5th International Symposium on Hydraulic Structures Brisbane, Australia, 25-27 June 2014
Hydraulic Structures and Society: Engineering Challenges and Extremes
ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

Probabilistic Collocation Method for Uncertainty Analysis of Soil Infiltration in Flood Modelling

Y. Huang[1,2], and X.S. Qin[1,2*]
[1] School of Civil & Environmental Engineering, Nanyang Technological University, Singapore, 639798
SINGAPORE
[2] Institute of Catastrophe Risk Management, Nanyang Technological University, Singapore, 639798
SINGAPORE
E-mail: xsqin@ntu.edu.sg

Abstract: *The probabilistic collocation method (PCM) based on the Karhunen-Loevè expansion (KLE) and Polynomial chaos expansion (PCE) is applied for uncertainty analysis of flood inundation modelling. The floodplain hydraulic conductivity (K_S) is considered as one of the important parameters in a 2-dimensional (2D) physical model FLO-2D (with Green-Ampt infiltration method) and has a nonlinear relationship with the flood simulation results, such as maximum flow depths (h_{max}). In this study, due to the spatial heterogeneity of soil, log-transformed K_s was assumed a random field in spatiality with normal distribution and decomposed with KLE in pairs of corresponding eigenvalues and eigenfuctions. The h_{max} random field is expanded by a second-order PCE approximation and the deterministic coefficients in PCE are solved by FLO-2D. To demonstrate this method, a simplified flood inundation case was used, where the mean and variance of the simulation outputs were compared with those from direct Monte Carlo Simulation. The comparison indicates that PCM could efficiently capture the statistics of flow depth in flood modelling with much less computational requirements.*

Keywords: *Karhunen-Loevè expansion, PCM, PCE, flood inundation modelling, Monte Carlo Simulation*

1. INSTRUCTION

Flood inundation modelling processes often involve a large amount of uncertainties, especially at topographical formation where insufficient data are available for model parameterization, calibration and validation. During the past decades, many studies were developed to do uncertainty quantification of flood modelling, during which two traditional approaches are widely used including Bayesian method and Monte Carlo simulations (MCS) (Beven and Binley, 1992, Van Vuren et al., 2005, Lu and Zhang, 2007, Reza Ghanbarpour et al., 2011, Jung and Merwade, 2012). However, most of these studies assumed homogenous values of parameter over the modelling domain, which would result in obvious deviations between the observed data and simulation results. It has been recognized that spatial heterogeneity of geological media could bring significant effects on fluid modelling system; however, for many cases, only a limited number of locations are taken with measurement or survey of soil media; consequently error occurs when the measured data is applied for the entire modelling system (Liu et al., 2006). During a two-dimensional (2D) fluid numerical modelling system, the 2D governing equations have to be converted from the deterministic equations into stochastic (Li and Zhang, 2009, Liu and Matthies, 2010). Traditional Monte Carlo Simulation (MCS) and other approaches based on the MCS are normally used for such a type of uncertainty quantifications (Huang et al., 2001, Ballio and Guadagnini, 2004, Zhang and Lu, 2004, Lu and Zhang, 2007). The MCS approach is easy to understand and simple to operate. However, a significant drawback is that the computation is intensive, especially for flood modelling system. In recent decades, new approaches based on the Karhunen-Loevè expansion (KLE) and Polynomial chaos expansion (PCE) are developed for uncertainty analysis in many fields such as subsurface modelling system (Liu et al., 2006, Li and Zhang, 2007). However, studies with consideration of spatial distribution are relatively limited for the flood modelling under uncertainty. One attempt was made by Liu and Matthies (2010), who developed a framework of uncertain analysis combing the KLE and PCE expansion (based on Galerkin projection) and examined three uncertain parameters including inflow, topography and floodplain manning's *n* over the entire flood modelling. In this study, we aim to apply the KLE-based PCM to analyze the uncertainty

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

propagation of floodplain hydraulic conductivity (K_S) which is deemed a key parameter of Green-Ampt method for modelling the infiltration process during the flood inundation modelling using FLO-2D, a two-dimensional flood model; no previous studies have specifically looked into such a parameter. A simplified inundation case is used for demonstrating the methodology.

2. METHODOLOGY

2.1. Stochastic Flood Inundation Model

To describe a 2D flood inundation process, shallow water equations can be used (Chow et al., 1988, FLO-2D Software, 2012):

$$\frac{\partial h(X)}{\partial t} + \nabla\left[h(X)V\right] = R - I \tag{1a}$$

$$S_f = S_o - \nabla h(X) - \frac{1}{g}(V \cdot \nabla)V - \frac{1}{g}\frac{\partial V}{\partial t} \tag{1b}$$

$$I = -nK_S(X)\left[\frac{\nabla\theta(\psi - h_0)}{F} + 1\right] \tag{1c}$$

where, $h(X, t)$ is defined as the flow depth; V is the averaged-in-depth velocity in each direction; $X = (x, y)$ represents 2D Cartesian coordinate spatially; t is the time; S_o is the bed slope, and S_f is the friction slope; R is rainfall intensity and I is surface infiltration rate; F is the total infiltration; n is the soil porosity; $K_S(X)$ is defined as saturated hydraulic conductivity; $\Delta\theta$ is the effective soil moisture; Ψ is the dry suction, generally in negative values, h_0 represents the ponded depth on the soil surface and can be negligible compared to Ψ. Equations (1) and (2) represent the 2D continuity equation and the momentum equation, both of which are the fundamental equations in the flood modelling. Equation 1(c) represents the Green-Ampt infiltration model.

2.2. Karhunen-Loevè Expansion (KLE) of K_f

In this study, $K_f(X)$ is assumed as a random field of floodplain hydraulic conductivity and $K(X) = \ln[K_f(X, \omega)]$, where $X \in D$ and $\omega \in \Omega$ (a probabilistic space), which means that lnK_f varies over the domain spatially with a bounded, symmetric, and positive covariance function $K(x,y) = \langle K'(X,\omega)\, K'(X,\omega)\rangle$, which can be expressed as (Ghanem and Spanos, 1991):

$$K(x,y) = \sum_{n=1}^{\infty} \lambda_n f_n(x) f_n(y) \tag{2}$$

where λ_n represent eigenvalues; $f_n(x)$ and $f_n(y)$ are eigenfunctions, which are orthogonal and determined by dealing with the Fredholm equation analytically or numerically (Courant and Hilbert, 1953):

$$\int_D K(x,y) f_n(x)\,dx = \lambda_n f_n(y) \tag{3}$$

where λ_n and $f_n(x)$ for some specific covariance functions could be solved analytically (Zhang and Lu, 2004). The KLE representation of floodplain roughness coefficients can be expressed as (Liu et al., 2006):

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

$$K(X,\omega) = u(X) + \sum_{n=1}^{\infty} \varsigma_n(\omega)\sqrt{\lambda_n} f_n(X) \tag{4}$$

where $u(X)$ represents the mean of K; and $\varsigma_n(\omega)$ represent the orthogonal random variables with standard Gaussian distribution. Here, ω is omitted in following related expressions and equations. The calculated λ_n can be ranked as monotonically decreasing series with index n. In order to make the approximation effective in both computational effort and accuracy, the KLE representation of lnK_f (x) can be written in a finite form as follows (Liu et al., 2006):

$$\tilde{K}(X) = u(X) + \sum_{n=1}^{m} \varsigma_n \sqrt{\lambda_n} f_n(X) \tag{5}$$

where m represents the number of the items saved in the equation (5). Generally, the more the items obtained in equation, the more accurate the random hydraulic conductivity filed is approximated, which implies that more random field energy is saved, but with an increasing significant computational efforts. For simplicity of expression, the tilde over the symbol K is omitted in the future expressions and functions.

2.3. Polynomial Chaos Expansion

Polynomial chaos expansion (PCE), develop by Wiener (1938), is applied to express the dependent random fields of the simulation results, which can be decomposed as (Wiener, 1938):

$$\begin{aligned} y(X,\omega) &= a_0(X) + \sum_{i_1}^{\infty} a_{i_1}(X)\Gamma_{i_1}(\varsigma_{i_1}) + \sum_{i_1}^{\infty}\sum_{i_2}^{i_1} a_{i_1 i_2}(X)\Gamma_2(\varsigma_{i_1},\varsigma_{i_2}) \\ &+ \sum_{i_1}^{\infty}\sum_{i_2}^{i_1}\sum_{i_3}^{i_2} a_{i_1 i_2 i_3}(X)\Gamma_3(\varsigma_{i_1},\varsigma_{i_2},\varsigma_{i_3}) + ... \end{aligned} \tag{6}$$

where $a_0(X)$ *and* $a_{i_1,\ldots,iN}$ (X) represents the deterministic coefficients. The basis Γ_m $(\varsigma_{i1}\ldots\varsigma_{im})$ are defined as a set of m-order polynomial chaos for the random variables $(\varsigma_{i1}\ldots\varsigma_{im})$. Here, random variables $(\varsigma_{i1}\ldots\varsigma_{im})$ are assumed as normal distribution and the Γ_m $(\varsigma_{i1}\ldots\varsigma_{im})$ are m-order N-dimensional Hermite Polynomials defined as (Wiener, 1938):

$$\Gamma_3(\varsigma_{i_1},\varsigma_{i_2},\varsigma_{i_3}) = (-1)^m e^{\frac{1}{2}\varsigma^T\varsigma} \frac{\partial^m}{\partial\varsigma_{i_1}\ldots\partial\varsigma_{i_m}}\left[e^{-\frac{1}{2}\varsigma^T\varsigma}\right] \tag{7}$$

where $\boldsymbol{\varsigma}$ representing a vector $(\varsigma_{i1}\ldots\varsigma_{im})^T$; it exists corresponding relationship between the items in equation (6). The total number of the items P *can be determined by m* and N: $P = (N+m)!/N!m!$. The Hermite polynomials can be used to build up the best orthogonal basis for the normal random variables $\boldsymbol{\varsigma}$ and therefore can be used to set up the random fields of h_{max} (Ghanem and Spanos, 1991). A second-order PCE approximation of $K(X)$ can be expressed as:

$$\hat{h}_{max}(X,\varsigma) = a_0(X) + \sum_{i}^{N} a_i(X)\varsigma_i + \sum_{i}^{N} a_{ii}(X)\left(\varsigma_i^2 - 1\right) + \sum_{i=1}^{N-1}\sum_{j>I}^{N} a_{ij}(X)\varsigma_i\varsigma_j \tag{8}$$

where N represents the total number of the random variables and, in this study, it is the number of pairs of the eigenvalues and eigenfuctions correspondingly, the tilde over the symbol h_{max} (x) is omitted in the future expressions and functions. Define the coefficients in the equation (8) as $c_1(x)$, $c_1(x)$,…, $c_P(x)$ *and then define a coefficients matrix as* $\mathbf{C}$(x)=[$c_1(x)$, $c_1(x)$,… $c_P(x)$]. Subsequently, we

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

can transform the expression of equation (8) into: Z**C**(x)=h(x), in which **Z** is PXP corresponding to Hermite polynomials in the equation (8) with the selected sets of collocation points and h(x) is the output matrix $[h_{max1}(x), h_{max2}(x), \ldots, h_{maxP}(x)]^T$ accordingly.

2.4. KLE-based Probabilistic Collocation Method (PCM)

Based on the PCE, pprobabilistic collocation method easily transfers the complicated nonlinear problems into to independent deterministic equations. The details of the application of second-order KLE-based PCM are listed as follows (Liu et al., 2006, Li and Zhang, 2007):

(i) Choose the *P* collocation points $\boldsymbol{\varsigma_i}\ (\varsigma_{i1}\ \varsigma_{i2,\ldots,}\varsigma_{iN})^T$, *i=1, 2, …P*, which have the highest probability in the probabilistic space. As one key procedure of influencing the performance of PCM, one typical way is to take the roots of the higher hermiter polynomials. For instance, the collocation points for second-order PCE expansion are chosen form the roots of the 3-order *N*-dimensional Hermite Polynomials $H_3(\varsigma)= \varsigma^3 - 3\varsigma$: $(0, -\sqrt{3}, \sqrt{3})$. For 0, it has the highest probability for the standard normal random distribution, so the first set of collocation set is (0, …, 0).
(ii) Substitute $\boldsymbol{\varsigma_i} = (\varsigma_{i1}\ \varsigma_{i2,\ldots,}\varsigma_{iN})^T$ into the KLE and obtain the corresponding realizations for the input *Ks* random field.
(iii) Put the *P* realizations of K_s into the 2D physical model FLO-2D, run them and extract the maximum flow depth h_{max} from the simulation results.
(iv) Build up the relationship between Z and h(x), calculate the coefficients matrix **C**(x).
(v) Evaluate the stochastic moments of the simulated outputs.

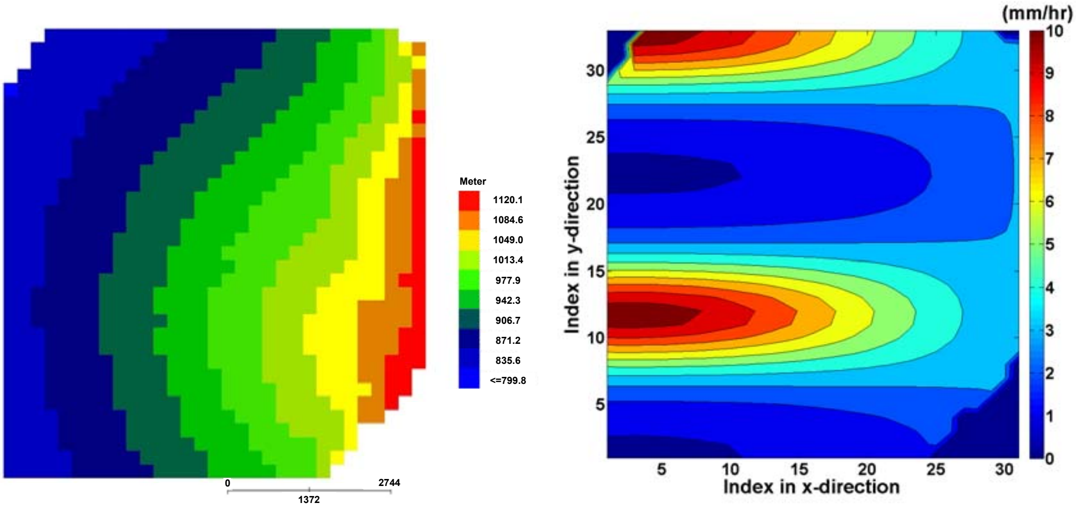

Figure 1 – (a) Ground surface elevation of modelling domain with 228.6-meter grid; (b) the 5th realization of floodplain hydraulic conductivity.

3. ILLUSTRATIVE EXAMPLE

3.1. Case Study

In this study, a simplified 2D alluvial flood inundation case located in Nevada, USA, is chosen to demonstrate the KLE-based PCM approach (FLO-2D Software, 2012). The average elapsed time is about 5.5 seconds based on the following computational platform: (i) operation system: Microsoft Windows 7 Professional edition (64 bits); (ii) CPU: AMD Phenom (tm) II×4 B95 @ 3.00GHz; (iiI) SIMM: 4GB (HLX DDR2 667MHZ). The basic settings for the case (as shown in Figures 1a and 1b) are: (i) the domain area is divided into 983 grid elements with 31 grid elements in x-direction and 33 grid elements in y-direction; (ii) the length of rectangular grid elements is 228.6 m (750 inch); (iii) and the elevation changes from 835.6 m to 1120.1 meter; (iv) m_x = 3, 4 and 5 and m_y = 3, 4 and 5 items are

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

retained for x-direction and y-direction separately and the total items are 3 x 3 = 9, 4 x 4 = 16 and 5 x 5 = 25; they are kept for the KLE decomposition with *K* random fields and the corresponding PCM items are 55, 153 and 351 according to equation (7); the *K* random field is in normal distribution with mean being 0.1145 mm/hour and standard deviation being 1.3931 mm/hour; Figure (1b) shows the 5th realization of the random field of floodplain hydraulic conductivity by KLE approximation; (v) the normalized correlation length is assumed as 0.5; (vi) to evaluate the performance of PCM, 10,000 runs of Monte Carlo simulation with the same random field of K are carried out.

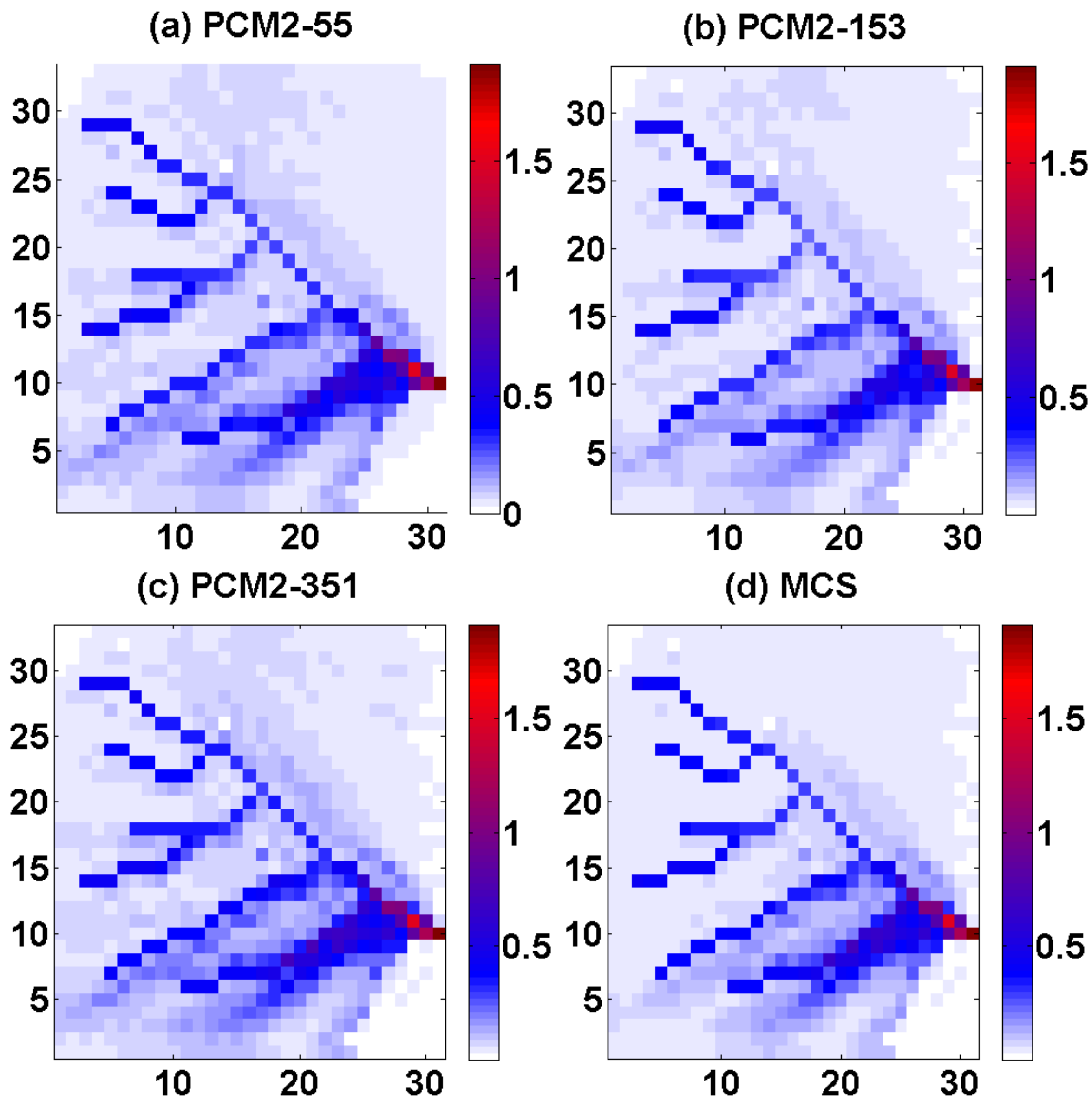

Figure 2 - Means of the maximum flow depth distributions from PCM and MCS simulations. PCM2-55 denotes the second-order PCM with 55 items; PCM2-153 denotes the second-order PCM with 153 items, PCM2-351 denotes the second-order PCM with 351 items and MCS denotes the Monte Carlo simulation.

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

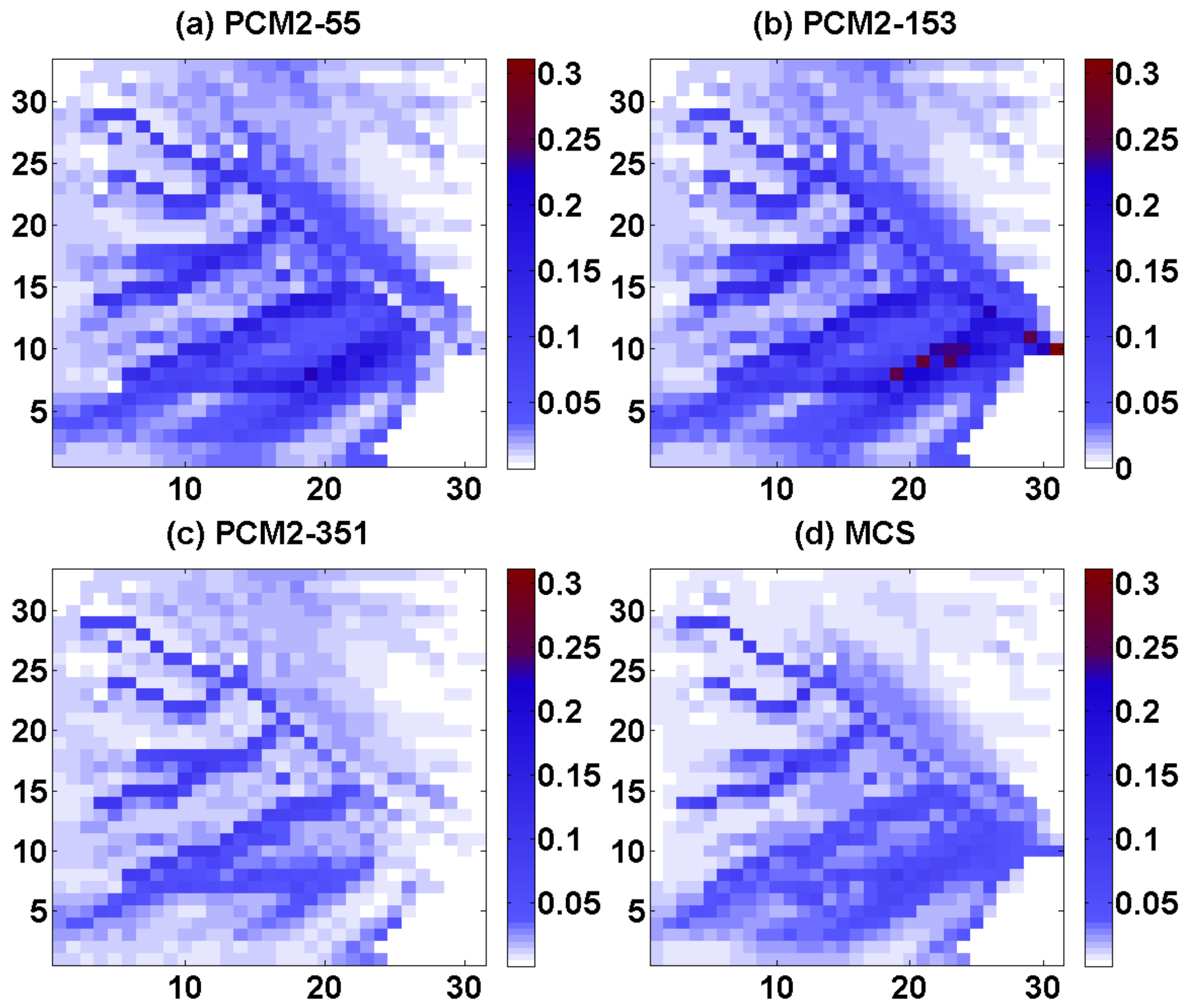

Figure 3 - Standard deviations of the maximum flow depth distributions from PCM and MCS simulations. PCM2-55 denotes the second- order PCM with 55 items; PCM2-153 denotes the second-order PCM with 153 items, PCM2-351 denotes the second-order PCM with 351 items and MCS denotes the Monte Carlo simulation results.

3.2. Comparison with MCS

Figure 2 presents the mean of maximum flow depth distribution calculated by second-order PCM with 55, 153 and 351 items and the MCS simulation results. The total number of runs of the flood physical models in PCM2-55, PCM2-153, and PCM2-351 are 55, 153, and 351 runs, respectively. It can be seen from the figures that the mean by second-order PCM could capture the mean of the MCS simulation results very well. Beside, for the same order of the PCM, the more items retained in the expansion the more accurate of the simulation results will be but the corresponding computation burden would increase significantly. Figure 3 shows the standard deviation (SD) of the simulated results by the second-order PCM with different items and MCS. The second-order PCM with 55 items and 153 items are much higher than that of 10,000 runs of MCS. The SD of the simulated results by the second-order PCM with 351 items (PCM2-351) fit the MC very well. Though the simulation runs have increased compared with those based on 55 and 153 items, the time of runs within the framework of PCM2-351 is much shorter.

Figure 4 demonstrates the mean and standard deviation of the maximum flow depth distributions for PCM and MCS simulation results crossing the 16th grid location expanding in x-direction. It seems that the second-order PCM with 55 items, 153 items and 531 items could well capture the mean of the maximum flow depth distributions crossing the 16th grid location expanding in x-direction calculated by

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

MCS. Furthermore, the second-order PCM with 351 items shows the best performance than PCMs with 55 items and 153 items (shown in Figure 4a). The SD of the results shown in the Figure (4b) indicates different divergences. The SD calculated by PCM with 55 items and 153 items are somewhat higher than that from MCS. For example, at the 9th grid element indexing in the x direction, the SD for the 55 items and 153 items are 0.13691 m and 0.1167 m, respectively; in comparison, the SD from MCS is 0.1076. From the comparison results, it is concluded that the PCM with 351 items fit the MCS better than those with 55 items and 153 items, although the computational time is higher; considering the time taken for the MCS (i.e. 10,000 runs), it is still efficient.

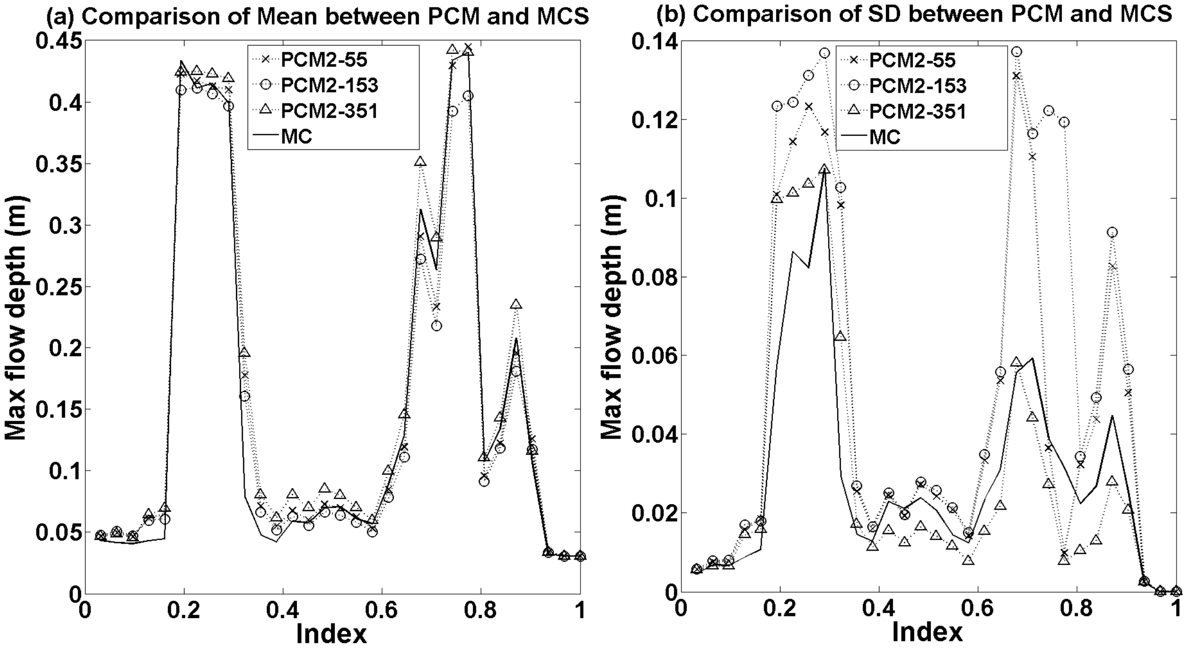

Figure 4 – Means and Standard deviations of the maximum flow depth distributions from PCM and MCS simulation crossing the 16th grid indexed in the y-direction. PCM2-55 denotes the second-order PCM with 55 items; PCM2-153 denotes the second-order PCM with 153 items, PCM2-351 denotes the second-order PCM with 351 items, and MCS denotes Monte Carlo simulation.

4. CONCLUSIONS

In this study, a second-order Karhunen-Loevè expansion (KLE) based probabilistic collocation method (PCM) was used to investigate the uncertainty propagation of uncertainty associated with floodplain hydraulic conductivity (K_s). The results including the mean and standard deviation were compared with those calculated from traditional Monte Carlo simulation (MCS). The 9, 16 and 25-items KLE decomposition were carried out and the corresponding PCM items were 55, 153 and 351 for the normally-distributed field of log-transformed *Ks (K)* considering the uncertain spatial distribution represented by pairs of eigenvalue and eigenfuctions. The results indicated that the mean of the maximum flow depth field by the second-order PCMs reproduced well the results calculated by MCS, but the standard deviation results were somewhat overestimated. Although the number of simulation runs by PCM2-351 was higher than those by PCM2-55 and PCM2-153, the computational effort was much lower than direct MCS (which took about 10,000 runs). A number of future works are expected: (i) compared with second-order KLE-based PCM, the 3-order expansion may bring about more accurate approximation of the random outputs fields, but the required computational effort would increase dramatically during the flood inundation modelling; improvement on the method is needed (ii) for our study, the simulation is taken for a single input random field, the PCM study for multi-input random field for flood inundation modelling could be potentially explored; (iii) in this study, the spatially-distributed uncertainty of the random input field was tackled; more complicated cases involving temporal uncertainties are desired to be examined.

ISBN 9781742721156 - DOI: 10.14264/uql.2014.40

5. ACKNOWLEDGMENTS

This research was supported by Singapore's Ministry of Education (MOM) AcRF Tier 1 Project (M4010973.030) and Tier 2 Project (M4020182.030).

6. REFERENCES

Ballio, F. and A. Guadagnini (2004). *Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology*. Water Resources Research 40(4): W04603.

Beven, K. and A. Binley (1992). *The future of distributed models：model calibration and uncertainty prediction*. Hydrological Processes 6: 279-298.

Chow, V. T., D. R. Maidment and L. W. Mays (1988). *Applied hydrology,* McGraw-Hill, New York, USA.

Courant, R. and D. Hilbert (1953). *Methods of Mathematical Physics*, Wiley-VCH, Hoboken .

FLO-2D Software, I. (2012). FLO-2D Reference Manual 2009. <https://www.flo-2d.com/download/>.

Ghanem, R. G. and P. D. Spanos (1991). Stochastic Finite Elements: A Spectral Approach, New York, NY : Springer New York, 1991.

Huang, S. P., S. T. Quek and K. K. Phoon (2001). *Convergence study of the truncated Karhunen–Loeve expansion for simulation of stochastic processes*. International Journal for Numerical Methods in Engineering, 52(9): 1029-1043.

Jung, Y. H. and V. Merwade (2012). *Uncertainty Quantification in Flood Inundation Mapping Using Generalized Likelihood Uncertainty Estimate and Sensitivity Analysis*. Journal of Hydrologic Engineering, 17(4): 507-520.

Li, H. and D. Zhang (2007). *Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods*. Water Resources Research, 43(9): W09409.

Li, H. and D. X. Zhang (2009). *Efficient and Accurate Quantification of Uncertainty for Multiphase Flow With the Probabilistic Collocation Method*. Spe Journal, 14(4): 665-679.

Liu, D. S. and H. G. Matthies (2010). *Uncertainty quantification with spectral approximations of a flood model*. 9th World Congress on Computational Mechanics and 4th Asian Pacific Congress on Computational Mechanics, WCCM/APCOM 2010, July 19 - July 23, 2010, Sydney NSW, Australia, Institute of Physics Publishing.

Liu, G., D. Zhang and Z. Lu (2006). *Stochastic uncertainty analysis for unconfined flow systems*. Water Resources Research, 42(9): W09412.

Lu, Z. and D. Zhang (2007). *Stochastic Simulations for Flow in Nonstationary Randomly Heterogeneous Porous Media Using a KL - Based Moment - Equation Approach*. Multiscale Modeling & Simulation, 6(1): 228-245.

Reza Ghanbarpour, M., S. Salimi, M. M. Saravi and M. Zarei (2011). *Calibration of river hydraulic model combined with GIS analysis using ground-based observation data*. Research Journal of Applied Sciences, Engineering and Technology, 3(5): 456-463.

Van Vuren, S., H. De Vriend, S. Ouwerkerk and M. Kok (2005). *Stochastic modelling of the impact of flood protection measures along the river waal in the Netherlands*. Natural Hazards, 36(1-2): 81-102.

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

Zhang, D. and Z. Lu (2004). *An efficient, high-order perturbation approach for flow in random porous media via Karhunen–Loève and polynomial expansions*. Journal of Computational Physics, 194(2): 773-794.