Environmental Modelling and Software 125 (2020) 104612

Contents lists available at ScienceDirect

Environmental Modelling and Software

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

Evaluation of parameter interaction effect of hydrological models using the sparse polynomial chaos (SPC) method

Heng Wang [a], Wei Gong [a], Qingyun Duan [a,b,*], Zhenhua Di [a]

[a] Institute of Land Surface System and Sustainable Development, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu, 210098, China

ARTICLE INFO

Keywords:
Hydrologic model
Sensitivity analysis
Interaction effect
Sparse polynomial chaos

ABSTRACT

Most of the commonly available sensitivity analysis methods cannot reliably compute the interaction effect. Even though the Sobol' type methods that use Monte Carlo simulation can evaluate the interaction effect, the result is either inaccurate or requires an extraordinary number of model runs to obtain a reasonable estimate. In this study, we evaluate the sparse polynomial chaos (SPC) method as a reasonable way to estimate the interaction effect. This method is evaluated on two mathematical test functions (Ishigami and Sobol' G) and two hydrologic models (HBV-SASK and SAC-SMA). Our results show the SPC method needs about a sample size of 30 to 70 times the number of dimensions of the parameter space to evaluate the interaction effects of hydrologic models. Our findings are significant for hydrologic simulation and model calibration, as we aim to improve the understanding of complex interactions among model components and to reduce model uncertainty.

1. Introduction

As more hydrological models are being developed, model structure is getting increasingly complex, and in many cases, those models are over-parameterized (Schoups et al., 2008). Sensitivity analysis (SA) is an important tool for assessing parametric uncertainty of hydrological models, as it can determine the degree of influence of parameters on model simulations, improve model identifiability and ultimately enhance model performance (Song et al., 2015).

There are a number of ways that model parameters influence model simulations. From a variance decomposition view angle (Razavi and Gupta, 2015), the impact of variation of a single parameter on model simulation is known as the first order effect, or the main effect. The effect of simultaneous variations of two or more parameters on model simulation is known as higher order effect. The sum of all the effects connected with a specific parameter is called the total effect. In practice, only the main effect, the total effect and the second order effect, which is also known as interaction effect, are evaluated. As parameter interaction effect plays a big role in hydrological simulation capability and in robust parameter optimization, accurate calculation of interaction effect of a hydrological model is vitally important.

The commonly used local SA (LSA) method can only be used to assess the main effect by varying a single model parameter at a time. LSA cannot consider higher order effect and its estimate of the main effect is also not reliable (Li et al., 2013). To assess higher order effects, global SA (GSA) methods are required. There are three types of GSA methods which that are widely applied in hydrological modeling: derivative-based GSA, regression-based GSA, and variance-based GSA (Wang et al., 2016). Most derivative based (e.g., Morris-One-At-a-Time, MOAT) (Morris, 1991) and regression based (e.g., Multivariate Adaptive Regression Splines, MARS) (Friedman, 1991) GSA methods are effective in assessing the main effect. However, they cannot be used to assess higher order effects.

Variance-based GSA methods are designed to compute the main effect as well as higher order effects. For example, the classical Sobol' method, one of the commonly used variance based GSA methods, does a good job of assessing the main effect in a relatively efficient manner (Gan et al., 2014). However, the Sobol' method usually requires an extra-ordinary number of model runs to obtain a reasonable estimate of the second order effect (typically, thousands to hundreds of thousands) (Tang et al., 2007; Wan et al., 2015). For large-scale hydrological model applications, the cost of performing second order or higher order sensitivity analysis using classical Sobol' method would be unbearable.

In order to reduce the number of model runs needed to calculate different sensitivity indices (i.e., the main effect and interaction effect), meta-model based SA methods have been proposed. Meta-model based approach is generally implemented in two steps. First, a surrogate model is constructed. Then SA is performed on the surrogate model.

* Corresponding author at: College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu, 210098, China.
E-mail address: qyduan@hhu.edu.cn (Q. Duan).

https://doi.org/10.1016/j.envsoft.2019.104612
Received 12 January 2019; Received in revised form 19 December 2019; Accepted 23 December 2019
Available online 23 December 2019

The advantage of a meta-model based approach is that it requires a significantly smaller number of model runs to perform SA than traditional SA methods. To construct a surrogate model, a design of experiment (DoE) approach is usually used to sample the parameter space. Then the model is run using those sampled parameter sets and the corresponding performance metrics (i.e., the output of interest) are computed. Finally, a surrogate model is constructed based on the parameter-objective function statistical relationship. There exist a plethora of statistical methods to construct a surrogate model (Razavi et al., 2012a), for example the aforementioned MARS method, SVM (support vector machine) method, GPR (Gaussian process regression) method, RF (random forest) and ANN (artificial neural network) method.

Meta-model based Sobol' method that uses Monte Carlo simulation has been used to perform SA of hydrological models by numerous researchers (Song et al., 2012a,b; Zhan et al., 2013; Wang et al., 2016; Gan et al., 2017). Because Monte Carlo approximations of the integrals required in Sobol' method can lead to some numerical errors, the estimate of the interaction effect based on meta-model based approach can be inaccurate. For example, the second order effect computed by the classical Sobol' method sometimes results in negative values of sensitivity indices, which defies the definition of variance terms, which must be positive (Tang et al., 2007). Furthermore, the estimation of interaction effect usually requires an extraordinary number of model runs to obtain a reasonable estimate of the interaction effect. For example, Zhan et al. (2013) studied the parameter interaction effects of the distributed time-variant gain model (DTVGM) with 14 parameters using the meta-model based Sobol' method and it required 2600 DTVGM model runs and 100,000 meta model runs (186 times the number of parameters) to obtain a reasonable estimate of interaction effect. Song et al. (2012a) studied Xinanjiang hydrological model (7 parameters) using the Sobol' method and it costs 1000 Xinanjiang model runs and 100,000 meta model runs (286 times the number of parameters) to avoid negative sensitivity indices. Further, there is no guarantee that the interaction effect computed by the meta-model based Sobol' method that uses Monte Carlo simulation is accurate, as we will show in this study.

Generally, SA based on meta-models involves two kinds of numerical errors, one is the error between the surrogate model and the real response surface of the physical models, the other is the Monte Carlo simulation error (The later error will converge to zero when the sample size is infinite). The superposition of the two errors will have a significant impact on the confidence of the calculation of Sobol' global sensitivity indices. On the contrary, if the Monte Carlo simulation error can be avoided, the accuracy of global sensitivity indices can be greatly improved. In this paper, we investigate the use of the sparse polynomial chaos (SPC) SA method (Sudret, 2008; Blatman, 2009; Blatman and Sudret, 2011; Fajraoui et al., 2012; Marelli and Sudret, 2014; Tang and Zhou, 2015; Hu and Zhang, 2016; Shao et al., 2017) to calculate the interaction effect of the hydrological model parameters. If polynomial chaos expansion is used as a surrogate model, the mean and variance of output can be calculated directly by the expansion coefficients, as well as the global Sobol' sensitivity indices can be calculated directly. In other words, the SPC method can avoid Monte Carlo simulation error directly.

The rest of paper is organized as follows: Section 2 introduces the SPC method. In Section 3, by using SPC method, the Ishigami function (3 tunable parameters), Sobol' function (30 tunable parameters) as the mathematical function model cases, and the HBV-SASK model (10 tunable parameters) and SAC-SMA model (13 tunable parameters) as the hydrological model cases are used for illustration. Section 4 presents a detailed discussion about the interaction effects. Finally, a short conclusion is given in Section 5.

2. Sparse polynomial chaos method for sensitivity analysis

2.1. Full polynomial chaos

Firstly, the full polynomial chaos method is reviewed. Considering a general second order stochastic process in the probability space (Ω, P, F) with space of events Ω, σ-algebra P and probability measure F. Let us consider a stochastic model $Y(\xi)$ with model input $\xi = (\xi_1, \xi_2, \ldots) \in \Omega$ and model output Y. For a stochastic analysis of Y, the model $Y(\xi)$ may be approached as follows:

$$Y(\xi) = a_0\Psi_0 + \sum_{i_1=1}^{\infty} a_{i_1}\Psi_1(\xi_{i_1}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Psi_2(\xi_{i_1}, \xi_{i_2}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Psi_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) + \cdots, \tag{2.1}$$

where $\Psi_i(\xi)$ is i order orthogonal polynomial basis functions in the variables (ξ). $a_0, a_{i_1}, \ldots$ is the approximation coefficient. Here the expansion bases $\Psi_i(\xi)$ are multi-dimensional orthogonal polynomials defined as tensor products of the corresponding one-dimensional polynomials bases $\{\phi_k\}_{k=0}^{\infty}$:

$$\Psi_i(\xi) := \prod_{k=1}^{n} \phi_{i_k}(\xi_k), i := \sum_{k=1}^{n} i_k, \tag{2.2}$$

where n is the number of random variables. In realistic applications, we usually consider a finite number of terms to truncate (2.1), and we have,

$$Y(\xi) = \sum_{\alpha \in \mathcal{A}^{p,n}} a_{\alpha}\Psi_{\alpha}(\xi), \mathcal{A}^{p,n} \equiv \{\alpha \in N^n : |\alpha| \le p\}, \tag{2.3}$$

where α is a multi-index that identifies the components of the multivariate polynomials, and $\mathcal{A}$ is the set of selected multi-indices of multivariate polynomials, and a_{α} is the corresponding polynomial chaos expansions coefficients. The number of the polynomial expansion term N is related to the number of random space independent variables used in the stochastic process system, and it is also related to the degree of freedom (i.e. the maximum number of orders of the polynomial basis). When the number of random variables and the degree of freedom are given, the number of the polynomial expansion term N in the random process is as follows:

$$N = \frac{(p+n)!}{p!n!} - 1, \tag{2.4}$$

where p is the degree of freedom, and n is the number of random variables. The expression Eq. (2.1) is called the full polynomial chaos.

For polynomial chaos methods, there are three key steps. They are respectively the determination of orthogonal polynomial basis of model input parameters, the calculation of the coefficients of polynomial chaos and estimators of accuracy of the polynomial chaos approximations. For the determination of orthogonal polynomial basis, according to the Wiener–Askey polynomial chaos (Xiu and Karniadakis, 2003), there exist different optimal polynomials for different probability density functions, e.g. normalized Legendre (resp. Hermite) polynomials can be associated to a uniform (resp. Gaussian) probability density functions. In this work, the model parameters are considered to be uniformly distributed and the Legendre polynomial chaos are chosen.

In the calculation process of coefficients of the polynomial expansion, two methods for intrusive method and non-intrusive method can be used to determine the coefficients of polynomial expansion (Xiu, 2010). The intrusive method needs to adjust the original model, and non-intrusive method does not need to adjust the original model code. In this way, the non-intrusive method is suitable for the study of most models. The non-intrusive method has two main methods: the regression method and the projection method (Blatman and Sudret, 2011). Here, the regression-based non-intrusive method is chosen in this work. The coefficients may be estimated by determining the

L_2-projection of the response $Y(\xi)$ onto the space spanned by the polynomials $\{\Psi_\alpha(\xi), |\alpha| \le p\}$ as follows:

$$\hat{a} = \arg\min_{a \in R^n} E[Y(\xi) - a^T \Psi(\xi)]. \tag{2.5}$$

Given a sampling of size N of the input random vector $\chi = \{\xi^{(1)}, \dots, \xi^{(N)}\}^T$ (the design of experimental) and the corresponding model responses $\mathcal{Y} = \{y^{(1)}, \dots, y^{(N)}\}^T$, the ordinary least-square solution of Eq. (2.5) is:

$$\hat{a} = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \mathcal{Y}, \tag{2.6}$$

where $\boldsymbol{A} = \{\boldsymbol{A}_{ij} = \Psi_j(\xi^{(i)}), i = 1, \dots, N; j = 1, \dots, card\,\mathcal{A}\}$ is the so-called experimental matrix that contains the values of all the basis polynomials in the experimental design points.

After the polynomial coefficients are computed, through the error estimation, the polynomial chaos-based model with the appropriate accuracy is successfully established. The leave-one-out (LOO) error estimation is chosen in this work. The leave-one-out (LOO) cross-validation error ϵ_{LOO} is designed by using cross-validation. It consists in building N meta-models $Y^{PC\setminus i}$, each one created on a reduced experimental design $\chi \setminus \xi^{(i)} = \{\xi^{(j)}, j = 1, \dots, N, j \neq i\}$ and comparing its prediction on the excluded point $\xi^{(i)}$ with the real value $y^{(i)}$ (Blatman and Sudret, 2010). The leave-one-out cross-validation error can be written as:

$$\epsilon_{LOO} = \frac{\sum_{i=1}^{N} (Y(\xi^{(i)}) - Y^{PC\setminus i}(\xi^{(i)}))^2}{\sum_{i=1}^{N} (Y(\xi^{(i)}) - \hat{\mu}_Y)^2}, \tag{2.7}$$

where $\hat{\mu}_Y$ is the sample mean of the experimental design response. In practice, when the results of a least-square minimization are available, there is no need to explicitly calculate N separate meta-models, after some algebra Eq. (2.7) can reduce to:

$$\epsilon_{LOO} = \frac{\sum_{i=1}^{N} (\frac{Y(\xi^{(i)}) - Y^{PC}(\xi^{(i)})}{1 - h_i})^2}{\sum_{i=1}^{N} (Y(\xi^{(i)}) - \hat{\mu}_Y)^2}, \tag{2.8}$$

where h_i is the ith diagonal term of matrix $\boldsymbol{A}(\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T$ (matrix $\boldsymbol{A}$ is defined in Eq. (2.6)) and $Y^{PC}(\cdot)$ is the PC expansion built up from the full experimental design χ.

However, note that the number of full PC coefficients N increases exponentially with n and p since Eq. (2.4). Thus the number of coefficients to be computed increases dramatically when n is large, say n > 10. This is known as the curse of dimensionality. For this issue, it is solved satisfactorily using specific methods to compute sparse polynomial chaos.

2.2. Sparse polynomial chaos

Suppose $\mathcal{A}$ be a non-empty finite subset of N^n, with which the truncated polynomial chaos can be defined by

$$Y(\xi) = \sum_{\alpha \in \mathcal{A}} a_\alpha \Psi_\alpha(\xi). \tag{2.9}$$

The common truncation scheme in Eq. (2.3) corresponds to the choice $\mathcal{A} = \mathcal{A}^{p,n}$. Since the curse of dimensionality problem, the determination of truncation sets $\mathcal{A}$ of small cardinality is of interest. Thus, we define that if the following condition is verified, the truncated PC Eq. (2.9) is sparse:

$$IN = \frac{card(\mathcal{A})}{card(\mathcal{A}^{p,n})} < 1. \tag{2.10}$$

Generally, polynomial chaos can be sparse from two aspects: hypothesis of the structure of polynomial chaos and the numerical algorithm for solving polynomial chaos. From the point of view of the structure of polynomial chaos, according to the sparsity-of-effects principle (Montgomery, 2004), the hyperbolic truncation or maximum interaction can be used. From the point of view of the solution algorithm of polynomial chaos, the least angle regression (LAR)-based or orthogonal matching pursuit (OMP)-based solution algorithm can be applied (Blatman and Sudret, 2011; Doostan and Owhadi, 2011).

2.3. Sparse polynomial chaos for sensitivity analysis

Finally, when we obtain the coefficients of Eq. (2.9) successfully, Eq. (2.9) can be interpreted as a model response surface for $Y(\xi)$, the uncertainty and sensitivity analysis can be calculated directly as simple analytical functions of the SPC coefficients. For the expectation and variance, we have

$$E[Y(\xi)] = a_0, Var[Y(\xi)] = \sum_{\alpha \in \mathcal{A}, \alpha \neq 0} a_\alpha^2. \tag{2.11}$$

For the sensitivity indices, the SPC can be rewritten in the form of the Sobol' decomposition:

$$Y(\xi) = a_0 + \sum_{i_1=1}^{N} \sum_{\alpha \in I_{i_1}} a_\alpha \Psi_\alpha(\xi_{i_1}) + \cdots + \sum_{i_s > \cdots > i_1}^{N} \sum_{\alpha \in I_{i_1, \dots, i_s}} a_\alpha \Psi_\alpha(\xi_{i_1}, \dots, \xi_{i_s}) + \cdots + \sum_{\alpha \in I_{1, \dots, N}} a_\alpha \Psi_\alpha(\xi), \tag{2.12}$$

where $I_{i_1, \dots, i_s} = \{\alpha \in (\alpha_1, \dots, \alpha_N) : \alpha_k = 0 \Leftrightarrow k \notin \{i_1, \dots, i_s\} \subset \{1, \dots, N\}\}$. Due to the orthogonal property of the polynomial basis, the partial variance can be derived analytically from the SPC coefficients as follows:

$$D_{i_1 \dots i_N} = \sum_{\alpha \in I_{i_1, \dots, i_N}} a_\alpha^2. \tag{2.13}$$

So the partial sensitivity indices for the subset of input variables $\{\xi_{i_1}, \dots, \xi_{i_N}\}$ is as follows:

$$S_{i_1 \dots i_N} = \frac{D_{i_1 \dots i_N}}{D}. \tag{2.14}$$

Note that once a SPC approximation of the model response has been built as a metamodel, compared with all Sobol' type SA methods which require Monte Carlo simulation to calculate the sensitivity indices, the SPC-based sensitivity indices are computed analytically from the SPC coefficients which is of a negligible computational cost.

3. Evaluation of parameter interaction effect of the analytical test problems and the hydrological models

In this section, the efficiency and effectiveness of the SPC SA method are evaluated. Firstly, two well-known mathematical functions are investigated: the 3-Dimensional Ishigami function and the 30-Dimensional Sobol' G function which have analytical expression for sensitivity indices. Then the SPC SA method is used to examine the parameter interaction effect of two hydrological models in real-world application settings, i.e., the 10-Dimensional Hydrologiska Byråns Vattenbalansavdelning-University of Saskatchewan (HBV-SASK) hydrological model and the 13-Dimensional Sacramento-Soil Moisture Accounting (SAC-SMA) hydrological model, both of which have no analytical expression for sensitivity indices.

3.1. Analytical test models: the Ishigami and Sobol' G functions

Firstly, the Ishigami function is chosen as our test case (Ishigami and Homma, 1990). Its expression is shown as follows:

$$Y(\boldsymbol{x}) = sinx_1 + asin^2 x_2 + bx_3^4 sinx_1, \tag{3.1}$$

where the input variables x_1, x_2, and x_3 are uniformly distributed over $[-\pi, \pi]$. The variance D of Y and the Sobol' sensitivity indices can be computed analytically as follows.

$$D = \frac{a^2}{8} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{1}{2}, D_1 = \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} + \frac{1}{2}, D_2 = \frac{a^2}{8},$$
$$D_3 = 0, D_{12} = D_{23} = 0, D_{13} = \frac{8b^2\pi^8}{225}. \tag{3.2}$$

In this example, $a = 7$ and $b = 0.1$. For the sparse polynomial chaos, the OMP-based calculation method of the coefficients is selected and

the adaptive polynomial degree is set as $2 < p < 15$ based on LOO cross-validation error estimates. In addition, the classical Sobol' method (see Appendix A), the Gaussian process regression-based (GPR) SA method (see Appendix B) and the polynomial chaos-Kriging-based (PC-Kring) SA method (see Appendix C) are used for comparison purpose. The Gaussian process regression and polynomial chaos-Kriging are the commonly used meta-models (Wang et al., 2014; Kersaudy et al., 2015; Gong et al., 2015, 2016). Table 1 shows the root mean square errors (RMSE) of the estimates of the total, first order and second order Sobol' indices of SPC, GPR and PC-Kring methods against theoretical values for sample sizes $n = 20$, $n = 40$ and $n = 80$ and Sobol' method for the sample size $n = 200{,}000$, respectively. The negative values calculated by the Sobol' method and meta-model based Sobol' method are all set to zero in this work.

The second analytical test problem is the Sobol' G function (Saltelli and Sobol', 1995), whose expression is as follows:

$$Y(\boldsymbol{x}) = \prod_{i=1}^{n} \frac{|4X_i - 2| + a_i}{1 + a_i}, \tag{3.3}$$

where the input variables $\boldsymbol{X}$ are uniformly distributed over $[0, 1]$ and $\boldsymbol{a}$ are nonnegative. The variance D of Y and the Sobol' sensitivity indices can be computed analytically as follows.

$$D = \prod_{i=1}^{n}(D_i + 1) - 1, D_i = \frac{1}{3(1 + a_i)^2}, D_{Ti} = D_i \prod_{j \neq i}(1 + D_j),$$
$$S_{i_1,\ldots,i_n} = \frac{1}{D}\prod_{j=1}^{n} D_{i_j}. \tag{3.4}$$

In this example, $n = 30$ and $a = [1, 2, 5, 10, 20, 50, 100, 500, 999, \ldots, 999]$. For the sparse polynomial chaos, the OMP-based calculation method of the coefficients is selected and the adaptive polynomial degree is set as $2 < p < 5$ based on LOO cross-validation error estimates. In addition, the classical Sobol' method, the Gaussian process regression-based SA method and the polynomial chaos-Kriging-based SA method are still chosen for comparison. Table 2 displays the RMSEs of the estimates of the total, first order and second order Sobol' indices of SPC, GPR and PC-Kring methods against theoretical values for sample sizes $n = 150$, $n = 300$ and $n = 600$ and Sobol' method for $n = 200{,}000$, respectively.

From the sensitivity indices results of the Ishigami (Table 1) and Sobol' G functions (Table 2), it is easy to see that the SPC method is more accurate in calculating all second order effects while identifying total and first order effects. In particular, when the sample size of the Ishigami function is 80 and the sample size of the Sobol' G function is 300, the RMSEs of their second order indices against theoretical indices reach 0. In addition, as in the case of the Ishigami test function, we can see the importance of interaction effect. The parameter x_3 is of zero main effect, but has strong interaction with parameter x_1, under which situation interaction effect should not be ignored.

It is essential to monitor and evaluate the convergence rate of the GSA methods using some efficient techniques (Nossent et al., 2011; Razavi et al., 2012b; Sarrazin et al., 2016; Harenberg et al., 2019; Sheikholeslami et al., 2019), as it can enable us to diagnose the convergence behavior of the GSA. Thus, in order to test the convergence behavior, the robustness analysis is designed to estimate first and second order indices for a set of increasingly larger experimental designs by means of bootstrap method (using 100 bootstrap replicates) (Efron, 1979). For the Ishigami function and the Sobol' G function, the results of the convergence of the estimates of the first order indices for two most sensitive parameters (e.g. parameters x_1 and x_2 of the Ishigami function, parameters X_1 and X_2 of the Sobol'G function) and the largest second order indices (e.g., parameter pair x_{13} of the Ishigami function, parameter pair X_{12} of the Sobol'G function) are shown in Figs. 1 and 2, respectively. The Sobol' total effect indices are not shown because their convergence behavior is essentially identical to that of the first order indices. By the convergence analysis, confirming that the estimation of interaction effects is accurate enough to make a solid conclusion.

3.2. Evaluation of parameter interaction effect of the HBV-SASK hydrological model

The HBV-SASK model (Fig. 3) is a conceptual rainfall-runoff model which was coded at the University of Saskatchewan for educational purposes, based on an interpretation of the Hydrologiska Byråns Vattenbalansavdelning model (Lindström et al., 1997). Here, ten of the HBV-SASK model parameters are considered tunable. These parameters are described in Table 3. The rainfall-runoff model (HBV-SASK) used in this study has 12 tunable parameters, as presented in (Gupta and Razavi, 2018; Razavi and Gupta, 2019; Razavi et al., 2019). Here we consider only 10 parameters. Because parameter 11 is the base of unit hydrograph for watershed routing in day and its default is 1 for small watersheds. Parameter 12 is the precipitation multiplier to address uncertainty in precipitation and its default is 1. Here we chose their default values in the HBV-SASK model.

The Oldman watershed is chosen as the study area. The 1434.73 km^2 Oldman watershed is located in the Rocky Mountains of Alberta, Canada. Historical data is available for the period 1979–2008, from which the average annual precipitation (rainfall+snowfall) is estimated to be 611 mm, and average annual streamflow to be 11.7 m^3/s at gauge 05AA023 on the Oldman River (runoff ratio = 0.42) (Razavi and Gupta, 2019).

To evaluate model responses as a function of different parameter values, the RMSE values between the simulated and observed daily streamflow discharge (m^3/s) is used as the objective function:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{s,t} - Q_{o,t})^2}, \tag{3.5}$$

where $Q_{s,t}$ and $Q_{o,t}$ are simulated and observed streamflow discharge values at time t, n is the total number of observations. There is no analytical values for the sensitivity indices of the HBV-SASK hydrology model in a real-world application setting. To obtain a reasonable estimate of the "true" sensitivity indices, we performed SA using the classical Sobol' method that employs 1000,000 Monte Carlo simulations (MCS) (see Appendix A for details) to approximate the "true" sensitivity indices. To demonstrate the convergence of the SPC method, a set of nested experimental designs of increasing sample sizes {100, 200, 300, 400, 500, 700, 1000} are generated based on the Sobol' sequence sampling. For the sparse polynomial chaos, selecting the adaptive polynomial degree $2 < p < 5$ and OMP-based calculation method of the coefficients. The selections of adaptive polynomial degree and calculation method are based on the leave-one-out error estimation of SPC method. The leave-one-out errors of the experimental designs {100, 200, 300, 400, 500, 700, 1000} are $5.96e^{-12}$, $4.45e^{-15}$, $1.54e^{-12}$, $5.02e^{-13}$, $1.34e^{-13}$, $4.38e^{-14}$ and $8.45e^{-14}$, respectively. Again, the GPR SA method and PC-Kriging SA method are used for comparison purpose.

3.2.1. Effectiveness of the SPC method in calculating interaction effects

Firstly, the effectiveness in estimating the total and first order SA indices by different meta-model based SA methods is investigated. Parameter sensitivity rankings of MCS-based Sobol' method with 1000,000 model evaluations and SPC-based Sobol' method with 1000 model evaluations are given in Fig. 4, which shows that the three most sensitive parameters are parameters 7, 5 and 2, respectively, based on both the first order and total SA indices. The SPC-based Sobol' method provides excellent matches with MCS-based Sobol' method. Fig. 5 displays the RMSE values of the estimates of the first order and total SA indices calculated by SPC, GPR and PC-Kriging SA methods. Compared with the GPR and PC-Kriging methods, the SPC method has much smaller RMSE values for both the first order and total effects than the other two methods.

Next, the effectiveness in estimating the second order SA indices is considered. The second order results for the MCS-based Sobol' and SPC-based Sobol' methods are given in Fig. 6, which shows that parameter

Table 1
The comparison of the RMSE values of SA indices for the 3-Dimensional Ishigami function based on Monte Carlo simulations and three meta-model based methods.

Root mean square error (RMSE)	MCS	SPC			GPR			PC-Kring		
	200,000	40	80	160	40	80	160	40	80	160
Total order indices	0.003	0.048	0	0	0.217	0.124	0.041	0.069	0.061	0.012
First order indices	0.002	0.034	0	0	0.061	0.044	0.017	0.032	0.027	0.005
Second order indices	0.006	0.020	0	0	0.170	0.091	0.031	0.050	0.054	0.010

Table 2
The comparison of the RMSE values of SA indices for the 30-Dimensional Sobol's G function based on Monte Carlo simulations and three meta-model based SA methods.

Root mean square error (RMSE)	MCS	SPC			GPR			PC-Kring		
	200,000	150	300	600	150	300	600	150	300	600
Total order indices	0.001	0.022	0.007	0.006	0.021	0.010	0.005	0.015	0.005	0.003
First order indices	0.006	0.010	0.006	0.005	0.023	0.014	0.009	0.006	0.008	0.007
Second order indices	0.001	0.001	0	0	0.004	0.005	0.005	0.003	0.005	0.005

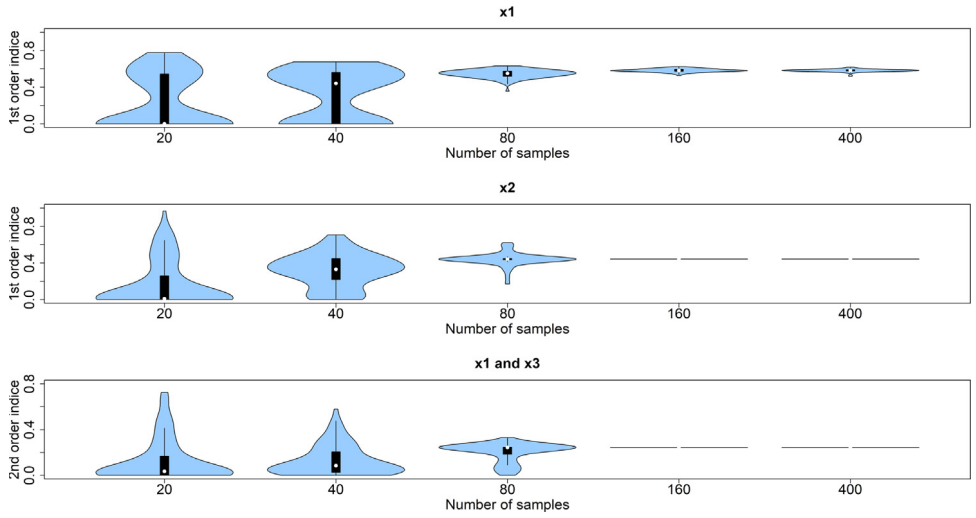

Fig. 1. Convergence of the estimates of the first and second order indices for the Ishigami function. Error bounds are calculated by bootstrap resampling.

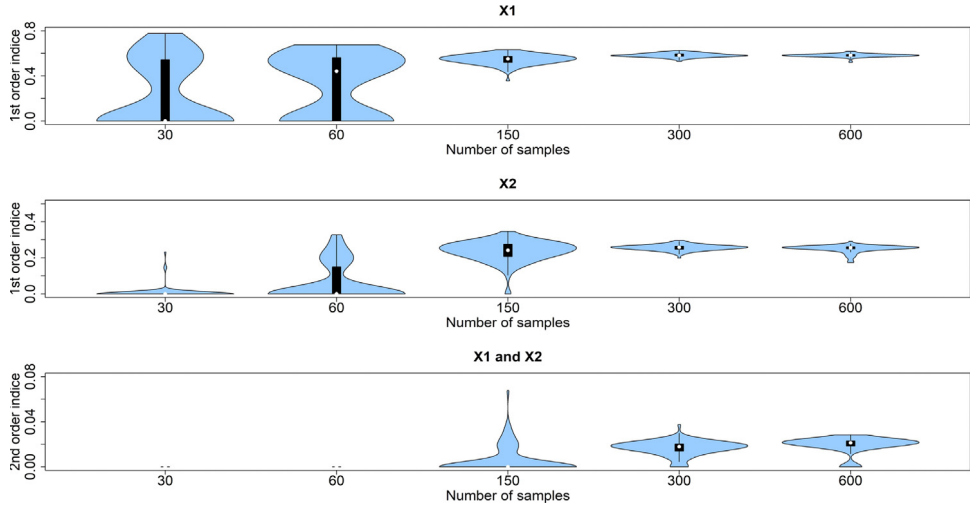

Fig. 2. Convergence of the estimates of the first and second order indices for the Sobol' G function. Error bounds are calculated by bootstrap resampling.

pairs $\{5,7\}, \{2,7\}, \{1,2\}$, and $\{2,5\}$ have the largest interaction effects. Note that when the sample size for the SPC method reaches 300 or more, the results of SPC-based and MCS-based Sobol' methods are very similar. Fig. 7 exhibits the RMSE values of the estimates of the second order sensitivity indices calculated by SPC, GPR and PC-Kriging methods, respectively. From Fig. 7, the SPC method shows a better convergence behavior than the GPR and PC-Kriging methods. In addition, for the training time of various methods, the time consumption of SPC method is the least. For sample sizes {100, 300, 500, 700, 1000}, the SPC method can save {10, 4, 5, 4, 2} times against the GPR method and {12, 9, 13, 7, 11} times against the PC-Kriging method, respectively.

The interaction between TT (parameter 1), the air temperature threshold about melting/freezing state of precipitation, and CO (parameter 2), the base melt factor of snow, is due to the melting/freezing

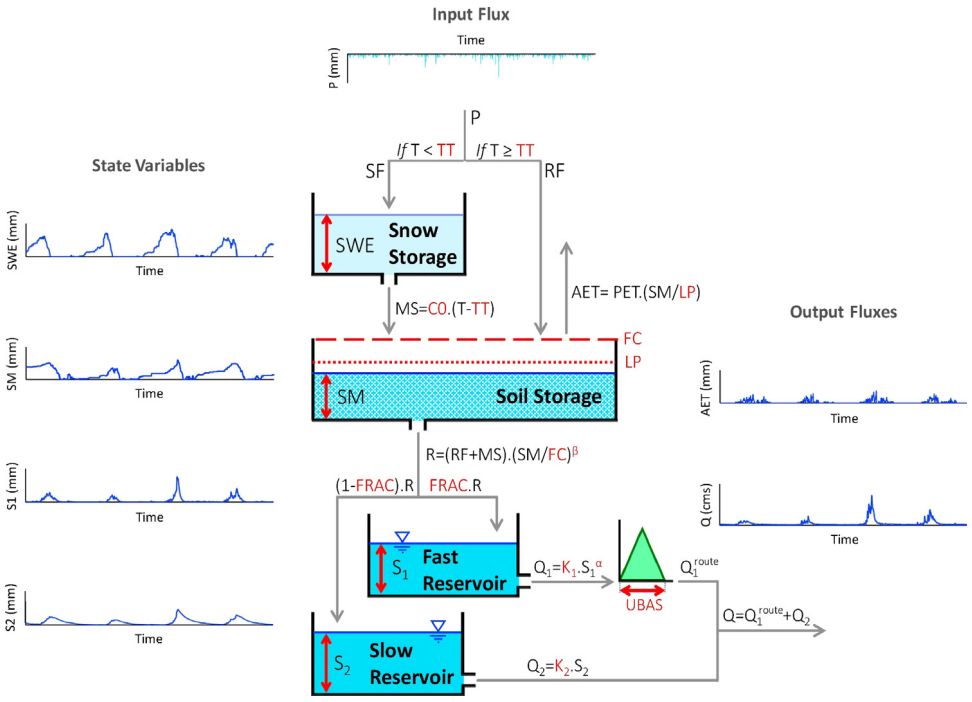

Fig. 3. System architecture of the HBV-SASK hydrologic model (Razavi and Gupta, 2019).

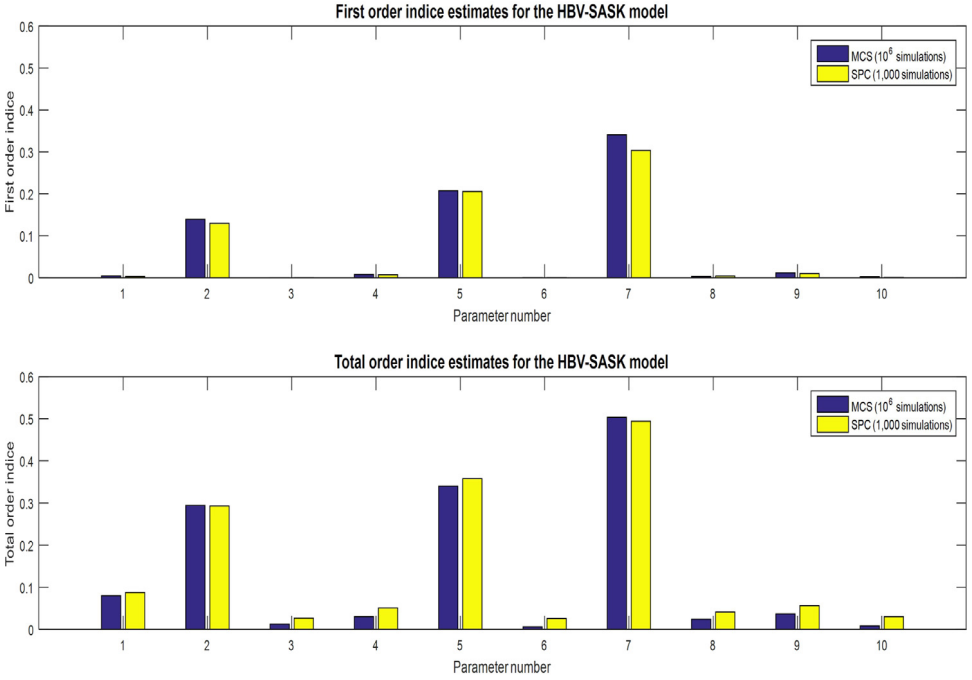

Fig. 4. The first and total order sensitivity indices obtained by MCS-based and SPC-based methods.

state of water before and after reaching ground. CO also has a strong correlation with FC (parameter 5), the soil field capacity, and FRAC (parameter 7), the fraction of water released from soil to fast reservoir. The reason of strong interaction between CO, FC and FRAC is that these three parameters control the state of water in this watershed: frozen in snowpack, stored in soil, or flowing in the river channel. These three parameters are also the most important parameters controlling the whole hydrological process in this watershed. Parameter TT only interacts with CO as both of them are related to the process in snow cover.

3.2.2. Robustness of the SPC method in calculating interaction effects

The robustness of the SPC method using different samples sizes in obtaining reliable estimates of different SA indices is analyzed. In order to test the convergence behavior, a robustness study is still designed to estimate first and second order indices for a set of increasingly larger experimental designs with a maximum of N = 1000. Confidence bounds for each estimate are calculated by means of a bootstrap method (using 100 bootstrap replicates). The results of the convergence of the estimates of the first order indices for three most sensitive parameters (e.g., parameters 7, 5, 2) are shown in Fig. 8 (violin graph). The Sobol' total effect indices are not shown because their convergence behavior

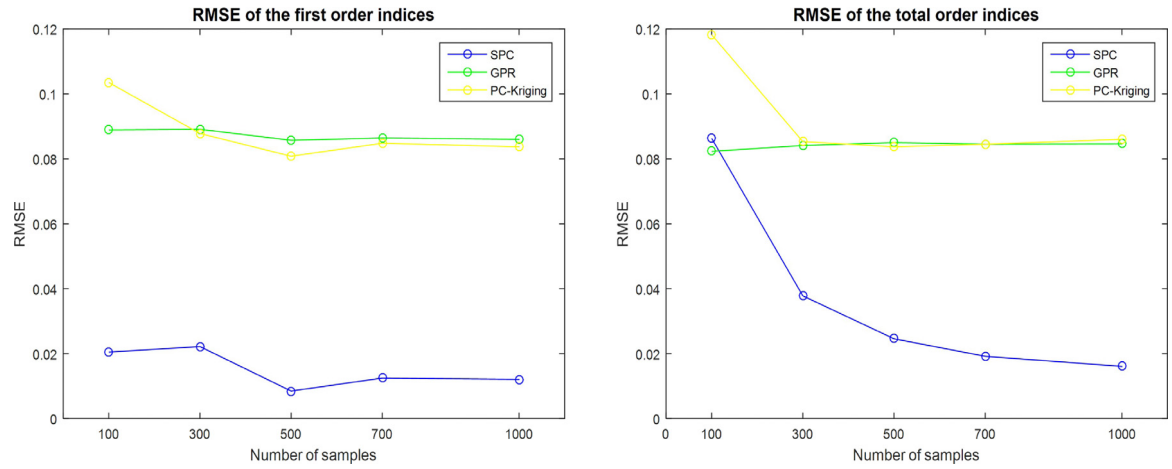

Fig. 5. The comparison of the RMSE values of the first and total order SA indices by three meta-model based SA methods.

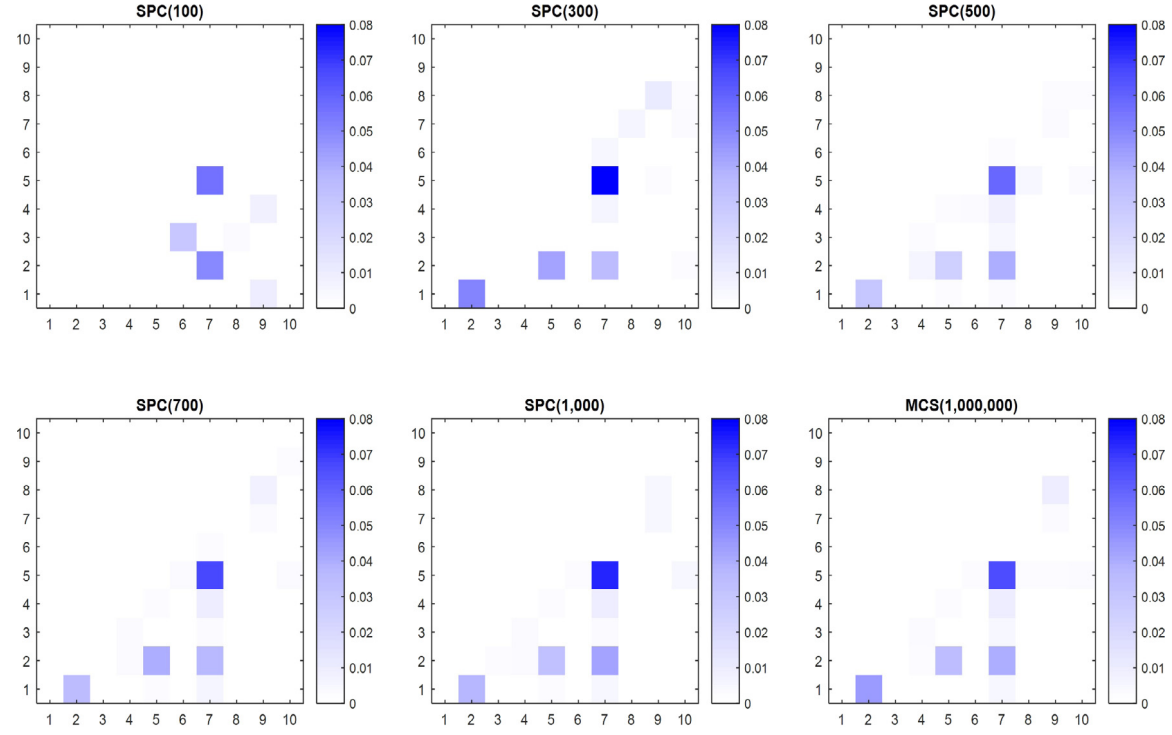

Fig. 6. The second order SA indices obtained by MCS-based with 1000,000 parameter samples and the SPC-based method with different sample sizes. The color legends and shading represent the Sobol' indices' magnitudes and ranges.

is essentially identical to that of the first order indices. From Fig. 8, it is clear that reasonably good estimates of the first order indices are already obtained with 200 model evaluations (i.e. the sample size is 20 times as large as the number of parameters).

For the robustness of the second order SA indices estimates, the convergence behavior of the estimates of the Sobol' second order indices as a function of the experimental design sizes is shown in Fig. 9 for the four parameter pairs with the largest second order Sobol' indices (e.g., pairs {5, 7}, {2, 7}, {1, 2}, {2, 5}). Fig. 9 shows clearly that reasonably good estimates of the second order indices can be obtained with 300 model evaluations (i.e. the sample size is 30 times as large as the number of parameters).

3.3. Evaluation of parameter interaction effect of the SAC-SMA hydrological model

The SAC-SMA hydrological model (Fig. 10) has a highly non-monotonic, non-linear input parameter-model output relationship. This model is the most widely used hydrological model by the River Forecast Centers of the U.S. National Weather Service for catchment modeling and flood forecasting (Burnash et al., 1973). Here, thirteen of the SAC-SMA model parameters are considered tunable (Brazil, 1988). These parameters are described in Table 4.

The South Branch Potomac River basin near Springfield, West Virginia in the U.S. is chosen as the study area. Historical precipitation, potential evapotranspiration and streamflow observations from January 1, 1960 to December 31, 1979 are obtained from the Model

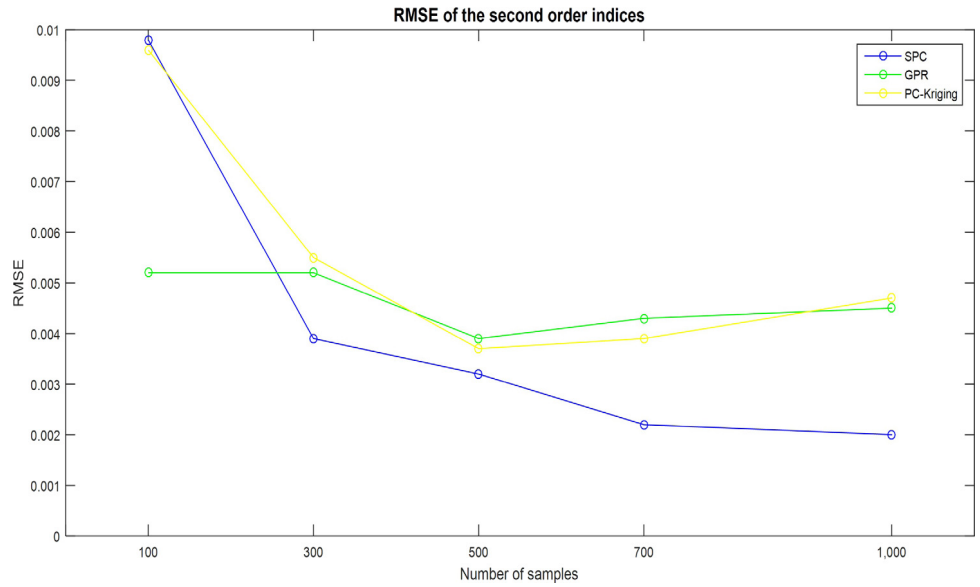

Fig. 7. The comparison of the RMSE values of the second order sensitivity indices by three meta-model based SA methods.

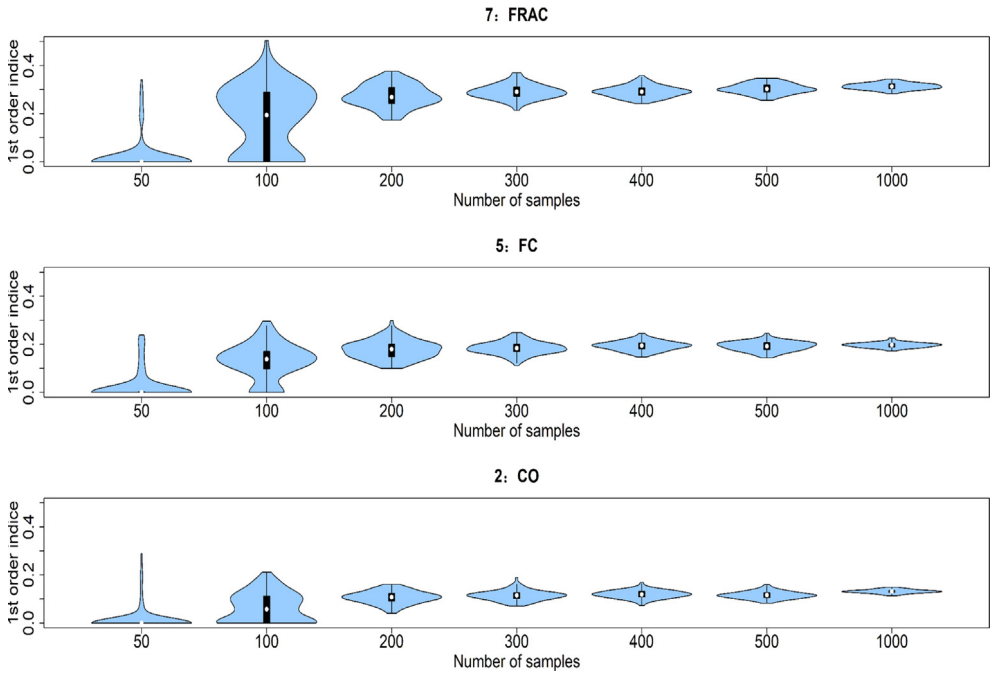

Fig. 8. Convergence of the estimates of the first order indices for different parameters. Error bounds are calculated by bootstrap resampling.

Parameter Estimation Experiment (MOPEX) database for this study (Duan et al., 2006). The hydrological simulations were run at a 6h time step over the entire data period. The average annual runoff is 39.5 m^3/s, average annual potential evapotranspiration is 762 mm, and average annual precipitation over this period is 1021 mm (Wang et al., 2014).

To evaluate model responses under different parameters, the root mean square error (RMSE) is chosen as the objective function. Similarly, we still performed SA using the classical Sobol' method that employs 1000,000 Monte Carlo simulations (MCS) to approximate the "true" sensitivity indices. To demonstrate the convergence of the SPC method, a set of nested experimental designs of increasing size {130, 260, 390, 520, 650, 910, 1300} is generated based on the Sobol' sequence sampling. For the sparse polynomial chaos, selecting the adaptive polynomial degree $2 < p < 5$ and OMP-based calculation method of the coefficients. The leave-one-out errors of the experimental designs {130, 260, 390, 520, 650, 910, 1300} are $1.28e^{-10}$, $3.68e^{-16}$, $3.71e^{-11}$, $1.52e^{-12}$, $2.16e^{-12}$, $6.41e^{-13}$ and $1.59e^{-12}$, respectively. Again, the GPR SA method and PC-Kriging SA method are still used for comparison purpose.

3.3.1. *Effectiveness of the SPC method in calculating interaction effects*

Firstly, the effectiveness in estimating the total and first order SA indices by different meta-model based SA methods is investigated. Parameter sensitivity rankings of MCS-based Sobol' method with 1000,000 model evaluations and SPC-based Sobol' method with 1300 model evaluations are given in Fig. 11, which shows that the two most sensitive parameters are parameters 5 and 4, respectively, based on both the first order and total SA indices. The SPC based Sobol' method provides excellent matches with MCS-based Sobol' method. Fig. 12 displays the

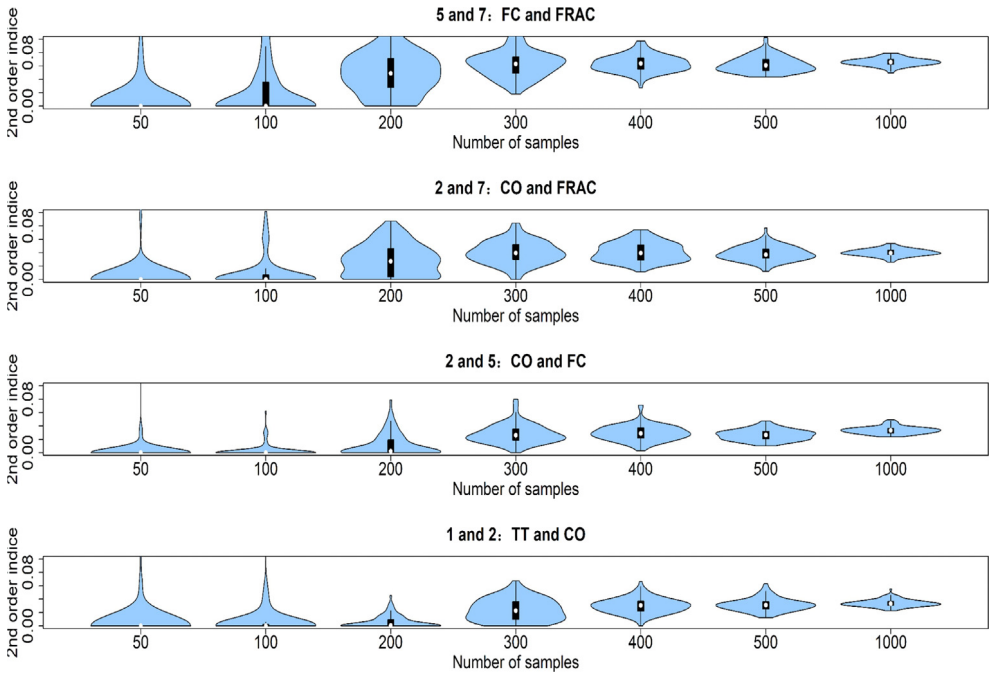

Fig. 9. Convergence of the estimates of the second order indices for different parameter pairs. Error bounds are calculated by bootstrap resampling.

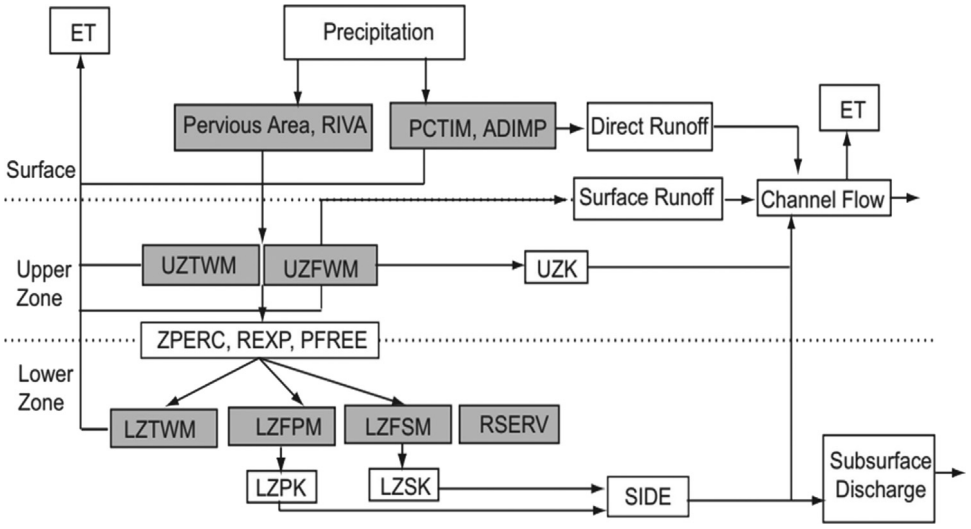

Fig. 10. A schematic of the SAC-SMA model (Tang et al., 2007).

RMSE values of the estimates of the first order and total SA indices calculated by SPC, GPR and PC-Kriging SA methods. Compared with the GPR and PC-Kriging methods, SPC method is also effective for identifying first and total order effects.

Next, the effectiveness in estimating the second order SA indices is considered. The second order results for the MCS-based Sobol' and SPC-based Sobol' methods are given in Fig. 13, which shows that parameter pairs {9, 11}, {10, 11} and {5, 8} have the largest interaction effects. Note that when the sample size for the SPC method reaches 910 or more, the results of SPC-based and MCS-based Sobol' methods are very similar. Fig. 14 exhibits the RMSE values of the estimates of the second order sensitivity indices calculated by SPC, GPR and PC-Kriging methods, respectively. Similarly, the SPC method shows a better convergence behavior than the GPR and PC-Kriging methods. In addition, for the training time of various methods, the time consumption of SPC is the least. For sample sets {130, 390, 650, 910, 1300}, the SPC method can save {2, 3, 2, 1, 1} times against the GPR method and {5, 5, 2, 2, 1} times against the PC-Kriging method, respectively.

The LZPK (parameter 11), which is about the lateral drainage rate of lower zone supplementary free water, has significant interaction with the LSFSM (parameter 9) and LZFPM (parameter 10), because the three parameters are about the free water in the lower zone. LZPK controls the drainage rate, while LZFPM controls the storage capacity of the lower zone primary free water (slower). LZPK also have interactions with LZFSM, which controls the storage of supplementary free water (faster). The strong interactions between the three parameters indicated that these processes are highly correlated and with only the observation of streamflow, it will be not easy to identify the values of them. The interaction between ADIMP (parameter 5) and LZTWM (parameter 8), as well as many other moderate interaction effects, indicated that in SAC-SMA there are many not overwhelmingly strong but nonnegligible interactions that make a great challenge to parameter calibration. The

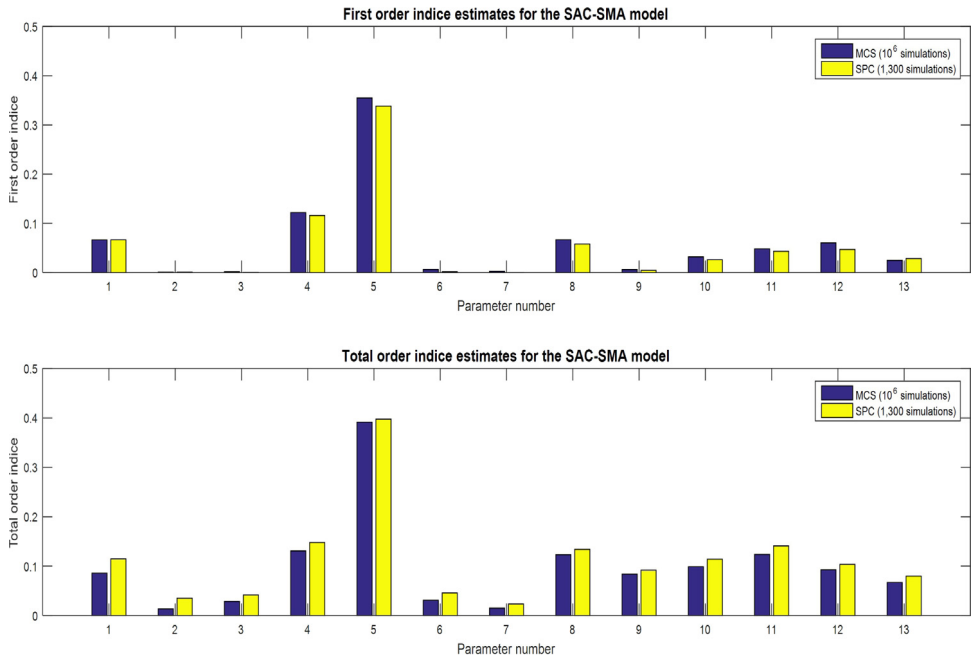

Fig. 11. The first and total order sensitivity indices obtained by MCS-based and SPC-based methods.

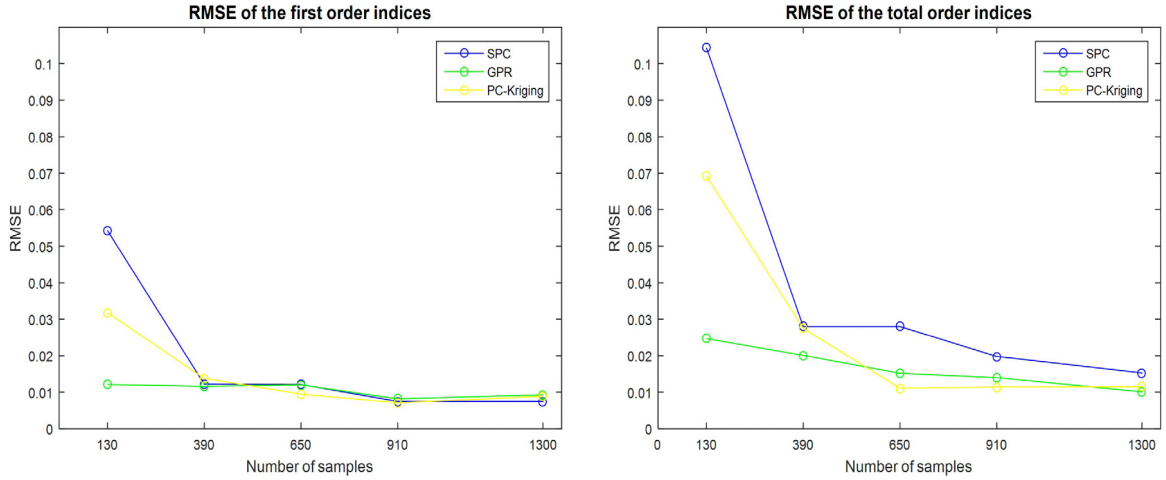

Fig. 12. The comparison of the RMSE values of the first and total order SA indices by three meta-model based SA methods.

interaction effects are the inherent reason of equifinality, and the quantification of interaction effects will provide a lot of useful information to understand the uncertainty of hydrological models.

3.3.2. *Robustness of the SPC method in calculating interaction effects*

The robustness of the SPC method using different samples sizes in obtaining reliable estimates of different SA indices is analyzed. In order to test the convergence behavior, a robustness study is designed to estimate first and second order indices for a set of increasingly larger experimental designs with a maximum of N = 2000. Confidence bounds for each estimate are calculated by means of bootstrap method (using 100 bootstrap replicates). The results of the convergence of the estimates of the first order indices for six most sensitive parameters (e.g., parameters 5, 4, 1, 8, 12 and 11) are shown in Fig. 15 (violin graph). Similarly, the Sobol' total effect indices are not shown because their convergence behavior is essentially identical to that of the first order indices. From Fig. 15, it is clear that reasonably good estimates of the first order indices are already obtained with 390 model evaluations (i.e. the sample size is 30 times as large as the number of parameters).

For the robustness of the second order SA indices estimates, the convergence behavior of the estimates of the Sobol' second order indices as a function of the experimental design sizes is shown in Fig. 16 for the three parameter pairs with the largest second order Sobol' indices (e.g., pairs {9, 11}, {10, 11} and {5, 8}). Fig. 16 shows clearly that reasonably good estimates of the second order indices can be obtained with 910 model evaluations (i.e. the sample size is 70 times as large as the number of parameters).

4. Discussion

As discussed in Razavi and Gupta (2015), the variance-based approach is the most commonly used approach that provides a global measure of interaction effects. The Sobol' type sensitivity analysis method is a variance-based SA method. However, the way that Sobol' SA indices are computed can lead to negative variances, as the formulation contains pluses and minuses terms. Due to sampling errors (especially when the sample size is relatively small), the minus terms can dominate and lead to negative values. This work conducted an

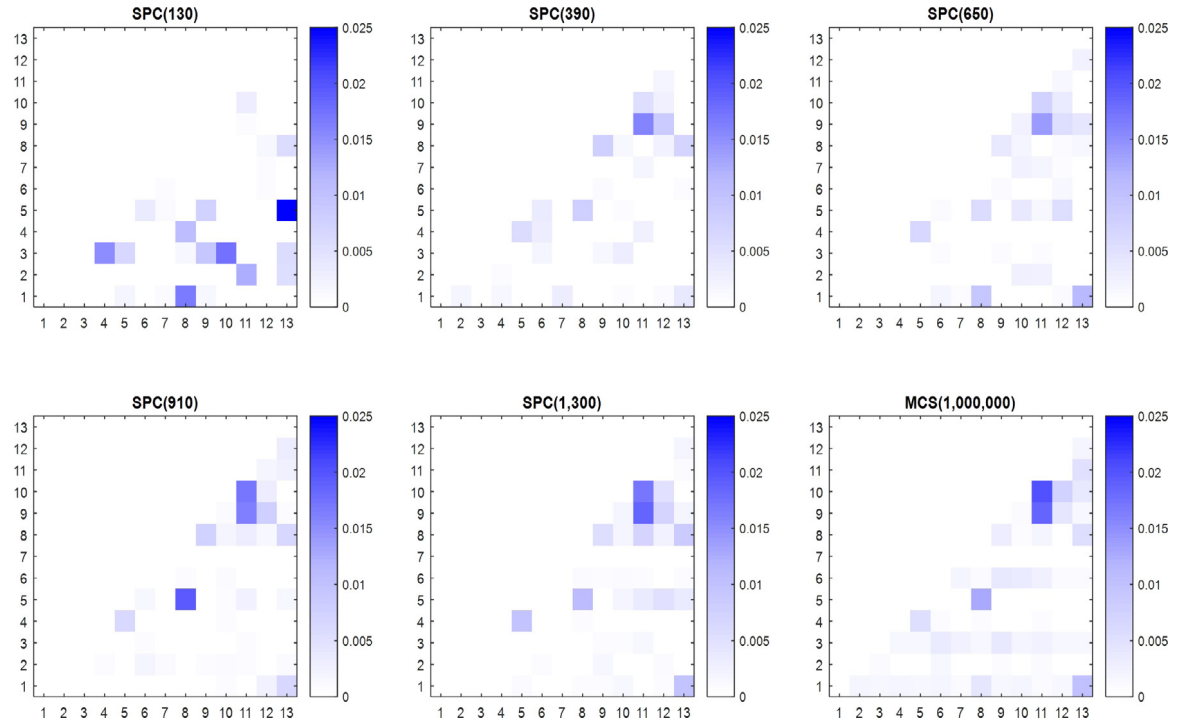

Fig. 13. The second order SA indices obtained by MCS-based with 1000,000 parameter samples and the SPC-based method with different sample sizes. The color legends and shading represent the Sobol' indices' magnitudes and ranges.

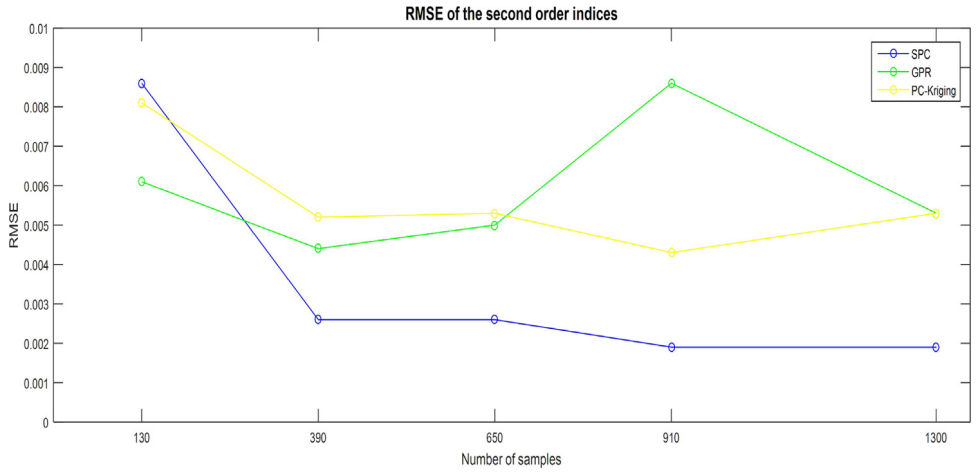

Fig. 14. The comparison of the RMSE values of the second order sensitivity indices by three meta-model based SA methods.

evaluation of parameter interaction effect based on Sobol's definition of SA indices. The second order interaction effects based on two meta-model based Sobol' methods evaluated in this study (i.e., GPR and PC-Kriging) can result in negative values in SA indices. This is not an artifact of the ANOVA decomposition, but rather is due to the sampling errors that are inherent in Monte Carlo simulations in those Sobol' methods. The SPC SA method presented in this study does not have this problem and provides more reliable estimates of the SA indices.

With the knowledge of not only the main effects but also the interaction effects, the parameters can be optimized more accurately. As in the case of the Ishigami test function in this paper, parameter x_3 is of zero main effect, but has strong interaction with parameter x_1, under which situation interaction effects should not be ignored. Shi et al. (2019) have shown that in analyzing the sensitivity of parameters of the Earth system model of intermediate complexity (LOVECLIM), the parameter interaction effects cannot be ignored in identifying the optimal model parameters. Their results showed that all ocean-related parameters have shown only a little sensitivity in the time scale of thousands of years. But as the ocean processes control the climate variability through their interactions with other earth system components such as land and atmospheric systems, the ocean-related parameters should be considered in parameter optimization. Huang et al. (2018), who examined the parameter optimization of a single column community atmosphere model, pointed out that the optimization metrics improved by 67% by considering interaction effects.

The sensitivities of model parameters are dependent on the choice of objective functions. Here the root mean square error (RMSE) is chosen as the objective function, which is commonly used in model calibration (parameter identification). The role of objective function's selection on the implementation and interpretation of parameter interaction evaluation is vital. RMSE depicts different behavior compared to other objective functions such as Nash–Sutcliffe efficiency coefficient (NSE) and Kling–Gupta efficiency coefficient (KGE) that would yield different results. In recent studies (Gupta and Razavi, 2018; Razavi and

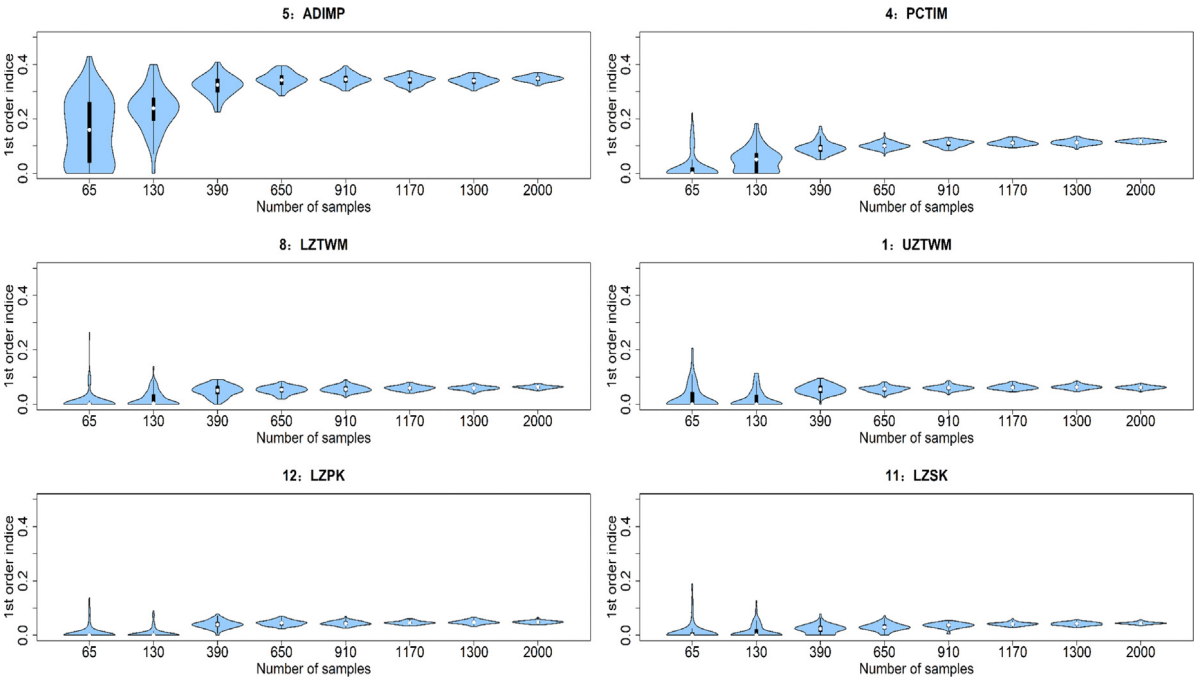

Fig. 15. Convergence of the estimates of the first order indices for different parameters. Error bounds are calculated by bootstrap resampling.

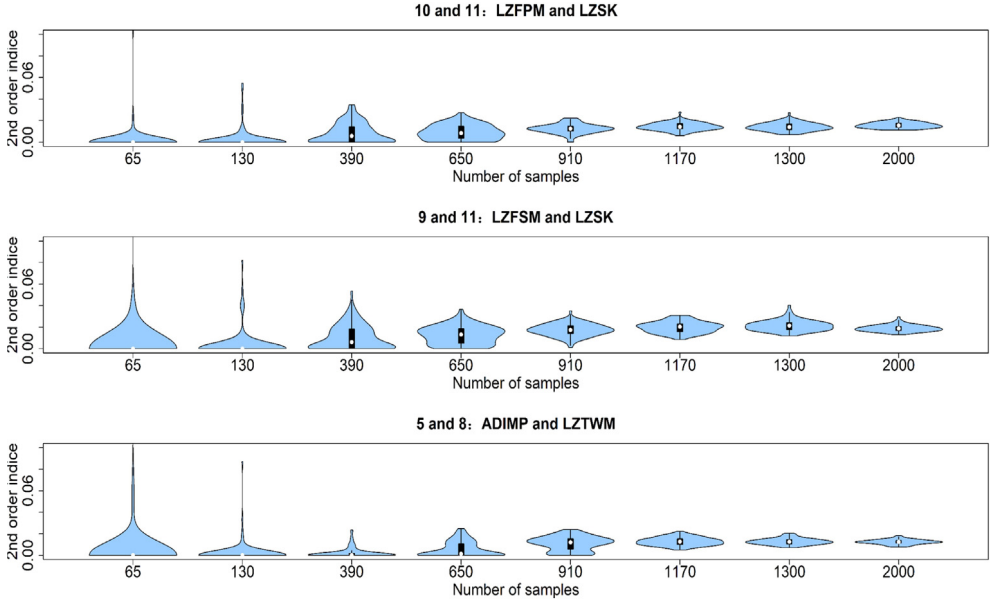

Fig. 16. Convergence of the estimates of the second order indices for different parameter pairs. Error bounds are calculated by bootstrap resampling.

Gupta, 2019), the filtering role of objective functions when used in sensitivity analysis was discussed. Because sensitivity analysis is not only for parameter optimization, it can have other applications. One of the applications is to support a better understanding of the model behaviors. For example, interaction effects can improve understanding of the interaction processes and mechanisms in model simulations.

As discussed in Razavi et al. (2012b) and Harenberg et al. (2019), due to the randomness inherent in DoEs, a robust numerical assessment is needed for a meta-modeling method. Therefore, for the robustness of the SPC method in calculating interaction effects, we apply the bootstrap resampling technique (Efron, 1979) to SPC method. Because the bootstrap resampling technique can easily be performed without adding to the overall computing cost of GSA, the efficiency analysis does not require extra model evaluations. Our results show that the SPC method needs about 300 samples (30 times the parameter dimension) and 910 samples (70 times the parameter dimension) to evaluate the interaction effects of the HBV-SASK and SAC-SMA models, respectively. The computational cost for interaction effects can depend on the number of interaction effects. For a n-dimensional model, its number of interaction effects is $n(n-1)/2$. The more the number of interaction effects, the more complex the model may be. The complexity of (the original) model structure is an important factor affecting the performance of meta-modeling. Therefore, the computational cost of interaction effects is closely related to the number of model parameters. From the results of robustness assessment, as the sample size increases, the confidence intervals of interaction effects narrow. It shows the robustness of our

Table 3
The description of the parameters of the HBV-SASK model.

No.	Random variables	Description	Range
1	TT	Air temperature threshold in °C for melting/freezing and separating rain and snow	[−4, 4]
2	CO	Base melt factor, in mm/°C per day	[0, 10]
3	ETF	Temperature anomaly correction in 1/°C of potential evapotranspiration	[0, 1]
4	LP	Limit for PET as a multiplier to FC, that is, soil moisture below which evaporation becomes supply limited	[0, 1]
5	FC	Field capacity of soil, in mm. The maximum amount of water that the soil can retain	[50, 500]
6	β	Shape parameter (exponent) for soil release equation (unitless)	[1, 3]
7	FRAC	Fraction of soil release entering fast reservoir (unitless)	[0.1, 0.9]
8	K1	Fast reservoir coefficient, which determines what proportion of the storage is released per day (unitless)	[0.05, 1]
9	α	Shape parameter (exponent) for fast reservoir equation (unitless)	[1, 3]
10	K2	Slow reservoir coefficient, which determines what proportion of the storage is released per day (unitless)	[0, 0.05]

Table 4
The description of the parameters of the SAC-SMA model.

No.	Random variables	Description	Range
1	UZTWM	Upper zone tension water maximum storage (mm)	[5.0, 300.0]
2	UZFWM	Upper zone free water maximum storage (mm)	[5.0, 150.0]
3	UZK	Upper zone free water lateral drainage rate (day^{-1})	[0.1, 0.75]
4	PCTIM	Impervious fraction of the watershed area (decimal fraction)	[0.0, 0.1]
5	ADIMP	Additional impervious area (decimal fraction)	[0.0, 0.4]
6	ZPERC	Maximum percolation rate (dimensionless)	[5.0, 350.0]
7	REXP	Exponent of the percolation equation (dimensionless)	[1.0, 5.0]
8	LZTWM	Lower zone tension water maximum storage (mm)	[10.0, 700.0]
9	LZFSM	Lower zone supplemental free water maximum storage (mm)	[5.0, 500.0]
10	LZFPM	Lower zone primary free water maximum storage (mm)	[100.0, 1200.0]
11	LZSK	Lower zone supplemental free water lateral drainage rate (day^{-1})	[0.01, 0.6]
12	LZPK	Lower zone primary free water lateral drainage rate (day^{-1})	[0.001, 0.05]
13	PFREE	Fraction of water percolating from upper zone directly to lower zone free water (decimal fraction)	[0.0, 0.9]

results. By the robustness assessment and convergence analysis, confirming that the estimation of interaction effects is accurate enough to make a solid conclusion. In addition, the impact of employing different sampling strategies on the results of sensitivity analysis is different. Here, the Sobol' sequence sampling is chosen to construct a surrogate model (Wang et al., 2014).

Finally, it must be noted that a limitation of Sobol' based method is that it only exists in the case of independent input variables. When the parameters are correlated, the so-called analysis of covariance (ANCOVA) method may be needed to generalize the Sobol' decomposition for models with correlated inputs to calculate sensitivity indices (Caniou, 2012).

5. Conclusion

In this work, the second order interaction effects of model parameters are studied. The SPC SA method was tested on four case studies, including two mathematical functions, namely the 3-Dimensional Ishigami and 30-Dimensional Sobol' G function and two hydrological models, namely 10-Dimensional HBV-SASK model and 13-Dimensional SAC-SMA model. Compared with the classical Sobol' and meta-model based Sobol' SA methods, the effectiveness and efficiency of the SPC method in the second-order sensitivity analysis are shown. The investigation we have carried out leads to reference takeaways for the analysis of hydrological models as well as for the practice of sensitivity analysis: (1) which method can calculate all the interaction effects effectively; and (2) how much computational cost is required to achieve the robust interaction effects.

When the expansion of SPC is available, the advantage of SPC method is that the Sobol' indices at any order may be computed analytically. Hence, as one of meta-model based SA, the approach not only avoids expensive computing, but also avoids sampling errors from the Monte Carlo simulation. The SPC method is shown to be one effective and efficient sensitivity analysis method for calculating interaction effects. It has much potential for further applications. For instance, the parameters may be optimized reasonably by combining main effects and interaction effects. Further studies should be as well focused on general application to other uncertainty quantification cases such as large complex dynamical system models, which would be of great significance on practical application.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the support of the National Basic Research Program of China (No. 2015CB953703), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDA19070104, XDA20060401), the State Key Laboratory of Earth Surface Processes and Resource Ecology (No. 2017-KF-05), the Fundamental Research Funds for the Central Universities-Beijing Normal University Research Fund (No. 2015KJJCA04) and the Special Fund for Meteorological Scientific Research in Public Interest (No. GYHY201506002). Mr. Heng Wang gratefully acknowledges the scholarship support provided by China Scholarship Council (No. 201806040132). We thank the editor Holger Maier for his work. We thank Razavi Saman and three anonymous reviewers for their insightful comments and suggestions that helped improve the manuscript. In this study, the experimental designs and sensitivity analysis methods we used come from UQ-PyL software package (http://www.uq-pyl.com/) and UQLab software package (https://www.uqlab.com/). The HBV-SASK rainfall-runoff model and data we used come from VARS-Tool software package (http://vars-tool.com/). The SAC-SMA rainfall-runoff model and data we used come from UQ-PyL.

Appendix A. Sobol' sensitivity analysis method

In Sobol' method (Sobol, 1993), the variance of the model output is decomposed into components that result from individual parameters as well as parameter interactions. Conventionally, the direct model output is replaced by a model performance measure such as RMSE as used in this study. The sensitivity of each parameter or parameter interaction is then assessed based on its contribution (measured as a percentage) to the total variance computed using a distribution of

model responses. Assuming the parameters are independent, the Sobol' variance decomposition is:

$$D(Y) = \sum_i D_i + \sum_{i<j} D_{ij} + \sum_{i<j<k} D_{ijk} + \cdots + D_{12\ldots n}, \tag{A.1}$$

where D_i is the measure of the sensitivity to model output y due to the ith component of the input parameter vector, D_{ij} is the portion of output variance that results due to the interaction between parameters. The variable n defines the total number of parameters. The sensitivity of parameters and interaction between parameters is obtained by normalizing the above formula,

$$1 = \sum_i \frac{D_i}{D(Y)} + \sum_{i<j} \frac{D_{ij}}{D(Y)} + \sum_{i<j<k} \frac{D_{ijk}}{D(Y)} + \cdots + \frac{D_{12\ldots n}}{D(Y)}, \tag{A.2}$$

the variance decomposition shown in Eq. (A.2) can be used to define the sensitivity indices of different orders as

$$S_i = \frac{D_i}{D}, S_{ij} = \frac{D_{ij}}{D}, S_{Ti} = 1 - \frac{D_{\sim i}}{D}, \tag{A.3}$$

where S_i is the first order sensitivity index (main effect), S_{ij} is the second order sensitivity index (interaction effect) and S_{Ti} is the total order sensitivity index (total effect).

The original Sobol' method required $n * (2m + 1)$ model runs to calculate all the first, second order and total order sensitivity indices. An enhancement of the method made by Saltelli (2002) provides the first, second and total order sensitivity indices using $n * (2m+2)$ model runs. In this study, this modified version of Sobol' methodology is chosen to compute the first order, second order and total order indices.

Appendix B. Gaussian process regression-based Sobol' sensitivity analysis method

Gaussian process regression (GPR) is a machine learning method based on statistical learning theory and Bayesian theory. A Gaussian process regression meta-model is described by the following equation (Santner et al., 2003; Rasmussen and Williams, 2006; Dubourg, 2011; Lataniotis et al., 2018)

$$Y(\boldsymbol{x}) = \boldsymbol{\beta}^T f(\boldsymbol{x}) + \sigma^2 Z(\boldsymbol{x}, \omega), \tag{B.1}$$

the first term in Eq. (B.1), $\boldsymbol{\beta}^T f(\boldsymbol{x})$, is the mean value of the Gaussian process (i.e. trend) and it consists of the regression coefficients $\{\beta_j, j = 1, \ldots, P\}$ and the basis functions $\{f_j, j = 1, \ldots, P\}$. The second term in Eq. (B.1) consists of σ^2, the variance of the Gaussian process and $Z(\boldsymbol{x}, \omega)$, a zero mean, unit variance, stationary Gaussian process. The $Z(\boldsymbol{x})$ is fully determined by the auto-correlation function between two input sample points $R(\boldsymbol{x}, \boldsymbol{x}') = R(|\boldsymbol{x} - \boldsymbol{x}'|; \boldsymbol{\theta})$ due to stationarity, where $\boldsymbol{\theta}$ are hyper-parameters to be computed.

The Gaussian assumption states that the vector formed by the true model responses, $\boldsymbol{y}$ and the prediction, $\hat{Y}(\boldsymbol{x})$, has a joint Gaussian distribution defined by

$$\begin{Bmatrix} \hat{Y}(\boldsymbol{x}) \\ \boldsymbol{y} \end{Bmatrix} \sim \mathcal{N}_{N+1}\left(\begin{Bmatrix} \boldsymbol{\beta} f(\boldsymbol{x})^T \\ \boldsymbol{\beta F} \end{Bmatrix}, \sigma^2 \begin{Bmatrix} 1 & \boldsymbol{r}^T(\boldsymbol{x}) \\ \boldsymbol{r}(\boldsymbol{x}) & \boldsymbol{R} \end{Bmatrix} \right) \tag{B.2}$$

where $\boldsymbol{F}$ is the information matrix of generic terms $F_{ij} = f_j(\boldsymbol{x_i}), i = 1, \ldots, N, j = 1, \ldots, P$. $\boldsymbol{r}(\boldsymbol{x})$ is the vector of cross-correlations between the prediction point x and each one of the observations whose terms read $\boldsymbol{r_i} = R(\boldsymbol{x}, \boldsymbol{x_i}; \boldsymbol{\theta}), i = 1, \ldots, N$. $\boldsymbol{R}$ is the correlation matrix whose terms read $R_{ij} = R(\boldsymbol{x_i}, \boldsymbol{x_j}; \boldsymbol{\theta}), i, j = 1, \ldots, N$.

Then the mean and variance of the Gaussian random variate Y(x) (a.k.a. mean and variance of the GPR predictor) can be calculated

$$\mu_{\hat{Y}}(\boldsymbol{x}) = \boldsymbol{\beta} f(\boldsymbol{x})^T + \boldsymbol{r}(\boldsymbol{x})^T \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{F\beta}) \tag{B.3}$$

$$\sigma^2_{\hat{Y}}(\boldsymbol{x}) = \sigma^2(1 - \boldsymbol{r}(\boldsymbol{x})^T \boldsymbol{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})^T (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{\mu}(\boldsymbol{x})) \tag{B.4}$$

where $\boldsymbol{\beta} = (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{y}$ is the generalized least-squares estimate of the underlying regression problem and $\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})$. Then predictions for new points can be made in terms of the mean and variance of $\hat{Y}(x)$, using Eqs. (B.3) and (B.4).

Finally, combining with Sobol' method, Gaussian process regression-based sensitivity analysis method can be carried out. The Sobol' sensitivity indices can be calculated by means of samples of input parameters and the mean values of GPR output. It needs to be estimated at the cost of $N = n * (2m + 2)$ model evaluations when using Sobol' sampling.

In this work, for the selection of correlation function and trend for Gaussian process regression, the Matérn-5/2 covariance kernel is considered with a constant yet unknown trend. The hyper-parameters are estimated with the maximum likelihood method and are solved by using the hybrid genetic algorithm. We select about 200,000 samples to calculate the sensitivity indices through Sobol' sampling method.

Appendix C. Polynomial Chaos Kriging-based Sobol' sensitivity analysis method

Kriging (a.k.a. Gaussian process regression) interpolates the local variations of Y as a function of the neighboring experimental design points, whereas polynomial chaos approximates well the global behavior of Y. By combining the global and local approximation of these techniques, the polynomial-chaos-Kriging metamodel is achieved. The polynomial-chaos-Kriging (PC-Kriging) is defined as a universal Kriging model the trend of which consists of a set of orthonormal polynomials (Kersaudy et al., 2015; Schobi et al., 2019):

$$Y(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) + \sigma^2 Z(\boldsymbol{x}, \omega), \tag{C.1}$$

where $\sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ is a weighted sum of orthonormal polynomials describing the trend of the PC-Kriging model, σ^2 and $Z(\boldsymbol{x}, \omega)$ denote the variance and the zero mean, unit variance, stationary Gaussian process, respectively, as introduced in Appendix B. Hence, PC-Kriging can be interpreted as a universal Kriging model with a specific trend.

Constructing a PC-Kriging model consists of two parts: the determination of the optimal set of polynomials contained in the trend and the calibration of the Kriging model. The two parts can be combined in various ways. Sequential PC-Kriging and optimal PC-Kriging can be implemented.

Finally, combining with Sobol' method, Gaussian process regression-based sensitivity analysis method can be carried out. The Sobol' sensitivity indices can be calculated by means of samples of input parameters and the mean values of PC-Kriging output. It needs to be estimated at the cost of $N = n * (2m + 2)$ model evaluations when using Sobol' sampling.

In this work, the sequential PC-Kriging is chosen. The optimal set of polynomials is determined by sparse polynomial chaos based on least angle regression (LAR), and the adaptive polynomial degree is set as $2 < p < 5$. For the selection of correlation function and trend for Gaussian process regression, the Matérn-5/2 covariance kernel is considered with a constant yet unknown trend. The hyper-parameters are estimated with the maximum likelihood method and are solved by using the hybrid genetic algorithm. Similarly, we select about 200,000 samples to calculate the sensitivity indices through Sobol' sampling method.

References

Blatman, G., 2009. Adaptive Sparse Polynomial Chaos Expansions for Uncertainty Propagation and Sensitivity Analysis. Bibliogr.

Blatman, G., Sudret, B., 2010. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. Probab. Eng. Mech. 25 (2), 183–197.

Blatman, G., Sudret, B., 2011. Adaptive sparse polynomial chaos expansion based on least angle regression. J. Comput. Phys. 230 (6), 2345–2367.

Brazil, L., 1988. Multilevel Calibration strategy for complex hydrologic simulation models (Ph.D. thesis). Department of Civil Engineering, Colorado State University, Fort Collins, CO.

Burnash, R., Ferral, R., McGuire, R., 1973. A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers. US Department of Commerce, National Weather Service, Sacramento, CA.

Caniou, Y., 2012. Global sensitivity analysis for nested and multiscale models (Ph.D. thesis). Universite Blaise Pascal, Clermont-Ferrand.

Doostan, A., Owhadi, H., 2011. A non-adapted sparse approximation of PDEs with stochastic inputs. J. Comput. Phys. 230 (8), 3015–3034.

Duan, Q., Schaake, J., et al., 2006. Model parameter estimation experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. J. Hydrol. 320 (1–2), 3–17.

Dubourg, V., 2011. Adaptive surrogate models for reliability analysis and reliability-based design optimization (Ph.D. thesis). Universite Blaise Pascal, Clermont-Ferrand, France.

Efron, B., 1979. Bootstrap methods: another look at the Jackknife. Ann. Statist. 7 (1), 1–26, 15.

Fajraoui, N., Mara, T., Younes, A., Bouhlila, R., 2012. Reactive transport parameter estimation and global sensitivity analysis using sparse polynomial chaos expansion. Water Air Soil Pollut. 223 (7), 4183–4197.

Friedman, J., 1991. Multivariate adaptive regression splines. Ann. Statist. 19 (1), 1–67.

Gan, Y., Duan, Q., Gong, W., et al., 2014. A comprehensive evaluation of various sensitivity analysis methods: a case study with a hydrological model. Environ. Model. Softw. 51 (1), 269–285.

Gan, Y., Liang, X., Duan, Q., et al., 2017. A systematic assessment and reduction of parametric uncertainties for a distributed hydrological model. J. Hydrol. 564, 697–711.

Gong, W., Duan, Q., Li, J., et al., 2015. Multi-objective parameter optimization of common land model using adaptive surrogate modeling. Hydrol. Earth Syst. Sci. 11 (5), 2409–2425.

Gong, W., Duan, Q., Li, J., et al., 2016. Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical models. Water Resour. Res. 52 (3).

Gupta, H., Razavi, S., 2018. Revisiting the basis of sensitivity analysis for dynamical earth system models. Water Resour. Res. 54 (11), 8692–8717.

Harenberg, D., Marelli, S., Sudret, B., Winschel, V., 2019. Uncertainty quantification and global sensitivity analysis for economic models. Quant. Econ. 10 (1), 1–41. http://dx.doi.org/10.3982/QE866.

Hu, J., Zhang, S., 2016. Global sensitivity analysis based on polynomial chaos. Computational Physics 33 (1), 1–14.

Huang, X., Li, Q., Luo, Y., et al., 2018. An effective parameter screening method based on sensitivity analysis. Appl. Electron. Tech. 44 (12), 106–110.

Ishigami, T., Homma, T., 1990. An importance quantification technique in uncertainty analysis for computer models. In: Proceedings of the ISUMA'90, First International Symposium on Uncertainty Modelling and Analysis, pp. 398-403.

Kersaudy, P., Sudret, B., Varsier, N., Picon, O., Wiart, J., 2015. A new surrogate modeling technique combining Kriging and polynomial chaos expansions-Application to uncertainty analysis in computational dosimetry. J. Comput. Phys. 286, 103–117.

Lataniotis, C., Marelli, S., Sudret, B., 2018. UQLab User Manual-Kriging (Gaussian Process Modelling). Report UQLab-V1.1-105, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich.

Li, J., Duan, Q., Gong, W., et al., 2013. Assessing parameter importance of the common land model based on qualitative and quantitative sensitivity analysis. Hydrol. Earth Syst. Sci. 17 (8), 3279–3293.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergstrom, S., 1997. Development and test of the distributed HBV-96 hydrological model. J. Hydrol. 201 (1–4), 272–288.

Marelli, S., Sudret, B., 2014. UQLab: A framework for uncertainty quantification in Matlab. In: Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom, pp. 2554–2563.

Montgomery, D., 2004. Design and Analysis of Experiments. John Wiley and Sons, New York.

Morris, M., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol' sensitivity analysis of a complex environmental model. Environ. Model. Softw. 26 (12), 1515–1525.

Rasmussen, C., Williams, C., 2006. Gaussian Processes for Machine Learning. MIT Press.

Razavi, S., Gupta, H., 2015. What do we mean by sensitivity analysis? The need for comprehensive characterization of global sensitivity in earth and environmental systems models. Water Resour. Res. 51, 3070–3092.

Razavi, S., Gupta, H., 2019. A multi-method generalized global sensitivity matrix approach to accounting for the dynamical nature of earth and environmental systems models. Environ. Model. Softw. 114, 1–11.

Razavi, S., Sheikholeslami, R., Gupta, H., 2019. VARS-tool: a toolbox for comprehensive, efcient, and robust global sensitivity analysis. Environ. Model. Softw. 112, 95–107.

Razavi, S., Tolson, B., Burn, D., 2012a. Review of surrogate modeling in water resources. Water Resour. Res. 48 (7).

Razavi, S., Tolson, B., Burn, D., 2012b. Numerical assessment of metamodelling strategies in computationally intensive optimization. Environ. Model. Softw. 34, 67–86.

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Comm. 145, 280–297.

Saltelli, A., Sobol', I., 1995. About the use of rank transformation in sensitivity analysis of model output. Reliab. Eng. Syst. Saf. 50, 225–239.

Santner, T., Williams, B., Notz, W., 2003. The design and analysis of computer experiments. In: The Design and Analysis of Computer Experiments. Springer.

Sarrazin, F., Pianosi, F., Wagener, T., 2016. Global sensitivity analysis of environmental models: Convergence and validation. Environ. Model. Softw. 79, 135–152.

Schobi, R., Marelli, S., Sudret, B., 2019. UQLAB user manual-Polynomial chaos Kriging. Report UQLab-V1.2-109, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich.

Schoups, G., d, Van., Savenije, H., 2008. Model complexity control for hydrologic prediction. Water Resour. Res. 44 (1), 370–380.

Shao, Q., Younes, A., Fahs, M., Mara, T., 2017. Bayesian sparse polynomial chaos expansion for global sensitivity analysis. Comput. Methods Appl. Mech. Engrg. 318, 474–496.

Sheikholeslami, R., Razavi, S., Gupta, H.V., Becker, W., Haghnegahdar, A., 2019. Global sensitivity analysis for high-dimensional problems: How to objectively group factors and measure robustness and convergence while reducing computational cost. Environ. Model. Softw. 111, 282–299.

Shi, Y., Gong, W., Duan, Q., et al., 2019. How parameter specification of an earth system model of intermediate complexity influences its climate simulations. Progr. Earth Planetary Sci. http://dx.doi.org/10.1186/s40645-019-0294-x.

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

Song, X., Kong, F., Zhan, C., Han, J., 2012a. Sensitivity analysis of hydrological model parameters using a statistical theory approach. Adv. Water Sci. 23 (5), 642–649.

Song, X., Zhan, C., Xia, J., Kong, F., 2012b. An efficient global sensitivity analysis approach for distributed hydrological model. J. Geogr. Sci. 22 (2), 209–222.

Song, X., Zhang, J., Zhan, C., et al., 2015. Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. J. Hydrol. 523 (523), 739–757.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. 93 (7), 964–979.

Tang, Y., Reed, P., Wagener, T., et al., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrol. Earth Syst. Sci. 3 (6), 793–817.

Tang, T., Zhou, T., 2015. Recent developments in high order numerical methods for uncertainty quantification. Sci. Sin. Math. 45, 891–928 (in Chinese).

Wan, H., Xia, J., Zhang, L., et al., 2015. Sensitivity and interaction analysis based on sobol' method and its application in a distributed flood forecasting model. Water 7 (12), 2924–2951.

Wang, C., Duan, Q., Gong, W., et al., 2014. An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. Environ. Model. Softw. 60 (76), 167–179.

Wang, C., Duan, Q., Tong, C., et al., 2016. A GUI platform for uncertainty quantification of complex dynamical models. Environ. Model. Softw. 76, 1–12.

Xiu, D., 2010. Numerical methods for stochastic computations: a spectral method approach. Commun. Comput. Phys. 5 (2–4), 242–272.

Xiu, D., Karniadakis, G., 2003. The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. 24 (2), 619–644.

Zhan, C., Song, X., Xia, J., et al., 2013. An efficient integrated approach for global sensitivity analysis of hydrological model parameters. Environ. Model. Softw. 41 (41), 39–52.