Journal of Hydrology 520 (2015) 101–122

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations

Mohammad Mahdi Rajabi [a,*], Behzad Ataie-Ashtiani [a], Craig T. Simmons [b]

[a] Department of Civil Engineering, Sharif University of Technology, PO Box 11155-9313, Tehran, Iran
[b] National Centre for Groundwater Research & Training and School of the Environment, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia

ARTICLE INFO

Article history:
Received 21 August 2014
Received in revised form 3 November 2014
Accepted 7 November 2014
Available online 20 November 2014
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Seawater intrusion
Uncertainty propagation
Sensitivity analysis
Monte Carlo methods
Polynomial chaos expansions

SUMMARY

Real world models of seawater intrusion (SWI) require high computational efforts. This creates computational difficulties for the uncertainty propagation (UP) analysis of these models due the need for repeated numerical simulations in order to adequately capture the underlying statistics that describe the uncertainty in model outputs. Moreover, despite the obvious advantages of moment-independent global sensitivity analysis (SA) methods, these methods have rarely been employed for SWI and other complex groundwater models. The reason is that moment-independent global SA methods involve repeated UP analysis which further becomes computationally demanding. This study proposes the use of non-intrusive polynomial chaos expansions (PCEs) as a means to significantly accelerate UP analysis in SWI numerical modeling studies and shows that despite the highly non-linear and non-smooth input/output relationship that exists in SWI models, non-intrusive PCEs provide a reliable and yet computationally efficient surrogate of the original numerical model. The study illustrates that for the considered two and six dimensional UP problems, PCEs offer a more accurate estimation of the statistics describing the uncertainty in model outputs compared to Monte Carlo simulations based on the original numerical model. This study also shows that the use of non-intrusive PCEs in the estimation of the moment-independent sensitivity indices (i.e. delta indices) decreases the computational time by several orders of magnitude without causing significant loss of accuracy. The use of non-intrusive PCEs for the generation of SWI hazard maps is proposed to extend the practical applications of UP analysis in coastal aquifer management studies.

1. Introduction

After several decades of research, the subject of seawater intrusion (SWI) numerical modeling has reached a level of maturity so as to allow for a more in depth evaluation of the critical issue of simulation under uncertain conditions. Addressing this subject involves, among other things, understanding and considering two key issues: first, how the uncertainty in model structure and inputs propagates through the model and leads to uncertainty of the output quantities. Quantifying this uncertainty is known as uncertainty propagation (UP) analysis (Brown and Heuvelink, 2006). Second, how the uncertainty in the outputs of the model can be allocated to different sources of uncertainty in the model inputs. This is known as sensitivity analysis (SA) (Saltelli, 2002).

UP analysis allows for possibilistic or probabilistic representations of SWI model outputs in the form of probability density functions (PDFs), prediction intervals (PIs), probability of exceedance of critical thresholds, fuzzy numbers, etc. Note that UP analysis is of particular importance in SWI modeling studies because the estimation of some SWI model inputs (such as longitudinal and transverse dispersivities, recharge rate and boundary conditions), is almost always accompanied by a relatively large level of uncertainty (Carrera et al., 2010; Rajabi and Ataie-Ashtiani, 2014). On the other hand, SWI models are extensively used in the framework of decision support systems, combined simulation–optimization approaches and risk analysis, as tools to derive optimal coastal aquifer management strategies (Werner et al., 2013; Ataie-Ashtiani et al., 2014; Ketabchi and Ataie-Ashtiani, 2014). Without proper UP analysis, models could lead to unrealistic predictions of the coastal aquifer system response and hence ineffective management strategies.

The UP analysis procedure is conceptually illustrated in Fig. 1. The procedure starts by the identification and characterization of model input uncertainties and subsequently propagates these

* Corresponding author.
E-mail addresses: mmrajabi@mehr.sharif.edu (M.M. Rajabi), ataie@sharif.edu (B. Ataie-Ashtiani), craig.simmons@flinders.edu.au (C.T. Simmons).

http://dx.doi.org/10.1016/j.jhydrol.2014.11.020
0022-1694/

Nomenclature

x, X	uncertain input variable(s)
N_{MC}	number of Monte Carlo realizations
$x_1, x_2, \ldots, x_{N_{MC}}$	N_{MC} realizations of x
$y(x)$	a random variable representing the output quantity of interest
$E()$	expected value
$\tilde{y}(x)$	Monte Carlo estimate of $E(y(x))$
$f_Y(y)$	(unconditional) probability density function of the output $y(x)$
$\hat{f}_Y(y)$	kernel density estimate of $f_Y(y)$
$K()$	kernel function
h	kernel bandwidth
a	a real number
$\sigma_{\tilde{y}}$	the standard deviation of $\tilde{y}(x)$
$Pr[]$	probability
ξ	a random variable
α_i	mode strength
ψ_i	mode function
p_ξ	probability density function of ξ
$H_{e_n(x)}$	Hermite polynomial of order n
$f_n(\xi)$	truncated polynomial chaos expansion representing $f(\xi)$
$\Psi_i(\xi)$	mode function for the multivariate case
m	number of components of ξ vector in the multivariate case
d	order of polynomial chaos expansions
D	number of polynomial chaos expansion terms
q	number of regression points
LC	least squares criterion
$\xi^{(k)}$	realizations of the random input variable ξ
$y(\xi^{(k)})$	exact solutions based on $\xi^{(k)}$
μ	mean
σ^2	variance
$f_{Y\|X=x_i}(y)$	conditional probability density function of the output $y(x)$ for $X = x_i$
$s(x_i)$	the difference between the areas under the conditional ($f_{Y\|X=x_i}(y)$) and unconditional ($f_Y(y)$) probability density functions
δ_X	delta indices
NL_1, NL_2	number of sample points in the two loops of the double-loop Monte Carlo method
n_p	number of uncertain inputs in the double-loop Monte Carlo method
NL	number of simulations in the single-loop Monte Carlo method
f_{Y,X_i}	joint probability density function of the model output and each of the input variables
k	permeability of the aquifer
Q	total constant fresh-water inflow on the inland boundary
$N_1, N_2, N_3, N_4, N_{Ref}, N_6$	monitoring points in the Henry problem domain
Q_f	freshwater inflow in the small island problem
C_f	freshwater inflow solute concentration in the small island problem
C_s	seawater solute concentration
KH_1	horizontal permeability of the upper layer in the small island problem
KV_1	vertical permeability of the upper layer in the small island problem
KH_2	horizontal permeability of the lower layer in the small island problem
KV_1	vertical permeability of the lower layer in the small island problem
LDH	longitudinal dispersivities for horizontal flows in the small island problem
LDV	longitudinal dispersivities for vertical flows in the small island problem
W_1, W_2, W_3, W_{Ref}	monitoring wells in the small island problem
μ_{Ref}	mean values in the reference solution
σ^2_{Ref}	variance values in the reference solution
$\epsilon(\mu)$	normalized deviations of mean estimates from the reference solutions
$\epsilon(\sigma^2)$	normalized deviations of variance estimates from the reference solutions
$std(\mu)$	normalized internal standard deviations of μ
$std(\sigma)$	normalized internal standard deviations of σ
κ	kurtosis
Pr_{exd}	the probability of exceedance of salinity concentration from 2000 mg/l
$N_{MC,(C>2000)}$	the number of Monte Carlo simulations resulting in salinity concentrations above 2000 mg/l
N_{rep}	number of repetitions of the uncertainty propagation methods

input uncertainties through the computational model in order to quantify their impact on the output quantities of interest (QoI). There are numerous methods for the propagation of uncertainty. These methods can be classified based upon how they characterize uncertainty, their input information requirements, the information that they provide regarding outputs, the necessity to reformulate the governing equations of the model and their mathematical basis. Based on these criteria, at least five different classifications of UP methods have been presented in the literature, which are reviewed in Table 1. The appropriate UP method that fits for a specific problem should be selected based on the characteristics of the uncertainty sources, the required level of uncertainty quantification, accuracy and confidence level, as well as computational cost or efficiency (Liu and Gupta, 2007; Lee and Chen, 2009). No approach to the propagation and analysis of uncertainty can be optimum for all needs (Helton and Davis, 2002). However, for several reasons reviewed by Rajabi and Ataie-Ashtiani (2014), Monte Carlo simulations (MCSs) are the most commonly used method of UP in mathematical and computational models used in almost every field of engineering and applied science, including SWI modeling. Examples of Monte Carlo based UP analysis in SWI modeling studies include Darvini et al. (2002), Lecca and Cao (2004, 2009), Prieto et al. (2006), Kerrou and Renard (2010), Kerrou et al. (2010), and Herckenrath et al. (2011).

SA is an important tool in model simplification, importance ranking, risk reduction, managing the data collection process and other purposes (Wei et al., 2013). There are several methods of SA, many of which are summarized in Fig. 2. A review of literature illustrates that SA in groundwater and SWI modeling studies has traditionally relied on local methods (examples in SWI modeling include Sanz and Voss (2006) and Ataie-Ashtiani et al. (2013)). This widespread use of local SA method is due to: (1) the high computational cost of global SA methods, (2) the simplicity of local methods, and (3) the capabilities of generic codes such as PEST (Doherty, 2005) and UCODE (Poeter and Hill, 1999) in the estimation of local sensitivity indices. A number of more recent studies in

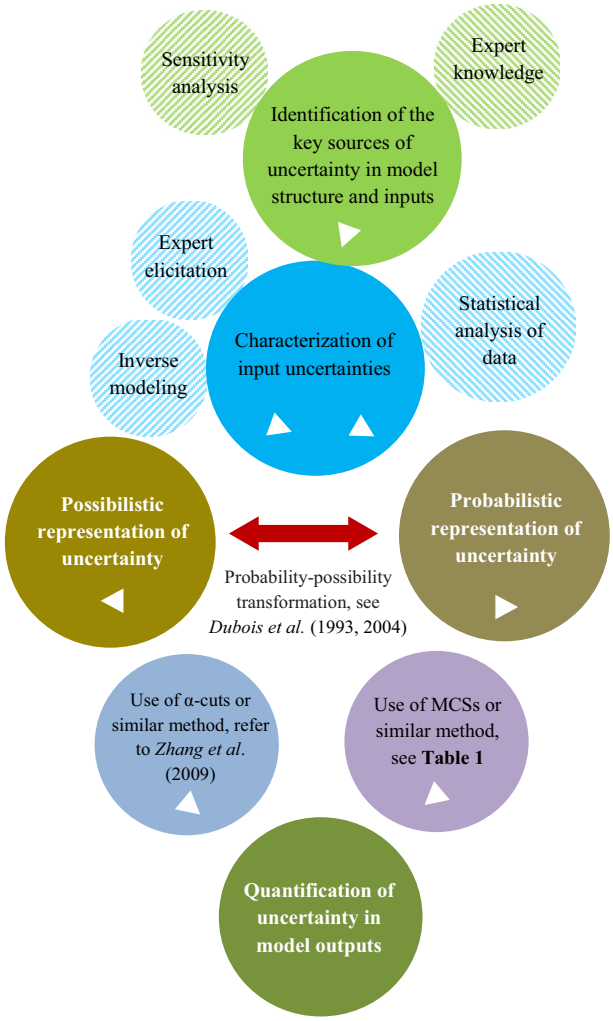

Fig. 1. The UP analysis procedure. (See above-mentioned references for further information.)

groundwater modeling have employed derivative-based (e.g., Malaguerra et al., 2013) and variance-based (e.g., Oladyshkin et al., 2012) global SA methods.

A basic question is that within the context of uncertainty analysis, which SA method is best suited for SWI modeling studies regardless of the computational and practical difficulties of some SA methods. It has been argued that SA methods used within the framework of uncertainty analysis should have the following four attributes: (1) they should be quantitative, (2) global, (3) model independent (i.e. include no assumptions on the functional relationship between model inputs and outputs), and (4) moment independent (Saltelli, 2002; Borgonovo, 2007). The fourth requirement implies that a sensitivity indicator should not be based on a single moment of the output distribution (such as the variance), because mapping the distribution into a single value leads to the inevitable loss of resolution (Helton and Davis, 2003). The second attribute leaves out the local SA methods. Note that some input parameters of SWI models (such as dispersivity and hydraulic conductivity) can vary within large ranges and it would be of great interest to consider the entire uncertainty range of these parameters. Moreover, possible interactions between different input parameters can be taken into account by evaluating the sensitivity indices of a given parameter while all other parameters are simultaneously varying (Haro Sandoval et al., 2012). Both of these objectives can be accomplished through the use of global SA. Within the framework of global SA, non-parametric methods often do not satisfy the model independence requirement because they are frequently bound to cases where a linear or monotone input/output relationship exists (Saltelli and Sobol', 1995; Borgonovo, 2007). Non-parametric methods are particularly not suitable for SWI modeling studies because in general, SWI model outputs are nonlinear and non-monotonic functions of the model inputs. Variance-based and derivative-based methods do not comply with the moment-independence requirement and hence, moment independent methods are generally the best option for SA in the framework of uncertainty analysis. This applies to SWI simulations as well. Moment independent SA methods (along with other global methods) involve the propagation of uncertainty through the model (Borgonovo, 2007) and therefore often incorporate MCSs. Note that due to the complex nature of SWI models, it is very difficult, if not impossible, to analytically derive the moment-independent sensitivity indices.

MCSs employed within the framework of both UP analysis and SA require a large ensemble of deterministic simulations to provide a reliable estimate of model uncertainties. Hence, they become computationally expensive when the computational demand of a single deterministic run is sufficiently high. This is particularly problematic in real world numerical models of SWI and must be somehow confronted in order to facilitate the successful implementation of UP analysis and SA in such models. A review of literature illustrates that four primary approaches have previously been employed in SWI modeling studies to tackle this problem: (1) parallelization and grid computing (e.g., Kerrou et al., 2010; Lecca and Cao, 2009), (2) using more efficient sampling

Table 1
Review of UP methods.

Categorization presented by		Category	Examples of methods
Mulani (2006)	1	Possibilistic methods	Alpha-cuts method
	2	Probabilistic methods	Monte Carlo methods
Lee and Chen (2009)	1	Simulation based methods	Monte Carlo methods
	2	Local expansion-based methods	Taylor series, Perturbation methods
	3	Most probable point based methods	First-order and second-order reliability methods
	4	Functional expansion-based methods	Neumann expansion, Polynomial chaos expansions
	5	Numerical integration-based methods	Dimension reduction method
Loeven et al. (2007)	1	Intrusive methods	Intrusive polynomial chaos expansions
	2	Non-intrusive methods	Monte Carlo methods, non-intrusive polynomial chaos expansions
Perez (2008)	1	Sampling based methods	Monte Carlo methods
	2	Spectral methods	Polynomial chaos expansions
Cullen and Frey (1999)	1	Analytical methods	Transformation of variables method
	2	Approximation methods	Taylor series
	3	Numerical methods	Monte Carlo methods

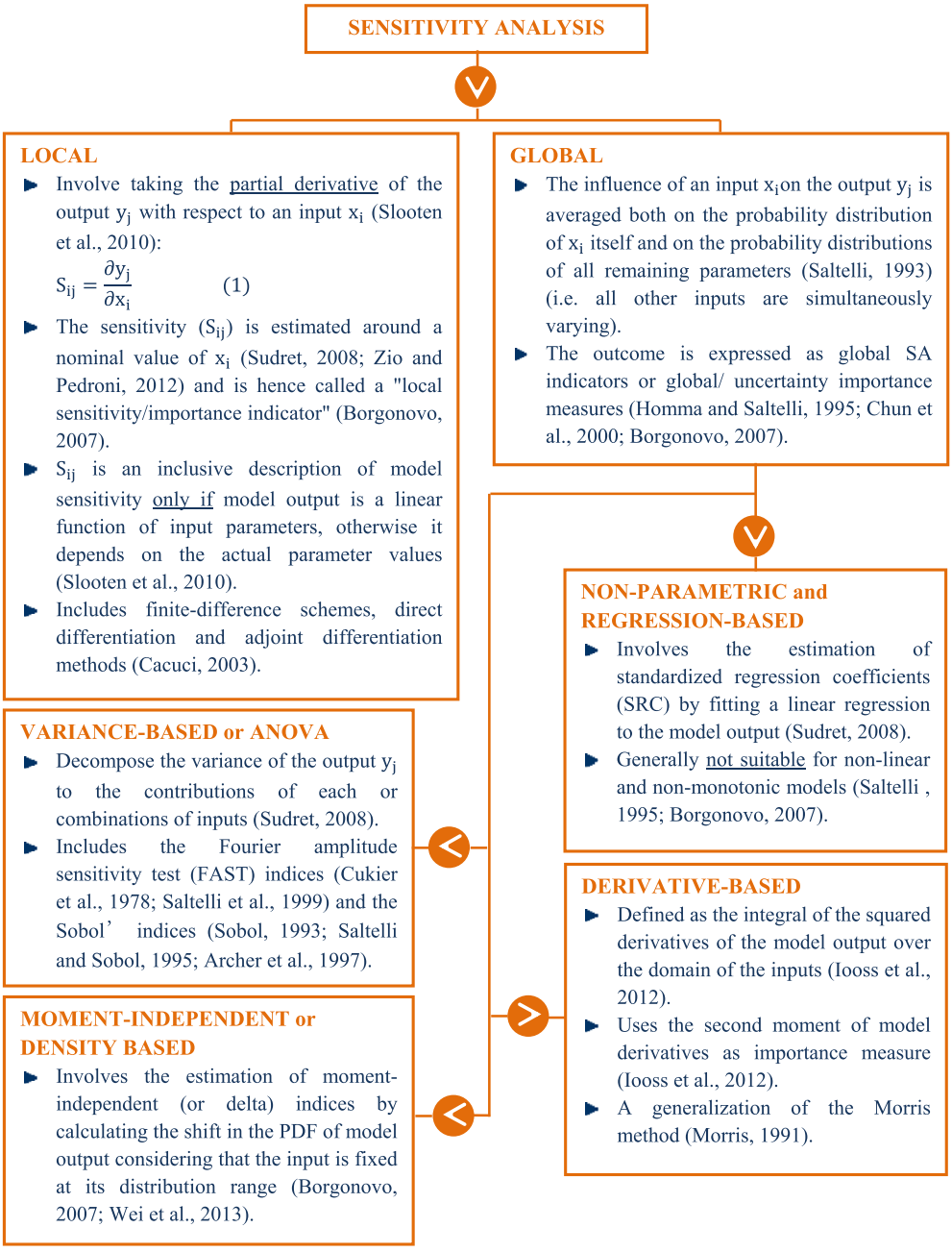

Fig. 2. A review of SA methods. (See above-mentioned references for further information.)

strategies such as Latin hypercube sampling (LHS) and optimized Latin hypercube sampling (OLHS) (e.g., Rajabi and Ataie-Ashtiani, 2014), (3) lower-fidelity modeling, that is, developing simplified and less-detailed versions of the original model while retaining the key physical characteristics (e.g., Kerrou and Renard, 2010), and (4) response surface surrogate modeling (i.e. developing data-driven and physics-free approximations of the model response (Razavi et al., 2012)). An example of the latter approach is the use of genetic programming by Sreekanth and Datta (2014).

An alternative surrogate modeling approach which has received no attention in the SWI modeling community is the use of non-intrusive polynomial chaos expansions (PCEs). Non-intrusive PCEs offer a number of benefits compared to other response surface surrogates previously used in SWI modeling studies such as artificial neural networks and genetic programming. First, PCEs are more than just an approximation of the simulator, they allow for a fully probabilistic prediction of what the simulator would produce. In fact, the full randomness of the response is contained within the set of the expansion coefficients (Haro Sandoval et al., 2012). Second, with the use of PCEs the mean and variance of the output QoI are available in closed-form (Oladyshkin and Nowak, 2012). Third, PCEs can be used with any second-order random process (i.e. processes with finite variance, which applies to most physical processes including SWI) (Xiu and Karniadakis, 2003a). Fourth, with the introduction of generalized PCEs (gPCE) (Xiu and Karniadakis, 2002, 2003a), multi-element gPCEs (Wan and Karniadakis, 2006; Prempraneerach et al., 2010) and arbitrary PCEs (aPCEs) (Oladyshkin and Nowak, 2012), PCEs can handle many probability distribution types (e.g., normal, gamma, beta, Poisson, etc.) as well as arbitrary distributions with arbitrary probability measures specified either analytically or numerically. Fifth, PCEs are transparent, simple to implement and have a strong mathematical basis (Cameron and Martin, 1947; Xiu and Karniadakis, 2003a; Lee and Chen, 2009). Finally, non-intrusive PCEs reach fast

Table 2
Review of the applications of non-intrusive PCEs in hydrogeology.

Reference	Objective of PCE application	Type of problem	Number/name of uncertain inputs	Uncertain output(s)	Type of polynomial	Order of PCE(s)	Method for selection of data points	Method for the estimation of PCE coefficients
Li and Zhang (2007)	UP	Flow in saturated media (1D, 2D/HE)	6, 7/hydraulic conductivity field (KLE)	Hydraulic heads	Hermite	2, 3, 4	Deterministic	PCM
Li et al. (2009)	UP	Flow in unsaturated media (2D/HE)	150/unsaturated conductivity (KLE)	Pressure heads, effective saturation	Hermite	2	Deterministic	PCM[a]
Fajraoui et al. (2011)	Global SA (VB)	Non-reactive solute transport in saturated media (2D/LE)	4, 6/inlet flow rates, permeability, transverse and longitudinal dispersivities	Solute concentrations	Legendre	4	Deterministic	PCM
Sochala and Le Maître (2013)	Global SA (VB)	Flow in unsaturated media (1D/HE)	3, 5/water content, hydraulic conductivity	Position and spreading of the wetting front	Legendre	1, 2, 3	Deterministic	SPM
Laloy et al. (2013)	BI	Flow in saturated media (3D/RP)	102/hydraulic conductivity field (KLE)	Hydraulic heads	Legendre, Hermite	1	Deterministic	PCM
Ciriello et al. (2013)	Global SA (VB)	Radionuclide transport in saturated media (2D/HE)	3/hydraulic conductivity, partitioning coefficient, dispersivity	Peak concentrations	Legendre	2, 3, 4	Deterministic	RM

UP: Uncertainty propagation.
SA: Sensitivity analysis.
BI: Bayesian inference.
VB: Variance based (Sobol indices).
HE: Hypothetical Example.
LE: Laboratory experiments.
RP: Real-world problem.
D: Dimensional.
KLE: Parameterized using the Karhunen–Loève expansion.
PCM: Probabilistic collocation method.
SPM: Spectral projection method.
RM: Regression method.
[a] Used leading term approximation to reduce the number of PCE coefficients.

convergence when the solutions are sufficiently smooth in the random space (Haro Sandoval et al., 2012).

PCEs have been previously used in UP studies in a variety of fields, such as structural dynamics (Sarkar and Ghanem, 2002), heat conduction (Xiu and Karniadakis, 2003b), air pollution dispersion (Konda et al., 2010) and fluid dynamics problems (Knioa and Le Maître, 2006). A number of studies related to hydrogeology have also used PCEs for UP analysis, which we will review in the following. The earliest example of the application of PCEs in problems involving transport in porous media is the study of Ghanem (1998). In his study, the Karhunen–Loève expansion (KLE) is used to represent the spatial randomness in the physical parameters of the porous medium (including dispersion and permeability), in a problem involving the transport of Trichloroethylene in the unsaturated zone. The resulting concentrations and effective heads throughout the domain are subsequently expanded by relying on the intrusive PCEs of the stochastic processes. Subsequent applications of PCEs in hydrogeology have mostly relied non-intrusive PCEs as a meta-modeling approach. Table 2 presents a review of such studies. The studies reviewed in Table 2 illustrate that non-intrusive PCEs can accurately estimate model output statistics and PDFs several orders of magnitude faster than conventional MCSs in applications involving flow and solute transport in saturated or unsaturated porous media. PCEs have not previously been used for UP analysis of density-dependent SWI numerical models. Note that the input–output relationship in density-dependent groundwater models is highly non-linear and non-smooth (Bear, 1999; Carrera et al., 2010; Werner et al., 2013) as compared to test case problems that have been previously used in the application of non-intrusive PCEs. This is important, because the PCE accuracy is known to crucially depend on the model smoothness and linearity with respect to its parameters, and building an accurate non-intrusive PCE for largely nonlinear models might be out of computational reach even in low dimensions (Chen et al., 2005; Oladyshkin and Nowak, 2012; Sochala and Le Maître, 2013). On the other hand, if the function to be represented with the non-intrusive PCE is smooth, the estimation of a small number of PCE coefficients and hence a small number of deterministic simulations, is enough to provide a reliable estimate of that function (Haro Sandoval et al., 2012). This is the case in many groundwater flow models. For example, Laloy et al. (2013) argued that a non-intrusive order-one PCE can adequately estimate the model output statistics of a three dimensional real-world groundwater flow model with 102 uncertain input variables, in an application involving screening of good candidates within a Bayesian inversion procedure. In such cases, non-intrusive PCEs provide the most advantage in UP analysis over MCSs based on the original model. However, for density-dependent SWI problems which are well known for their high non-linearity and non-smoothness, the extent to which non-intrusive PCEs can accurately estimate the output statistics and the degree to which they reduce the computational time is unknown.

PCEs were introduced to global sensitivity analysis by Sudret (2008). He proved that the variance-based Sobol' sensitivity indices can be computed analytically from the expansion coefficients, after a proper PCE has been constructed for the desired model. Thus the computational cost is transferred to the estimation of the PCE coefficients, leaving the subsequent SA almost computationally costless. Since then, a number of other studies have used PCEs for the estimation of variance-based sensitivity indices (e.g., Blatman and Sudret, 2010; Sochala and Le Maître, 2013; Garcia-Cabrejo and Valocchi, 2014). The possibility of employing non-intrusive PCEs to provide a computationally efficient estimation of moment-independent sensitivity indices has not previously been assessed. This is important as it paves the way for the use of moment-independent method in the SA of computationally demanding models such as numerical models of SWI.

In view of these facts, the current study tends to answer the following questions: (1) Can non-intrusive PCEs provide a reliable estimate of model uncertainties in SWI numerical modeling studies? (2) To what extent do non-intrusive PCEs decrease the computational efforts of UP analysis in SWI numerical models compared to MCSs based on the original models? As discussed in the paper, the answer to this question depends on a number of factors such as the order of polynomial which could accurately estimate the model output statistics in density-dependent SWI models, the number of data points needed for the estimation of PCE coefficients, and the sampling method used in MCSs. (3) Can non-intrusive PCEs provide a computationally efficient estimation of moment-independent sensitivity indices without significant loss of accuracy? Even though our focus is on SWI modeling, the answer to the third question is of interest in a much wider spectrum of applications involving computationally extensive simulations.

A number of points should be addressed here to clarify the study objectives and scope: (1) we assume that the predictive uncertainty is governed by the uncertainty of input parameters and hence the structural uncertainty is neglected. (2) Unlike methodologies such as the generalized likelihood uncertainty estimation (GLUE) (Beven and Binley, 1992), the outcome of the UP analysis is not constrained to any observed system responses. (3) Here, we focus on the key uncertain input parameters of SWI models, namely permeability, dispersivity and recharge rate. In real world problems, the estimation of these parameters is almost always faced with aleatory and epistemic uncertainty. Furthermore, it is assumed that the uncertainty of these parameters can be represented by log-normal distributions, and so the Hermite polynomial basis is employed for PCEs (refer to Section 2.2 for more details). It is common practice to employ a log-normal distribution to represent uncertainty in permeability in relatively homogeneous aquifers (Meerschaert et al., 2013). It has also been shown that the uncertainty in the recharge rate can be characterized by log-normal distributions (e.g., Hassan et al., 2009). Moreover, some studies have employed log-normal distributions to characterize uncertainty in dispersivities (e.g., Chaudhuri and Sekhar, 2005). Hence, many previous studies involving UP analysis in groundwater and SWI modeling applications are based on log-normal probability assumptions for the uncertain inputs (examples include Prieto et al. (2006), Lecca and Cao (2009), and Rajabi and Ataie-Ashtiani (2014)). Nevertheless, the basic findings of these studies are equally applicable for other uncertain input parameters, probability distribution types and related polynomial basis functions. (4) There are several methods for estimating the non-intrusive PCE coefficients which are reviewed in Section 2.2. The aim here is not to address all of them or present a comparative study of their performance. This has been previously done, to some extent, by a number of studies including Eldred et al. (2008). We have chosen a rather simple method of PCE construction (in terms of the involved coding effort) to show the practical advantages of PCEs in SWI UP and SA studies.

We will start by reviewing the theoretical background of MCSs, PCEs and moment-independent SA in Section 2. Two test cases of SWI are introduced in Section 3 which form the basis of the subsequent analysis presented in Section 4.

2. Theoretical framework

2.1. UP using MCSs

MCSs rely on repeated random sampling from the hypothetical PDFs of the uncertain input variables. These random samples are then employed to build and run an ensemble of deterministic simulations. These simulations result in a number of realizations of the output QoI which are subsequently used for the approximation of the output statistics. For example, if we denote the uncertain input variable(s) by x, the N_{MC} random samples of x by $(x_1, x_2, \ldots, x_{N_{MC}})$ and the output QoI by $y(x)$, the expected value of $y(x)$ is approximated by (Dimov, 2008):

$$E(y(x)) \cong \tilde{y}(x) = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} y(x_i) \tag{1}$$

The PDF of the output QoI (denoted here as $f_Y(y)$)) can also be estimated by its Monte Carlo realizations through a number of methods. In the current study, we use the non-parametric kernel density estimation (KDE) method. Given N_{MC} realizations of the output QoI $(y_1, y_2, \ldots, y_{N_{MC}})$, the basic KDE $\hat{f}_Y(y)$ of $f_Y(y)$ is given by (Scott, 1992):

$$\hat{f}_Y(y) = \frac{1}{N_{MC} h} \sum_{i=1}^{N_{MC}} K\left(\frac{y - y_i}{h}\right) \tag{2}$$

In Eq. (2), $K()$ is the kernel function and h is the bandwidth. We choose our kernel function to be Gaussian. In this study, a Matlab toolbox developed by Botev (2011) is used for KDE.

The statistics (such as mean and variance), computed on the basis of a finite number of realizations of the uncertain output QoI are themselves uncertain quantities. This implies that we can define confidence intervals for mean, variance and other output statistics in order to provide the order of magnitude of uncertainty associated with them. For example, the confidence interval for the mean estimates ($\tilde{y}(x)$) can be approximated by the Chebyshev inequality assuming that $\tilde{y}(x)$ has an unknown probability distribution (Ballio and Guadagnini, 2004):

$$Pr\left[\tilde{y}(x) - a\frac{\sigma_{\tilde{y}}}{\sqrt{N_{MC}}} \leqslant E(y(x)) \leqslant \tilde{y}(x) + a\frac{\sigma_{\tilde{y}}}{\sqrt{N_{MC}}}\right] \geqslant 1 - \frac{1}{a^2} \tag{3}$$

where $\sigma_{\tilde{y}}$ is the standard deviation of $\tilde{y}(x)$ for several repetitions of MCSs and $(1 - 1/a^2)$ is an estimation of the probability that $E(y(x))$ lies within the confidence interval around $\tilde{y}(x)$. The Chebyshev inequality can also be used to estimate the confidence interval for variance and higher moments. These confidence intervals can be used to analyze the convergence of Monte Carlo estimates of mean and variance, and to identify a criterion for the reliable choice of N_{MC} for a desired level of accuracy (Ballio and Guadagnini, 2004).

In this study, the implementation of MCSs is based on the capabilities of the SENSAN code (within the framework of the PEST suite (Doherty, 2005)) to carry out multiple model runs without user intervention.

2.2. PCEs

PCEs provide the means to represent a random variable (y) as a function ($f()$) of another random variable (ξ) with a "predefined probability distribution":

$$y = f(\xi) \tag{4}$$

Note that Eq. (4) cannot be merely interpreted as an equation relating specific values of the two random variables. What it really implies is that y has the same probability distribution as $f(\xi)$ (in statistical notation: $y \sim f(\xi)$). The PCE seeks an appropriate function $f()$, by decomposing y into separate deterministic and stochastic components as follows (Ghanem, 1998):

$$y = f(\xi) = \sum_{i=0}^{\infty} \alpha_i \psi_i(\xi) \tag{5}$$

In Eq. (5), α_i is the deterministic component or the mode strength, and ψ_i is the stochastic component which is also called the mode function. ψ_i is a polynomial of order i that satisfies the orthogonality condition:

$$\langle \psi_j, \psi_k \rangle = \int \psi_j(\xi)\psi_k(\xi)p_\xi(\xi)d\xi = 0 \quad j \neq k \tag{6}$$

where $\langle \psi_j, \psi_k \rangle$ is the inner product of ψ_j and ψ_k, and p_ξ is the PDF of ξ. Based on the Wiener theory (Wiener, 1938) and the generalized Cameron–Martin theorem (Cameron and Martin, 1947), any second order random process y can be expanded by polynomial function series in the form of Eq. (5), and the resulting expansion is convergent in the mean square sense. The choice for the type of orthogonal polynomial used in the construction of PCEs is dictated by the probability distribution of ξ. The classification tree that maps the distribution of ξ to the corresponding "optimal" orthogonal polynomial is called the Askey scheme (Askey and Wilson, 1985) and is presented in Table 3. Note that sub-optimal convergence rates will result when the optimal Askey polynomials are not employed (Xiu and Karniadakis, 2002). In case the uncertain inputs have different probability distributions, PCEs can still be used for the hybrid propagation of uncertainty (see Ayres et al. (2014)). In order to use the PCEs, the uncertain input parameters must be statistically independent. Linear correlation between the uncertain input variables can be removed prior to the construction of the PCE by employing adequate linear transformations, for example through the use of principal component analysis (Oladyshkin and Nowak, 2012).

This study is based on the assumption of lognormal distributions for the uncertain input variables. These log-normal distributions are converted to the standard normal distribution, prior to employing the Hermite polynomials. The Hermite polynomials are described as (Xiu and Karniadakis, 2003a):

$$\begin{aligned} &H_{e_0}(x) = 1, \quad H_{e_1}(x) = x, \\ &H_{e_{n+1}}(x) = xH_{e_n(x)} - nH_{e_{n-1}}(x) \quad \text{for } n = 1, 2, \ldots \end{aligned} \tag{7}$$

For practical reasons, Eq. (5) is often truncated and only a limited number of terms are considered (Li and Zhang, 2007):

$$y \cong f_n(\xi) = \sum_{i=0}^{D} \alpha_i \psi_i(\xi) \tag{8}$$

For the multivariate situation, ξ is a vector and the polynomial $\psi_i(\xi)$ (usually denoted as $\Psi_i(\xi)$ in the multivariate case) is a tensor product of the polynomial bases for each component of ξ, given that the components of ξ are independent and identically distributed (i.i.d.) random variables:

$$\Psi_i(\xi) = \prod_{j=1}^{m} \psi_{i_j}(\xi) \tag{9}$$

In Eq. (9), m is the number of components of ξ. Hence, the number of order d multivariate polynomials of m variables is $D = \binom{m+d}{d} = \frac{(m+d)!}{m!d!}$.

In order to implement the PCEs, the mode strengths should be estimated by intrusive or non-intrusive (i.e. black box) methods (Perez, 2008; Oladyshkin and Nowak, 2012). In the intrusive method, the random model inputs are replaced with their PCEs in the governing equations of the model. The resulting stochastic equations are then solved by the same numerical methods applied to the original deterministic system. However, the procedure may become analytically cumbersome (Oladyshkin and Nowak, 2012). On the contrary, the non-intrusive methods do not require any modifications to the deterministic code and hence give rise to the use of PCEs as meta-models. Non-intrusive methods may be classified as projection methods (Ghiocel and Ghanem, 2002; Le Maitre et al., 2002), stochastic or probabilistic collocation methods (Xiu and Hesthaven, 2005), gradient-based methods also known as collocation methods coupled with sensitivity derivatives (Perez, 2008), and regression methods (Berveiller et al., 2006). In the current study we focus on the regression method. This method consists of the following steps: (1) choosing a set of q regression points in the probability space of the random input variable(s) ($\xi^{(k)}$, $k = 1, \ldots, q$), through deterministic or random sampling methods. In the current study, the choice of regression points is based on random sampling. Random selection of regression points has the advantage of being more flexible in selecting the number of required simulations (Eldred et al., 2008). Random sampling is performed by the use of OLHS based on the centered L_2-discrepency (CLD) criterion and the enhanced stochastic evolutionary (ESE) optimization algorithm (here denoted as the CLD–ESE sampling strategy). This choice has been made due to the superior space-filling properties of sample designs based on the CLD–ESE strategy. For more details the reader is referred to Rajabi and Ataie-Ashtiani (2014). This study employs the R open source statistics software (http://www.r-project.org/) and the software package 'DiceDesign' (see http://cran.r-project.org/web/packages/DiceDesign/index.html) by Franco, Dupuy and Roustant to implement the CLD–ESE sampling strategy. (2) Using these regression points to perform q deterministic simulations of the model. The resulting exact solutions are denoted here as $y\left(\xi^{(k)}\right)$. (3) Estimating the non-intrusive PCE coefficients by minimizing the following least square criterion:

$$LC = \sum_{k=1}^{q} \left[y\left(\xi^{(k)}\right) - \sum_{i=0}^{d} \alpha_i \psi_i\left(\xi^{(k)}\right) \right]^2 \tag{10}$$

Note that Eq. (10) is the regression of the exact solution with respect to the PC basis (Sudret, 2008). Eq. (10) can be solved by optimization algorithms such as pattern search (Hooke and Jeeves, 1961). Alternatively, Eq. (10) can also be written as a system of q nonlinear equations. The latter approach becomes computationally more efficient as the dimensions of the problem increase with increasing values of q and d, and hence this approach has been employed in the current study. The resulting systems of nonlinear equations are solved using the Levenberg–Marquardt (Levenberg, 1944) algorithm.

If necessary, the parameters should be mapped onto the required intervals given in Table 3. There are many possible PCEs of a given y with respect to a given ξ, which differ only in the series of mode strengths. Once the PCEs are built, the mean (μ) and variance (σ^2) of y can be obtained in closed-form using properties of the orthogonal polynomials (Oladyshkin and Nowak, 2012):

$$\mu(Y) = \alpha_0 \tag{11-a}$$

$$\sigma^2(Y) = \sum_{i=1}^{n} \alpha_i^2 \tag{11-b}$$

Higher moments and the PDF of y can be estimated by applying MCSs with the non-intrusive PCEs as the surrogate model.

2.3. Moment-independent SA

As previously described, the PDF of an output QoI can be estimated through UP analysis. If all defined uncertain input variables are varied simultaneously over their range of uncertainty, the resulting PDF of the output QoI is called the unconditional PDF and is denoted here as $f_Y(y)$. Now consider the case in which one of the uncertain input variables (X) is fixed at a specific value ($X = x_i$) while the other uncertain inputs vary over their respective ranges. The UP analysis procedure can still be used to estimate the PDF of the output QoI, which in this case is called the conditional PDF. We denote the conditional PDF as $f_{Y|X=x_i}(y)$, and estimate the difference between the areas under the two PDF curves as follows (Liu and Homma, 2009; Wei et al., 2013):

Table 3
The relationship between the probability distribution of the uncertain inputs and the type of Wiener–Askey polynomial (Xiu and Karniadakis, 2003a).

Distribution		Polynomial type	Support
Continuous distribution	Normal[a]	Hermite	$(-\infty, +\infty)$
	Log-normal[b]		
	Gamma (exponential)	Laguerre	$[0, +\infty)$
	Beta	Jacobi	$[a, b]$
	Uniform	Legendre	$[a, b]$
Discrete distribution	Poisson	Chalier	$\{0, 1, 2, \ldots\}$
	Binomial	Krawtchouk	$\{0, 1, 2, \ldots, N^c\}$
	Negative binomial	Meixner	$\{0, 1, 2, \ldots\}$
	Hypergeometric	Hahn	$\{0, 1, 2, \ldots, N^c\}$

[a] Gaussian anamorphosis or normal score transformation can be used to convert other distribution types to normal distribution (Wackernagel, 1998), so that the Hermite polynomial can be used. However, this could lead to slow convergence of the expansion (Oladyshkin and Nowak, 2012).
[b] The log-normal distribution should be transformed to normal distribution using exponential transformation prior to the successful application of the Hermite polynomials.
[c] N is a finite integer.

$$s(x_i) = \int |f_Y(y) - f_{Y|X=x_i}(y)| dy \tag{12}$$

By re-calculating $s(x_i)$ for different values of x_i over the entire range of X, the expected shift between the unconditional and conditional PDFs can be estimated (Liu and Homma, 2009; Wei et al., 2013):

$$E_X[s(x_i)] = \int f_X(x_i) s(x_i) dx_i \tag{13}$$

Finally, the delta indices for the uncertain input X can be estimated by (Liu and Homma, 2009; Wei et al., 2013):

$$\delta_X = \frac{1}{2} E_X[s(x_i)] \tag{14}$$

The following is always true for δ_X (Wei et al., 2013):

$$\delta_X \in [0, 1] \tag{15-a}$$

$$\text{If } y \text{ is independent of } X \Rightarrow \delta_X = 0 \tag{15-b}$$

In practice, delta indices are estimated by a number of numerical methods including the PDF-based method (Borgonovo, 2007), the CDF-based method (Liu and Homma, 2009) and the double-loop and single-loop Monte Carlo methods proposed by Wei et al. (2013). These methods are all, to different extents, computationally extensive because they employ repeated MCSs. For example, Wei et al. (2013) employed 6,001,000 simulations within the framework of the double-loop Monte Carlo methods, and 3000 simulations within the context of the single-loop Monte Carlo methods, to estimate the delta indices of six uncertain input parameters of a cantilever beam structure problem.

In order to solve this computational problem, Borgonovo et al. (2012) proposed the meta-modeling approach for accelerating the computation of delta indices. Their study was based on the application of the state dependent parameter (Ratto et al., 2007) and kriging emulator. Here we propose the use of non-intrusive PCEs to accelerate the calculation of delta indices. The methodology is applied to the double-loop and single-loop Monte Carlo methods of Wei et al. (2013), although it can similarly be applied to other methods of calculating the delta indices. The algorithm behind the double loop Monte Carlo method is illustrated in Fig. 3. The method involves two ensembles with NL_1 and NL_2 sample points. The total number of simulations in the double loop Monte Carlo method (as well as the PDF-based and CDF-based methods) is $(n_p NL_2 + 1)NL_1$, where n_p is the number of uncertain inputs (Wei et al., 2013). Compared with the PDF-based and CDF-based methods, the double-loop MCS method does not require the computation of the intersection points between $f_Y(y)$ and $f_{Y|X=x_i}(y)$, and is therefore more accurate (Wei et al., 2013).

The algorithm of the single-loop Monte Carlo method is demonstrated in Fig. 4. The number of simulations required by the single-loop Monte Carlo method (denoted here by NL), does not increase with respect to the number of the input variables. This method has been shown by Wei et al. (2013) to be much more efficient than the double-loop Monte Carlo method, most notably in problems with high number of uncertain parameters. An important aspect of the single-loop method is the use of bivariate KDE.

3. Test cases

Two hypothetical test cases of SWI are described in this section. The first test case, based on the Henry problem, involves SWI in a two-dimensional cross section of a confined coastal aquifer under steady-state conditions. The second test case is a three dimensional transient problem involving SWI due to over-pumping in a small island aquifer system.

3.1. The Henry problem

The Henry problem (Henry, 1964) is a classic example of variable density SWI. The simplicity and minute simulation time of the Henry problem numerical model has made it an ideal test case for studies involving large numbers of SWI numerical simulations. As a result, it has been previously applied to the study of SA (e.g., Sanz and Voss, 2006) and Monte Carlo based UP analysis (e.g., Kerrou and Renard, 2010; Rajabi and Ataie-Ashtiani, 2014) in SWI simulations. The Henry problem (illustrated in Fig. 5) considers a two-dimensional cross section of a rectangular confined coastal aquifer system. The inland boundary condition consists of freshwater inflow, while SWI occurs from the seaward boundary on the other side. The top and bottom of the aquifer are assumed to be impermeable. Here, the 2×1 units of length model domain is discretized into 20×10 square elements using a nodal spacing of 0.1 units of length. The system is initially filled with freshwater and the simulation is continued long enough for the concentration distribution to become steady. In the numerical simulations, the system reaches steady state after 100 one minute time steps. The numerical simulation of the Henry problem is carried out by employing the USGS SUTRA finite element code (Voss and Provost, 2010). The input parameter values are listed in Table 4.

Two of the input parameters of the Henry problem are assumed to be uncertain: (1) the permeability of the aquifer (k), and (2) total constant fresh-water inflow on the inland boundary (Q). The uncertainty of k and Q are characterized by log-normal distributions described in Table 5. These probability distributions are purely hypothetical and serve for illustrative purposes. The output QoI are the concentration (of total dissolved solids) and pressure in the monitoring points which are illustrated in Fig. 5.

3.2. The small island problem

The second test case involves SWI into the unconfined aquifer system of a small oceanic island. In such islands, fresh groundwater usually forms a convex-shaped fresh groundwater lens, which floats on the underlying seawater. This freshwater lens is highly vulnerable to salinization from SWI due to excessive groundwater extraction and climate variability (Falkland, 1991; Mahmoodzadeh et al., 2014). Many real cases of small islands involve two-layer hydrogeological conceptualizations (Ketabchi et al., 2014;

Mahmoodzadeh et al., 2014) and this concept has been incorporated into the current case study by using two different permeabilities for the two layers. Fig. 6 illustrates the hydrogeologic layering in a radial cross-section of the small island problem. The permeability is anisotropic in both layers. We also assume that the permeability of the upper layer is greater than the base layer which is often the case in sandy islands (Falkland, 1991). Here, we simulate a quadrant of a circular island undergoing SWI as a result of over-pumping. Simulations are performed in transient mode using the USGS SUTRA finite element code. The shape of the model, numerical discretization, boundary conditions and constant input parameter values are adapted from Voss and Provost (2010). The model proposed by Voss and Provost (2010) has been modified to include the described two-layer hydrogeology, and four pumping wells illustrated in Fig. 7. The input parameter values for the small island problem are presented in Table 6. In order to generate the initial conditions for the transient simulations, we start with saline water conditions everywhere beneath the land surface and simulate the formation of the freshwater lens as a result of saline water being flushed out by freshwater recharge from the island's surface. The simulation continues to run for a time that is long enough to reach equilibrium. The resulting

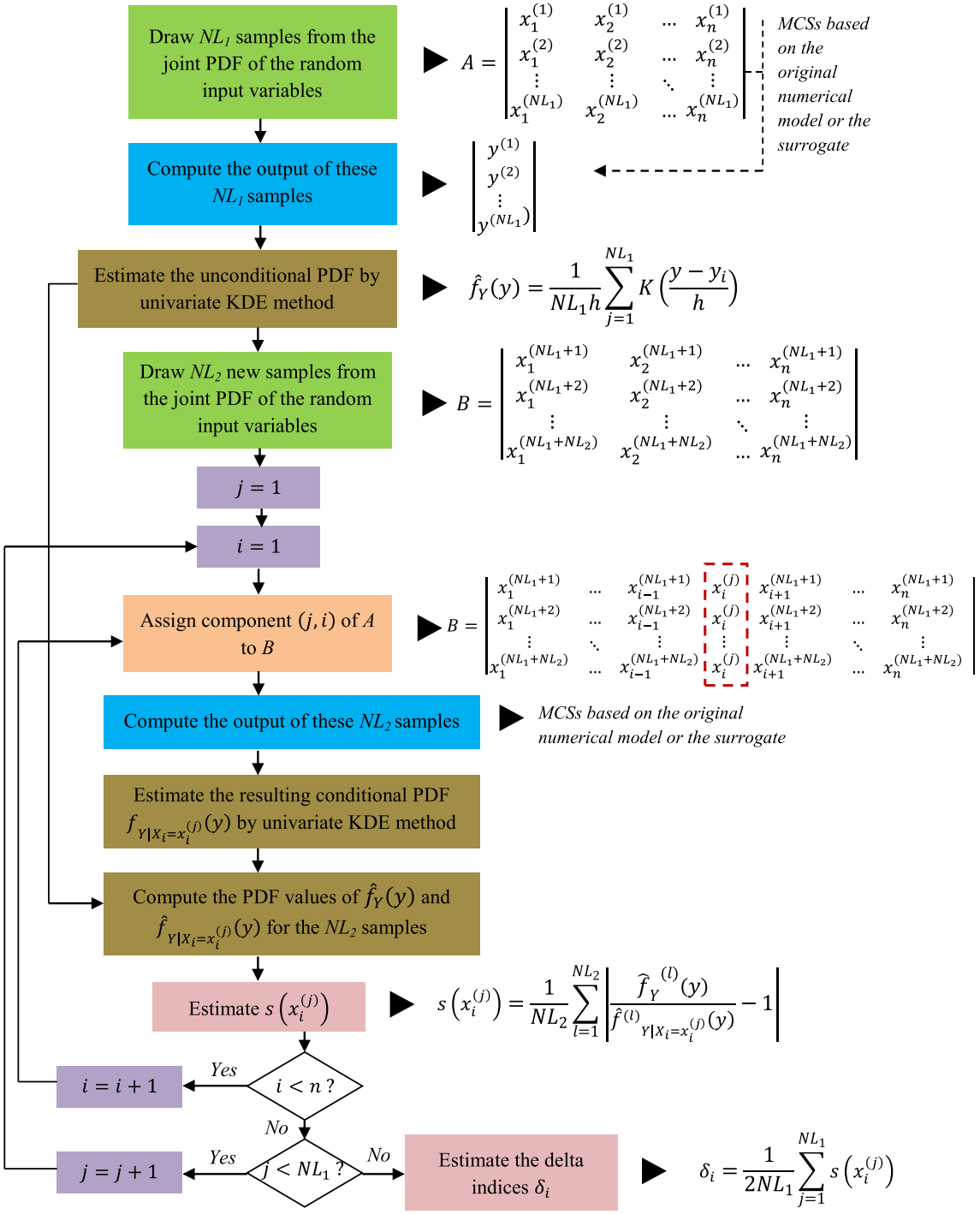

Fig. 3. The double-loop Monte Carlo method for the estimation of the moment independent delta indices (Liu and Homma, 2009; Wei et al., 2013).

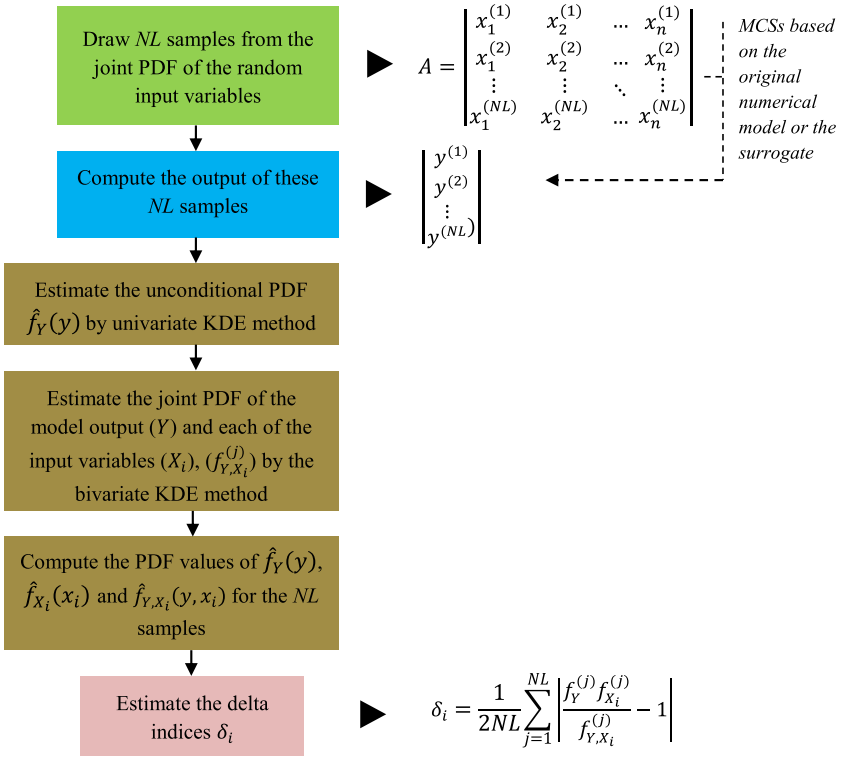

Fig. 4. The single-loop Monte Carlo method for the estimation of the moment independent delta indices (Wei et al., 2013).

pressure and salinity distributions (refer to Fig. 8) are used as the initial conditions for the transient simulations in which groundwater extraction starts through four pumping wells and continues for a period of 15 years. In the transient simulation, the following six input parameters are assumed to be uncertain: the horizontal and vertical permeabilities of the upper layer (KH_1 and KV_1 respectively), the horizontal and vertical permeabilities of the lower layer (KH_2 and KV_2 respectively), and longitudinal dispersivities for horizontal and vertical flows (*LDH* and *LDV* respectively). The uncertainties of these six inputs are characterized by log-normal distributions described in Table 7. These probability distributions are purely hypothetical and serve illustrative purposes. The output QoI are the concentration (of total dissolved solids) and pressure in the monitoring wells illustrated in Fig. 7. The numerical model of the small island problem has 42,432 nodes and 80 time steps and so the computational time is much higher than the Henry problem.

4. Results and discussion

4.1. Monte Carlo based UP analysis

In the first step, MCSs based on the original numerical model of the two test case problems are used for UP analysis. The aim is twofold: (1) to provide a reference solution in order to assess the accuracy and computational efficiency of non-intrusive PCEs, and (2) to study the characteristics of the PDFs of the output QoI in the test case problems. In order to address the first objective, a convergence analysis of MCSs has been performed. Accordingly, the number of numerical simulations within the context of MCSs (N_{MC}) has been steadily increased from 10 to 20,000 for the Henry problem, and from 10 to 1000 for the small island problem. This has been repeated in 12 parallel sets of independent chains to allow for the reliable estimation of confidence intervals. These MCSs have been performed using the simple random sampling (SRS) method. Fig. 9a and b depicts the Monte Carlo estimates of mean (μ) and standard deviation (σ) for pressure and concentration in the N_{Ref} monitoring point of the Henry problem against the size of the Monte Carlo samples. Fig. 9c and d shows the same graph for the W_{Ref} monitoring well of the small island problem. In these figures, the bold lines illustrate an average of μ and σ for the 12 repetitions of MCSs of a specific sample size. The dotted lines demonstrate the respective 96% estimated confidence intervals. The results have been plotted in semi-log scale to provide a better visual assessment of convergence. It is observed that after $N_{MC} = 1000$ the μ and σ estimates for pressure and concentration solutions relatively stabilize. In the Henry problem test case, the width of the confidence interval decreases to less than 10^{-3} times the μ and σ estimates after $N_{MC} = 10{,}000$. Consequently, the results of MCSs based on 10,000 numerical simulations are chosen as the reference solution in the Henry problem test case for all subsequent comparisons. In the small island problem, the computational cost of numerical simulations is several orders of magnitude larger than the Henry problem. Hence, we suffice to 1000 MCSs for the reference solutions. In the majority of UP studies involving MCSs, convergence analysis is either entirely omitted (e.g., Prieto et al., 2006; Kerrou et al., 2010; Herckenrath et al., 2011) or assessed qualitatively by means of visual inspection of sample moments vs. number of simulations graphs plotted in natural scale (e.g., Li et al., 2003; Lecca and Cao, 2009). However as indicated by Ballio and Guadagnini (2004), a convergence analysis of MCSs must be performed both qualitatively and quantitatively in order to make sure that the results have converged to a stable solution, and to identify the order of magnitude of uncertainty associated with Monte Carlo estimates. In terms of the number of MCSs required to reach stable solutions and the magnitude of uncertainty associated with Monte Carlo estimates, the outcome of this study is comparable to those obtained by Ballio and Guadagnini (2004) for the MCS of a two dimensional problem involving horizontal

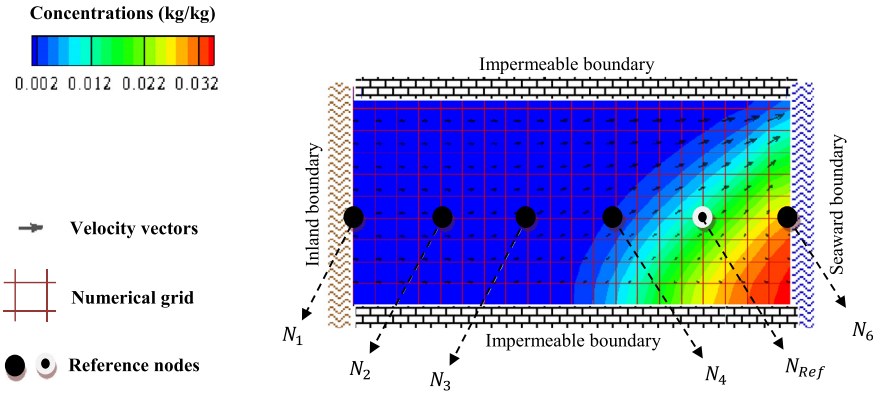

Fig. 5. Problem domain, boundary conditions, numerical grid, reference nodes, concentrations and velocity vector field for the steady state solution of the Henry problem (Henry, 1964).

Table 4
Constant input parameter values for the Henry problem (Henry, 1964; Sanz and Voss, 2006; Rajabi and Ataie-Ashtiani, 2014).

Parameter	Value	Unit
Longitudinal and transverse dispersivities	0.0	m
Porosity	0.35	–
Molecular diffusion	1.88571×10^{-5}	m^2/s
Viscosity	1×10^{-3}	kg/m s
Freshwater solute concentration	0.0	kg/kg
Total dissolved solute concentration of seawater	0.0357	kg/kg
Density of freshwater	1000	kg/m^3
Density of seawater	1024.99	kg/m^3

groundwater flow to a well. Similar to the Monte Carlo convergence analysis of this study, Ballio and Guadagnini (2004) used SRS for the generation of sample points.

The normalized PDFs of pressure and concentration solutions are illustrated in Fig. 10a and b for the six monitoring points of the Henry problem, and in Fig. 10c and d for the four monitoring wells of the small island problem. These PDFs are derived from the reference MCSs. A number of conclusions can be drawn with respect to Fig. 10: (1) the results of Lilliefors test (Lilliefors, 1967) (with 90% confidence) show that none of the illustrated PDFs for the Henry problem can be considered normally or log-normally distributed from a statistical point of view. This conclusion can be extended to pressure and concentration solutions of all 231 nodes of the Henry problem numerical grid. In the small island problem, the PDFs for concentration solutions are not normal or log-normal. However, the PDFs for pressure solutions are close to the normal distribution. Some of the pressure PDFs, such as the one for W_{Ref}, can be considered normally distributed with respect to the Lilliefors test. These comparisons show that in UP analysis of SWI numerical models, the PDFs of pressure and concentration solutions may not have the same shape as the PDFs of the uncertain inputs. (2) Qualitatively speaking, the PDFs of pressure solutions are generally smoother than concentration solutions most notably for points within the SWI transition zone. (3) The shape of the PDFs for pressure and concentration solutions varies significantly through the problem domain in both test cases.

Table 5
Characteristics of the hypothetical PDF for the uncertain input parameters of the Henry problem (Rajabi and Ataie-Ashtiani, 2014).

Parameter	PDF	Mean	Variance	Unit
Isotropic permeability	Log-normal	1.020408×10^{-9}	9.5×10^{-19}	m^2
Total constant freshwater inflow on the inland boundary	Log-normal	0.06	0.0008	kg/s

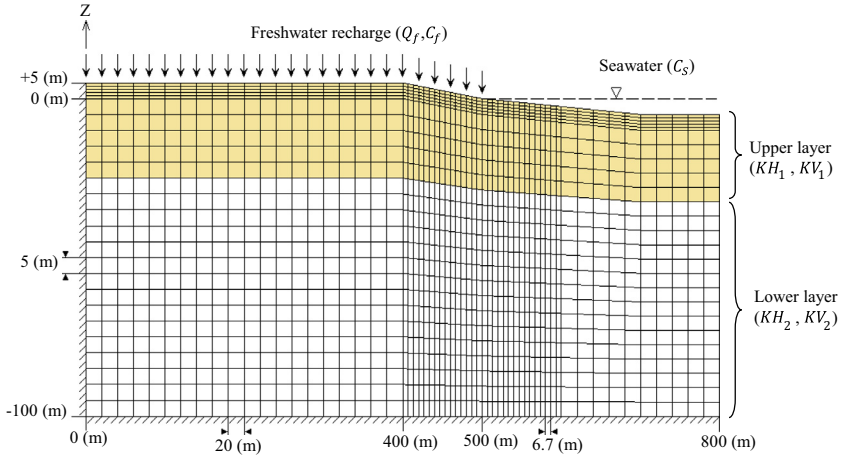

Fig. 6. Numerical grid (Voss and Provost, 2010) and hydrogeologic layering in a radial crosssection of the small island problem.

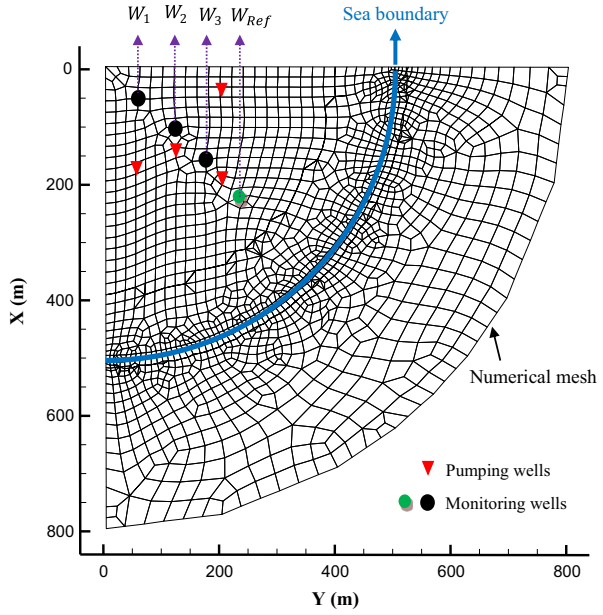

Fig. 7. The location of the pumping and monitoring wells in the plan view of the small island problem. The depth of the pumping and monitoring wells are 20 and 30 m respectively.

Table 6
Constant input parameter values for the small island problem (Voss and Provost, 2010).

Parameter	Value	Unit
Transverse dispersivities	0.1	m
Freshwater inflow	75	cm/yr
Porosity	0.1	–
Molecular diffusion	1×10^{-9}	m^2/s
Viscosity	1×10^{-3}	kg/m s
Freshwater solute concentration	0.0	kg/kg
Total dissolved solute concentration of seawater	0.0357	kg/kg
Density of freshwater	1000	kg/m^3
Density of seawater	1024.99	kg/m^3

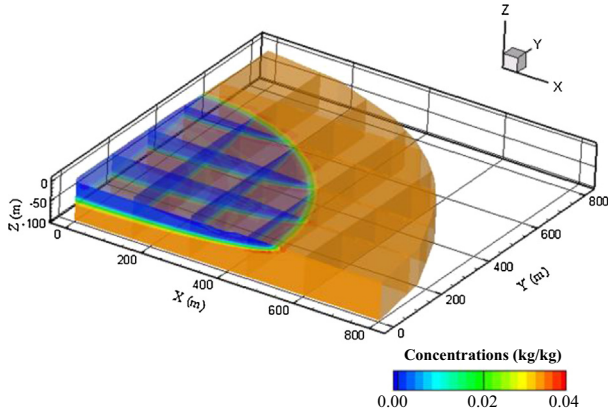

Fig. 8. The problem domain and initial concentrations in the small island problem.

4.2. PCE based UP analysis

Based on the steps described in Section 2.2, non-intrusive PCE are constructed for the pressure and concentration solutions of the two test case problems. Two key questions arise here. First, what is the proper order for the PCEs to obtain the most accurate results? Second, how many regression points should be used in construction of the PCE to obtain the optimal results in terms of accuracy. As described in the following paragraphs, the answers to these questions are not straightforward. We will focus our analysis here to the N_{Ref} monitoring point of the Henry problem and the W_{Ref} monitoring well of the small island problem. We denote polynomial order as d and the number of regression points as q. The accuracy is expressed as the normalized deviations from the reference solutions. The normalized deviations of μ and σ estimates of the output QoI are calculated from (Janssen, 2013; Rajabi and Ataie-Ashtiani, 2014):

$$\epsilon_i(\mu) = \frac{|\mu_i - \mu_{Ref}|}{\mu_{Ref}} \tag{16-a}$$

$$\epsilon_i(\sigma) = \frac{|\sigma_i - \sigma_{Ref}|}{\sigma_{Ref}} \tag{16-b}$$

where μ_{Ref} and σ_{Ref} are the mean and standard deviation of the reference solutions and μ_i and σ_i are the mean and standard deviation derived from a desired UP method. The results of an UP method are considered to be more accurate when their normalized deviations from the reference solution are smaller. Moreover, for a specified number of deterministic simulations, the UP method with smaller normalized deviations from the reference solutions is more efficient. As previously described, the current study employs random sampling for the selection of regression points. Hence, in order to dampen the effect of stochastic variations in the generation of random numbers, N_{rep} distinct PCEs have been built for each d and q, based on different sets of regression points derived from the CLD–ESE sampling strategy. The arithmetic mean of $\epsilon_i(\mu)$ and $\epsilon_i(\sigma)$ for the N_{rep} PCEs subsequently represents the average normalized deviations of PCE outputs for a given d and q. These arithmetic means are denoted here as $\epsilon(\mu)$ and $\epsilon(\sigma)$. The value of N_{rep} is eight for the Henry problem and four for the small island problem.

In order to additionally compare the robustness of different UP methods, the normalized internal standard deviations of μ and σ have been estimated using the following equations (Janssen, 2013; Rajabi and Ataie-Ashtiani, 2014):

$$std(\mu) = \sigma\sqrt{\frac{1}{n_r}} \tag{17-a}$$

$$std(\sigma) = \frac{\sigma}{2}\sqrt{\frac{2}{n_r - 1} + \frac{\kappa}{n_r}} \tag{17-b}$$

μ, σ and κ are the mean, standard deviation and excess kurtosis of N_{rep} repetitions of each UP strategy respectively. For a specified number of sample points, the sampling strategy that results in μ and σ with smaller normalized internal standard deviations is more robust.

Figs. 11 and 12 compare the accuracy and robustness of μ and σ estimates of pressure and concentration solutions for different orders of PCEs (from $d = 1$ to $d = 6$) and different numbers of regression points (for $q = 10$, 30, 60 and 100) in the Henry problem test case. The numbers of PCE coefficients for the bivariate polynomial orders $d = 1$, 2, 3, 4, 5 and 6 are 3, 6, 10, 15, 21 and 28, respectively. These figures also show the accuracy of the corresponding estimates obtained from MCSs based on the original numerical model with three different sampling strategies, namely SRS, LHS and CLD–ESE. Figs. 11 and 12 highlight a number of interesting results which we will review in the following.

(1) The relative deviations of μ and σ estimates from the reference solutions highly increase when the number of regression points is smaller than the number of PCE coefficients.

Table 7
Characteristics of the hypothetical PDF for the uncertain input parameters of the small island problem.

Parameter	PDF	Mean	Variance	Unit
Upper layer horizontal permeability	Log-normal	5×10^{-12}	1×10^{-24}	m^2
Upper layer vertical permeability	Log-normal	5×10^{-13}	1×10^{-26}	m^2
Lower layer horizontal permeability	Log-normal	5×10^{-13}	1×10^{-26}	m^2
Lower layer vertical permeability	Log-normal	5×10^{-14}	1×10^{-28}	m^2
Longitudinal dispersivity for horizontal flow	Log-normal	15	4	m
Longitudinal dispersivity for vertical flow	Log-normal	5	0.4	m

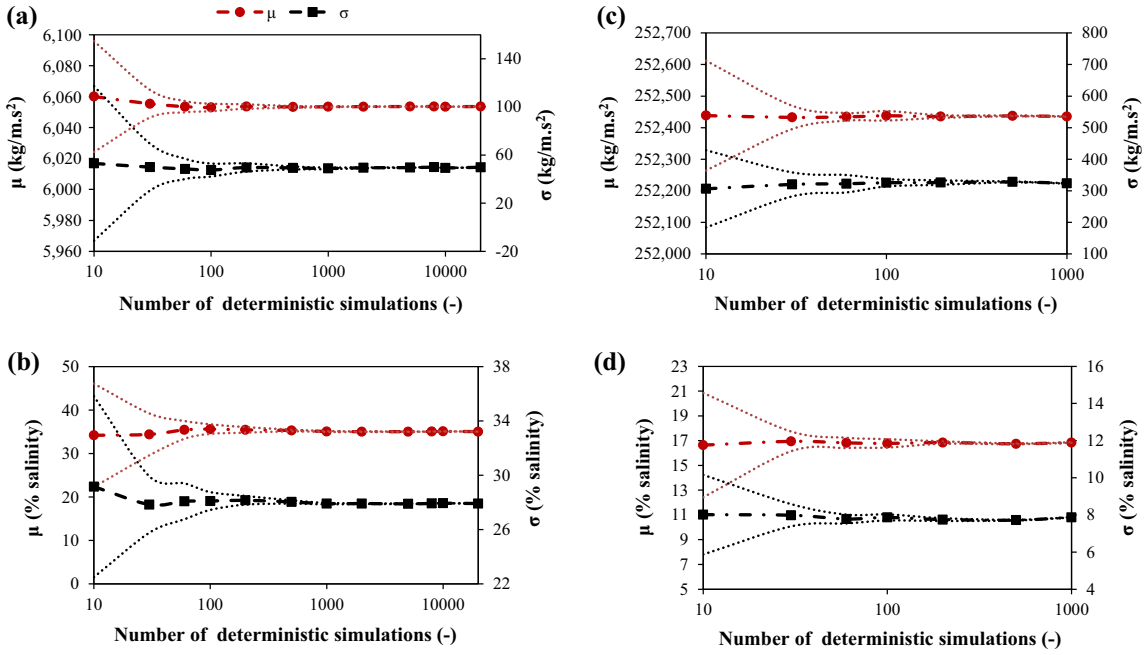

Fig. 9. Convergence analysis of MCSs for (a) pressure and (b) concentration solutions in the Henry problem, and (c) pressure and (d) concentration solutions in the small island problem.

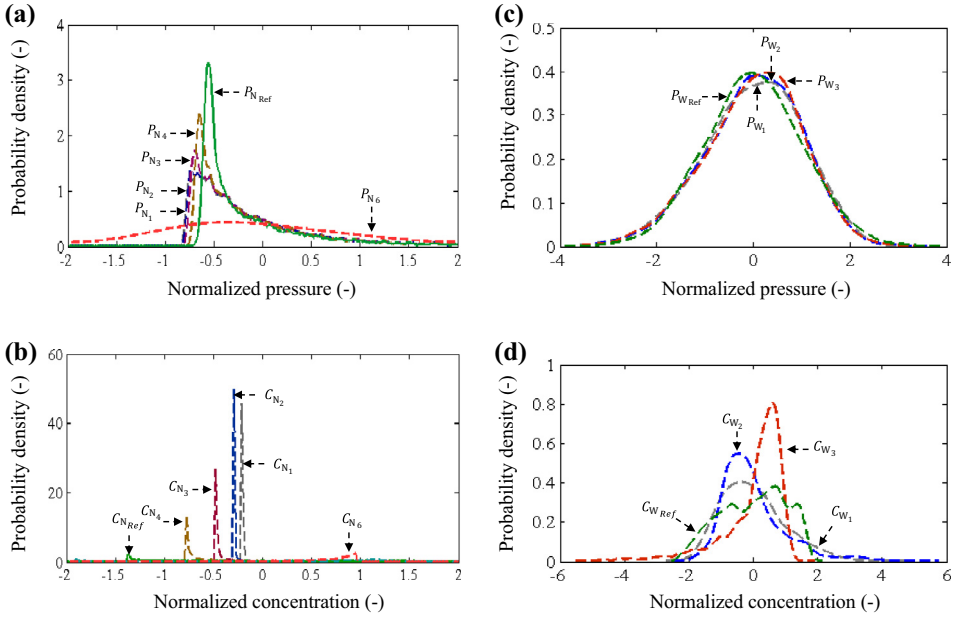

Fig. 10. Normalized PDFs of (a) pressure and (b) concentration solutions in the six monitoring points of the Henry problem, and (c) pressure and (d) concentration solutions in the four monitoring wells of the small island problem.

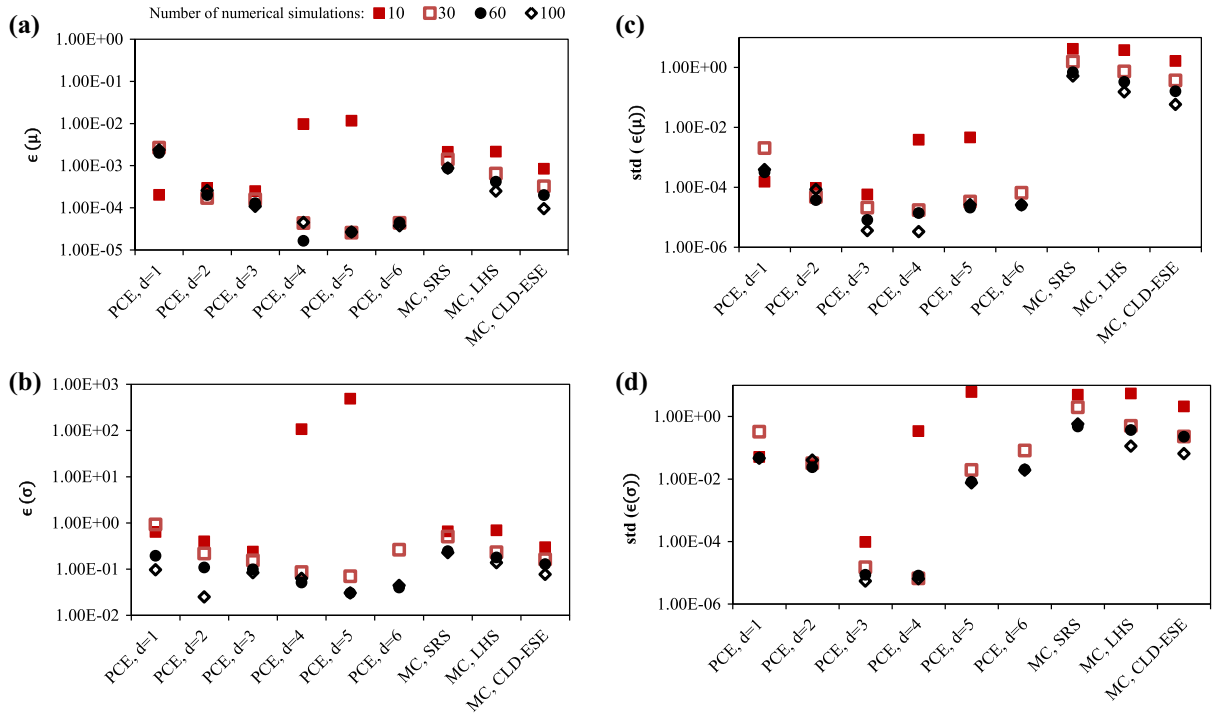

Fig. 11. Average normalized deviations from the reference solutions of (a) mean pressure, (b) standard deviation of pressure, and normalized internal deviations of (c) mean pressure, (d) standard deviation of pressure for various UP methods in the Henry problem.

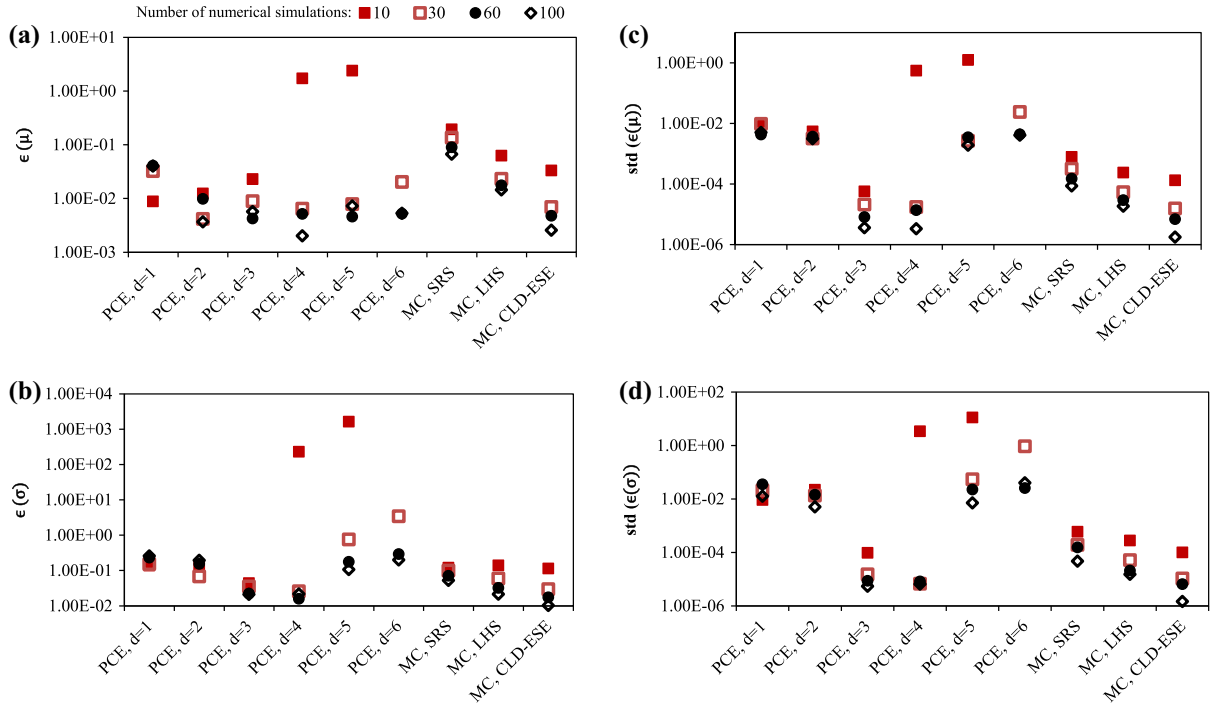

Fig. 12. Average normalized deviations from the reference solutions of (a) mean concentration, (b) standard deviation of concentration, and normalized internal deviations of (c) mean concentration, (d) standard deviation of concentration for various UP methods in the Henry problem.

This is also true for the normalized internal standard deviations of μ and σ estimates. Hence, the system of equations for the estimation of PCE coefficients must be over-determined. This conclusion is in accordance with previous studies such as Karagiannis and Lin (2014) and Hosder et al. (2007) which recommends that the number of regression points be at least two times the number of PCE coefficients.

(2) After the system of equations becomes over-determined, further increasing the number of regression points mostly results in lower accuracy and robustness for PCE of $d = 1$. For PCE of higher orders, the accuracy and robustness generally (but not necessarily) improves by increasing the number of regression points. Increasing the number of regression points not only requires further simulations using

Table 8
μ and σ and their confidence intervals estimates in the Henry problem test case: MCSs vs. PCEs.

QoI	UP method	μ estimate	96% confidence interval for μ	σ estimate	96% confidence interval for σ
$P_{N_{Ref}}$ (kg/m s^2)	MC (N_{MC} = 10,000)	6053.483	6053.472–6053.495	49.308	49.274–49.343
	PCE (d = 4, q = 30)	6053.513	6053.391–6053.635	48.796	48.479–49.112
$C_{N_{Ref}}$ (%salinity)	MC (N_{MC} = 10,000)	35.107	35.104–35.110	27.944	27.943–27.946
	PCE (d = 4, q = 30)	35.183	34.941–35.424	27.676	27.584–27.768

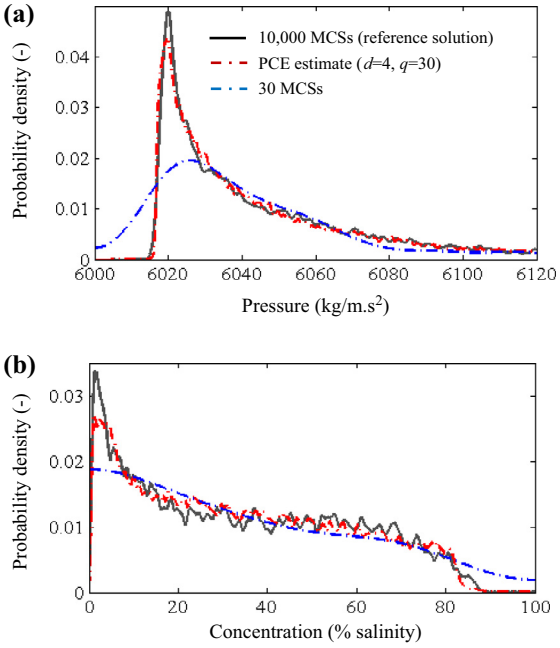

Fig. 13. PDFs of the (a) pressure and (b) concentration solutions in the monitoring point N_{Ref} of the Henry problem: Monte Carlo vs. non-intrusive PCE estimates.

the original numerical model, but also exponentially increases the computational time for the estimation of the PCE coefficients. The number of regression points should therefore be selected to reflect accuracy needs and computational constraints.

(3) The relative deviations from the reference solutions and the internal standard deviations generally decrease with increasing PCE order from d = 1 to d = 4 for both pressure and concentration. Higher orders of PCEs (d = 6 for pressure solutions and d = 5 and 6 for concentration solutions) illustrate higher relative deviations. Hence, as stated by some previous studies such as Eldred et al. (2008), higher order expansions can in fact be less accurate and robust than the lower order expansions. In theory, higher degree expansions should capture more model nonlinearities compared to smaller degree expansions and therefore be more accurate. However, two factors prohibit this: first, constructing an accurate higher order expansion may require much more data points than a lower order expansion. When constructed using the same given set of data points, there will always be an expansion order for which the associated expansion performs worse than lower order expansions. Second, the dimension of the optimization problem (or the associated system of non-linear equations depending on how Eq. (10) is solved in the regression method) increases with increase in the order of the PCE, leading to a potential increase in estimation errors.

(4) The relative deviations for pressure solutions (most notably for μ estimates) are in most cases smaller than concentration solutions. This could reflect the smoother nature of pressure PDFs.

(5) Polynomial orders equal to 3 or 4 permit to finely obtain the first two moments of pressure and concentration solutions in the Henry problem. Table 8 compares the μ and σ estimates of pressure and concentration as well as their confidence intervals for the reference solutions and the PCEs (d = 4 and q = 30). The table demonstrates that the outcomes of PCEs are very much identical to those of the reference solutions. Note that previous studies such as Laloy et al. (2013) have shown that PCEs of order one can accurately represent the output statistics of pressure solutions in conventional groundwater flow problems with high number of uncertain inputs. However as illustrated by this study, the highly non-linear nature of the input–output relationships in density-dependent SWI simulation requires higher order PCEs for accurate estimation of the output statistics of pressure solutions.

(6) In terms of accuracy and robustness, the non-intrusive PCEs of order 3 and 4 generally outperform MCSs based on all three sampling strategies. However, the difference between the accuracy and robustness of PCEs and MCSs decreases with increasing number of deterministic simulations. For 100 deterministic simulations, the accuracy of non-intrusive PCEs and MCSs based on the CLD–ESE sampling strategy are nearly identical. According to Figs. 11 and 12, the efficiency improves by 79–98% with the use of non-intrusive PCEs (with d = 4) compared with MCSs based on SRS, by 49–95% compared with MCSs based on LHS, and −6% to 89% compared with MCSs based on CLD–ESE (the minus sign illustrates decreased efficiency).

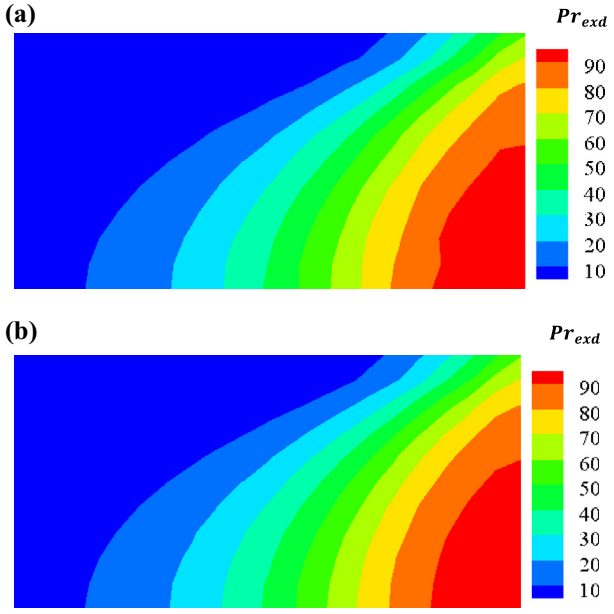

Fig. 14. SWI hazard map derived from MCSs based on (a) non-intrusive PCEs, (b) the numerical model in the Henry problem test case.

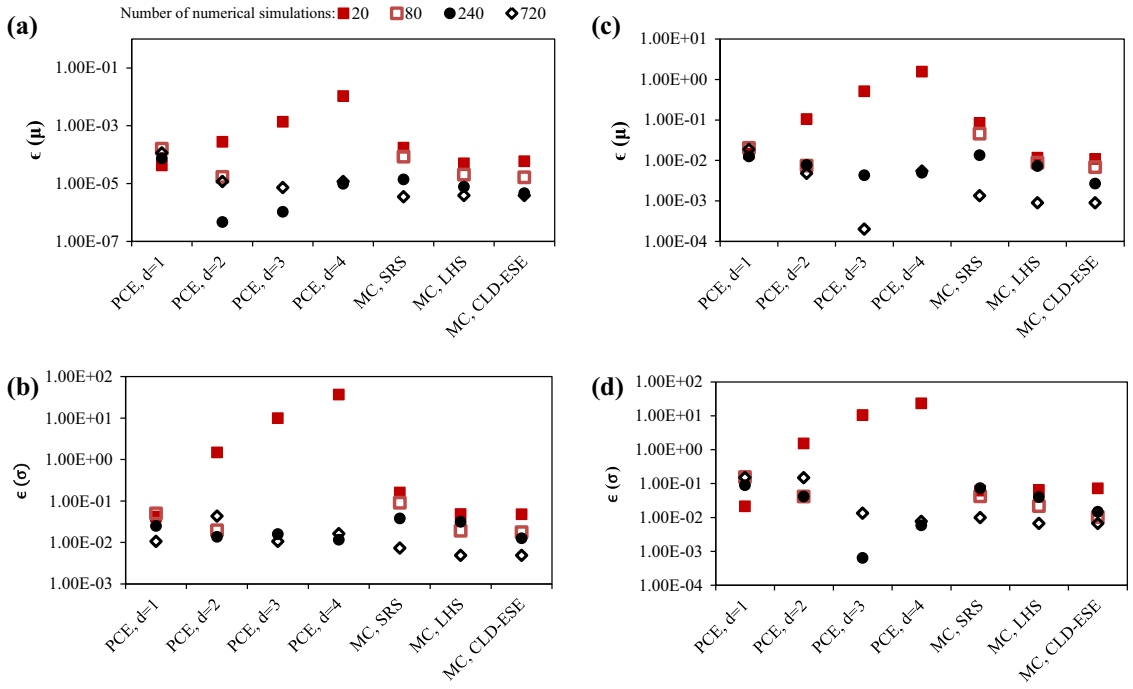

Fig. 15. Average normalized deviations from the reference solutions of (a) mean pressure, (b) standard deviation of pressure, (c) mean concentration, and (d) standard deviation of concentration for various UP methods in the small island problem.

Hence, the type of sampling strategy used in MCSs significantly affects the degree to which PCEs improve efficiency as compared to MCSs. Fig. 13a and b compares the reference PDFs of the pressure and concentration solutions with PDFs derived from non-intrusive PCEs (with $d = 4$ and $q = 30$) and also the data set used for the construction of the PCE surrogates which is based on 30 MCSs with CLD–ESE sampling. All three PDFs are built using KDE. As illustrated, the PDFs resulting from the PCEs compare well with the reference solutions. Interestingly, the PCE density functions comply better with the reference solutions compared with PDFs of the data set used for their construction.

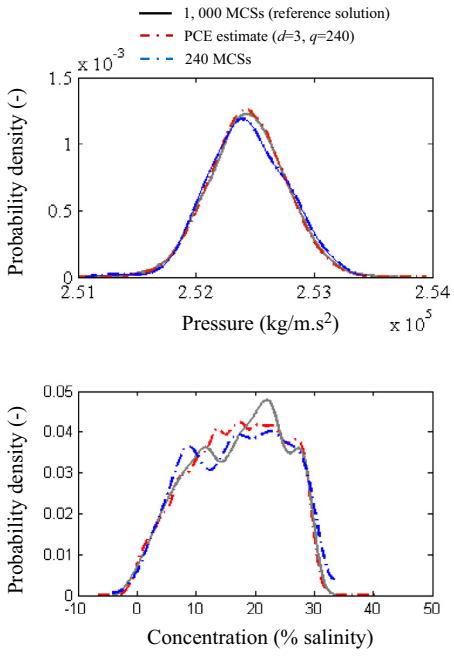

Fig. 16. PDFs of the (a) pressure and (b) concentration solutions in the monitoring well W_{Ref} of the small island problem: Monte Carlo vs. non-intrusive PCE estimates.

To make sure that non-intrusive PCEs can provide reliable solutions for the entire domain of the test case problem, PCEs (with $d = 4$ and $q = 30$) have been constructed for all 231 nodes of the Henry problem numerical grid. These PCEs were then used to generate the SWI hazard map of the Henry problem. We define the SWI hazard as the probability of exceedance of salinity concentrations from 2000 mg/l. We denote this probability by Pr_{exd}, the total number of MCSs (based either on the numerical model or the non-intrusive PCE surrogate) by N_{MC}, and the number of MCSs resulting in salinity concentrations above 2000 mg/l in a specific node as $N_{MC,(C>2000)}$. Pr_{exd} (expressed in percentage) can be estimated using the following equation:

$$Pr_{exd} = \frac{N_{MC,(C>2000)}}{N_{MC}} \times 100 \tag{18}$$

Fig. 14a illustrates the SWI hazard map derived from 10,000 MCSs based on the non-intrusive PCE surrogate model. Fig. 14b demonstrates the same SWI hazard map obtained from 10,000 MCSs based on the original numerical model. These two figures are very much identical, illustrating the ability of PCEs to map the uncertainty of the entire domain without introducing any bias.

We now move to the second test case, namely the small island problem. Sets of 20, 80, 240 and 720 regression points are first generated using the CLD–ESE sampling strategy. These regression points are then used for constructing PCEs of order 1–4. The numbers of PCE coefficients for the six variable polynomial orders $d = 1$, 2, 3 and 4 are 7, 28, 84 and 210 respectively. Fig. 15 compares the accuracy of μ and σ estimates of pressure and concentration solutions for different orders of PCEs and different numbers of regression points in the small island problem. The accuracy of the corresponding estimates obtained from MCSs based on the original numerical model with SRS, LHS and CLD–ESE sampling strategies are also shown in Fig. 15. As described in the beginning of this section, $\epsilon(\mu)$ and $\epsilon(\sigma)$ estimates are based on four repetitions of each UP method in the small island problem. The normalized internal

Table 9
μ and σ and their confidence intervals estimates in the small island problem: MCSs vs. PCEs.

QoI	UP method	μ estimate	96% confidence interval for μ	σ estimate	96% confidence interval for σ
$P_{W_{Ref}}$ (kg/m s^2)	MC (N_{MC} = 1000)	252435.11	252,434.45–252,435.78	323.34	322.72–323.97
	PCE (d = 3, q = 240)	252435.00	252434.72–252435.27	320.97	320.87–321.07
$C_{W_{Ref}}$ (%salinity)	MC (N_{MC} = 1000)	16.85	16.83–16.86	7.87	7.86–7.88
	PCE (d = 3, q = 240)	16.75	16.74–16.76	7.89	7.88–7.90

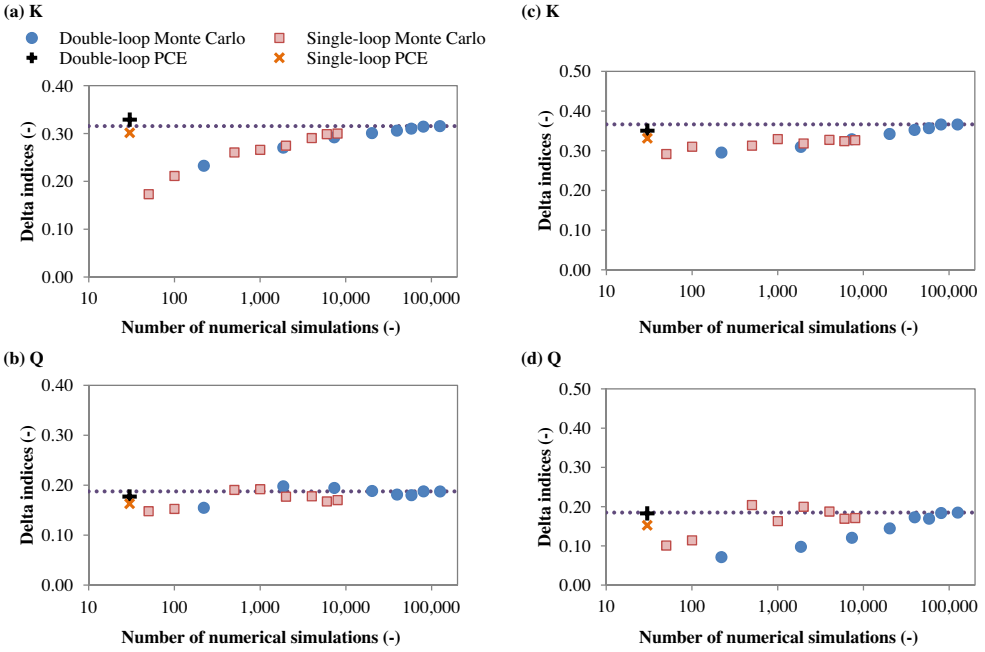

Fig. 17. Comparison of various method for the estimation of the delta indices for (a) K and (b) Q with respect to pressure solutions, and (c) K and (d) Q with respect to concentration solutions in the Henry problem test case.

deviations are not calculated for the small island problem due to the relatively small number of repetitions. The conclusions drawn for the second test case are very much similar to the Henry problem. We again see that polynomial orders equal to 3 or 4 can finely capture the first two moments of pressure and concentration solutions in the small island problem. This can be well understood from a comparison of reference PDFs with PDFs derived from non-intrusive PCEs (with d = 3 and q = 240), which is illustrated in Fig. 16. As demonstrated, the PDFs resulting from the PCEs compare well with the reference solutions. This conclusion can also be drawn from an analysis of Table 9 which compares the μ and σ estimates of pressure and concentration and their confidence intervals for the reference solutions and the PCEs (d = 3 and q = 240).

Similar to the Henry problem, the accuracy of pressure estimates obtained through the use of PCEs are in most cases higher than concentration estimates due to the smoother nature of pressure PDFs. Moreover, despite the fact that the small island problem is three dimensional, conceptually more complex and has more uncertain inputs than the Henry problem; the pressure and concentration PDFs in the small island problem are relatively smoother than the Henry problem (refer to Fig. 10), thus allowing for a more accurate estimation of the first two moments using PCE methodology.

We can now answer the first two objective questions of this study:

- Can non-intrusive PCEs provide a reliable estimate of model uncertainties in SWI numerical modeling studies?

Based on the two test cases the answer is "yes", provided that sufficient number of regression points (that is if the regression method is employed) and a proper order of PCE is used. A similar conclusion has been reported in previous studies involving groundwater flow in saturated and unsaturated porous media (e.g., Li and Zhang, 2007; Sochala and Le Maître, 2013; Laloy et al., 2013). However as illustrated, due to the highly non-linear and non-smooth input–output relationship in SWI models, higher order polynomials are needed for accurate estimation of model output statistics as compared to conventional groundwater flow problems.

As demonstrated by this study and a number of previous studies such as Xiu and Karniadakis (2003a), the accuracy of PCEs can in fact be higher than MCSs based on the original numerical model if the number of uncertain inputs is small. Our study shows that this conclusion is most notably true when only a very limited number of model simulations are affordable. As the number of simulations increases, MCSs based on non-intrusive PCEs and the numerical model illustrate nearly identical accuracies.

- To what extent do non-intrusive PCEs decrease the computational cost of UP analysis in SWI numerical models compared to MCSs based on the original models?

PCE simulations are almost computationally costless compared to numerical models of SWI, and the computational cost associated with the use of PCEs in UP analysis is due to the generation of regression points. For example our studies show that a MCS that

Table 10
Comparison of the delta indices estimates and their respective standard deviations for the Henry problem.

QoI	Input parameter	Double-loop MCSs based on the original model		Double-loop MCSs based on PCE ($d = 4$, $q = 30$)		Single-loop MCSs based on the original model		Single-loop MCSs based on PCE ($d = 4$, $q = 30$)	
		Estimate	σ	Estimate	σ	Estimate	σ	Estimate	σ
$P_{N_{Ref}}$	k	0.3145	0.0012	0.3291	0.0056	0.3000	0.0076	0.3016	0.0051
	Q	0.1877	0.0072	0.1772	0.0062	0.1702	0.0029	0.1630	0.0017
$C_{N_{Ref}}$	k	0.3663	0.0073	0.3507	0.0086	0.3267	0.0048	0.3315	0.0019
	Q	0.1837	0.0026	0.1828	0.0093	0.1709	0.0094	0.1524	0.0025

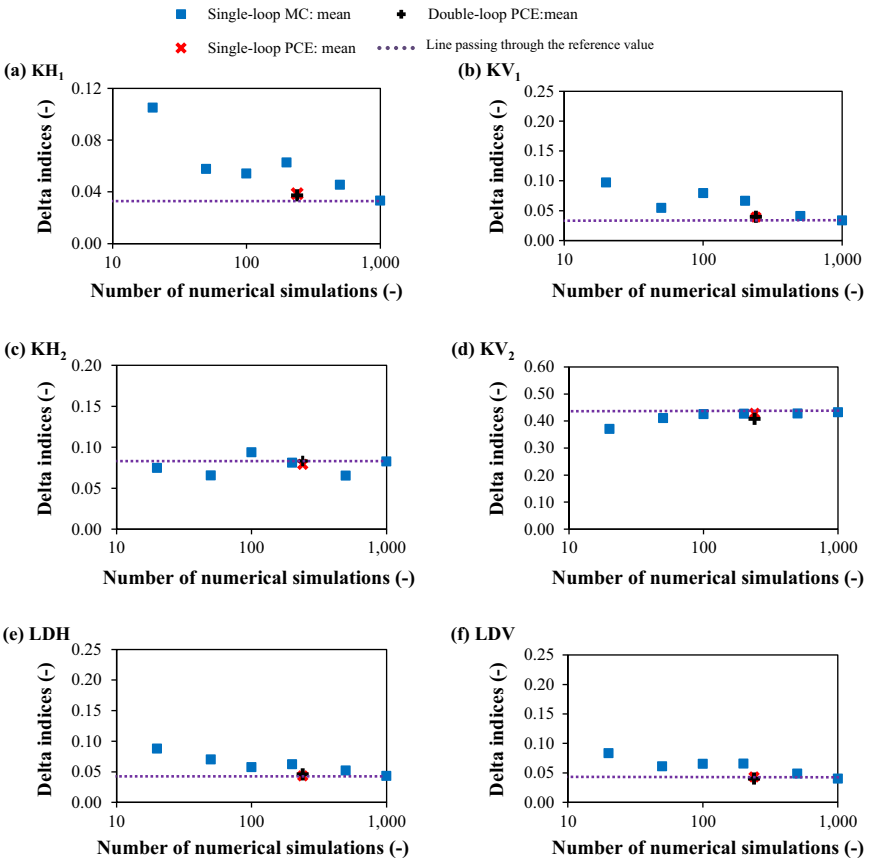

Fig. 18. Comparison of various method for the estimation of the delta indices for (a) KH_1, (b) KV_1, (c) KH_2, (d) KV_2, (e) LDH and (f) LDV with respect to pressure in the small island problem.

involves 10,000 simulations by a non-intrusive PCE surrogate model (with $d = 4$ and $q = 30$), takes less than 15 s on a 2.17 GHz dual-core personal computer (PC). The computational time required by MCSs (with a comparable level of accuracy) relying on the original numerical model may take several days for a complex three dimensional model of SWI such as the small island problem presented in this paper. Hence, PCEs can decrease the computational time of UP by several orders of magnitude depending on the computational cost of a single deterministic simulation by the groundwater model. This conclusion is similar to those reported by previous studies involving groundwater modeling applications (e.g., Li and Zhang, 2007; Sochala and Le Maître, 2013). Moreover, as illustrated by this study, the type of sampling strategy used in MCSs significantly affects the degree to which PCEs improve efficiency as compared to MCSs.

4.3. Moment Independent SA

In this subsection, we apply PCEs to accelerate the estimation of delta indices by the single-loop and double-loop Monte Carlo methods (Wei et al., 2013). Reference solutions based on the numerical models are initially produced in order to provide benchmarks for the assessment of the accuracy of PCE based solutions. In the Henry problem, we apply the double-loop Monte Carlo methods to generate this reference solution. A convergence analysis is done in which the values of NL_1 and NL_2 in the double-loop Monte Carlo method are gradually increased until the estimated delta indices stabilize. The sample points are selected by the CLD–ESE sampling strategy and the estimation of delta indices has been repeated 8 times for each value of NL_1 and NL_2. The resultant mean values of the delta indices for the 8 repetitions are illustrated in Fig. 17. The figure shows that for $NL_1 = NL_2 = 250$ the resulting delta indices have converged to the correct output. This choice of NL_1 and NL_2 leads to a total of 125,250 deterministic simulations with a computational time of around 70 h on a 2.17 GHz dual-core PC. Note that this rather long computational time is associated with a simple two-dimensional test case problem, and the computational time of complex three dimensional numerical models may be several orders of magnitude higher, which becomes prohibitive unless massive parallel computing facilities are available. In a

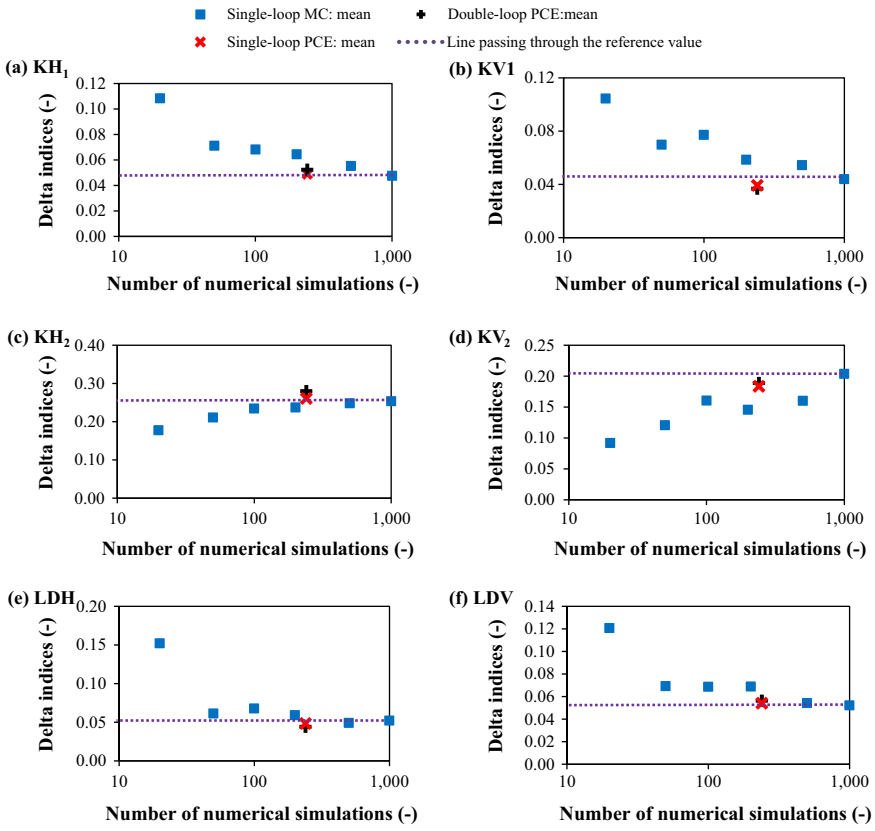

Fig. 19. Comparison of various method for the estimation of the delta indices for (a) KH_1, (b) KV_1, (c) KH_2, (d) KV_2, (e) *LDH* and (f) *LDV* with respect to concentration in the small island problem.

similar convergence analysis of delta indices estimates by Wei et al. (2013), a value of 1000 was chosen for NL_1 andNL_2. However they used the less efficient SRS method for the generation of sample points.

We have also estimated the delta indices for the Henry problem using the single-loop Monte Carlo method for $NL = 50–8000$. It is observed in Fig. 17 that the single-loop method arrives at nearly the same delta indices (the difference being less than 10%), with less than 0.064 the number of simulations required by the double-loop method. This comparison confirms the conclusions put forward by Wei et al. (2013) regarding the superior efficiency of the single-loop Monte Carlo method.

The numerical model is subsequently replaced by the non-intrusive PCE surrogate models (with $d = 4$ and $q = 30$) in both the double-loop and single-loop Monte Carlo algorithms. In this case, only 30 deterministic simulations are needed for the estimation of the PCE coefficients and the subsequent simulations required for the calculation of the delta indices are done by use of the computationally cheap PCEs. For example, the computational time of the double-loop Monte Carlo method with non-intrusive PCEs as the surrogate model and $NL_1 = NL_2 = 1000$ decreases to around 80 min on a 2.17 GHz dual-core PC. The delta indices estimates obtained through the use of PCEs are demonstrated in Fig. 17. Moreover, Table 10 compares the delta indices estimates along with their respective standard deviations across 8 replicates for the double-loop and single-loop Monte Carlo methods based on the original numerical model and the non-intrusive PCEs. Fig. 17 and Table 10 illustrate that: (1) the difference amongst the mean estimates of delta indices for the two strategies (i.e. using the original numerical model vs. the PCE surrogate) in the double-loop Monte Carlo method is small (less than 6%) and they lead to the same importance ranking ($k > Q$) for both pressure and concentration solutions. (2) The average deviation of delta indices obtained through the use of PCEs with the single-loop Monte Carlo method from the respective solution obtained from the double-loop and single-loop Monte Carlo methods with the original numerical model is 11% and 4% respectively, with the same importance ranking. (3) As illustrated in Table 10, using the original numerical model or the PCE surrogate model with both the double-loop and single-loop Monte Carlo methods lead to small standard deviations for the delta indices and the standard deviations are of the same order of magnitude. (4) When using PCE surrogates, the difference between the robustness and computational efficiency of the double-loop and single-loop Monte Carlo methods are much less significant. The reason is that the computational time of PCE simulations are so small that even the computationally less efficient double-loop Monte Carlo method wraps up in a small time frame.

For the small island problem, the reference solutions for the delta indices are obtained through the use of the original numerical model with the single-loop Monte Carlo method. The use of double-loop method is not computationally feasible for this test case. A convergence analysis is done in which the values of *NL* is gradually increased until the estimated delta indices stabilize. Similar to the Henry problem, the sample points are selected by the CLD–ESE sampling strategy and the estimation of delta indices has been repeated 8 times for each value of *NL*. The resultant mean values of the delta indices for the 8 repetitions are illustrated in Figs. 18 and 19. We take the delta indices obtained from $NL = 1000$ as the reference solution. The delta indices are subsequently estimated by using PCEs (with $d = 3$ and $q = 240$) with both the double-loop and single-loop Monte Carlo methods. We observe that the average deviations of the resulting delta indices from the respective reference solutions are around 7% for both pressure and for concentration solutions (see Table 11).

Table 11
Comparison of the delta indices estimates and their respective standard deviations for the small island problem.

QoI	Input parameter	Single-loop MCSs based on the original model		Double-loop MCSs based on PCE ($d = 3$, $q = 240$)		Single-loop MCSs based on PCE ($d = 3$, $q = 240$)	
		Estimate	σ	Estimate	σ	Estimate	σ
$P_{W_{Ref}}$	KH_1	0.0333	0.0035	0.0372	0.0013	0.0385	0.0007
	KV_1	0.0337	0.0040	0.0394	0.0025	0.0393	0.0017
	KH_2	0.0827	0.0036	0.0831	0.0007	0.0791	0.0015
	KV_2	0.4321	0.0036	0.4073	0.0039	0.4284	0.0028
	LDH	0.0433	0.0026	0.0456	0.0023	0.0432	0.0008
	LDV	0.0402	0.0093	0.0403	0.0038	0.0429	0.0017
$C_{W_{Ref}}$	KH_1	0.0475	0.0184	0.0462	0.0037	0.0487	0.0076
	KV_1	0.0439	0.0036	0.0354	0.0085	0.0392	0.0032
	KH_2	0.2533	0.0138	0.2689	0.0102	0.2607	0.0083
	KV_2	0.2037	0.0067	0.1922	0.0093	0.1832	0.0148
	LDH	0.0521	0.0199	0.0447	0.0064	0.0483	0.0046
	LDV	0.0521	0.0094	0.0565	0.0119	0.0540	0.0088

We now address the third objective question of this study: can non-intrusive PCEs provide a computationally efficient estimation of moment-independent sensitivity indices without significant loss of accuracy? With respect to the outcome of this study, we can conclude that replacing the numerical model with the non-intrusive PCE surrogate model in the estimation of the delta indices does not affect the resulting importance ranking and only introduces a small deviation in the estimated indices. However, estimating the delta indices through the use of PCEs is several orders of magnitude (depending on the computational cost of the deterministic model) faster than the case in which the numerical model is employed. Furthermore, using the double-loop Monte Carlo method for computational demanding models can be made possible with the use of PCEs.

5. Conclusions

Real world models of SWI are characterized by high computational demands. This, in addition to the inherently complex, nonlinear and non-monotonic input/output relationship of these models, creates computational difficulties in UP analysis. These difficulties arise due to the need for extensive repeated simulations by the numerical model in order to adequately capture the underlying statistics that describe the uncertainty in model outputs. Moreover, despite the obvious advantages of moment-independent global methods, these methods have rarely been employed for the SA of SWI and other complex groundwater models. The reason is that moment-independent global SA methods involve repeated UP analysis which further becomes computationally problematic unless massive parallel computing facilities are available. This study proposes the use of non-intrusive PCEs as a means to significantly accelerate UP analysis in SWI numerical modeling studies. The study shows that non-intrusive PCEs can provide a reliable and yet computationally efficient surrogate of the original numerical model which can be used in UP analysis, SA, simulation–optimization schemes and stochastic inverse modeling applications. The study also illustrates that for the considered two and six dimensional UP problems, PCEs offer a more accurate estimation of the statistics describing the uncertainty in model outputs compared to Monte Carlo simulations based on the original numerical model. We also proposed the use of non-intrusive PCEs for the generation of SWI hazard maps, based on the ability of PCEs to map the uncertainty of the entire domain without introducing any bias. The proposed methodology can extend the practical applications of UP analysis in coastal aquifer management studies.

This study further illustrated that the use of non-intrusive PCEs in the estimation of the moment-independent sensitivity indices can decrease the computational time by several orders of magnitude without causing significant loss of accuracy. The proposed methodology is based on replacing the numerical model with the non-intrusive PCE surrogate model in the single-loop and double-loop Monte Carlo methods for the estimation of the delta sensitivity indices. This methodology can make the use of double-loop Monte Carlo method for computational demanding SWI models possible.

The two case studies in this paper were chosen to be rather simple SWI models with affordable simulation time in order to provide the tools for the more than 2,000,000 and 28,000 numerical simulations (for the first and second test case respectively) of the reference solutions used in the evaluation of the results obtained by PCEs. However, the true benefits of PCEs are obtained when real world SWI models with extremely high computationally demands are evaluated.

Acknowledgments

The last two authors appreciate the funding of National Centre for Groundwater Research and Training, a collaborative initiative of the Australian Research Council and the National Water Commission, Australia. The authors appreciate the comments of anonymous reviewers, associate Editor, and editor in chief Geoff Syme on this manuscript, which helped improve the final manuscript.

References

Archer, G.E.B., Saltelli, A., Sobol, I.M., 1997. Sensitivity measures, ANOVA like techniques and the use of bootstrap. J. Stat. Comput. Simul. 58, 99–120. http://dx.doi.org/10.1080/00949659708811825.

Askey, R., Wilson, J., 1985. Some Basic Hypergeometric Polynomials that Generalize Jacobi Polynomials, Memoirs of the American Mathematical Society. AMS, Providence, RI, pp. 319.

Ataie-Ashtiani, B., Rajabi, M.M., Ketabchi, H., 2013. Inverse modeling for freshwater lens in small islands: Kish Island, Persian Gulf. Hydrol. Process. 27 (19), 2759–2773. http://dx.doi.org/10.1002/hyp.9411.

Ataie-Ashtiani, B., Ketabchi, H., Rajabi, M.M., 2014. Optimal management of freshwater lens in a small island using surrogate models and evolutionary algorithms. J. Hydrol. Eng. 19 (2), 339–354. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000809.

Ayres, D., Park, S., Eaton, M.D., 2014. Propagation of input model uncertainties with different marginal distributions using a hybrid polynomial chaos expansion. Ann. Nucl. Energy 66, 1–4. http://dx.doi.org/10.1016/j.anucene.2013.11.041.

Ballio, F., Guadagnini, A., 2004. Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology. Water Resour. Res. 40, W04603. http://dx.doi.org/10.1029/2003WR002876.

Bear, J. (Ed.), 1999. Seawater Intrusion in Coastal Aquifers: Concepts, Methods and Practices. Springer.

Berveiller, M., Sudret, B., Lemaire, M., 2006. Stochastic finite elements: a non-intrusive approach by regression. Eur. J. Comput. Mech. 15 (1–3), 81–92. http://dx.doi.org/10.3166/remn.15.81-92.

Beven, K.J., Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298. http://dx.doi.org/10.1002/hyp.3360060305.

Blatman, G., Sudret, B., 2010. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. Reliab. Eng. Syst. Saf. 95 (11), 1216–1229. http://dx.doi.org/10.1016/j.ress.2010.06.015.

Borgonovo, E., 2007. A new uncertainty importance measure. Reliab. Eng. Syst. Saf. 92, 771–784. http://dx.doi.org/10.1016/j.ress.2006.04.015.

Borgonovo, E., Castaings, W., Tarantola, S., 2012. Model emulation and moment-independent sensitivity analysis: an application to environment modeling. Environ. Modell. Softw. 34, 105–115. http://dx.doi.org/10.1016/j.envsoft.2011.06.006.

Botev, Z.I., 2011. Kernel Density Estimation Using Matlab. <http://www.mathworks.us/matlabcentral/fileexchange/authors/27236>.

Brown, J.D., Heuvelink, G.B.M., 2006. 79° Assessing uncertainty propagation through physically based models of soil water flow and solute transport. In: Anderson, MG. (Ed.), Encyclopedia of Hydrological Sciences. John Wiley & Sons, pp. 1181–1195. http://dx.doi.org/10.1002/0470848944.hsa081.

Cacuci, D.G., 2003. Sensitivity and Uncertainty Analysis: Theory, vol. 1. Chapman & Hall/CRC, Boca Raton.

Cameron, R., Martin, W., 1947. The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals. Ann. Math. 48, 385–392.

Carrera, J., Hidalgo, J.J., Slooten, L.J., Vázquez-Suñé, E., 2010. Computational and conceptual issues in the calibration of seawater intrusion models. Hydrogeol. J. 18, 131–145.

Chaudhuri, A., Sekhar, M., 2005. Analytical solutions for macrodispersion in a 3D heterogeneous porous medium with random hydraulic conductivity and dispersivity. Transp. Porous Media 58 (3), 217–241. http://dx.doi.org/10.1007/s11242-004-6300-8.

Chen, Q.Y., Gottlieb, D., Hesthaven, J.S., 2005. Uncertainty analysis for the steady-state flow in a dual nozzle. J. Comput. Phys. 204 (1), 378–398.

Chun, M.H., Han, S.J., Tak, N.I., 2000. An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. Reliab. Eng. Syst. Saf. 70 (3), 313–321. http://dx.doi.org/10.1016/S0951-8320(00)00068-5.

Ciriello, V., Di Federico, V., Riva, M., Cadini, F., De Sanctis, J., Zio, E., Guadagnini, A., 2013. Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer. Stoch. Environ. Res. Risk Assess. 27, 945–954. http://dx.doi.org/10.1007/s00477-012-0616-7.

Cukier, H., Levine, R.I., Shuler, K., 1978. Nonlinear sensitivity analysis of multiparameter model systems. J. Comput. Phys. 26, 1–42. http://dx.doi.org/10.1016/0021-9991(78)90097-9.

Cullen, A.C., Frey, H.C., 1999. Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Springer.

Darvini, G., Spendolini, L., Salandin, P., 2002. Saltwater intrusion for finite Peclet numbers in random permeability aquifers. Dev. Water Sci. 47, 523–530. http://dx.doi.org/10.1016/S0167-5648(02)80104-6.

Dimov, IT., 2008. Monte Carlo Methods for Applied Scientists. World Scientific Publishing Co., Pte. Ltd, UK.

Doherty, J., 2005. PEST: Model Independent Parameter Estimation, User Manual, fifth ed. Watermark Numerical Computing.

Dubois, D., Prade, H., Sandri, S., 1993. On possibility/probability transformations. In: Lowen, R., Roubens, M. (Eds.), Fuzzy Logic. Kluwer Academic Publishers, Netherlands.

Dubois, D., Foulloy, L., Mauris, G., Prade, H., 2004. Probability–possibility transformations, triangular fuzzy sets and probabilistic inequalities. Reliable Comput. 10, 273–297. http://dx.doi.org/10.1023/B:REOM.0000032115.22510.b5.

Eldred, M.S., Webster, C.G., Constantine, P.G., 2008. Evaluation of non-intrusive approaches for Wiener-Askey generalized polynomial chaos. Proceedings of the 10th AIAA Non-Deterministic Approaches Conference, Number AIAA-2008-1892, vol. 117. Schaumburg, Illinois, pp. 189.

Fajraoui, N., Ramasomanana, F., Younes, A., Mara, T.A., Ackerer, P., Guadagnini, A., 2011. Use of global sensitivity analysis and polynomial chaos expansion for interpretation of nonreactive transport experiments in laboratory-scale porous media. Water Resour. Res. 47, W02521. http://dx.doi.org/10.1029/2010WR009639.

Falkland, A., 1991. Hydrology and Water Resources of Small Islands: A Practical Guide. A Contribution to the International Hydrological Programme. United Nations Educational, Scientific and Cultural Organization, Paris, France.

Garcia-Cabrejo, O., Valocchi, A., 2014. Global sensitivity analysis for multivariate output using polynomial chaos expansion. Reliab. Eng. Syst. Saf. 126, 25–36. http://dx.doi.org/10.1016/j.ress.2014.01.005.

Ghanem, R., 1998. Probabilistic characterization of transport in heterogeneous media. Comput. Methods Appl. Mech. Eng. 158 (3), 199–220. http://dx.doi.org/10.1016/S0045-7825(97)00250-8.

Ghiocel, D., Ghanem, R., 2002. Stochastic finite element analysis of seismic soil structure interaction. J. Eng. Mech. 128, 66–77. http://dx.doi.org/10.1061/(ASCE)0733-9399(2002)128:1(66).

Haro Sandoval, E., Anstett-Collin, F., Basset, M., 2012. Sensitivity study of dynamic systems using polynomial chaos. Reliab. Eng. Syst. Saf. 104, 15–26. http://dx.doi.org/10.1016/j.ress.2012.04.001.

Hassan, A.E., Bekhit, H.M., Chapman, J.B., 2009. Using Markov Chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model. Environ. Modell. Softw. 24 (6), 749–763. http://dx.doi.org/10.1016/j.envsoft.2008.11.002.

Helton, J.C., Davis, F.J., 2002. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. Sandia National Laboratories, New Mexico, USA, SAND2001-0417.

Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Saf. 81, 23–69. http://dx.doi.org/10.1016/S0951-8320(03)00058-9.

Henry, H.R., 1964. Effects of Dispersion on Salt Encroachment in Coastal Aquifers. USGS Water-Supply Paper 1613-C. Sea Water Coastal Aquifers, C71-84.

Herckenrath, D., Langevin, C.D., Doherty, J., 2011. Predictive uncertainty analysis of a saltwater intrusion model using null-space Monte Carlo. Water Resour. Res. 47 (5), W05504. http://dx.doi.org/10.1029/2010WR009342.

Homma, T., Saltelli, A., 1995. Use of Sobol' quasirandom sequence generator for integration of modified uncertainty importance measure. J. Nucl. Sci. Technol. 32 (11), 1164–1173. http://dx.doi.org/10.1080/18811248.1995.9731832.

Hooke, R., Jeeves, T.A., 1961. Direct search solution of numerical and statistical problems. J. ACM (JACM) 8 (2), 212–229.

Hosder, S., Walters, R.W., Balch, M., 2007. Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables. In: Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, No. AIAA-2007-1939, Honolulu, HI.

Iooss, B., Popelin, A.L., Blatman, G., Ciric, C., Gamboa, F., Lacaze, S., Lamboni, M., 2012. Some new insights in derivative-based global sensitivity measures. In: Proceedings of PSAM, vol. 11, pp. 1094–1104.

Janssen, H., 2013. Monte-Carlo based uncertainty analysis: sampling efficiency and sampling convergence. Reliab. Eng. Syst. Saf. 109, 123–132. http://dx.doi.org/10.1016/j.ress.2012.08.003.

Karagiannis, G., Lin, G., 2014. Selection of polynomial chaos bases via Bayesian model uncertainty methods with applications to sparse approximation of PDEs with stochastic input. J. Comput. Phys. 259, 114–134. http://dx.doi.org/10.1016/j.jcp.2013.11.016.

Kerrou, J., Renard, P., 2010. A numerical analysis of dimensionality and heterogeneity effects on advective dispersive seawater intrusion processes. Hydrogeol. J. 18 (1), 55–72. http://dx.doi.org/10.1007/s10040-009-0533-0.

Kerrou, J., Renard, P., Lecca, G., Tarhouni, J., 2010. Grid-enabled Monte Carlo analysis of the impacts of uncertain discharge rates on seawater intrusion in the Korba aquifer (Tunisia). Hydrol. Sci. J. – J. Sci. Hydrol. 55 (8), 1325–1336. http://dx.doi.org/10.1080/02626667.2010.519706.

Ketabchi, H., Ataie-Ashtiani, B., 2014. Evolutionary algorithms for the optimal management of coastal groundwater: A comparative study toward future challenges. J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2014.11.043.

Ketabchi, H., Mahmoodzadeh, D., Ataie-Ashtiani, B., Werner, A.D., Simmons, C.T., 2014. Sea-level rise impact on fresh groundwater lenses in two-layer small islands. Hydrol. Process. 28, 5938–5953. http://dx.doi.org/10.1002/hyp.10059.

Knioa, O.M., Le Maître, O.P., 2006. Uncertainty propagation in CFD using polynomial chaos decomposition. Fluid Dyn. Res. 38, 616–640. http://dx.doi.org/10.1016/j.fluiddyn.2005.12.003.

Konda, U., Singh, T., Singla, P., Scott, P., 2010. Uncertainty propagation in puff-based dispersion models using polynomial chaos. Environ. Modell. Softw. 25, 1608–1618. http://dx.doi.org/10.1016/j.envsoft.2010.04.005.

Laloy, E., Rogiers, B., Vrugt, J.A., Mallants, D., Jacques, D., 2013. Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. Water Resour. Res. 49, 2664–2682. http://dx.doi.org/10.1002/wrcr.20226.

Le Maitre, O., Reagan, M., Najm, H., Ghanem, R., Knio, O., 2002. A stochastic projection method for fluid flow – II. Random process. J. Comput. Phys. 181, 9–44. http://dx.doi.org/10.1006/jcph.2002.7104.

Lecca, G., Cao, P., 2004. Stochastic modeling for seawater intrusion risk assessment in exploited coastal aquifers: the Oristano (Sardinia, Italy) case study. Dev. Water Sci. 55 (2), 1619–1628. http://dx.doi.org/10.1016/S0167-5648(04)80171-0.

Lecca, G., Cao, P., 2009. Using a Monte Carlo approach to evaluate seawater intrusion in the Oristano coastal aquifer: a case study from the AQUAGRID collaborative computing platform. Phys. Chem. Earth, Parts A/B/C 34 (10), 654–661. http://dx.doi.org/10.1016/j.pce.2009.03.002.

Lee, S.H., Chen, W., 2009. A comparative study of uncertainty propagation methods for black-box-type problems. Struct. Multidisc. Optim. 37, 239–253. http://dx.doi.org/10.1007/s00158-008-0234-7.

Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. 2, 164–168.

Li, H., Zhang, D., 2007. Probabilistic collocation method for flow in porous media: comparisons with other stochastic methods. Water Resour. Res. 43 (9), W09409. http://dx.doi.org/10.1029/2006WR005673.

Li, L., Tchelepi, H.A., Zhang, D., 2003. Perturbation-based moment equation approach for flow in heterogeneous porous media: applicability range and analysis of high-order terms. J. Comput. Phys. 188, 296–317. http://dx.doi.org/10.1016/S0021-9991(03)00186-4.

Li, W., Lu, Z., Zhang, D., 2009. Stochastic analysis of unsaturated flow with probabilistic collocation method. Water Resour. Res. 45, W08425. http://dx.doi.org/10.1029/2008WR007530.

Lilliefors, H., 1967. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. J. Am. Stat. Assoc. 62, 399–402.

Liu, Y., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. Water Resour. Res. 43, W07401. http://dx.doi.org/10.1029/2006WR005756.

Liu, Q., Homma, T., 2009. A new computational method of a moment-independent uncertainty importance measure. Reliab. Eng. Syst. Saf. 94, 1205–1211. http://dx.doi.org/10.1016/j.ress.2008.10.005.

Loeven, G.J.A., Witteveen, J.A.S., Bijl, H., 2007. Probabilistic collocation: an efficient non-intrusive approach for arbitrarily distributed parametric uncertainties. In: Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit, vol. 317, AIAA Paper.

Mahmoodzadeh, D., Ketabchi, H., Ataie-Ashtiani, B., Simmons, C.T., 2014. Conceptualization of a fresh groundwater lens influenced by climate change: a modeling study of an arid-region island in the Persian Gulf, Iran. J. Hydrol. 519, 399–413. http://dx.doi.org/10.1016/j.jhydrol.2014.07.010.

Malaguerra, F., Albrechtsen, H.J., Binning, P.J., 2013. Assessment of the contamination of drinking water supply wells by pesticides from surface water resources using a finite element reactive transport model and global sensitivity analysis techniques. J. Hydrol. 476, 321–331. http://dx.doi.org/10.1016/j.jhydrol.2012.11.010.

Meerschaert, M.M., Dogan, M., Dam, R.L., Hyndman, D.W., Benson, D.A., 2013. Hydraulic conductivity fields: Gaussian or not? Water Resour. Res. 49 (8), 4730–4737. http://dx.doi.org/10.1002/wrcr.20376.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33, 161–174. http://dx.doi.org/10.1080/00401706.1991.10484804.

Mulani, S.B., 2006. Uncertainty Quantification in Dynamic Problems with Large Uncertainties. Doctoral Dissertation. Virginia Polytechnic Institute and State University.

Oladyshkin, S., Nowak, W., 2012. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. Reliab. Eng. Syst. Saf. 106, 179–190. http://dx.doi.org/10.1016/j.ress.2012.05.002.

Oladyshkin, S., de Barros, F.P.J., Nowak, W., 2012. Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. Adv. Water Resour. 37, 10–22. http://dx.doi.org/10.1016/j.advwatres.2011.11.001.

Perez, R.A., 2008. Uncertainty Analysis of Computational Fluid Dynamics via Polynomial Chaos. Ph.D. Thesis. Virginia Polytechnic Institute and State University, Virginia.

Poeter, E.P., Hill, M.C., 1999. UCODE, a computer code for universal inverse modeling. Comput. Geosci. 25, 457–462. http://dx.doi.org/10.1016/S0098-3004(98)00149-6.

Prempraneerach, P., Hover, F.S., Triantafyllou, M.S., Karniadakis, G.E., 2010. Uncertainty quantification in simulations of power systems: multi-element polynomial chaos methods. Reliab. Eng. Syst. Saf. 95, 632–646. http://dx.doi.org/10.1016/j.ress.2010.01.012.

Prieto, C., Kotronarou, A., Destouni, G., 2006. The influence of temporal hydrological randomness on seawater intrusion in coastal aquifers. J. Hydrol. 330 (1–2), 285–300. http://dx.doi.org/10.1016/j.jhydrol.2006.03.024.

Rajabi, M.M., Ataie-Ashtiani, B., 2014. Sampling efficiency in Monte Carlo based uncertainty propagation strategies: application in seawater intrusion simulations. Adv. Water Resour. 67, 46–64. http://dx.doi.org/10.1016/j.advwatres.2014.02.004.

Ratto, M., Pagano, A., Young, P.C., 2007. State dependent parameter metamodelling and sensitivity analysis. Comput. Phys. Commun. 177, 863–876. http://dx.doi.org/10.1016/j.cpc.2007.07.011.

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48, W07401. http://dx.doi.org/10.1029/2011WR011527.

Saltelli, A., 1993. Sensitivity analysis of model output an investigation of new techniques. Comput. Stat. Data Anal. 15 (2), 211–238. http://dx.doi.org/10.1016/0167-9473(93)90193-W.

Saltelli, A., 2002. Sensitivity analysis for importance assessment. Risk Anal. 22, 579–590. http://dx.doi.org/10.1111/0272-4332.00040.

Saltelli, A., Sobol', I.M., 1995. About the use of rank transformation in sensitivity of model output. Reliab. Eng. Syst. Saf. 50, 225–239. http://dx.doi.org/10.1016/0951-8320(95)00099-2.

Saltelli, A., Tarantola, S., Chan, K., 1999. A quantitative, model independent method for global sensitivity analysis of model output. Technometrics 41 (1), 39–56. http://dx.doi.org/10.1080/00401706.1999.10485594.

Sanz, E., Voss, C.I., 2006. Inverse modeling for seawater intrusion in coastal aquifers: insights about parameter sensitivities, variances, correlations and estimation procedures derived from the Henry problem. Adv. Water Resour. 29, 439–457. http://dx.doi.org/10.1016/j.advwatres.2005.05.014.

Sarkar, A., Ghanem, R., 2002. Mid-frequency structural dynamics with parameter uncertainty. Comput. Meth. Appl. Mech. Eng. 191 (47), 5499–5513. http://dx.doi.org/10.1016/S0045-7825(02)00465-6.

Scott, D.W., 1992. Multivariate Density Estimation: Theory, Practice, and Visualization. Wiley, New York.

Slooten, L.J., Carrera, J., Castro, F., Fernandez-Garcia, G., 2010. A sensitivity analysis of tide-induced head fluctuations in coastal aquifers. J. Hydrol. 393, 370–380. http://dx.doi.org/10.1016/j.jhydrol.2010.08.032.

Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–414.

Sochala, P., Le Maître, O.P., 2013. Polynomial chaos expansion for subsurface flows with uncertain soil parameters. Adv. Water Resour. 62, 139–154. http://dx.doi.org/10.1016/j.advwatres.2013.10.003.

Sreekanth, J., Datta, B., 2014. Stochastic and robust multi-objective optimal management of pumping from coastal aquifers under parameter uncertainty. Water Resour. Manage 28 (7), 2005–2019. http://dx.doi.org/10.1007/s11269-014-0591-5.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. 93, 964–979. http://dx.doi.org/10.1016/j.ress.2007.04.002.

Voss, C.I., Provost, A.M., 2010. SUTRA, A Model for Saturated–Unsaturated Variable-density Ground-water Flow with Solute or Energy Transport. U.S. Geological Survey, Water-Resources Investigations, Open-File Report 02-4231.

Wackernagel, H., 1998. Multivariate Geostatistics, An Introduction with Applications, second ed. Springer, Berlin, ISBN: 3-540-64721-X.

Wan, X., Karniadakis, G.E., 2006. Multi-element generalized polynomial chaos for arbitrary probability measures. SIAM J. Sci. Comput. 28 (3), 901–928. http://dx.doi.org/10.1137/050627630.

Wei, P., Lu, Z., Yuan, X., 2013. Monte Carlo simulation for moment-independent sensitivity analysis. Reliab. Eng. Syst. Saf. 110, 60–67. http://dx.doi.org/10.1016/j.ress.2012.09.005.

Werner, A.D., Bakker, M., Post, V.E.A., Vandenbohede, A., Lu, C., Ataie-Ashtiani, B., Simmons, C.T., Barry, D.A., 2013. Seawater intrusion processes, investigation and management: recent advances and future challenges. Adv. Water Resour. 51, 3–26. http://dx.doi.org/10.1016/j.advwatres.2012.03.004.

Wiener, N., 1938. The homogeneous chaos. Am. J. Math. 60, 897–936.

Xiu, D., Hesthaven, J., 2005. High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. 27 (3), 1118–1139. http://dx.doi.org/10.1137/040615201.

Xiu, D., Karniadakis, G.E., 2002. The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. 24 (2), 619–644. http://dx.doi.org/10.1137/S1064827501387826.

Xiu, D., Karniadakis, G.E., 2003a. Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comput. Phys. 187, 137–167. http://dx.doi.org/10.1016/S0021-9991(03)00092-5.

Xiu, D., Karniadakis, G.F., 2003b. A new stochastic approach to transient heat conduction modeling with uncertainty. Int. J. Heat Mass Transfer 46 (24), 4681–4693. http://dx.doi.org/10.1016/S0017-9310(03)00299-0.

Zhang, K., Li, H., Achari, G., 2009. Fuzzy-stochastic characterization of site uncertainty and variability in groundwater flow and contaminant transport through a heterogeneous aquifer. J. Contam. Hydrol. 106 (1), 73–82. http://dx.doi.org/10.1016/j.jconhyd.2009.01.003.

Zio, E., Pedroni, N., 2012. Monte Carlo simulation-based sensitivity analysis of the model of a thermal–hydraulic passive system. Reliab. Eng. Syst. Saf. 107, 90–106. http://dx.doi.org/10.1016/j.ress.2011.08.006.