning that the offline phase involves the collection of snapshots from the high-fidelity solutions and building the POD reduced basis, learning, and computing coefficients by regression. The proposed reduced-order model (POD-BSBEM) has a lower computational cost in both its offline and online phases in comparison with the POD-ANN technique. It also presents a much better performance than the Full-order model polynomial chaos expansion (Full-PCE).

Tableau 4.3 Comparison of the computational cost required by the LHS ($N_s = 50\,000$ realizations), Full-PCE ($p = 13$, $n_p = 2$), POD-ANN ($H_{i\in\{1,2,3\}} = 25$) and POD-BSBEM ($p_{i\in\{1,2,3\}} = 2$, $nx_{i\in\{1,2,3\}} = 5$) with $\epsilon_s = 10^{-10}$ ($L = 14$). The time cost unit is a second (s).

	POD-ANN	**POD-BSBEM**	**Full-PCE**	**LHS**
Offline	3 988.51	141.02	-	-
Online	264.50	20.45	478.94	330.15

4.3.2 Burgers' equation test case

The second test case concerns the well-known viscous Burgers' equation, which describes the standard nonlinear time-dependent advection-diffusion problems in one spatial dimension. Its dimensionless form with a parameterized diffusion coefficient is given by (Burgers, 1948 ; Guo & Hesthaven, 2019) :

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \frac{1}{Re}\frac{\partial^2 u}{\partial x^2}\,, \qquad x \in [0,\ 1]\,, \qquad t \in [0,\ 1] \tag{4.27}$$

with the following initial and Dirichlet boundary conditions : $u(x,0) = \frac{x}{1+exp\left(\frac{Re}{16}(4x^2-1)\right)}$ and $u(0,t) = u(1,t) = 0$, respectively (Ahmed, San, Rasheed & Iliescu, 2020). Re denotes the Reynolds number, considered as an uncertain input parameter describing the variability surrounding the viscous effects. The aforementioned partial differential equation admits an exact analytical solution of the unknown field variable $u(x,t)$ expressed as (Maleewong & Sirisup, 2011 ; San, Maulik & Ahmed, 2019) :

$$u(x,t) = \frac{\frac{x}{t+1}}{1 + \sqrt{\frac{t+1}{t_0}}exp(Re\frac{x^2}{4t+4})} \tag{4.28}$$

with $t_0 = exp(\frac{Re}{8})$. The 1D computational space domain $\Omega = [0, 1]$ is discretized with a mesh of $N_e = 1000$ nodes in which the above analytical solution is evaluated. The discretization over the time domain $\mathcal{T} = [0, 1]$ is performed with $N_t = 50$ time instances, leading to a uniform time step $\Delta t = 0.02$.

The test case described above represents a challenging time-dependent problem for uncertainty propagation analysis. The analytical solution of the viscous Burgers' equation, given by Eq.(4.28), is adopted as a deterministic model through which uncertainties, translated in terms of the variability surrounding the diffusion coefficient, are propagated into the output quantity of interest. Thus, the Reynolds number is considered as an uncertain input parameter described by uniform distribution with two nominal values $\mu_{Re} = 200$ and $\mu_{Re} = 800$ and a coefficient of variation of $cv = 25\,\%$, which represent the standard deviation values of 50 and 200, respectively. The values of the input random Reynolds number are selected within its plausible variability ranges $Re_{\mu=200,\,\sigma=50} \in \mathcal{U}\,[114,\ 287]$ and $Re_{\mu=800,\,\sigma=200} \in \mathcal{U}\,[454,\ 1\,146]$, which correspond to a smooth and a very steep solution, respectively, of the Burgers' equation. This test case is of particular interest because it allows an efficient analysis of the efficacy of the proposed stochastic non-intrusive reduced order model compared to the other reduced and full-order models for both smooth and steep output responses with even discontinuous shock. The most relevant results are presented below and discussed in terms of statistical moments, ℓ^2-error, and probability density function profiles.

Fig.4.5 shows the contour plots of the mean (top row) and standard deviation (bottom row) of the Burgers' solution in the space-time plan, obtained with the Monte Carlo method for two variability ranges of the uncertain input parameter corresponding to $\mu_{Re} = 200$ (first column) and $\mu_{Re} = 800$ (second column) with the same coefficient of variation of $cv = 25\,\%$. These contour plots are presented for a better overall visualization of the statistical moments corresponding to smooth and step variation of the u-velocity, as can be observed in Figs.4.5a and 4.5b, respectively. The two vertical dashed lines denote the t locations for which the variation of the statistical moments' profiles obtained from the POD-BSBEM, POD-ANN, and Full-PCE models as a function of the x-space coordinate are compared to those from the MC method with $N_s = 10^6$ realizations, which is considered as a high-fidelity reference solution. It is worth mentioning that the feed-forward neural network architecture retained in the POD-ANN approach to deal with the Burgers' time-dependent test case is constituted by three hidden layers, each of which has 50 neurons ($H_{1,2,3} = 50$). The constructed networks were trained using the scaled conjugate gradient backpropagation algorithm within the MATLAB neural network training toolbox (*nntraintool*) with its available default options.

A comparison of the mean and standard deviation profiles as a function of the x-coordinate at different times ($t \approx 0.3,\ 1$) obtained with POD-BSBEM (with $p = 2$, $nx = 10$), POD-ANN (with $H_{1,2,3} = 50$, $N_s = 300$), and Full-PCE (with $p = 6$, $n_p = 2$) with those from the MC reference solution (with $N_s = 10^6$ realizations) for $\mu_{Re} = 200$ and $cv = 25\,\%$ is depicted in Fig.4.6. In addition to the high-fidelity MC method, the Full-PCE approach, which is considered as a stochastic expansion full-order model, is introduced in order to assess the contribution of the reduced-order aspect of the proposed POD-BSBEM in comparison with that of the POD-ANN. It can be observed that the statistical moments' distributions of the u-velocity obtained by the proposed non-intrusive reduced order model (POD-BSBEM) are in excellent agreement with the distributions from the full-order PCE and the high-fidelity MC reference solution. The variability of the relatively smooth output response of the Burgers' solution is localized around the moving shock wave, whose amplitude decreases as a function of time. Slight deviations in this output response can be observed in the POD-ANN solution as time evolves, conversely to the POD-BSBEM predictions, which present profiles nearly indistinguishable from those of the reference MC solution, as shown by the close-up views of the standard deviation profiles (Figs.4.6b and 4.6d).

A similar comparison is presented in Fig.4.7 for relatively high values of the uncertain input parameter with $\mu_{Re} = 800$. This increase in the mean value of the Reynolds number associated

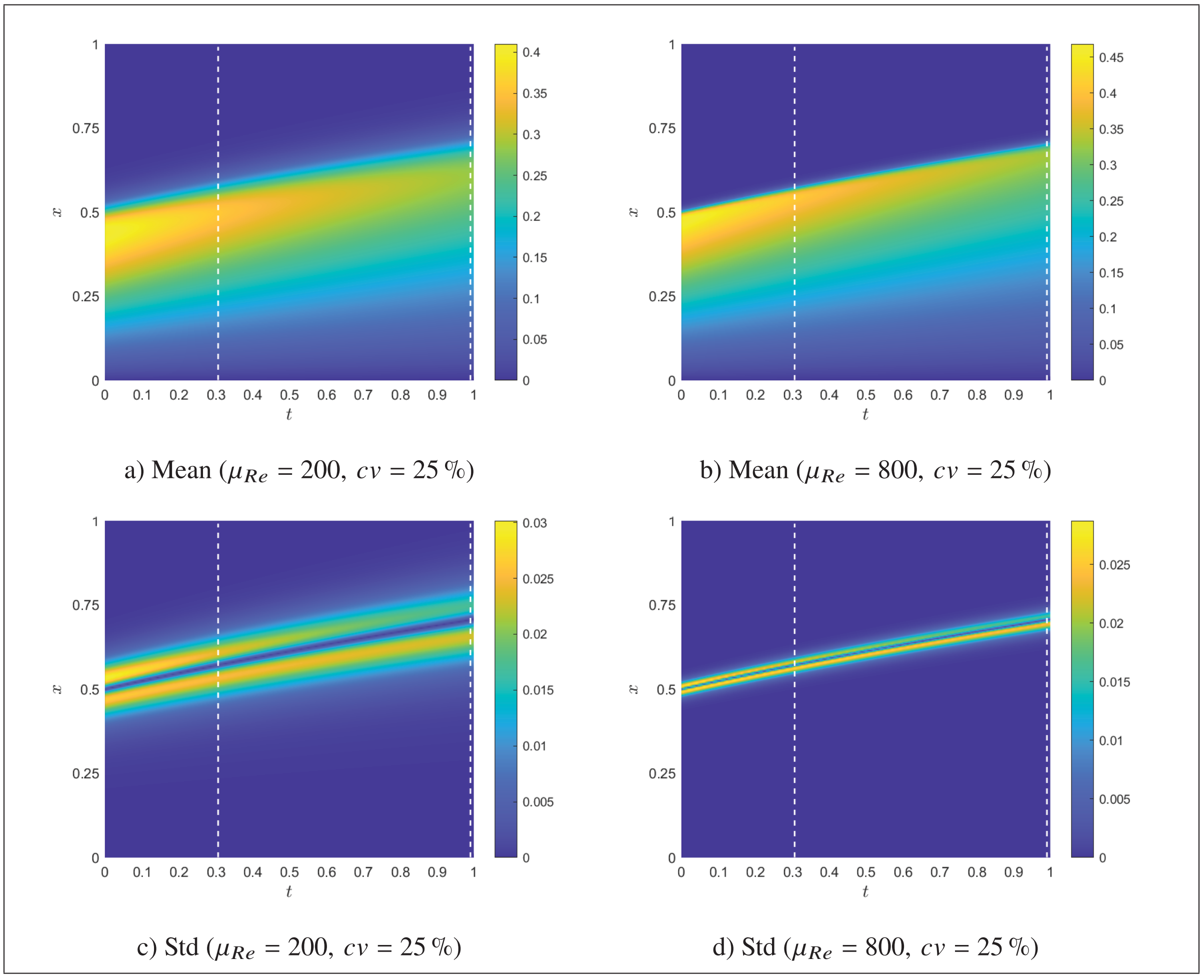


Figure 4.5 Contour plots of the mean (top row) and standard deviations (bottom row) of the Burgers' analytical solution obtained from 10^6 Monte Carlo realizations. (a and c) : $Re = 200$, (b and d) : $Re = 800$. The two vertical white lines indicate the times for which the statistics of the Full-PCE, POD-ANN and POD-BSBEM are compared to those of the MC reference solutions. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article).

with a wide variability range ($cv = 25$ %) leads to a visible transformation of the output response in the Burgers' solution, where a more pronounced steep slope can be observed throughout the velocity front wave propagation. These conditions represent a real challenge for the reduced-order models to faithfully reproduce the output response of the quantities of interest. While both reduced-order models (POD-BSBEM and POD-ANN) accurately estimate the mean of the Burgers' solution where the profiles are in good accordance with those from Full-PCE and

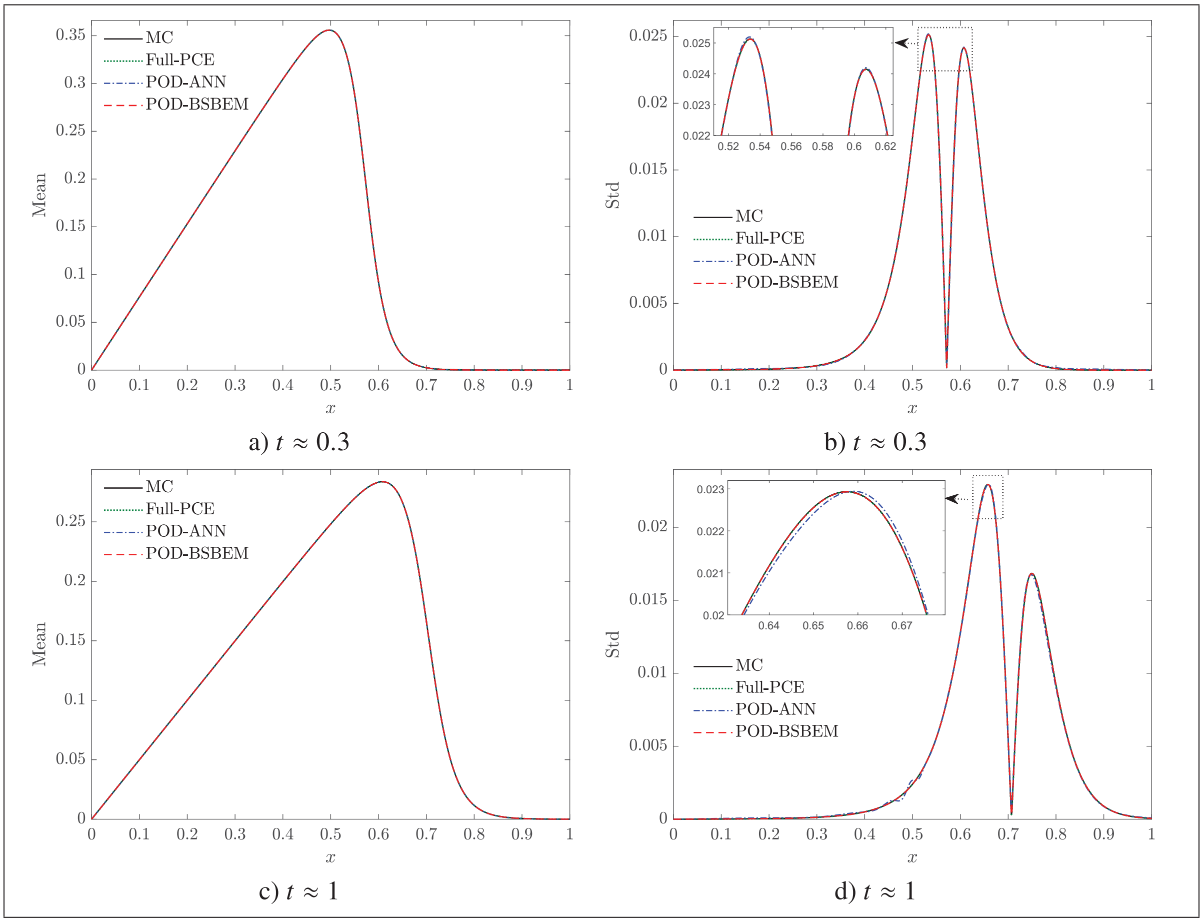


Figure 4.6 The mean (first column) and standard deviations (second column) of the Burgers' equation obtained with the POD-BSBEM ($p = 2,\ nx = 10$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and Full-PCE ($p = 6,\ n_p = 2$) compared to those of the MC solution (with 10^6 realizations) for $\mu_{Re=200}$ and $cv = 25\ \%$ at different times. (a and b) : $t \approx 0.3$ and (c and d) : $t \approx 1$.

MC, as depicted in the first column of Fig.4.7, it is clear that the POD-ANN approach shows deviations in the estimation of the peak value of the standard deviation profile where oscillations, which gain in intensity over time, can be observed. In contrast, the predictions obtained by the proposed POD-BSBEM are in excellent accordance with those from the Full-PCE and the high-fidelity MC reference solution, as shown by the close-up views in the second column of Fig.4.7.

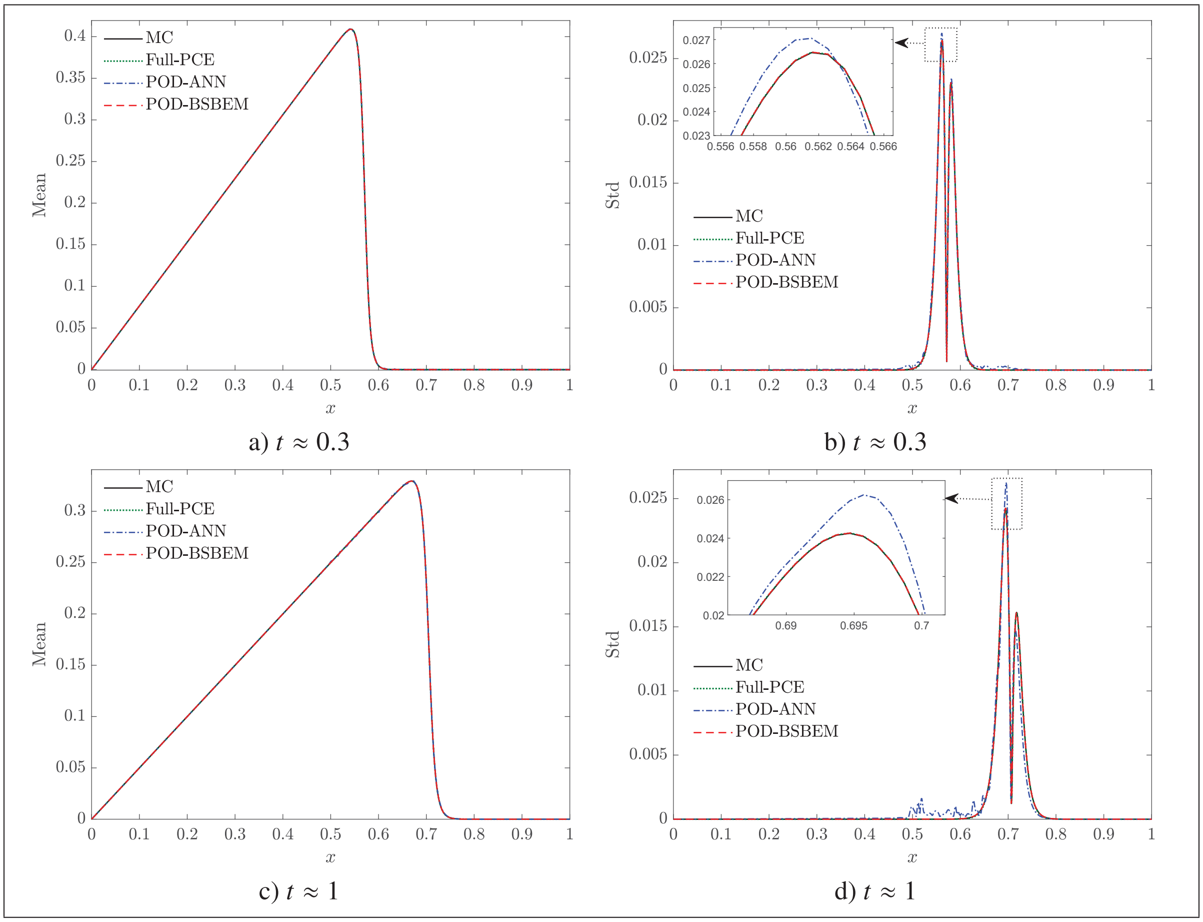
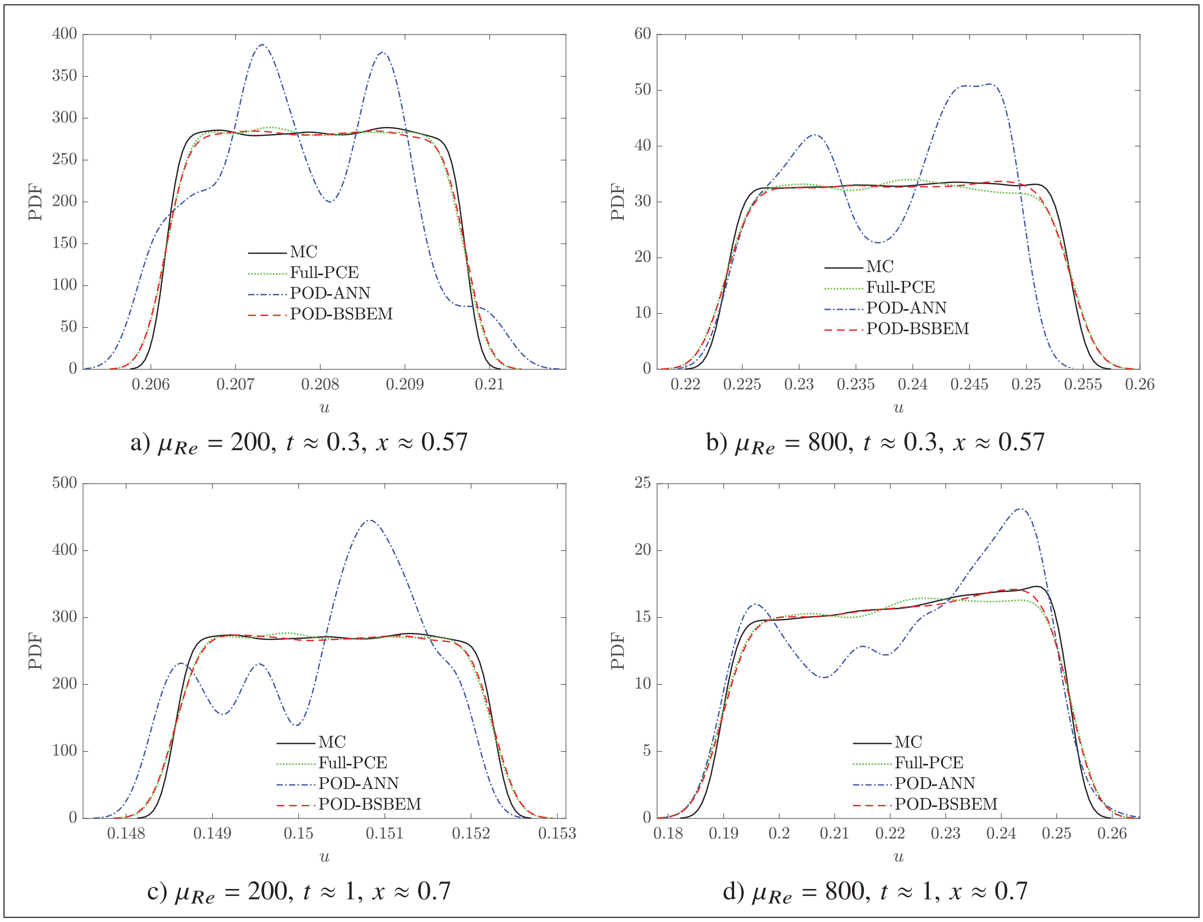


Figure 4.7 The mean (first column) and standard deviations (second column) of the Burgers' equation obtained with the POD-BSBEM (p = 2, nx = 10), POD-ANN ($H_{1,2,3}$ = 50, N_s = 300) and Full-PCE (p = 6, n_p = 2), compared to those of the MC solution (with 10^6 realizations) for $\mu_{Re=800}$ and cv = 25 % at different times. (a and b) : $t \approx 0.3$, and (c and d) : $t \approx 1$.

Another meaningful aspect for which the efficacy of the proposed POD-BSBEM is assessed, in addition to the estimation of the first two statistical moments, is its ability to accurately estimate the probability density function of the output response of quantities of interest comparison to the estimations of the other aforementioned techniques. To visualize this aspect, the distributions of the estimated kernel density function obtained by POD-BSBEM (p = 2, nx = 10), POD-ANN ($H_{1,2,3}$ = 50, N_s = 300) and Full-PCE (p = 6, n_p = 2) using the constructed surrogate models are compared with those from the 10^5 Monte Carlo realizations at different spatio-temporal locations for μ_{Re} = 200 and 800 with cv = 25 %, are depicted in Fig.4.8. It should be noted that the spatial locations where the pdfs are estimated correspond approximately to the positions of the front wave of the Burgers' solution for a given t-value. From Fig.4.8, it is obvious that all the pdf plots resulting from the proposed POD-BSBEM follow nearly the same trend of the pdfs obtained from the Full-PCE and the MC high-fidelity reference solution, whose shapes are almost uniform, unlike the POD-ANN approach which presents clear discrepancies in its pdf plots for all the represented x and t locations. The effect of increasing the mean value of the input Reynolds number from $\mu_{Re} = 200$ to $\mu_{Re} = 800$ with a coefficient of variation of $cv = 25\,\%$, which is traduced physically by a steep solution of the Burgers' equation, is visible in the variability range of the u-output response that widens as time evolves, as can be seen in the second column of Fig.4.8.

Fig.4.9 presents a complementary quantitative assessment of the efficiency of the proposed POD-BSBEM technique was performed via the ℓ^2-error analysis. The variation of the ℓ^2-error profiles of the mean and standard deviation, whose mathematical expression in its time-dependent form is given by the right term of Eq.(4.29), resulting from the POD-BSBEM, POD-ANN and Full-PCE for $\mu_{Re} = 200$ and $\mu_{Re} = 800$ with $cv = 25\,\%$. The relative error in the ℓ^2 norm is computed with respect to the MC high fidelity reference solution for each time step and over the N_e nodes that contains the computational domain. It can be seen that the predicted errors of the POD-BSBEM are small and have the same order of magnitude ($\approx 10^{-6}$ and 10^{-3} for the mean and standard deviation, respectively) as the Full-PCE technique, unlike the predicted errors from the POD-ANN, particularly for the standard deviation, which are on the order of 10^{-1} when the input Reynolds number takes relatively high values. This indicates that the proposed POD-BSBEM yields accurate predictions of statistical moments of both smooth and steep time-dependent Burgers' solutions.

The effect of the number of POD modes on the performance of the proposed POD-BSBEM approach is assessed through the maximum value of the time-dependent relative error in the ℓ^2-norm, retained as a criterion for evaluating and comparing the accuracy of the different



a) $\mu_{Re} = 200,\ t \approx 0.3,\ x \approx 0.57$ b) $\mu_{Re} = 800,\ t \approx 0.3,\ x \approx 0.57$

c) $\mu_{Re} = 200,\ t \approx 1,\ x \approx 0.7$ d) $\mu_{Re} = 800,\ t \approx 1,\ x \approx 0.7$

Figure 4.8 Kernel density estimated probability density function (PDF) of Burgers' solutions obtained with POD-BSBEM ($p = 2,\ nx = 10$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$ and Full-PCE ($p = 6,\ n_p = 2$) compared to the MC reference solution (with 10^6 realizations) for $\mu_{Re=200}$ (first column) and $\mu_{Re=800}$ (second column). (a and b) : $t \approx 0.3,\ x \approx 0.57$, and (c and d) : $t \approx 1,\ x \approx 0.7$.

models. Its mathematical formulation is expressed as follows :

$$Err_{\ell^2}^{max} = max\left(Err_{\ell^2,\Phi}^{Surr}(t) = \sqrt{\frac{\sum_{i=1}^{N_e} \left(\Phi_{i,Surr}(t) - \Phi_{i,MC}(t)\right)^2}{\sum_{i=1}^{N_e} (\Phi_{i,MC}(t))^2}} \right) \tag{4.29}$$

The ℓ^2-error mentioned above is evaluated with respect to the MC reference solution over all the computational nodes for each time step. The variation of the $Err_{\ell^2}^{max}$ error as a function of the number of POD modes L for two variability ranges of the uncertain input Reynolds

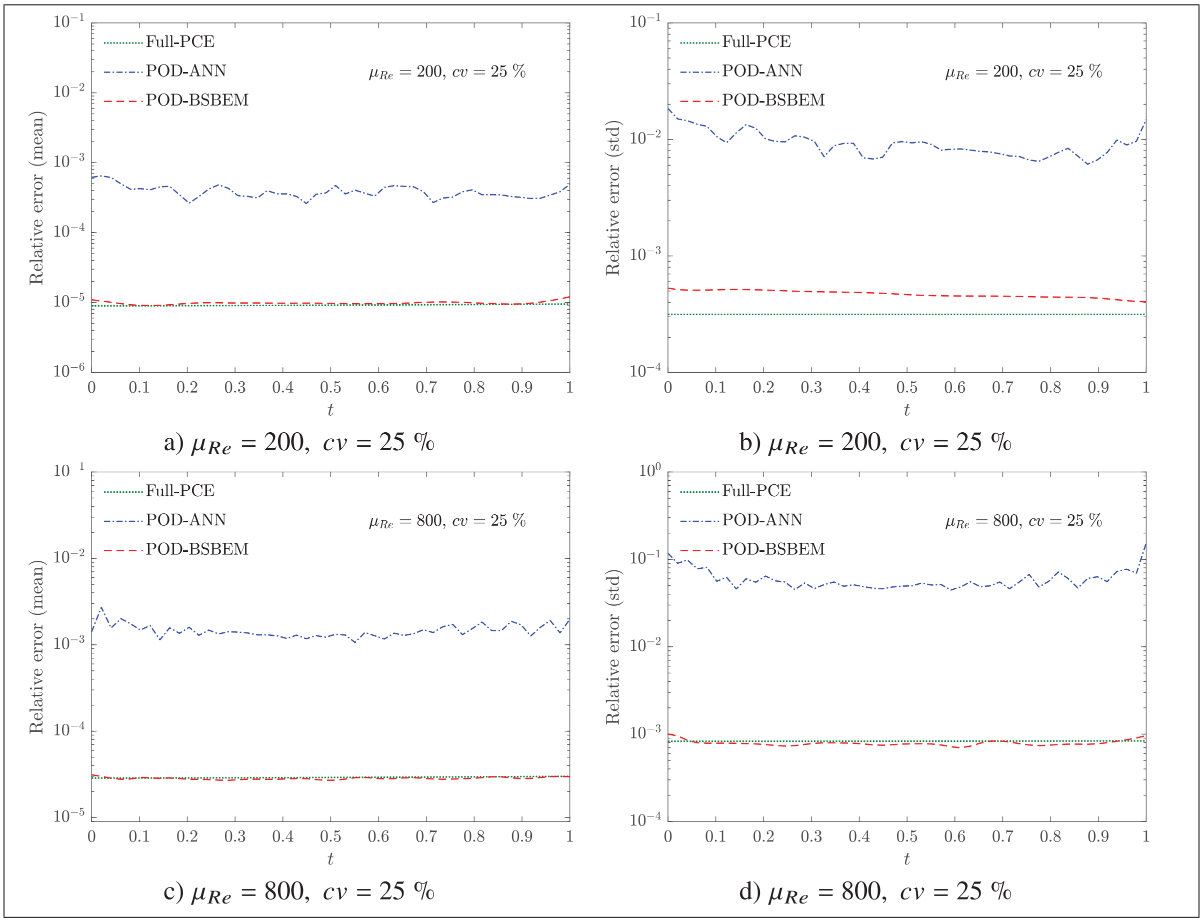


Figure 4.9 Comparison of the relative ℓ^2-error profiles of the mean (first column) and standard deviations (second column) as a function of time, obtained with the Full-PCE ($p = 6,\ n_p = 6$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and POD-BSBEM ($p = 2,\ nx = 10$). Errors are computed with respect to the MC reference solution (with 10^6 realizations). (a and b) : $\mu_{Re} = 200,\ cv = 25$ %, (c and d) : $\mu_{Re} = 800,\ cv = 25$ %.

number is reported in Tables 4.4 and 4.5, respectively. It should be mentioned that the $Err_{\ell^2}^{max}$ of the Full-PCE approach is not affected by the variation of the POD modes' number since it represents a full-order model ; this variation was included in the aforementioned tables to serve as a reference with which to compare the accuracy of the reduced-order models.

Thus, for relatively low values of the input Reynolds number, the results reported in Table 4.4 show that errors in both statistical moments decrease as the number of POD modes increases (as a result of decreasing the threshold $\epsilon_s = \epsilon_t$), where very low values can be observed, particularly in the mean error. However, a clear discrepancy can be noted in the standard deviation errors obtained with $L = 34$ modes (which corresponds to $\epsilon_s = \epsilon_t = 10^{-10}$) between the POD-ANN and POD-BSBEM techniques, which are on the order of 10^{-2} and 10^{-4}, respectively. The same trend can be observed for relatively high values of the Reynolds number with a wide variability range, where the number of POD modes, for the same threshold value, is higher than in the former case (low Reynolds number values). The predicted errors obtained from the POD-BSBEM, which are of the same degree of magnitude as those of the Full-PCE, are much lower than the predicted errors of the POD-ANN approach for both the mean and the standard deviation, as reported in Table 4.5. This quantitative analysis clearly illustrates the ability of the proposed model to accurately estimate the time-dependent stochastic output response of quantities of interest.

Tableau 4.4 Effect of ϵ_s and ϵ_t on the maximum relative error over time in the ℓ^2-norm for the mean and standard deviation obtained from Full-PCE ($p = 6$, $n_p = 2$), POD-ANN ($H_{i\in\{1,2,3\}} = 50$, $N_s = 300$) and POD-BSBEM ($p = 2$, $nx = 10$). Errors are computed with respect to the MC reference solution (with $N_s = 10^6$ realizations) for $Re_{\mu=200,\sigma=50} \in \mathcal{U}\,[114,\ 287]$.

$\epsilon_s,\ \epsilon_s$	L	$\mathbf{Err}^{\mathbf{max}}_{\ell^2,\,\mathbf{Mean}}$		$\mathbf{Err}^{\mathbf{max}}_{\ell^2,\,\mathbf{Std}}$	
		POD-ANN	**POD-BSBEM**	**POD-ANN**	**POD-BSBEM**
1E-03	7	0.008882	0.009098	0.211378	0.409364
1E-05	13	8.4564E-04	5.3330E-04	0.024265	0.023550
1E-08	24	6.1322E-04	1.1664E-04	0.017521	7.9154E-04
1E-10	34	8.9175E-04	1.2082E-05	0.021630	5.2967E-04
Full-PCE		1.3006E-05		5.4106E-04	

As with the Ackley test case, the estimation of the computational costs of the proposed approach are compared to those of other techniques for this time-dependent test case ; these are presented in Table 4.6. The computations were carried out on a personal computer : an Intel (R) Xeon (R) CPU E3-125 v6 @3.30 GHz with 16 GB of memory. The results reveal that the online estimation of the statistical moments of the output response of quantities of interest by the POD-BSBEM can be made very quickly. Indeed, the associated speed-up factors of the present time-dependent

Tableau 4.5 Effect of ϵ_s and ϵ_t on the maximum relative error over time in the ℓ^2-norm for the mean and standard deviation obtained from Full-PCE ($p = 6$, $n_p = 2$), POD-ANN ($H_{i\in\{1,2,3\}} = 50$, $N_s = 300$) and POD-BSBEM ($p = 2$, $nx = 10$). Errors are computed with respect to the MC reference solution (with $N_s = 10^6$ realizations) for $Re_{\mu=800,\,\sigma=200} \in \mathcal{U}\,[454,\ 1\,146]$.

ϵ_s, ϵ_s	**L**	$\mathbf{Err}^{\mathbf{max}}_{\ell^2,\,\mathbf{Mean}}$		$\mathbf{Err}^{\mathbf{max}}_{\ell^2,\,\mathbf{Std}}$	
		POD-ANN	**POD-BSBEM**	**POD-ANN**	**POD-BSBEM**
1E-03	12	0.020304	0.019633	0.653037	0.707108
1E-05	26	0.002284	0.001225	0.109855	0.003519
1E-08	53	0.002546	4.5725E-05	0.118874	0.002541
1E-10	76	0.002233	3.0940E-05	0.100247	0.001056
Full-PCE		3.0616E-05		0.001110	

test case, evaluated with respect to the time required by the crude Monte Carlo method, are about 885 for POD-BSBEM, 164 for Full-PCE, and 39 for POD-ANN, thus demonstrating the efficiency of the proposed approach compared to other techniques.

Tableau 4.6 Comparison of the time costs of the Full-PCE ($p = 6$, $n_p = 2$), POD-ANN ($H_{1,2,3} = 50$), POD-BSBEM ($p = 2$, $nx = 10$) and the MC (with $N_s = 10^6$ realizations), with $\epsilon_s = \epsilon_t = 10^{-10}$ ($L = 76$) and $Re_{\mu=800,\,\sigma=200} \in \mathcal{U}\,[454,\ 1\,146]$. The unit of the time cost is second (s).

	POD-ANN	**POD-BSBEM**	**Full-PCE**	**MC**
Offline	4 513.07	5.73	-	-
Online	98.93	4.32	23.29	3 824.83

4.3.3 Application to a hypothetical dam-break with real terrain data

In addition to the two benchmark test cases presented above for steady and time-dependent problems, the aforementioned modeling approaches were applied to uncertainty propagation analysis regarding a hypothetical dam-break with real terrain data within a reach of the Mille Iles river, which takes its source from the Lake of Two Mountains and discharging downstream into the

Des Prairies river near its confluence with the Saint-Laurent river (province of Québec, Canada). The data related to the bathymetry and the Manning roughness coefficient of the studied reach were provided by the Communauté Métropolitaine de Montréal (CMM). The hydraulic behavior of the dam-break flow is described by a deterministic in-house finite volume CuteFlow solver for shallow water equations in its multi-GPU version (Delmas & Soulaïmani, 2020 ; Zokagoa & Soulaïmani, 2012). The discretized free surface equations were solved through an unstructured mesh composed of 243 161 nodes connected to 481 930 triangular elements, representing the whole computational domain, half of which is shown in Fig.4.10a. A part of the computational domain, delineated by a sub-domain surrounding the location of the dam as shown in Fig.4.10b, composed of $N_e = 10\,200$ nodes and 16 763 elements was considered as a domain of study from which snapshots of the quantities of interest were extracted to construct the reduced-order models.

This test case describes a fictitious sudden rupture of a dam located within the domain of the study area, as shown by the line indicating its position in Fig.4.10b. The initial conditions are defined by unequal water levels on both sides of the dam. The downstream part is considered as dry ($\eta_{ds} = b$, where b denotes the bathymetry of the domain), while the initial upstream water surface elevation (η_{up}) is considered a uniformly-sampled uncertain input parameter with a given sample size (N_s) within its plausible variability range ($\eta_{up} \in \mathcal{U}\,[29,\ 32]\ m$). A single value is selected from the generated random space to run the CuteFlow deterministic solver for a simulation time of 50 s. The obtained high-fidelity solutions of quantities of interest over the N_e computational nodes for each time-step, defined as a time-parameter dependent snapshot, were recorded to construct the snapshot matrix. The number of time points considered is $N_t = 100$, which is much lower than the value from the adaptive time-steps used in the deterministic numerical solver.

The results of the dam-break test case are presented mainly in terms of the mean and standard deviation profiles of the water surface level and the confidence interval distribution contours of the flooding lines accompanying the propagation of the shock wave over the whole computational domain. The solutions obtained from the LHS method with $N_s = 2000$ are considered as reference solutions with which to assess the accuracy of the proposed approaches. The comparisons of the obtained statistical moments are performed over a cross-section line in the domain of study at different simulation times, as well as at different gauging points as a function of time, as shown in Fig.4.10b. The construction of the reduced-order models (POD-BSBEM and POD-ANN) was performed by a series of computational tests to select the most dominate modes constituting the POD basis. Thus, a dual-POD was performed for both reduced models, with $\epsilon_t = 10^{-7}$ and $\epsilon_s = 10^{-8}$ for POD-BSBEM, and $\epsilon_t = \epsilon_s = 10^{-8}$ for POD-ANN, producing $L = 824$ modes (with $N_s = 60$) and $L = 3387$ modes (with $N_s = 300$), respectively. The Artificial Neural Network used in the POD-ANN model is composed of feedforward neural networks which each have three hidden layers with 50 neurons in each ($H_{1,2,3} = 50$). As for the former time-dependent test case, the constructed networks were trained using the scaled conjugate gradient backpropagation algorithm within the MATLAB neural network training toolbox (*nntraintool*) with its available default options. The data set used in the learning process was obtained from the $N_s = 300$ high fidelity solutions of the deterministic model that were used for the snapshot matrix. These were divided into training, validation and testing subsets using the MATLAB *dividerand* function with its default parameters.

Fig.4.11 shows the mean profiles of the water surface level over the cross-section line x_{line} obtained with the Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and POD-BSBEM ($p = 2,\ nx = 20$) at different simulation times, compared with those from the LHS reference solution ($N_s = 2\,000$ realizations). The bathymetry of the terrain is also represented in the figure in order to show the propagation of the flooding wave over the selected cross-section as the simulation time evolves. In general, the mean distribution profiles of all methods appear to be in good concordance with those from the reference LHS solution at the represented simulation times. However, the comparison of the standard deviation profiles, depicted in Fig.4.12, shows clear deviations between the POD-ANN and the LHS results, unlike those obtained by the POD-BSBEM method where an excellent superposition is observed with the LHS reference results for all simulation times. In addition to the relatively significant learning time required

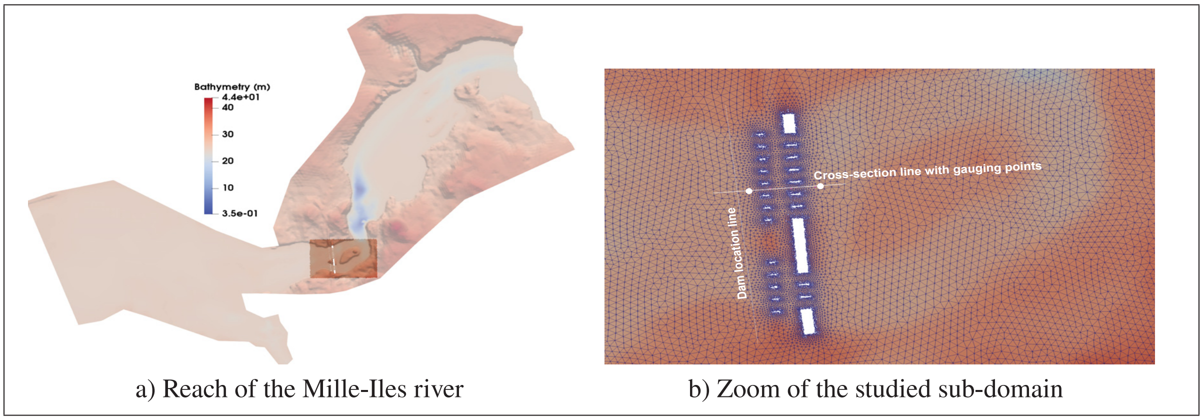


Figure 4.10 Sketch of the reach of the Mille Iles River. (a) : Bathymetry of the reach of the Mille Iles River ; (b) : Zoom of the sub-domain of study ; the horizontal line (x_{line}) indicates the cross-section over which results are presented, whereas the vertical line delineates the position of the dam from which the breaching process was initiated. The two points (from point 1 near the dam location to point 2 downstream) represent the gauging positions where results are depicted as a function of time (for interpretation of the references to color in this figure, please refer to the web version of this article).

by the POD-ANN technique in its offline stage, a total of 30 000 ($N_s = 300 \times N_t = 100$) high fidelity solutions were collected for computing the POD basis, versus 6 000 (60×100) for the proposed POD-BSBEM, which represents a reduction in the number of deterministic solver calls and thus in the computational effort required for the construction of the reduced basis space. Another point of comparison concerns the reduced-order aspect that characterizes the POD-BSBEM compared to the Full-order polynomial chaos expansion (Full-PCE). Indeed, while both techniques show a good approximation of the statistical moments, the Full-PCE is used to build a stochastic expansion that approximates the output response in each node over the whole computational space-time domain ($N_e \times N_t$), and the proposed POD-BSBEM only needs the construction of $\sum_{\ell}^{L} \mathcal{K}_{\ell}$ expansions over the L modes of the reduced basis.

In Fig.4.13, the temporal evolution of the mean profiles of the water level is depicted at two gauging points. These points are selected from the sub-domain of study near the cross-section line to show the abilities of the proposed POD-BSBEM to accurately approximate the statistical moments of the output responses with hyperbolic behavior as the simulation time evolves. The

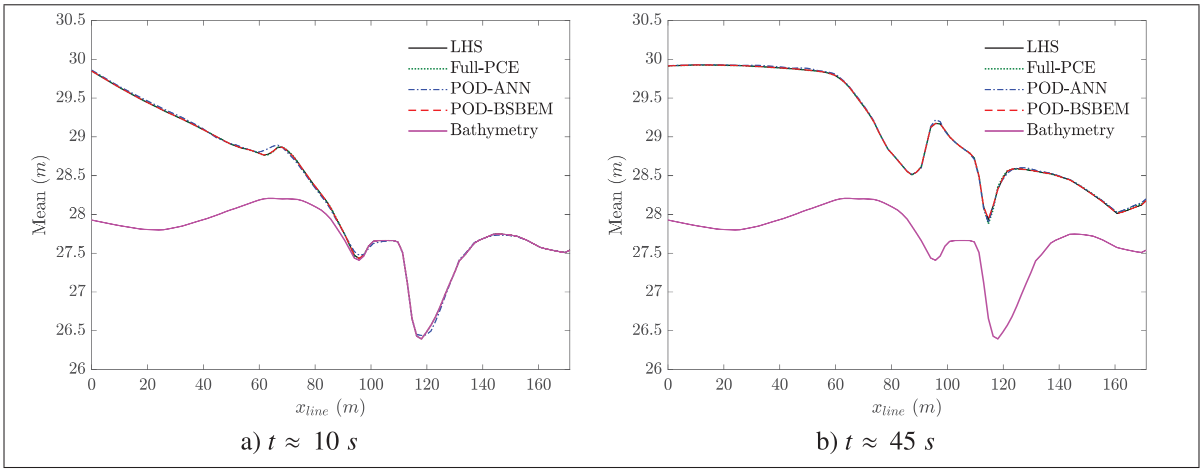


Figure 4.11 The mean profiles of the water levels over the cross-section line at different simulation times, obtained with the Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and POD-BSBEM ($p = 2,\ nx = 20$), compared to those of the LHS reference solution ($N_s = 2\,000$). (a) : $t \approx 10\ s$, (b) : $t \approx 45\ s$

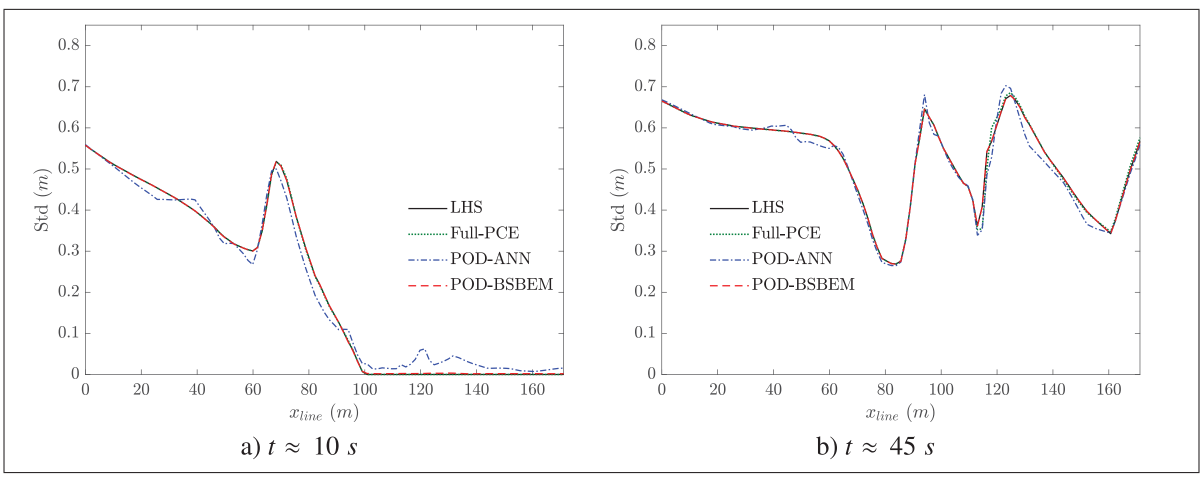


Figure 4.12 The standard deviation profiles of the water levels over the cross-section line at different simulation times, obtained with the Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and POD-BSBEM ($p = 2,\ nx = 20$), compared to those of the LHS reference solution ($N_s = 2\,000$). (a) : $t \approx 10\ s$, (b) : $t \approx 45\ s$

results are compared to those of the POD-ANN and Full-PCE techniques, as well as to the results of the reference LHS solutions. These plots reveal that both the POD-ANN and the Full-PCE techniques present some discrepancies in their profiles ; a slight oscillatory behavior can be observed, particularly at gauging point $P2$ as shown in Fig.4.13b. In contrast, the POD-BSBEM gives particularly good predictions in the mean-field reconstruction of the water level, where an excellent agreement can be observed between the proposed POD-BSBEM and LHS profiles at all gauging points over the whole temporal domain. This trend is confirmed by Fig.4.14 where the temporal evolution of the standard deviation of the water level is represented, and which seems to be much more sensitive than the mean profile. Indeed, it is clear that both the POD-ANN and the Full-PCE techniques present oscillations, and that their predictions of the standard deviation are less satisfactory, revealing non-physical behavior at some gauging points Fig.4.14b. The PCE basis functions are known to be less accurate than other approaches at estimating the strong hyperbolic behavior that characterizes the propagation of a flooding wave. On the other hand, an excellent agreement between the POD-BSBEM and LHS results can be observed for all gauging points. This gives a clear idea about the abilities of the POD-BSBEM to capture the temporal dynamics of the output quantities of interest and also indicates the potential that the piecewise B-splines basis functions offer to alleviate oscillatory behavior due to the mathematical features they procure and the multi-elements' aspect of the proposed approach.

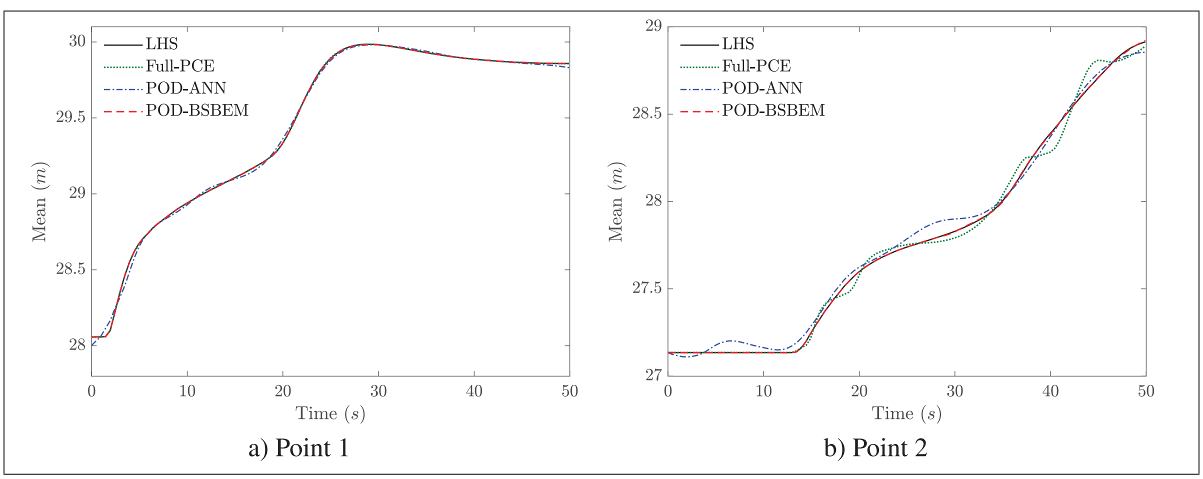
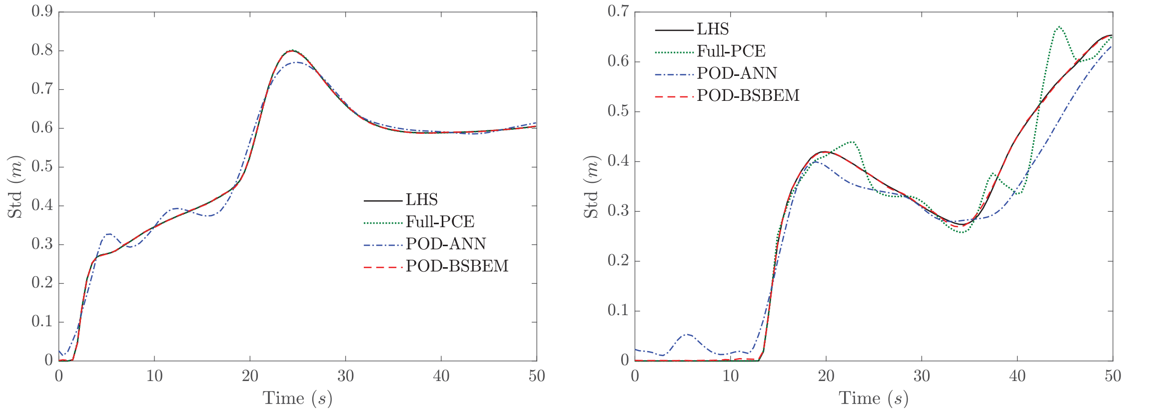


Figure 4.13 The mean water level profiles at gauging points as a function of time, obtained with the Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and the POD-BSBEM ($p = 2,\ nx = 20$), compared to those from the LHS reference solution ($N_s = 2\,000$). (a) : Point 1, (b) : Point 2.



a) Point 1 b) Point 2

Figure 4.14 The standard deviation profiles of the water levels at gauging points as a function of time, obtained with Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and POD-BSBEM ($p = 2,\ nx = 20$), compared to those from the LHS reference solution ($N_s = 2\,000$). (a) : Point 1, (b) : Point 2.

Fig.4.15 illustrates the 95 % confidence interval (95 % CI)-bound contours of the water depth, representing the probabilistic inundation map obtained by the POD-BSBEM and the POD-ANN, compared to the contours of the Full-PCE and the LHS reference solutions at different simulation times. The CI is quantified in each node of the computational domain by the mean, plus and minus two standard deviations of the water depth. The flooding line is selected by the threshold water depth of 5 cm defining the borderline between the wet and the dry areas. Thus, the projection of the lower and upper bounds of the CI (represented by the dashed and solid lines, respectively) corresponding to the aforementioned threshold of 5 cm onto the topography of the studied sub-domain delineate the variability range surrounding the flooding line. These plots show the temporal evolution of the variability range surrounding the flooding line, which becomes widespread as the simulation time evolves. This evolution traces the propagation of the uncertainty stemming from the initial input water level where the break was initiated. The POD-BSBEM clearly yields accurate predictions of the CI contours of the flooding line for all simulation times. Indeed, the superimposition of the flooding line contours illustrates an excellent agreement with the results from the Full-PCE and the reference LHS solutions (with 2 000 realizations). However, obvious deviations can be seen in the predictions from the POD-ANN technique, particularly at times $t = 10$ and $25\ s$, as shown in Figs.4.15a and 4.15b, where the obtained variability range contours do not match with the contour from the reference LHS solution.

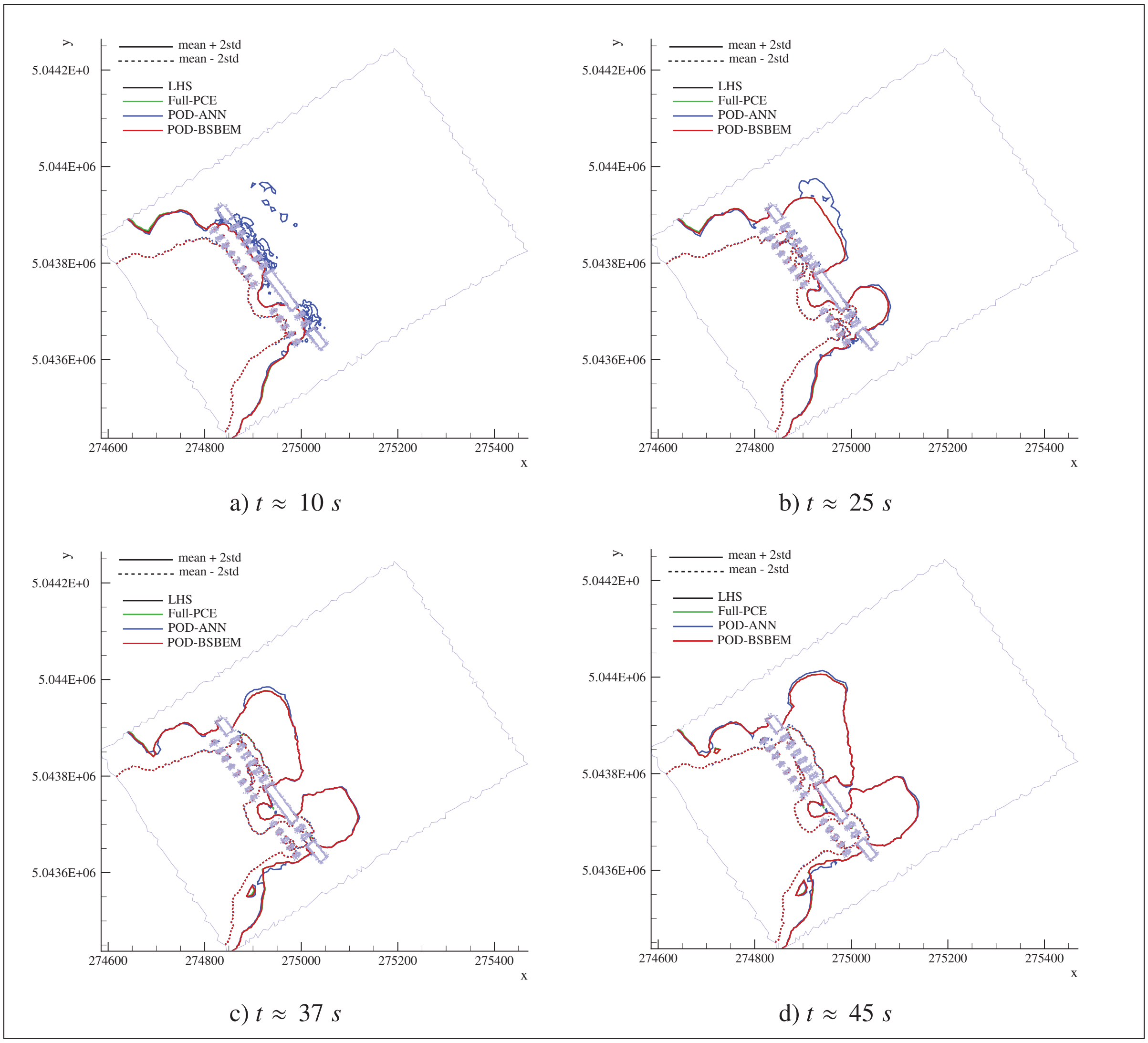


Figure 4.15 The 95 % confidence interval contours of flooding lines ($h = 0.05\,m$) at different simulation times, obtained with the Full-PCE ($p = 6,\ n_p = 2$), POD-ANN ($H_{1,2,3} = 50,\ N_s = 300$) and the POD-BSBEM ($p = 2,\ nx = 20$), compared to those from the LHS reference solution ($N_s = 2\,000$). The dashed and solid lines delineate the lower and upper bounds, respectively, of the 95 % CI of the flooding line. (a) : $t \approx 10\ s$, (b) : $t \approx 25\ s$, (c) : $t \approx 37\ s$, (d) : $t \approx 45\ s$. (For interpretation of the references to color in this figure, please refer to the web version of this article).

To illustrate the computational efficiency of the proposed method, its time cost is compared with the costs of other techniques, as presented in Table 4.7. The comparison mainly concerns the reduced-order models in their offline and online phases. The high-fidelity solutions of all these methods were obtained separately by running the numerical solver CuteFlow on a $V100$ GPU whose average computation time in the half-domain of study is almost 30 s. Therefore, this aspect is not included in the offline phase of the POD-ANN and POD-BSBEM approaches. Instead, it is considered separately by comparing the number of numerical solver calls (N_s) of each technique, as reported in the third row of Table 4.7. The POD-ANN implementation was performed on a V100 GPU with 32 GB of memory, with the MATLAB toolbox utilized to train the Artificial Neural Network with the aforementioned architecture, while the POD-BSBEM and Full-PCE were implemented serially and their computations were performed on a personal computer : an Intel (R) Xeon (R) CPU E3-125 v6 @3.30 GHz with 16 GB of memory. The data in Table 4.7 indicates that the POD-BSBEM presents a much lower time cost for both offline and online stages than the POD-ANN. The number of high-fidelity solutions required to build the surrogate reduced-order models also shows a significant difference between the two techniques : $N_s = 300$ for the POD-ANN and only $N_s = 60$ for the POD-BSBEM.

Tableau 4.7 Comparison of the time cost required by the POD-ANN ($H_{1,2,3} = 50,\ \epsilon_t = \epsilon_s = 10^{-8}$), POD-BSBEM ($p = 2,\ nx = 20,\ \epsilon_t = 10^{-7},\ \epsilon_s = 10^{-8}$), Full-PCE ($p = 6,\ n_p = 2$) and the LHS ($N_s = 2\,000$). The unit of time cost is the second (s).

	POD-ANN	**POD-BSBEM**	**Full-PCE**	**LHS**
Offline (s)	5 839.93	198.64	-	-
Online (s)	1 612.04	382.21	-	-
Number of solver calls (N_s)	300	60	14	2 000
Time cost per snapshot (s)	1.612	0.350	0.144	30

The time cost required for each snapshot reported in the last row of Table 4.7 shows that the use of the constructed surrogate models significantly reduces the computational cost compared to the high-fidelity numerical solver. The POD-BSBEM is more efficient than the POD-ANN, reducing the computation effort while maintaining good accuracy in the prediction of the output response. The results also reveal that the Full-PCE requires a lower number of high-fidelity solutions and therefore less computation cost per snapshot, despite its less accurate predictions with oscillatory behavior. This lower requirement is due to the univariate basis functions of polynomial chaos (one input random variable), which involves a minimal number of expansion terms. Indeed, the total number of expansion terms (coefficients) that have to be computed over the $N_e \times N_t = 10200 \times 100$ space-time nodes is only on the order of $M = p + 1 = 7$. However, this number grows significantly once the number of random variables increases, expressed as of $M = \frac{(p+m)!}{p!m!}$, and therefore the time required to evaluate each expansion increases as well (as shown in Table 4.3 for three random variables), particularly for very large computational domains with millions of meshing nodes. In contrast, the POD-BSBEM approach only constructs expansions over the L-modes of the reduced basis, which offers a significant gain in the computational cost, in addition to procuring accurate predictions for time-dependent problems with strong hyperbolic behavior.

4.4 Conclusion

This paper proposes a non-intrusive reduced-order model-based B-splines Bézier elements method (POD-BSBEM) for uncertainty propagation analysis for stochastic steady-state and time-dependent problems. The method combines the advantages of both the model reduction technique based on proper orthogonal decomposition (POD) and a stochastic approach, the B-splines Bézier elements method (BSBEM), which is based on splitting the random parameter domain into subspaces called Bézier elements. The construction of the reduced basis is performed from a collection of high-fidelity solutions through a two-level POD procedure by compressing the time-trajectory snapshots in the first POD level and then extracting the final reduced basis via a second POD level. Each associated projection coefficient dataset, which is randomly parametrized and time-dependent, is expressed as a regression constituted as a series of time-dependent reduced modes, generated via a third POD level, and their associated unknown stochastic parameter-dependent coefficients. These coefficients are then approximated as a local stochastic expansion using the local B-splines basis functions defined within the local

Bézier element constituting a part of the random parametric domain.

Two benchmark test cases, representing the steady-state Ackley function, with three random input parameters , and a univariate time-dependent viscous Burgers' equation, were used to investigate the performance and accuracy of the proposed method with respect to the non-intrusive reduced and full-order techniques POD-ANN and Full-PCE, respectively. The numerical results presented in this paper demonstrate the abilities of POD-BSBEM to predict the statistics and the estimated probability density function of the output quantities of interest, particularly for the time-dependent viscous Burgers' equation where a smooth profile of the standard deviation is obtained, unlike the POD-ANN approach which shows an oscillatory behavior. The results also highlighted the computational efficiency of the proposed approach, as they reveal a significant reduction in the computational effort required for both offline and online stages compared to the other techniques.

The proposed approach was then applied to investigate uncertainty propagation through a numerical model describing the highly non-linear flood wave stemming from a hypothetical dam-break with real terrain data of a reach of the Mille-Iles river located in the province of Québec. The results reveal that predictions from POD-BSBEM provide smooth profiles of the standard deviation as a function of time at the gauging points, where excellent agreements with the reference LHS profiles were observed, unlike the predictions from POD-ANN and Full-PCE that present a strong oscillatory behavior. Another meaningful finding is the ability of POD-BSBEM to provide accurate probabilistic inundation maps that consider the uncertainty stemming from the inputs for better flood hazard predictions while considerably reducing the computation time. Indeed, the proposed technique presents a much lower computational effort for both offline and online stages, with a significant speed-up ratio compared to the POD-ANN method. Moreover, the introduced method only requires the construction of expansions over a few limited numbers of modes of the reduced basis, unlike the full-order Full-PCE model that needs to construct an expansion over each node of the computational domain, and must do so for each time step, a cumbersome computational effort particularly for high-dimensional stochastic problems, i.e., with a high number of input random parameters. Thus, the proposed non-intrusive reduced-order model-based B-splines Bézier element method presents a promising tool for uncertainty propagation analysis for stochastic time-dependent problems with a strong hyperbolic behavior.

CONCLUSION ET RECOMMANDATIONS

Le travail présenté dans le cadre de cette thèse a porté sur l'analyse de la propagation des incertitudes paramétriques en modélisation numérique de l'hydraulique des ruptures de barrages. La thématique traitée est plus que jamais d'actualité au regard de la recrudescence des événements liés à des inondations, en partie générés par les ruptures de barrages hydrauliques. L'établissement de plans pour pallier à toute situation d'urgence s'avère, de ce fait, une nécessité absolue. Ces mesures préventives se basent sur des prédictions fiables obtenues par des outils appropriés qui tiennent compte des incertitudes émanant des paramètres d'entrée. Ainsi, les apports de ce travail sont multiples, en proposant un cadre détaillé pour l'analyse de la propagation des incertitudes, et ce en implémentant des techniques efficaces, déjà existantes, mais aussi en proposant de nouvelles approches dont l'efficacité est démontrée et dont l'usage peut être étendu à d'autres domaines de l'ingénierie.

Dans un premier temps, les techniques classiques de propagation d'incertitudes comme Monte Carlo et LHS ont été utilisées sur des modèles numériques décrivant les écoulements de ruptures de barrages. Une étude minutieuse de convergence a été menée afin de déterminer la taille optimale de l'échantillonnage assurant une convergence et une précision acceptables. Malgré l'efficacité de ces approches, qui sont aussi caractérisées par leur simplicité à mettre en œuvre, leur utilité pour des problèmes exigeant un temps de simulation important, comme c'est le cas des écoulements à surface libre, s'avère limitée. Ceci est la conséquence des grandes tailles des échantillons requises (N_s) pour atteindre un taux de convergence adéquat, qui est de l'ordre de $\frac{1}{\sqrt{N_s}}$. Il n'en demeure pas moins que ces approches sont largement utilisées comme solutions de références avec lesquelles toutes les autres méthodes ont été testées et validées.

Les méthodes des polynômes du chaos ont été introduites pour remédier aux limitations des méthodes d'échantillonnage. En effet, les techniques du chaos polynomial, considérées comme relativement récentes, permettent d'approximer efficacement la variabilité des réponses de sortie des modèles numériques avec un nombre de points d'échantillonnage considérablement réduit. La technique dite du point de collocation (Pcol) a été implémentée et testée à travers une série d'exemples numériques de références décrivant des ruptures de barrages unidimensionnels avec un lit sec ou mouillé, représentant, respectivement, les solutions de Ritter et de Stoker. La performance de cette méthode a été analysée en combinant différentes valeurs des hyperparamètres, comme l'ordre polynomial et le coefficient de suréchantillonnage, et les résultats obtenus ont montré un fort comportement oscillatoire dans les profiles des moments statistiques des variables de sortie (hauteur d'eau et la vitesse du front de l'onde), particulièrement pour le cas de la solution de Stoker où des oscillations sont plus perceptibles au voisinage de la discontinuité accompagnant la propagation de l'onde de submersion. Le même comportement oscillatoire a été aussi observé dans le cas des prédictions des moments statistiques concernant une rupture de barrage avec des données réelles.

Le comportement oscillatoire des techniques des polynômes du chaos pour les problèmes à fort comportement hyperbolique représente un inconvénient pour la fiabilité des approximations, qui dans certains cas conduisent à des valeurs qui n'ont pas d'interprétation physique (valeurs négatives de la hauteur d'eau). C'est dans ce contexte que la méthode B-Splines Bézier Elements Method (BSBEM) a été proposée, qui repose, d'une part, sur une approximation locale de la réponse de sortie avec un aspect multi-éléments dans le domaine paramétrique, et d'autre part, sur les caractéristiques mathématiques attractives des fonctions de base de type B-Splines comme le support compact et le haut niveau de continuité quelles offrent. La fiabilité et la performance de la méthode BSBEM ont été analysées à travers une série d'exemples analytiques à l'instar de la fonction d'Ishigami, connue pour être fortement non-linéaire. La solution analytique de Stoker décrivant l'évolution de la hauteur d'eau et la vitesse du front de l'onde résultant d'une rupture de barrage a été aussi considérée. Les résultats obtenus ont montré l'efficacité de l'approche proposée à approximer les moments statistiques avec précision où les profiles sont en bonne concordance avec ceux des solutions de référence avec l'absence d'un comportement oscillatoire, contrairement aux prédictions des polynômes du chaos.

La méthode BSBEM développée a été par la suite appliquée pour l'analyse de propagation d'incertitudes à travers un modèle numérique décrivant l'onde de submersion générée par une rupture hypothétique d'un barrage réel situé dans la rivière de Batiscan (province du Québec). La variabilité des paramètres d'entrée à savoir le débit amont, le coefficient de frottement de Strickler et la largeur moyenne de la brèche, a été définie suivant la fonction de distribution appropriée de chaque paramètre. Les résultats obtenus ont été représentés en termes de variation du débit et de la hauteur d'eau en fonction du temps au niveau de quelques sections d'intérêt préalablement sélectionnées. Les profils obtenus ont montré l'efficacité de l'approche BSBEM avec une bonne concordance avec ceux des solutions de référence de l'approche LHS (avec 10 000 réalisations), contrairement aux profils de la méthode des polynômes du chaos où des déviations prononcées avec de fortes oscillations ont été observées. L'autre aspect de la comparaison concerne la capacité de BSBEM à reproduire fidèlement les distributions de probabilité des réponses de sortie incluant le temps d'arrivée du front de l'onde. Ces distributions estimées par BSBEM donnent un aperçu réaliste de la variabilité entourant ces prédictions, qui est d'une importance capitale dans toute démarche d'établissement de mesure d'urgence, à l'inverse des prédictions des polynômes du chaos où des écarts apparents ont été observés. Les résultats ont aussi mis en exergue les performances de la méthode BSBEM en termes d'effort de calcul dans l'obtention de prédictions fiables avec un nombre de points de collocation substantiellement inférieur comparativement à la méthode LHS.

Une autre nouvelle approche non-intrusive de réduction de modèles (POD-BSBEM) a été proposée dans le cadre de cette thèse. Cette technique requiert la construction d'une décomposition en fonctions des paramètres d'entrée uniquement sur un nombre réduit de modes constituant la base réduite, au lieu de l'ensemble des nœuds que contient le domaine de calcul. La méthode proposée combine la techniques POD à deux niveaux et l'approche BSBEM, qui est associée à un troisième niveau POD pour découpler la dépendance paramétrique et temporelle des coefficients de projections. L'efficacité de l'approche POD-BSBEM a été comparée à celle de la technique non-intrusive de réduction de modèles basée sur les réseaux de neurones artificiels (POD-ANN) à travers des exemples numériques comme la fonction d'Ackley, fortement non-linaire, et l'équation paramétrique instationnaire de Burgers. Les résultats ont montré que l'approche POD-BSBEM permet une estimation efficace des moments statistiques dont les profils concordent avec ceux des solutions de références obtenues par la méthode de Monte Carlo, contrairement aux prédictions de la méthode POD-ANN où des écarts ont été observés avec l'apparition d'oscillations, particulièrement dans les profils de l'écart-type. L'approche a été ensuite appliquée à un cas de rupture de barrage situé sur une portion de la rivière des Milles Iles, et les résultats obtenus ont clairement montré son efficacité dans l'établissement de carte d'inondations probabilistes fiables avec un effort de calcul considérablement réduit en comparaison avec l'approche POD-ANN qui nécessite un temps d'apprentissage relativement élevé.

Ainsi, les méthodes proposées, en l'occurrence BSBEM et POD-BSBEM, qui sont des contributions originales dans le cadre de cette thèse, ont démontré leur efficacité dans l'analyse de propagation d'incertitudes à travers les modèles numériques décrivant des phénomènes physiques avec un fort comportement hyperbolique, ou présentant des discontinuités, comme c'est le cas des écoulements de ruptures de barrages. Néanmoins, ces approches présentent quelques limitations notamment pour les problèmes faisant intervenir un nombre relativement élevé de variables aléatoires où le processus de décomposition de l'espace paramétrique en éléments de Bézier, et la construction des fonctions de base multivariées, par produit tensoriel, peut conduire à une augmentation substantielle du nombre de points de collocations requis et par conséquent le coût de calcul. Ces limitations ouvrent ainsi le champ à des axes d'investigation afin d'améliorer davantage les performances et efficacités de ces méthodes. L'une des pistes à envisager est, entre autres, l'introduction des fonctions de base non structurées de type T-splines, et aussi l'utilisation d'un raffinement local ciblé qui pourrait contribuer à améliorer la précision sans pour autant alourdir l'effort de calcul. Un autre champ d'investigation prometteur est l'intégration des différentes variantes des algorithmes de l'apprentissage profond pour la construction de modèles réduit non-intrusifs pour les analyses stochastiques, qui présentent tout de même l'avantage de la simplicité de mise en œuvre. Il y a lieu de mentionner aussi l'utilité de prendre en considération les incertitudes qui peuvent émaner de ces différents modèles de substitution construits pour une analyse stochastique complète.

RÉFÉRENCES

Abdedou, A. & Soulaïmani, A. (2019). A non-intrusive B-splines Bézier elements-based method for uncertainty propagation. *Computer Methods in Applied Mechanics and Engineering*, 345, 774–804.

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

Abraham, S., Tsirikoglou, P., Miranda, J., Lacor, C., Contino, F. & Ghorbaniasl, G. (2018). Spectral representation of stochastic field data using sparse polynomial chaos expansions. *Journal of Computational Physics*, 367, 109–120.

Ahmadisharaf, E., Kalyanapu, A. J., Thames, B. A. & Lillywhite, J. (2016). A probabilistic framework for comparison of dam breach parameters and outflow hydrograph generated by different empirical prediction methods. *Environmental modelling & software*, 86, 248–263.

Ahmed, S. E., San, O., Rasheed, A. & Iliescu, T. (2020). A long short-term memory embedding for hybrid uplifted reduced order models. *Physica D : Nonlinear Phenomena*, 409, 132471.

Audouze, C., De Vuyst, F. & Nair, P. B. (2013). Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations*, 29(5), 1587–1628.

Aureli, F., Maranzoni, A. & Mignosa, P. (2014). A semi-analytical method for predicting the outflow hydrograph due to dam-break in natural valleys. *Advances in water resources*, 63, 38–44.

Avdonin, A., Jaensch, S., Silva, C. F., Češnovar, M. & Polifke, W. (2018). Uncertainty quantification and sensitivity analysis of thermoacoustic stability with non-intrusive polynomial chaos expansion. *Combustion and Flame*, 189, 300–310.

Ballarin, F., Manzoni, A., Quarteroni, A. & Rozza, G. (2015). Supremizer stabilization of POD–Galerkin approximation of parametrized steady incompressible Navier–Stokes equations. *International Journal for Numerical Methods in Engineering*, 102(5), 1136–1161.

Ballio, F. & Guadagnini, A. (2004). Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology. *Water resources research*, 40(4).

Baroth, J., Bodé, L., Bressolette, P. & Fogli, M. (2006). SFE method using Hermite polynomials : An approach for solving nonlinear mechanical problems with uncertain parameters. *Computer Methods in Applied Mechanics and Engineering*, 195(44-47), 6479–6501.

Baroth, J., Bressolette, P., Chauvière, C. & Fogli, M. (2007). An efficient SFE method using Lagrange polynomials : application to nonlinear mechanical problems with uncertain parameters. *Computer Methods in Applied Mechanics and Engineering*, 196(45-48), 4419–4429.

Barth, T. (2013). Non-intrusive uncertainty propagation with error bounds for conservation laws containing discontinuities. Dans *Uncertainty Quantification in Computational Fluid Dynamics* (pp. 1–57). Springer.

Beck, J., Tamellini, L. & Tempone, R. (2019). IGA-based Multi-Index Stochastic Collocation for random PDEs on arbitrary domains. *Computer Methods in Applied Mechanics and Engineering*.

Bellos, V. & Tsakiris, G. (2015). Comparing various methods of building representation for 2D flood modelling in built-up areas. *Water Resources Management*, 29(2), 379–397.

Berkooz, G., Holmes, P. & Lumley, J. L. (1993). The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25(1), 539–575.

Berveiller, M., Sudret, B. & Lemaire, M. (2006). Stochastic finite element : a non intrusive approach by regression. *European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique*, 15(1-3), 81–92.

Bērziņš, A., Helmig, J., Key, F. & Elgeti, S. (2020). Standardized non-intrusive reduced order modeling using different regression models with application to complex flow problems. *arXiv preprint arXiv :2006.13706*.

Bharath, R. & Elshorbagy, A. (2018). Flood mapping under uncertainty : a case study in the Canadian prairies. *Natural Hazards*, 94(2), 537–560.

Bijl, H., Lucor, D., Mishra, S. & Schwab, C. (2013). *Uncertainty quantification in computational fluid dynamics*. Springer Science & Business Media.

Blatman, G. (2009). *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*. (Thèse de doctorat, Clermont-Ferrand 2).

Blatman, G. & Sudret, B. (2010a). An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25(2), 183–197.

Blatman, G. & Sudret, B. (2010b). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering & System Safety*, 95(11), 1216–1229.

Blatman, G. & Sudret, B. (2011). Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6), 2345–2367.

Borden, M. J., Scott, M. A., Evans, J. A. & Hughes, T. J. (2011). Isogeometric finite element data structures based on Bézier extraction of NURBS. *International Journal for Numerical Methods in Engineering*, 87(1-5), 15–47.

Buffa, A. & Sangalli, G. (2016). *IsoGeometric Analysis : A New Paradigm in the Numerical Approximation of PDEs : Cetraro, Italy 2012*. Springer.

Burgers, J. M. (1948). A mathematical model illustrating the theory of turbulence. Dans *Advances in applied mechanics* (vol. 1, pp. 171–199). Elsevier.

Butt, M. J., Umar, M. & Qamar, R. (2013). Landslide dam and subsequent dam-break flood estimation using HEC-RAS model in Northern Pakistan. *Natural hazards*, 65(1), 241–254.

Cannata, M. & Marzocchi, R. (2012). Two-dimensional dam break flooding simulation : a GIS-embedded approach. *Natural hazards*, 61(3), 1143–1159.

Capillon, R., Desceliers, C. & Soize, C. (2016). Uncertainty quantification in computational linear structural dynamics for viscoelastic composite structures. *Computer Methods in Applied Mechanics and Engineering*, 305, 154–172.

Chang, T.-J., Kao, H.-M., Chang, K.-H. & Hsu, M.-H. (2011). Numerical simulation of shallow-water dam break flows in open channels using smoothed particle hydrodynamics. *Journal of Hydrology*, 408(1-2), 78–90.

Chatterjee, A. (2000). An introduction to the proper orthogonal decomposition. *Current science*, 808–817.

Chen, X., Yang, T., Wang, X., Xu, C.-Y. & Yu, Z. (2013). Uncertainty intercomparison of different hydrological models in simulating extreme flows. *Water Resources Management*, 27(5), 1393–1409.

Chen, X., Qiu, Z., Wang, X., Li, Y. & Wang, R. (2017). Uncertain reduced-order modeling for unsteady aerodynamics with interval parameters and its application on robust flutter boundary prediction. *Aerospace Science and Technology*, 71, 214–230.

Chen, X. & Qiu, Z. (2018). A novel uncertainty analysis method for composite structures with mixed uncertainties including random and interval variables. *Composite Structures*, 184, 400–410.

Cheng, K. & Lu, Z. (2018). Adaptive sparse polynomial chaos expansions for global sensitivity analysis based on support vector regression. *Computers & Structures*, 194, 86–96.

Chouvion, B. & Sarrouy, E. (2016). Development of error criteria for adaptive multi-element polynomial chaos approaches. *Mechanical Systems and Signal Processing*, 66, 201–222.

Cottrell, J. A., Hughes, T. J. & Bazilevs, Y. (2009). *Isogeometric analysis : toward integration of CAD and FEA*. John Wiley & Sons.

Couplet, M., Basdevant, C. & Sagaut, P. (2005). Calibrated reduced-order POD-Galerkin system for fluid flow modelling. *Journal of Computational Physics*, 207(1), 192–220.

da Veiga, L. B., Buffa, A., Sangalli, G. & Vázquez, R. (2016). An introduction to the numerical analysis of isogeometric methods. Dans *Numerical simulation in physics and engineering* (pp. 3–69). Springer.

Dai, F., Lee, C. F., Deng, J. & Tham, L. (2005). The 1786 earthquake-triggered landslide dam and subsequent dam-break flood on the Dadu River, southwestern China. *Geomorphology*, 65(3-4), 205–221.

De Almeida, G. A. & Bates, P. (2013). Applicability of the local inertial approximation of the shallow water equations to flood modeling. *Water Resources Research*, 49(8), 4833–4844.

De Rocquigny, E., Devictor, N. & Tarantola, S. (2008). *Uncertainty in industrial practice : a guide to quantitative uncertainty management*. John Wiley & Sons.

Debusschere, B. J., Najm, H. N., Pébay, P. P., Knio, O. M., Ghanem, R. G. & Le Maıtre, O. P. (2004). Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM journal on scientific computing*, 26(2), 698–719.

Delestre, O., Lucas, C., Ksinant, P.-A., Darboux, F., Laguerre, C., Vo, T.-N.-T., James, F. & Cordier, S. (2013). SWASHES : a compilation of shallow water analytic solutions for hydraulic and environmental studies. *International Journal for Numerical Methods in Fluids*, 72(3), 269–300.

Delmas, V. & Soulaïmani, A. (2020). Multi-GPU implementation of a time-explicit finite volume solver for the Shallow-Water Equations using CUDA and a CUDA-Aware version of OpenMPI. *arXiv preprint arXiv :2010.14416*.

Dertimanis, V. K., Spiridonakos, M. D. & Chatzi, E. N. (2017). Data-driven uncertainty quantification of structural systems via B-spline expansion. *Computers & Structures*.

Dertimanis, V. K., Spiridonakos, M. D. & Chatzi, E. N. (2018). Data-driven uncertainty quantification of structural systems via B-spline expansion. *Computers & Structures*, 207, 245–257.

Dinescu, C., Smirnov, S., Hirsch, C. & Lacor, C. (2010). Assessment of intrusive and non-intrusive non-deterministic CFD methodologies based on polynomial chaos expansions. *International Journal of Engineering Systems Modelling and Simulation*, 2(1-2), 87–98.

Ditkowski, A., Fibich, G. & Sagiv, A. (2018). A Spline-Based Approach to Uncertainty Propagation and Density Estimation. *arXiv preprint arXiv :1803.10991*.

El Moçayd, N., Mohamed, M. S., Ouazar, D. & Seaid, M. (2020). Stochastic model reduction for polynomial chaos expansion of acoustic waves using proper orthogonal decomposition. *Reliability Engineering & System Safety*, 195, 106733.

Eldred, M. & Burkardt, J. (2009). Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification. *47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition*, pp. 976.

Erpicum, S., Dewals, B. J., Archambeau, P. & Pirotton, M. (2010). Dam break flow computation based on an efficient flux vector splitting. *Journal of Computational and Applied Mathematics*, 234(7), 2143–2151.

Faghih, M., Mirzaei, M., Adamowski, J., Lee, J. & El-Shafie, A. (2017). Uncertainty Estimation in Flood Inundation Mapping : An Application of Non-parametric Bootstrapping. *River Research and Applications*, 33(4), 611–619.

Fajraoui, N., Fahs, M., Younes, A. & Sudret, B. (2017). Analyzing natural convection in porous enclosure with polynomial chaos expansions : Effect of thermal dispersion, anisotropic permeability and heterogeneity. *International Journal of Heat and Mass Transfer*, 115, 205–224.

Fang, F., Pain, C. C., Navon, I., Elsheikh, A. H., Du, J. & Xiao, D. (2013). Non-linear Petrov–Galerkin methods for reduced order hyperbolic equations and discontinuous finite element methods. *Journal of Computational Physics*, 234, 540–559.

Ferrari, A., Dumbser, M., Toro, E. F. & Armanini, A. (2009). A new 3D parallel SPH scheme for free surface flows. *Computers & Fluids*, 38(6), 1203–1217.

Foo, J., Wan, X. & Karniadakis, G. E. (2008). The multi-element probabilistic collocation method (ME-PCM) : Error analysis and applications. *Journal of Computational Physics*, 227(22), 9572–9595.

Froehlich, D. C. (2008). Embankment dam breach parameters and their uncertainties. *Journal of Hydraulic Engineering*, 134(12), 1708–1721.

Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Elsevier.

Gao, H., Wang, J.-X. & Zahr, M. J. (2020). Non-intrusive model reduction of large-scale, nonlinear dynamical systems using deep learning. *Physica D : Nonlinear Phenomena*, 412, 132614.

Gao, L., Zhou, Z.-F. & Huang, Q.-A. (2017). A Generalized Polynomial Chaos-Based Approach to Analyze the Impacts of Process Deviations on MEMS Beams. *Sensors*, 17(11), 2561.

Georgaka, S., Stabile, G., Star, K., Rozza, G. & Bluck, M. J. (2020). A Hybrid Reduced Order Method for Modelling Turbulent Heat Transfer Problems. *Computers & Fluids*, 104615.

Ghanem, R. (1999). Ingredients for a general purpose stochastic finite elements implementation. *Computer Methods in Applied Mechanics and Engineering*, 168(1-4), 19–34.

Ghanem, R. G. & Spanos, P. D. (1991). Stochastic Finite Element Method : Response Statistics. Dans *Stochastic Finite Elements : A spectral approach* (pp. 101–119). Springer.

Ghiocel, D. M. & Ghanem, R. G. (2002). Stochastic finite-element analysis of seismic soil–structure interaction. *Journal of Engineering Mechanics*, 128(1), 66–77.

Goodall, C. (1988). Principal component analysis. Taylor & Francis Group.

Goutal, N. & Maurel, F. (2002). A finite volume solver for 1D shallow-water equations applied to an actual river. *International Journal for Numerical Methods in Fluids*, 38(1), 1–19.

Goutal, N., Goeury, C., Ata, R., Ricci, S., El Mocayd, N., Rochoux, M., Oubanas, H., Gejadze, I. & Malaterre, P.-O. (2018). Uncertainty quantification for river flow simulation applied to a real test case : The garonne valley. Dans *Advances in Hydroinformatics* (pp. 169–187). Springer.

Guinot, V. (2012). Multiple porosity shallow water models for macroscopic modelling of urban floods. *Advances in Water Resources*, 37, 40–72.

Guo, M. & Hesthaven, J. S. (2018). Reduced order modeling for nonlinear structural analysis using gaussian process regression. *Computer methods in applied mechanics and engineering*, 341, 807–826.

Guo, M. & Hesthaven, J. S. (2019). Data-driven reduced order modeling for time-dependent problems. *Computer methods in applied mechanics and engineering*, 345, 75–99.

Hadrich, A., Zribi, M. & Masmoudi, A. (2016). Bayesian expectation maximization algorithm by using B-splines functions : Application in image segmentation. *Mathematics and Computers in Simulation*, 120, 50–63.

Halder, Y. V., Sanderse, B. & Koren, B. (2019). An adaptive minimum spanning tree multielement method for uncertainty quantification of smooth and discontinuous responses. *SIAM Journal on Scientific Computing*, 41(6), A3624–A3648.

Haltas, I., Elçi, S. & Tayfur, G. (2016). Numerical simulation of flood wave propagation in two-dimensions in densely populated urban areas due to dam break. *Water resources management*, 30(15), 5699–5721.

Hammersley, J. & Handscomb, D. (1964). Methuen's Monographs on Applied Probability and Statistics. *Monte Carlo Methods*.

Hesthaven, J. S. & Ubbiali, S. (2018). Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363, 55–78.

Hien, T. D. & Noh, H.-C. (2017). Stochastic isogeometric analysis of free vibration of functionally graded plates considering material randomness. *Computer Methods in Applied Mechanics and Engineering*, 318, 845–863.

Hijazi, S., Stabile, G., Mola, A. & Rozza, G. (2020). Non-intrusive polynomial chaos method applied to full-order and reduced problems in computational fluid dynamics : a comparison and perspectives. Dans *Quantification of Uncertainty : Improving Efficiency and Technology* (pp. 217–240). Springer.

Homma, T. & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1), 1–17.

Hosder, S. & Walters, R. (2010). Non-intrusive polynomial chaos methods for uncertainty quantification in fluid dynamics. *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, pp. 129.

Hosder, S., Walters, R. & Balch, M. (2007). Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables. *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pp. 1939.

Hu, J. (2015). Forecast Uncertainty Quantification using Monte Carlo, Polynomial Chaos Expansion and Unscented Transformation Methods.

Hughes, T. J. (2012). *The finite element method : linear static and dynamic finite element analysis*. Courier Corporation.

Hughes, T. J., Cottrell, J. A. & Bazilevs, Y. (2005). Isogeometric analysis : CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194(39-41), 4135–4195.

Hughes, T. J., Reali, A. & Sangalli, G. (2010). Efficient quadrature for NURBS-based isogeometric analysis. *Computer methods in applied mechanics and engineering*, 199(5-8), 301–313.

Jacquier, P., Abdedou, A., Delmas, V. & Soulaïmani, A. (2021). Non-intrusive reduced-order modeling using uncertainty-aware Deep Neural Networks and Proper Orthogonal Decomposition : Application to flood modeling. *Journal of Computational Physics*, 424, 109854.

Jakeman, J. D., Narayan, A. & Xiu, D. (2013). Minimal multi-element stochastic collocation for uncertainty quantification of discontinuous functions. *Journal of Computational Physics*, 242, 790–808.

Johnson, P. A. (1996). Uncertainty of hydraulic parameters. *Journal of hydraulic engineering*, 122(2), 112–114.

Kalinina, A., Spada, M., Vetsch, D. F., Marelli, S., Whealton, C., Burgherr, P. & Sudret, B. (2020). Metamodeling for Uncertainty Quantification of a Flood Wave Model for Concrete Dam Breaks. *Energies*, 13(14), 3685.

Kani, J. N. & Elsheikh, A. H. (2019). Reduced-order modeling of subsurface multi-phase flow models using deep residual recurrent neural networks. *Transport in Porous Media*, 126(3), 713–741.

Kim, B. & Sanders, B. F. (2016). Dam-break flood model uncertainty assessment : case study of extreme flooding with multiple dam failures in Gangneung, South Korea. *Journal of hydraulic engineering*, 142(5), 05016002.

Kwon, H.-H., Moon, Y.-I. & Khalil, A. F. (2007). Nonparametric Monte Carlo Simulation for Flood Frequency Curve Derivation : An Application to a Korean Watershed 1. *JAWRA Journal of the American Water Resources Association*, 43(5), 1316–1328.

Lai, J.-S., Guo, W.-D., Lin, G.-F. & Tan, Y.-C. (2010). A well-balanced upstream flux-splitting finite-volume scheme for shallow-water flow simulations with irregular bed topography. *International journal for numerical methods in fluids*, 62(8), 927–944.

Le Maître, O., Knio, O., Najm, H. & Ghanem, R. (2004a). Uncertainty propagation using Wiener–Haar expansions. *Journal of Computational Physics*, 197(1), 28–57.

Le Maître, O., Najm, H. N., Ghanem, R. & Knio, O. (2004b). Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *Journal of Computational Physics*, 197(2), 502–531.

Leendertse, J. J. & Gritton, E. C. (1971). A water-quality simulation model for well mixed estuaries and coastal seas : vol. II, Computation Procedures.

Lempérière, F. (2017). Dams and floods. *Engineering*, 3(1), 144–149.

Li, H. & Zhang, D. (2007). Probabilistic collocation method for flow in porous media : Comparisons with other stochastic methods. *Water Resources Research*, 43(9).

Li, K., Gao, W., Wu, D., Song, C. & Chen, T. (2018). Spectral stochastic isogeometric analysis of linear elasticity. *Computer Methods in Applied Mechanics and Engineering*, 332, 157–190.

Li, K., Wu, D., Gao, W. & Song, C. (2019). Spectral stochastic isogeometric analysis of free vibration. *Computer Methods in Applied Mechanics and Engineering*.

Liang, Y., Lee, H., Lim, S., Lin, W., Lee, K. & Wu, C. (2002). Proper orthogonal decomposition and its applications—Part I : Theory. *Journal of Sound and vibration*, 252(3), 527–544.

Liberge, E. (2008). *Modèles réduits obtenus par la méthode de POD-Galerkin pour les problèmes d'interaction fluide structure*. (Thèse de doctorat).

Loeven, G., Witteveen, J. & Bijl, H. (2007). Probabilistic collocation : an efficient non-intrusive approach for arbitrarily distributed parametric uncertainties. *45th AIAA Aerospace Sciences Meeting and Exhibit*, pp. 317.

Loukili, Y. & Soulaimani, A. (2007). Numerical tracking of shallow water waves by the unstructured finite volume WAF approximation. *International Journal for Computational Methods in Engineering Science and Mechanics*, 8(2), 75–88.

Lumley, J. L. (1967). The structure of inhomogeneous turbulent flows. *Atmospheric turbulence and radio wave propagation*.

Malakiyeh, M. M., Shojaee, S. & Javaran, S. H. (2018). Development of a direct time integration method based on Bezier curve and 5th-order Bernstein basis function. *Computers & Structures*, 194, 15–31.

Maleewong, M. & Sirisup, S. (2011). On-line and off-line POD assisted projective integral for non-linear problems : A case study with Burgers' equation. *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, 5(7), 984–992.

Marsooli, R. & Wu, W. (2014). 3-D finite-volume model of dam-break flow over uneven beds based on VOF method. *Advances in water resources*, 70, 104–117.

Mathelin, L., Hussaini, M. Y. & Zang, T. A. (2005). Stochastic approaches to uncertainty quantification in CFD simulations. *Numerical Algorithms*, 38(1-3), 209–236.

Matthies, H. G. & Keese, A. (2005). Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16), 1295–1331.

McKay, M. D., Beckman, R. J. & Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239–245.

Miller, K., Berg, S., Davison, J., Sudicky, E. & Forsyth, P. (2018). Efficient uncertainty quantification in fully-integrated surface and subsurface hydrologic simulations. *Advances in Water Resources*, 111, 381–394.

Mishra, S., Schwab, C. & Šukys, J. (2016). Multi-level Monte Carlo finite volume methods for uncertainty quantification of acoustic wave propagation in random heterogeneous layered medium. *Journal of Computational Physics*, 312, 192–217.

Munoz, D. H. & Constantinescu, G. (2020). 3-D dam break flow simulations in simplified and complex domains. *Advances in Water Resources*, 137, 103510.

Newman, A. J. (1996). *Model reduction via the Karhunen-Loeve expansion part i : an exposition*.

Oladyshkin, S. & Nowak, W. (2012). Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, 106, 179–190.

Ozmen-Cagatay, H. & Kocaman, S. (2011). Dam-break flow in the presence of obstacle : experiment and CFD simulation. *Engineering Applications of Computational Fluid Mechanics*, 5(4), 541–552.

Palar, P. S., Tsuchiya, T. & Parks, G. T. (2016). Multi-fidelity non-intrusive polynomial chaos based on regression. *Computer Methods in Applied Mechanics and Engineering*, 305, 579–606.

Palar, P. S., Zuhal, L. R., Shimoyama, K. & Tsuchiya, T. (2018). Global sensitivity analysis via multi-fidelity polynomial chaos expansion. *Reliability Engineering & System Safety*, 170, 175–190.

Pan, Q. & Dias, D. (2017). Sliced inverse regression-based sparse polynomial chaos expansions for reliability analysis in high dimensions. *Reliability Engineering & System Safety*, 167, 484–493.

Petaccia, G. & Natale, E. (2013). ORSADEM : a one-dimensional shallow water code for flood inundation modelling. *Irrigation and Drainage*, 62(S2), 29–40.

Pettersson, P., Iaccarino, G. & Nordström, J. (2013). An intrusive hybrid method for discontinuous two-phase flow under uncertainty. *Computers & Fluids*, 86, 228–239.

Pettersson, P., Iaccarino, G. & Nordström, J. (2014). A stochastic Galerkin method for the Euler equations with Roe variable transformation. *Journal of Computational Physics*, 257, 481–500.

Piegl, L. & Tiller, W. (1997). *The NURBS book : monograph in visual communication*. New York : Springer.

Prempraneerach, P., Hover, F. S., Triantafyllou, M. S. & Karniadakis, G. E. (2010). Uncertainty quantification in simulations of power systems : Multi-element polynomial chaos methods. *Reliability Engineering & System Safety*, 95(6), 632–646.

Quicken, S., Donders, W. P., van Disseldorp, E. M., Gashi, K., Mees, B. M., van de Vosse, F. N., Lopata, R. G., Delhaas, T. & Huberts, W. (2016). Application of an adaptive polynomial chaos expansion on computationally expensive three-dimensional cardiovascular models for uncertainty quantification and sensitivity analysis. *Journal of Biomechanical Engineering*, 138(12), 121010.

Rahman, S. (2011). Global sensitivity analysis by polynomial dimensional decomposition. *Reliability Engineering & System Safety*, 96(7), 825–837.

Raisee, M., Kumar, D. & Lacor, C. (2015). 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.

Raisee, M., Kumar, D. & Lacor, C. (2019). Non-intrusive uncertainty quantification by combination of reduced basis method and regression-based polynomial chaos expansion. Dans *Uncertainty Management for Robust Industrial Design in Aeronautics* (pp. 169–184). Springer.

Reagana, M. T., Najm, H. N., Ghanem, R. G. & Knio, O. M. (2003). Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. *Combustion and Flame*, 132(3), 545–555.

Ritter, A. (1892). Die fortpflanzung der wasserwellen. *Zeitschrift des Vereines Deutscher Ingenieure*, 36(33), 947–954.

Rogers, D. F. (2001). *An Introduction to NURBS : With Historical Perspective*. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc.

Saikia, M. & Sarma, A. (2010). Numerical Simulation of Flow and Morphological Evaluations in Rivers Under Dam-Break Flows. *International Journal of Modelling and Simulation*, 30(3), 289–297.

Salehi, S., Raisee, M., Cervantes, M. J. & Nourbakhsh, A. (2017). Efficient uncertainty quantification of stochastic CFD problems using sparse polynomial chaos and compressed sensing. *Computers & Fluids*, 154, 296–321.

Salehi, S., Raisee, M., Cervantes, M. J. & Nourbakhsh, A. (2018). An efficient multifidelity l1-minimization method for sparse polynomial chaos. *Computer Methods in Applied Mechanics and Engineering*, 334, 183–207.

Sammen, S. S., Mohamed, T., Ghazali, A., Sidek, L. M. & El-Shafie, A. (2017). An evaluation of existent methods for estimation of embankment dam breach parameters. *Natural Hazards*, 87(1), 545–566.

San, O., Maulik, R. & Ahmed, M. (2019). An artificial neural network framework for reduced order modeling of transient flows. *Communications in Nonlinear Science and Numerical Simulation*, 77, 271 - 287.

Sanderse, B. (2020). Non-linearly stable reduced-order models for incompressible flow with energy-conserving finite volume methods. *Journal of Computational Physics*, 421, 109736.

Sangalli, G., Takacs, T. & Vázquez, R. (2016). Unstructured spline spaces for isogeometric analysis based on spline manifolds. *Computer Aided Geometric Design*, 47, 61–82.

Schilders, W. H., Van der Vorst, H. A. & Rommes, J. (2008). *Model order reduction : theory, research aspects and applications*. Springer.

Schuß, S., Dittmann, M., Wohlmuth, B., Klinkel, S. & Hesch, C. (2019). Multi-patch isogeometric analysis for Kirchhoff–Love shell elements. *Computer Methods in Applied Mechanics and Engineering*, 349, 91–116.

Scott, M. A., Borden, M. J., Verhoosel, C. V., Sederberg, T. W. & Hughes, T. J. (2011). Isogeometric finite element data structures based on Bézier extraction of T-splines. *International Journal for Numerical Methods in Engineering*, 88(2), 126–156.

Sederberg, T. W., Zheng, J., Bakenov, A. & Nasri, A. (2003). T-splines and T-NURCCs. *ACM transactions on graphics (TOG)*, 22(3), 477–484.

Seyedashraf, O. & Akhtari, A. A. (2017). Two-dimensional numerical modeling of dam-break flow using a new TVD finite-element scheme. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39(11), 4393–4401.

Seyedashraf, O., Mehrabi, M. & Akhtari, A. A. (2018). Novel approach for dam break flow modeling using computational intelligence. *Journal of hydrology*, 559, 1028–1038.

Shields, M. D. & Zhang, J. (2016). The generalization of Latin hypercube sampling. *Reliability Engineering & System Safety*, 148, 96–108.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quarterly of applied mathematics*, 45(3), 561–571.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3), 271–280.

Sobol, I. & Levitan, Y. L. (1999). A pseudo-random number generator for personal computers. *Computers & Mathematics with Applications*, 37(4-5), 33–40.

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

Stoker, J. J. (1957). Water waves : The mathematical theory with applications. *Interscience, New York*, 2, 5.

Sudret, B. (2007). Uncertainty propagation and sensitivity analysis in mechanical models–Contributions to structural reliability and stochastic spectral methods. *Habilitationa diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France*, 18.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7), 964–979.

Sudret, B. (2014). Polynomial chaos expansions and stochastic finite element methods. *Risk and reliability in geotechnical engineering*, 265–300.

Sun, X., Pan, X. & Choi, J.-I. (2019). A non-intrusive reduced-order modeling method using polynomial chaos expansion. *arXiv preprint arXiv :1903.10202*.

Swischuk, R., Mainini, L., Peherstorfer, B. & Willcox, K. (2019). Projection-based model reduction : Formulations for physics-based machine learning. *Computers & Fluids*, 179, 704–717.

Tatang, M. A., Pan, W., Prinn, R. G. & McRae, G. J. (1997). An efficient method for parametric uncertainty analysis of numerical geophysical models. *Journal of Geophysical Research : Atmospheres*, 102(D18), 21925–21932.

Thapa, M., Mulani, S. B. & Walters, R. W. (2018). A new non-intrusive polynomial chaos using higher order sensitivities. *Computer Methods in Applied Mechanics and Engineering*, 328, 594–611.

Thapa, M., Mulani, S. B. & Walters, R. W. (2019). Stochastic multi-scale modeling of carbon fiber reinforced composites with polynomial chaos. *Composite Structures*, 213, 82–97.

Thomas, D., Engvall, L., Schmidt, S., Tew, K. & Scott, M. (2018). U-splines : Splines over unstructured meshes. *Preprint, https ://coreform. com/usplines*.

Tsai, C. W., Yeh, J.-J. & Huang, C.-H. (2019). Development of probabilistic inundation mapping for dam failure induced floods. *Stochastic Environmental Research and Risk Assessment*, 33(1), 91–110.

van Halder, Y., Sanderse, B. & Koren, B. (2018). An adaptive minimum spanning tree multi-element method for uncertainty quantification of smooth and discontinuous responses. *arXiv preprint arXiv :1803.06833*.

Vorogushyn, S., Apel, H. & Merz, B. (2011). The impact of the uncertainty of dike breach development time on flood hazard. *Physics and Chemistry of the Earth, Parts A/B/C*, 36(7-8), 319–323.

Wahl, T. L. (2004). Uncertainty of predictions of embankment dam breach parameters. *Journal of hydraulic engineering*, 130(5), 389–397.

Walton, S., Hassan, O. & Morgan, K. (2013). Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. *Applied Mathematical Modelling*, 37(20-21), 8930–8945.

Wan, X. & Karniadakis, G. E. (2005). An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2), 617–642.

Wang, Q., Hesthaven, J. S. & Ray, D. (2019). Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem. *Journal* *of computational physics*, 384, 289–307.

Weaver, A. B., Alexeenko, A. A., Greendyke, R. B. & Camberos, J. A. (2011). Flowfield Uncertainty Analysis for Hypersonic Computational Fluid Dynamics Simulations. *Journal of Thermophysics and Heat Transfer*, 25(1), 10–20.

Weitz, S., Blanco, S., Charon, J., Dauchet, J., El Hafi, M., Eymet, V., Farges, O., Fournier, R. & Gautrais, J. (2016). Monte Carlo efficiency improvement by multiple sampling of conditioned integration variables. *Journal of Computational Physics*, 326, 30–34.

Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics*, 60(4), 897–936.

Willcox, K. (2006). Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. *Computers & fluids*, 35(2), 208–226.

Wu, S.-J. & Hsu, C.-T. (2018). Modeling of uncertainty for flood wave propagation induced by variations in initial and boundary conditions using expectation operator on explicit numerical solutions. *International Journal for Numerical Methods in Engineering*, 113(9), 1447–1465.

Xiao, D., Fang, F., Buchan, A., Pain, C., Navon, I. & Muggeridge, A. (2015a). Non-intrusive reduced order modelling of the Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 293, 522 - 541.

Xiao, D., Fang, F., Buchan, A., Pain, C., Navon, I. & Muggeridge, A. (2015b). Non-intrusive reduced order modelling of the Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 293, 522–541.

Xiao, D., Heaney, C., Fang, F., Mottet, L., Hu, R., Bistrian, D., Aristodemou, E., Navon, I. & Pain, C. (2019). A domain decomposition non-intrusive reduced order model for turbulent flows. *Computers & Fluids*, 182, 15–27.

Xiao, D., Yang, P., Fang, F., Xiang, J., Pain, C. C. & Navon, I. M. (2016a). Non-intrusive reduced order modelling of fluid–structure interactions. *Computer Methods in Applied Mechanics and Engineering*, 303, 35–54.

Xiao, D. et al. (2016b). *Non-intrusive reduced order models and their applications*. (Thèse de doctorat, Imperial College London).

Xiao, M., Breitkopf, P., Coelho, R., Knopf-Lenoir, C., Sidorkiewicz, M. & Villon, P. (2010). Model reduction by CPOD and Kriging : Application to the shape optimization of an intake port. *Structural and Multidisciplinary Optimization*, 41(4), 555 - 74.

Xiong, Y. F. (2011). A dam break analysis using HEC-RAS. *Journal of Water Resource and Protection*, 3(6), 370.

Xiu, D. (2007). Efficient collocational approach for parametric uncertainty analysis. *Commun. Comput. Phys*, 2(2), 293–309.

Xiu, D. (2010). *Numerical methods for stochastic computations : a spectral method approach*. Princeton University Press.

Xiu, D. & Karniadakis, G. E. (2002). The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2), 619–644.

Ying, X., Khan, A. A. & Wang, S. S. (2004). Upwind conservative scheme for the Saint Venant equations. *Journal of hydraulic engineering*, 130(10), 977–987.

Yu, J., Yan, C. & Guo, M. (2019). Non-intrusive reduced-order modeling for fluid problems : A brief review. *Proceedings of the Institution of Mechanical Engineers, Part G : Journal of Aerospace Engineering*, 233(16), 5896–5912.

Yucel, A. C., Bagci, H. & Michielssen, E. (2013). An adaptive multi-element probabilistic collocation method for statistical EMC/EMI characterization. *IEEE Transactions on Electromagnetic Compatibility*, 55(6), 1154–1168.

Zhang, S., Xia, Z., Yuan, R. & Jiang, X. (2014). Parallel computation of a dam-break flow model using OpenMP on a multi-core computer. *Journal of hydrology*, 512, 126–133.

Zhiwei, S., Chen, W., Zheng, Y., Junqiang, B., Zheng, L., Qiang, X. & Qiujun, F. (2020). Non-intrusive reduced-order model for predicting transonic flow with varying geometries. *Chinese Journal of Aeronautics*, 33(2), 508–519.

Zokagoa, J.-M. (2011). *Modélisation numérique des écoulements à surface libre avec bancs couvrants-découvrants par les volumes finis et la décomposition orthogonale aux valeurs propres*. (Thèse de doctorat, École de technologie supérieure).

Zokagoa, J.-M. & Soulaïmani, A. (2010). Modeling of wetting–drying transitions in free surface flows over complex topographies. *Computer Methods in Applied Mechanics and Engineering*, 199(33-36), 2281–2304.

Zokagoa, J.-M. & Soulaïmani, A. (2012). A POD-based reduced-order model for free surface shallow water flows over real bathymetries for Monte-Carlo-type applications. *Computer Methods in Applied Mechanics and Engineering*, 221, 1–23.

Zokagoa, J.-M. & Soulaïmani, A. (2018). A POD-based reduced-order model for uncertainty analyses in shallow water flows. *International Journal of Computational Fluid Dynamics*, 32(6-7), 278–292.