



Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
 www.scientiairanica.com



Proposing a global sensitivity analysis method for linear models in the presence of correlation between input variables

Y. Daneshbod and M.J. Abedini*

Department of Civil and Environmental Engineering, Shiraz University, Shiraz, Iran.

Received 25 June 2012; received in revised form 21 June 2015; accepted 15 September 2015

KEYWORDS

Sensitivity analysis;
 OAT;
 Correlation;
 Local SA;
 Monte Carlo
 simulation.

Abstract. Sensitivity analysis is considered to be an important part of evaluating the performance of mathematical or numerical models. One-factor-at-a-time (OAT) and differential methods are among the most popular Sensitivity Analysis (SA) schemes employed in the literature. Two major limitations of the above methods are lack of addressing the correlation between model factors and being a local method. Given these limitations, its extensive use among modelers raises concern over the credibility of the associated sensitivity analyses. This paper proposes proof for the inefficiency of the aforementioned methods drawn from experimental designs, and provides a novel technique based on Principal Component Analysis (PCA) to address the issue of the correlation between input factors. In addition, proper guidelines are suggested to handle other conditions.

© 2016 Sharif University of Technology. All rights reserved.

1. Introduction

Sensitivity Analysis (SA) is the study of how the change in the output of a model can be attributed, qualitatively or quantitatively, to variation of different input variables, and of how the given model reacts upon the information passed into it. Based on this description, it is safe to conclude that SA is a necessary ingredient of model building in any setting, either diagnostic or prognostic, and in any discipline where models are called upon for design purposes [1]. Though this is a correct definition of SA, one may make false conclusion of thinking that SA is a tool specific for modelers. As a matter of fact, engineers as an end user of the developed models can benefit from it as well. For example, consider an environmental engineer who intends to apply rainfall-runoff model for a real-life problem. If this engineer knows which model input

parameter has more impact on model output results, a wise decision would be to spend much of the time and capital to define carefully the most important parameters to predict the runoff, precisely. Similar examples in engineering are easy to find where SA could be helpful in decision-making.

There are a number of different methods of sensitivity analysis with each method having a unique flexibility and functionality. As a result, different scientists have different ideas about categorizing SA methods. Frey and Patil [2] concentrate on methodology, classifying sensitivity analysis into mathematical, statistical, and graphic methods, while Ascough et al. [3] focus on capability, adopting the following classifications; screening, local, and global SA methods. From a functional standpoint, sensitivity analysis can be divided into local and global SAs or, according to Morgan's definition, into deterministic and probabilistic [3,4].

Local SA concentrates on the local impact of the factors on the model. Local SA is usually carried

*. Corresponding author.

E-mail address: abedini@shirazu.ac.ir (M.J. Abedini)

out by computing partial derivatives of the output function with respect to the input variables. The term ‘local’ means that all derivatives are taken at a unique and well-defined point, known as ‘baseline value’ or ‘nominal value’, in the domain of the input factors. One can see local SA as a particular case of one-factor-at-a-time (OAT) method, whereby one factor is varied while all others are held constant at their respective nominal values. In contrast, the global sensitivity techniques examine the global response (averaged over the variation of all the parameters) of model output(s) by exploring a finite (or even an infinite) region. Global sensitivity analyses are shown to be more effective when the predictor variable is influenced by simultaneous changes in seemingly independent variables. This simultaneous variation would take into account the interaction among factors. Though global SA methods enjoy not having any of the local methods’ limitations, they are much more complex and computationally expensive. The difference between local and global SAs can be best examined in a simple nonlinear example. Let us examine the sensitivity of independent variables of X_1 and X_2 in a nonlinear function, $Y = X_1^2 + X_2$. For local SA, the sensitivity index of the two independent variables changes as the SA domain changes. The dependent variable is more sensitive to the variable X_2 in the domain belonging to $[0 - 1]$, while for all other values, X_1 is the dominant variable. In global SA, regardless of domain, sensitivity index for X_1 is always greater.

In theory, local sensitivity analyses cannot be used to uncover the robustness of model based on inference unless the model is proven to be linear for the case of first-order derivatives or, at least, additive for the case of higher and mixed-order derivatives. In other words, derivatives are informative at the nominal point where they are computed, but do not provide any additional information for the rest of the domain of input factors, except when some conditions, such as linearity or additivity, are provided in the mathematical model [5]. When the properties of the models are not known in advance, a global SA is recommended. This is why global SA methods are often called model-free by practitioners.

By reviewing recent literature associated with SA, it can be said that practitioners have acknowledged the importance of SA in approving or refuting a model-based analysis. Yet, regardless of its limitations, most of the sensitivity analyses were performed using an OAT approach [6]. Speaking of limitations, there are two concerns regarding OAT approach. First, OAT is non-explorative, which means as the number of the factors increases, the exploration feature decreases. As an example in Figure 1, it can be seen, clearly, that OAT explores only 5 points forming a cross, out of total 9 points [7]. By generalizing this example with a

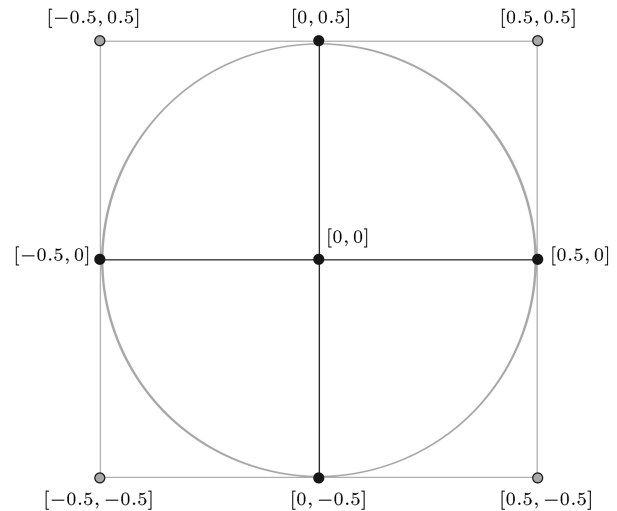


Figure 1. The OAT method can only reach few points of the entire domain.

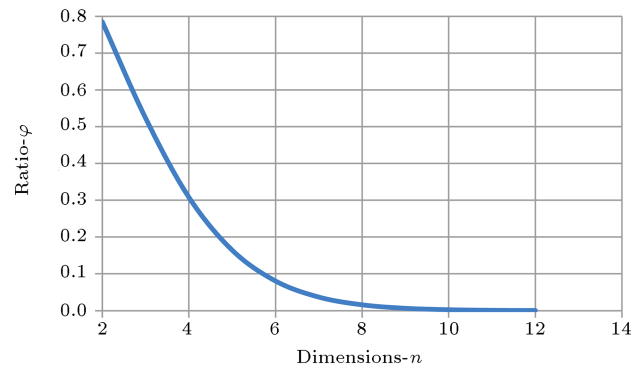


Figure 2. The curse of dimensionality; the volume ratio, φ , of hyper-sphere to hypercube decreases quickly to zero as n number of dimensions increase.

geometrical explanation of the functional domain in n -dimensions, the n -cross will unavoidably be inscribed in a hyper-sphere. The problem is that this hyper-sphere represents a small percentage of the total functional domain defined by the hypercube. This is illustrated in Figure 1, where the explored cross is inscribed in the circle of center $[0, 0]$ and radius of 0.5. In this 2-dimensional example, the ratio of the partially explored area to the total area is $\varphi = 0.78$. The exploratory domain quickly decreases as the number of dimensions increases (Figure 2). In 3 dimensions, it is $\varphi = 0.52$.

The general formula for the volume of the hyper-sphere of radius R in n dimensions is [8]:

$$V_n(R) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} R^n \text{ or } V_n(R) = C_n R^n, \quad (1)$$

where Γ is the gamma function. For even n , C_n reduces to $C_n = \frac{\pi^{\frac{n}{2}}}{(\frac{n}{2})!}$ and since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, for odd n , $C_n = \frac{2^{\frac{n+1}{2}} \pi^{\frac{n-1}{2}}}{n!!}$, where $n!!$ denotes double factorial.

For a hyper-sphere with $R = 0.5$ and $n = 13$, the

ratio is $\varphi = 0.000197$. In the literature, this is known as curse of dimensionality or OAT paradox [5,9].

This limitation has been addressed by several practitioners; Saltelli [10] believes that “A good OAT would be one where after having moved one step in one direction, say along X_1 , one would directly take another step along X_2 and so on, until all factors up to X_k are moved of one step each during the first stage.” This type of trajectory is known as ‘Elementary Effects’ (EE) [10].

Another shortcoming related to OAT is that in the presence of correlation, OAT cannot detect interactions between factors, because this feature is based on the simultaneous movements of more than one factor. If factors move one at a time, the interactions are not triggered and, subsequently, not detected [5].

So, why modelers still prefer OAT? Arguments which might justify nominating OAT as a favorite candidate are:

- (a) The SA is carried out in reference to baseline value, which is usually the most probable occurrence of the model in real world;
- (b) This method requires much less computational efforts (i.e., CPU time) than any global SA method;
- (c) Movement of one factor at a time means that whatever effect is observed on the output is due, exclusively, to that factor [5];
- (d) It never assesses a non-important factor as important. This is also known as type I error [5];
- (e) Since it deals with change of one parameter at a time, the chance of model failure decreases, as opposed to changing all factors. In case of model failure, the approach helps modeler to identify the factor responsible for model crash [5];
- (f) The OAT approach is consistent with the modeler’s way of changing one parameter at a time as he/she wants to verify, systematically, the effect of parameter variation.

To the best of the authors’ knowledge, there is no deterministic SA method to address the interaction between factors in a comprehensive way (Morris’s method is known to be global [1]). All of the SA methods, which produce reliable ranking in presence of correlation, are of probabilistic nature. These methods are not favored by end users due to their complexity and high computational cost. There is an ample need to have a simple method with no computational burden, yet powerful enough to address the correlation issue among the input variables involved. This study proposes a methodology based on standard OAT to extend its capability to address one of the limitations, which has long been subjected to extensive criticism.

Beyond the limitations of local SA methods, there are still situations where these methods fail to determine the correct order of importance among variables. Traditional modelers usually fail to notice them and perform perfunctory SA, clearly raising concerns regarding the developed models. This research also tends to address the common mistakes that may occur while performing OAT and differential SA for linear models. Some guidelines will be suggested so that practitioners avoid these common pitfalls of the discussed SA methods.

2. Methodology

The most popular SA methods among scientists are differential and OAT. Both are local SA methods, which explore sensitivity of the model around a narrow region of the feature space. First, methods are briefly explained in the following sections. Later on, the pitfalls of the methods are revealed by performing a series of experiments. Suitable modifications are introduced so that the practitioners avoid the common mistakes.

The benchmark for all of the experiments is Monte Carlo-based global SA, wherein 1000 samples are generated using random sampling. For examples, with correlated variables, the procedure proposed by Iman and Conover’s method [11] is used to generate correlated samples. The SA index is Pearson product moment correlation coefficient (PEAR), which is the usual linear correlation coefficient computed among a pair of variables. For comparison purposes, the SA measures resulted from different techniques are normalized.

2.1. Standard OAT-SA

There are different flavors of OAT according to Daniel [12], OAT methods can be classified into five types:

- (a) Strict OAT: This perturbs one factor from the condition of the last preceding model run;
- (b) Paired OAT: This produces two observations and, therefore, one simple comparison at a time;
- (c) Free OAT: In this type, each run can be performed under a new condition;
- (d) Curved OAT: This produces a subset of results by perturbing only one easy-to-vary factor while others are held constant;
- (e) Standard OAT: This perturbs one factor from a standard condition.

This research focuses on standard OAT, in which the standard condition is usually the average of input factors over their domain. This type of SA only addresses the sensitivity of model output relative to

Table 1. Summary of sensitivity indexes.

Factors	Differential	OAT	Benchmark	<i>n</i> -differential ¹
X_1	1.000	0.555	0.499	0.555
X_2	0.400	0.666	0.705	0.666
X_3	0.200	1.000	1.000	1.000

¹ Normalized differential method.

a point where the parameters are held constant. The sensitivity measure is determined by calculating the ratio of model output while perturbing each parameter's value about the baseline value, in turn, to the nominal value of model output. The amount of parameter perturbation can either be a plus/minus percentage of their baseline value or one standard deviation of their distribution. The latter case has the advantage of taking into account the parameter's variability and the associated influence on model output [13]. The main advantage of such methods is their low computational cost. Considering two values for each perturbed factor and an extra run for the nominal value of the model output, a model with K parameters requires $2K + 1$ runs.

2.2. Differential SA

Differential analysis, also known as the direct method, is based on the behavior of the model while all parameters are set to their mean value. It is based on partial differentiation of the combined model. When a model is described by an explicit, closed algebraic equation, the sensitivity measure for a specific independent variable is estimated by partial derivative of the dependent variable with respect to the independent variable [13]. In the following subsections, a few case studies are introduced to highlight the pitfalls associated with traditional methods and, subsequently, a remedial measure is suggested for each approach.

2.3. Case 1: Where differential SA fails

A simple test model [1] shows that partial derivative of the output Y with respect to an input factor X_i at a fixed point fails to properly address the sensitivity of linear models. This study assumes that X_1 , X_2 , and X_3 (the independent variables) are mutually independent and uniformly distributed in their respective range. The functional relationship between the dependent variable Y and the independent variables is:

$$Y = \sum_{i=1}^n a_i X_i, \quad (2)$$

where $X_i \sim U(\bar{x}_i - \sigma_i, \bar{x}_i + \sigma_i)$, $\bar{x}_i = 3^{i-1}$, and $\sigma_i = 0.5\bar{x}_i$ are uniformly distributed in their range. Taking $k = 3$, the independent variables, constant coefficients, and their computational domains are computed as follow:

$$X_1 \sim U(0.5, 1.5),$$

$$X_2 \sim U(1.5, 4.5),$$

$$X_3 \sim U(4.5, 13.5), \quad (3)$$

$$Y = 5X_1 + 2X_2 + X_3. \quad (4)$$

Now, let us discuss the sensitivity of the dependent variable Y to each independent variable. According to the sensitivity analysis conducted based on the differential method [14], all partial derivatives are constant as independent variables are mutually independent:

$$\frac{\partial Y}{\partial X_1} = 5, \quad \frac{\partial Y}{\partial X_2} = 2, \quad \frac{\partial Y}{\partial X_3} = 1. \quad (5)$$

The sensitivities of the dependent variable to the independent variables are ranked as X_1 , X_2 , and X_3 , and the value of each independent variable has no impact on the sensitivity coefficient, i.e. ranking. Sensitivity analysis for the test case is carried out by two methods based on OAT and Monte Carlo sampling techniques. The results are summarized in Table 1.

As can be seen, the differential method fails to rank the factors correctly. Even the simple OAT method is able to capture the ranks correctly. This is due to the fact that the differential method does not consider the variation of input variables. An available practice to address these misleading results is to normalize the derivatives by standard deviations (reference). For a general linear model of Eq. (2), the sensitivity measure is:

$$S_{X_i}^\sigma = (\sigma_{X_i/\sigma_Y}) (\partial Y / \partial X_i). \quad (6)$$

Since for a linear model $\sum_{i=1}^n (S_{X_i}^\sigma)^2 = 1$, $(S_{X_i}^\sigma)^2$ gives how much each individual factor contributes to the variance of the output of interest. This turns $S_{X_i}^\sigma$ to be a hybrid local-global measure. After incorporating the suggested modification to the test case, the results improve as can be seen in Table 1. Though, quantitatively, it is not in agreement with the benchmark values, it is now in a correct order.

2.4. Case 2: OAT and the range effect

Due to the local nature of OAT, the range and distribution of factors have no effect on the results obtained by the method. This will raise concerns when the differences between ranges of factors are significant or

Table 2. Summary of sensitivity measures.

Factors	S-OAT	Benchmark	OAT-SD ¹
X_1	1	0.155	0.117
X_2	0.8	1	1

¹ OAT standard deviation perturbation.

the factors have different distributions. A test case has been set up to demonstrate this shortcoming.

$$Y = X_1 + X_2, \quad (7)$$

$$X_1 \sim U(9, 11),$$

$$X_2 \sim U(1, 15). \quad (8)$$

As it can be seen in Table 2, OAT method fails to rank the factors properly. This is because the sensitivity measure was determined by variation of input parameters with $\pm 20\%$. This methodology is unable to capture the effects of the computational domain and distribution of factor. One way to overcome this limitation is to vary the input factors by one standard deviation of their input distribution rather than by 20%. The SA for Eq. (7) is repeated with the mentioned modification. Results have been improved. As can be seen, it is quite consistent with the benchmark ranking.

2.5. Case 3: The correlation effect (multi-collinearity)

When a question arises about which component of a linear model has more contribution to the model output, any SA method discussed earlier, depending on its conditions, is applicable. But when the linear independent assumptions are contravened, there will be doubt on the interpretability of regression coefficients, which were decisive in earlier methods. Thus, performing SA on a linear model based on an ordinary least square will produce unlikely results.

A workaround to the stated problem is to substitute least square with a reliable technique to build the desired linear model so that standardized regression coefficients can be used for SA. The suggested technique is based on component analysis, which converts a set of correlated variables into a set of linearly uncorrelated variables by an orthogonal transformation. The method will be discussed in detail in the following section.

As the first step, we state our linear model of Eq. (2) as a relationship of y to φ .

$$y = \alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \alpha_3 \varphi_3 + \dots + \alpha_k \varphi_k, \quad (9)$$

in which φ_k are a set of orthogonal variates (principal components) from linear combination of x_i s, where $k < n$; k is the number of principal components; y and x_i

are standardized values of the original variables (Y, X_i) with zero mean and unit variance:

$$\varphi_1 = m_{11}x_1 + m_{21}x_2 + m_{31}x_3 + \dots + m_{n1}x_n,$$

$$\varphi_2 = m_{12}x_1 + m_{22}x_2 + m_{32}x_3 + \dots + m_{n2}x_n,$$

$$\varphi_3 = m_{13}x_1 + m_{23}x_2 + m_{33}x_3 + \dots + m_{n3}x_n,$$

$$\vdots$$

$$\varphi_k = m_{1k}x_1 + m_{2k}x_2 + m_{3k}x_3 + \dots + m_{nk}x_n, \quad (10)$$

where m_{ij} is a typical i component of j eigenvector corresponding to the predictor variates correlation matrix.

Substituting Eq. (10) into Eq. (9) results in:

$$\begin{aligned} y = & (\alpha_1 m_{11} + \alpha_2 m_{12} + \alpha_3 m_{13} + \dots + \alpha_k m_{1k}) x_1 \\ & + (\alpha_1 m_{21} + \alpha_2 m_{22} + \alpha_3 m_{23} + \dots + \alpha_k m_{2k}) x_2 \\ & + (\alpha_1 m_{31} + \alpha_2 m_{32} + \alpha_3 m_{33} + \dots + \alpha_k m_{3k}) x_3 + \dots \\ & + (\alpha_1 m_{n1} + \alpha_2 m_{n2} + \alpha_3 m_{n3} + \dots + \alpha_k m_{nk}) x_n. \end{aligned} \quad (11)$$

If the numerical values of the parameter in the parenthesis could be calculated, Eq. (11) expresses a linear relation between y and x_i . We have already discussed how to compute m_{ij} . As for α_i , Eq. (9) is a regression equation with independent variables, α_k could be computed by ordinary least square. Simultaneous normal equations resulted from OLS are given by:

$$\begin{aligned} & \alpha_1 \sum \varphi_1^2 + \alpha_2 \sum \varphi_1 \varphi_2 + \alpha_3 \sum \varphi_1 \varphi_3 \\ & + \dots + \alpha_k \sum \varphi_1 \varphi_k = \sum \varphi_1 y, \\ & \alpha_1 \sum \varphi_1 \varphi_2 + \alpha_2 \sum \varphi_2^2 + \alpha_3 \sum \varphi_2 \varphi_3 \\ & + \dots + \alpha_k \sum \varphi_2 \varphi_k = \sum \varphi_2 y, \\ & \alpha_1 \sum \varphi_1 \varphi_3 + \alpha_2 \sum \varphi_3 \varphi_2 + \alpha_3 \sum \varphi_3^2 \\ & + \dots + \alpha_k \sum \varphi_3 \varphi_k = \sum \varphi_3 y, \\ & \vdots \\ & \alpha_1 \sum \varphi_1 \varphi_k + \alpha_2 \sum \varphi_k \varphi_2 + \alpha_3 \sum \varphi_k \varphi_3 \\ & + \dots + \alpha_k \sum \varphi_k^2 = \sum \varphi_k y. \end{aligned} \quad (12)$$

Since φ_i are orthogonal, all the $\sum \varphi_i \varphi_k$ terms with $i \neq k$ are zero. An important property of φ_i s is that their

respective variances are equal to eigenvalue λ_i , which in turn is equivalent to $\sum \varphi_i^2$. After some manipulation, α_k are:

$$\begin{aligned}\alpha_1 &= \frac{\sum \varphi_1 y}{\sum \varphi_1^2} = \frac{\sum \varphi_1 y}{\lambda_1}, \\ \alpha_2 &= \frac{\sum \varphi_2 y}{\sum \varphi_2^2} = \frac{\sum \varphi_2 y}{\lambda_2}, \\ \alpha_3 &= \frac{\sum \varphi_3 y}{\sum \varphi_3^2} = \frac{\sum \varphi_3 y}{\lambda_3}, \\ &\vdots \\ \alpha_k &= \frac{\sum \varphi_k y}{\sum \varphi_k^2} = \frac{\sum \varphi_k y}{\lambda_k}.\end{aligned}\quad (13)$$

Multiplying both sides of Eqs. (11) by y and adding them up results in:

$$\begin{aligned}\sum \varphi_i y &= m_{1k} \sum x_{1y} + m_{2k} \sum x_{2y} + m_{3k} \sum x_{3y} \\ &+ \dots + m_{nk} \sum x_{ny}.\end{aligned}\quad (14)$$

Making use of Eqs. (14) and (15), and noting that for standardized variate, $\sum x_i y$ equals their correlation coefficient $r_{x_i y}$, as can be simply computed from the following equations:

$$\begin{aligned}\alpha_1 &= (1/\lambda_1)(m_{11}r_{x_1 y} + m_{21}r_{x_2 y} + m_{31}r_{x_3 y} \\ &+ \dots + m_{n1}r_{x_n y}), \\ \alpha_2 &= (1/\lambda_2)(m_{12}r_{x_1 y} + m_{22}r_{x_2 y} + m_{32}r_{x_3 y} \\ &+ \dots + m_{n2}r_{x_n y}), \\ \alpha_3 &= (1/\lambda_3)(m_{13}r_{x_1 y} + m_{23}r_{x_2 y} + m_{33}r_{x_3 y} \\ &+ \dots + m_{n3}r_{x_n y}), \\ &\vdots \\ \alpha_k &= (1/\lambda_k)(m_{1k}r_{x_1 y} + m_{2k}r_{x_2 y} + m_{3k}r_{x_3 y} \\ &+ \dots + m_{nk}r_{x_n y}).\end{aligned}\quad (15)$$

Now that the coefficients of the linear model in Eq. (11) can be calculated, these coefficients can be attributed to their sensitivity ranking.

It is worth mentioning that selecting the right number of principal components in Eq. (10), i.e. k , is an important step in this method. Small eigenvalues corresponding to the last few principal components

cause high variance in regression coefficient, which, in turn, results in unstable coefficients [15]. Repeating the calculation similar to Eq. (11) by using all of the λ_i s is equivalent to performing OLS (ordinary least square), for which its consequences were mentioned and assessed earlier. There are different criteria to choose with regard to which and how many principal components should be retained. We chose a criterion whereby the set of largest k contributors were selected to achieve and meet the following inequality:

$$1.0 > \frac{\sum_{i=1}^k \lambda_i}{n} \geq 0.85. \quad (16)$$

Note that λ_i is the eigenvalues and n is total number of eigenvalues.

To justify the robustness of the proposed method, a linear model is fit on rainfall-runoff data for the White Hollow watershed from years 1935 to 1959 presented in Table 3 [16]. The rainfall data from three consecutive months October, November, and December of each year and the associated December runoff are considered. Furthermore, an assumptive weighted cross product of rainfall for the given months is introduced into the model as a strong correlated predictive variable.

The correlation matrix for the predictive variables is as follows:

$$\begin{bmatrix} 1 & 0.14 & 0.17 & 0.70 \\ 0.14 & 1 & -0.06 & 0.69 \\ 0.17 & -0.06 & 1 & 0.34 \\ 0.70 & 0.69 & 0.34 & 1 \end{bmatrix}$$

As can be seen, there is a strong positive correlation between variable 4 and others.

Three SA methods are preformed: differential analysis on OLS, the proposed method, and the benchmark. The results are summarized in Table 4.

As can be seen in Table 4, differential analysis fails to capture the interaction between the factors when the model is based on OLS. As a result, sensitivity measure is not consistent with the benchmark ranking. On the other hand, the proposed method demonstrates promising results, which are in good agreement with benchmark results. The main feature of this method that may have a great appeal for practitioners is that it can be performed systematically with much low computational burden compared to Monte Carlo-based simulation methods.

3. Conclusion

Complex SA methods probably redound to more accurate conclusions, but their complexity discourages the modelers from implementing them; so, they are lured to use less credible methods such as OAT and differential

Table 3. Monthly rainfall-runoff, white hollow watershed.

Year	Rainfall, inches			Weighted product	Runoff, inches December
	October	November	December		
	X_1	X_2	X_3	X_4	Y
1935	2.17	7.35	1.81	18.31	0.76
1936	8.17	2.42	2.61	32.6	1.21
1937	3.27	1.56	6.08	17.04	1.05
1938	2.62	4.98	0.16	13.31	0.57
1939	2.89	1.49	0.86	5.81	0.34
1940	2.55	2.29	2.01	8.92	0.41
1941	2.6	2.62	3.01	11.52	0.48
1942	8.68	1.79	1.77	24.74	2.62
1943	2.21	1.77	3.18	8.13	0.54
1944	3.86	3.78	0.92	16.71	1.49
1945	4.31	4.83	2.32	26.81	1.56
1946	3.25	3.6	4.19	19.86	0.86
1947	2.27	3.11	1.16	8.65	0.47
1948	5.07	10.02	1.32	54.81	1.82
1949	4.29	1.8	4.99	20.55	0.94
1950	3.37	3.03	1.07	12.37	1.03
1951	8.38	7.52	3.85	82.38	6.05
1952	4.1	5.11	1.04	23.49	0.65
1953	3.88	1.78	0.54	8.15	0.41
1954	8.66	2.88	2.16	36.2	1.39
1955	3.41	2.91	1.96	13.95	0.55
1956	6.25	2.43	1.84	22.06	1.13
1957	6.19	6.88	4.12	57.89	3.99
1958	2.98	3.01	1.5	11.65	0.43
1959	5.5	4.18	4.72	38.56	1.3
Sums	110.93	93.14	59.19	594.47	32.05
Means	4.4372	3.7256	2.3676	23.7788	1.282
Stdev	2.126383	2.200493	1.534822	18.33808	1.28163

Table 4. Summary of sensitivity measures.

Factors	OLS based model	Benchmark	Proposed method
X_1	0.58	0.84	0.90
X_2	1	0.62	0.57
X_3	0.52	0.38	0.36
X_4	0.37	1	1

analysis. The limitations of these methods have been clearly pointed out in the literature. First, being a local method, meaning that the sensitivity of the model output is analyzed only at a single point and second, lack of ability to address the interaction among input factors.

A novel global SA method has been presented in

this paper to deal with the interaction limitation. The interaction among the correlated data is captured in a linear model by taking advantage of principal components. The method can be implemented systematically with low computational cost. An example shows that the proposed method succeeds where common methods fail to justify the proper ranking.

Besides these limitations, there are few other situations, where if not paid enough attention, common SA methods result in false analysis. This paper designed a number of experiments to demonstrate the inefficiency of these methods.

In Case 1, when distribution of predictor variables is not taken into account, differential SA method fails to rank the factors correctly. The proper adjustment has been advised to overcome this concern.

Case 2 was designed to demonstrate the inefficiency of OAT method in common practice, especially when the predictor variables have high proportions of variance.

Case 3 was designed to demonstrate the ability of the proposed method to handle the correlation between predictor variables in a linear model, where the reviewed methods failed to capture. Although it is proposed for linear models, its implementation for nonlinear cases should not be ruled out. Research is underway to address this issue in subsequent publications.

References

1. Saltelli, A., Chan, K. and Scott, E.M., *Sensitivity Analysis*, In Wiley Series in Probability and Statistics, John Wiley and Sons, New York (2008).
2. Frey, C.H. and Patil, S.R. "Identification and review of sensitivity analysis methods", *Risk Analysis*, **22**(3), pp. 553-578 (2002).
3. Ascough II, J.C., Green, T.R., Ma, L. and Ahjua, L.R. "Key criteria and selection of sensitivity analysis methods applied to natural resource models", *Modism International Congress on Modeling and Simulation, Modeling and Simulation Society of Australia and New Zealand*, Melbourne, pp. 2463-2469 (2005).
4. Morgan, M.G., Henrion, M. and Small, M., *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press (1992).
5. Saltelli, A. and Annoni, P. "How to avoid a perfunctory sensitivity analysis", *Environ. Model. Softw.*, **25**(12), pp. 1508-1517 (2010).
6. Saltelli, A., Ratto, M., Tarantola, S. and Campolongo, F. "Sensitivity analysis practices: Strategies for model-based inference", *Reliability Engineering & System Safety*, **91**(10-11), pp. 1109-1125 (2006).
7. Vidal, C., *Artificial Cosmogenesis: A New Kind of Cosmology*, in *Irreducibility and Computational Equivalence: Wolfram Science 10 Years After the Publication of A New Kind of Science*, Zenil, H. (Ed.), arXiv:1205.1407v1 (2012).
8. Huber, G., *Gamma Function Derivation of n-Sphere Volumes*, The American Mathematical Monthly, **89**(5), pp. 301-302 (1982).
9. Asserina, O., Loredob, A., Petelet, M. and Iooss, B. "Global sensitivity analysis in welding simulations-What are the material data you really need?", *Finite Elements in Analysis and Design*, **47**(9), pp. 1004-1016 (2011).
10. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S., *Global Sensitivity Analysis: The Primer*, John Wiley and Sons, New York (2008).
11. Iman, R.L. and Conover, W.J. "A distribution-free approach to inducing rank correlation among input variables", *Communications in Statistics, Simulation and Computation*, **11**(3), pp. 311-334 (1982).
12. Daniel, C. "One-at-a-time plans", *Journal of the American Statistical Association*, **68**(342), pp. 353-360 (1973).
13. Hamby, D.M. "A comparison of sensitivity analysis techniques", *Health Phys.*, **68**(2), pp. 195-204 (1995).
14. Turanyi, T. and Rabitz, H. "Local methods and their applications", in *Sensitivity Analysis*, Saltelli, A., Chan, K. and Scott, E.M., Editors, John Wiley and Sons, New York (2008).
15. Rencher, A.C., *Multivariate Statistical Inference and Applications*, In Wiley Series in Probability and Statistics, John Wiley and Sons, New York (1998).
16. Snyder, W.M. "Some possibilities for multivariate analysis in hydrologic studies", *J. Geophys. Res.*, **67**(2), pp. 721-72 (1962).

Biographies

Younes Daneshbod was born in Shiraz, Iran, in 1974. He received a BS degree (1996) in Civil Engineering and an MS degree (2000) in Hydraulic Structures both from Shiraz University, Iran. He is currently a PhD candidate in the department of Civil and Environmental Engineering at Shiraz University. In 2004, he joined the Department of Civil Engineering at the Islamic Azad University, Arsanjan branch (IAUA), as a Lecturer and, subsequently, served as the head of the Department of Civil Engineering for two consecutive terms. His research interests include sensitivity analysis, uncertainty analysis, and artificial intelligence applied to civil engineering problems. He is the Life Member of the Iran Construction Engineering Organization (IRCEO) and a recipient of the IAUA Distinguished Researcher Award (2008).

Mohammad Javad Abedini was born in 1958, in Shiraz, Iran. He obtained his BS and MS degrees in Civil Engineering and Hydraulic Structures from Shiraz University, Iran, in 1985 and 1991, respectively, and his PhD degree in Water Resources Engineering from the University of Guelph, Canada, in 1998. He is currently Professor in the Department of Civil and Environmental Engineering at Shiraz University, Iran. His research interests include computational hydrology and hydraulics, analysis of temporal and spatial series, and data assimilation.