

## Making sense of global sensitivity analyses

Haruko M. Wainwright\*, Stefan Finsterle, Yoojin Jung, Quanlin Zhou, Jens T. Birkholzer

*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley 94720, CA, USA*



### ARTICLE INFO

#### Article history:

Received 13 March 2013

Received in revised form

10 June 2013

Accepted 13 June 2013

Available online 28 June 2013

#### Keywords:

Global sensitivity analysis

Morris OAT method

Sobol' index

Variance-based sensitivity indices

### ABSTRACT

This study presents improved understanding of sensitivity analysis methods through a comparison of the local sensitivity and two global sensitivity analysis methods: the Morris and Sobol'/Saltelli methods. We re-interpret the variance-based sensitivity indices from the Sobol'/Saltelli method as difference-based measures. It suggests that the difference-based local and Morris methods provide the effect of each parameter including its interaction with others, similar to the total sensitivity index from the Sobol' /Saltelli method. We also develop an alternative approximation method to efficiently compute the Sobol' index, using one-dimensional fitting of system responses from a Monte-Carlo simulation. For illustration, we conduct a sensitivity analysis of pressure propagation induced by fluid injection and leakage in a reservoir-aquitard-aquifer system. The results show that the three methods provide consistent parameter importance rankings in this system. Our study also reveals that the three methods can provide additional information to improve system understanding.

© 2013 Elsevier Ltd. All rights reserved.

## 1. Introduction

Hydrogeological modeling under uncertainty—from site characterization to prediction—requires not only the numerical or analytical models of flow and transport but also various statistical analyses, including parameter estimation (PE), uncertainty analysis (UA), sensitivity analysis (SA), data worth analysis, and experimental design (e.g., Van Griensven et al., 2006; Hill and Tiedeman, 2007; Tang et al., 2007; Finsterle et al., 2012). SA is a key component in this process, since there is a strong interaction between SA and other components. The objectives of SA include (1) to check whether a given dataset has sufficient information to determine a parameter given uncertainty of other parameters in PE, (2) to determine how to allocate limited resources to estimate each parameter as a part of data worth analysis, and (3) to reduce the number of parameters to be varied or estimated and hence to reduce computational burden in PE and UA.

SA started initially with the derivative-based local sensitivity method (Cacuci, 2003), but the global sensitivity analysis (GSA) methods (e.g., Morris 1991; Sobol', 2001; and Saltelli et al., 2008) have been increasingly applied in recent years. GSAs explore the parameter space so that they provide robust sensitivity measures in the presence of nonlinearity and interactions among the parameters compared to the local sensitivity analysis. The Morris one-at-a-time (OAT) method is a computationally frugal method that changes one parameter at a time from randomly generated reference parameter

sets, and computes the difference in the outputs. The Sobol'/Saltelli method provides the variance-based sensitivity indices that quantify the relative contribution of each parameter to the uncertainty in outputs.

GSAs, however, can be computationally intensive, since they require sampling parameter sets. Although several approximation methods have been developed to reduce the computational cost, such methods introduce additional model assumptions and response surface fittings, which are not universally applicable (Marrel et al., 2009; Oladyshkin et al., 2012). There is also an argument that local sensitivity analysis is sufficient, and GSAs do not provide additional information to justify the large computational cost (e.g., Foglia et al., 2009). Such arguments could be attributed to the fact that the value of GSAs has not been fully appreciated. The use of GSA is often limited to ranking parameter importance, even though GSAs can provide additional information for improving the system understanding.

The objective of this study is to improve the understanding of GSA through comparing the three SA methods, and to show more usages of GSAs for better system understanding. We re-interpret the Sobol'/Saltelli sensitivity indices as difference-based measures so that direct comparisons with the Morris and local methods are possible. We also show additional information and interpretations obtained from each method. Second, we compare the interpretation and computational cost of the SA methods and discuss what causes the differences. Finally, we propose an alternative approximation method to efficiently compute the first-order sensitivity index (i.e., Sobol' index).

We demonstrate our approach on a problem of pressure propagation induced by fluid injection and leakage in the CO<sub>2</sub> storage system. For CO<sub>2</sub> storage, there are potential risks associated with

\* Corresponding author. Tel.: +1 510 495 2038.

E-mail address: [hmwainwright@lbl.gov](mailto:hmwainwright@lbl.gov) (H.M. Wainwright).

leakage of injected CO<sub>2</sub> and resident brine. Since the pressure perturbations travel much faster and farther than the CO<sub>2</sub> plume, this pressure disturbance can be used to detect leaky pathways before CO<sub>2</sub> actually reaches the shallow aquifer. [Jung et al. \(in press\)](#) conducted a GSA to evaluate the sensitivity of leakages signals to model parameters, using the Morris method. In this study, we conduct a more detailed analysis, using the same semi-analytical model developed by [Cihan et al. \(2011\)](#).

## 2. Methodology

In this section, we discuss three SA methods. An additional interpretation is provided for the Sobol'/Saltelli method, so that similarities and differences among the three methods can be better compared. In all three methods, we consider a set of  $k$  parameters denoted by  $\{x_i | i=1, \dots, k\}$  and a scalar output  $y=f(x_i)$ , where  $f$  represents a hydrogeological forward model.

### 2.1. Local sensitivity method

The local sensitivity index for Parameter  $i$  is defined as the scaled partial derivative of  $y$  with respect to  $x_i$ . It is computed by changing each parameter by a small increment  $\Delta x_i$  from the reference parameter values  $\{x_i^*\}$  and computing the difference in  $y$

$$S_i^{local} = \frac{\tau_{x,i} \frac{\partial y}{\partial x_i}}{\tau_y} \Big|_{x_i^*} = \frac{\tau_{x,i} f(x_1^*, \dots, x_i^* + \Delta x_i, \dots, x_k^*) - f(x_1^*, \dots, x_k^*)}{\Delta x_i} \quad (1)$$

where  $\tau_{x,i}$  is the parameter-scaling factor, and  $\tau_y$  is the output-scaling factor.  $\tau_{x,i}$  can be the standard deviation or range of the parameter that represents the parameter variability or uncertainty. It can also be thought of as the amount by which the parameter would be changed in the sensitivity analysis, where the parameter is perturbed from its base-case value by an amount considered “reasonable” to examine its impact on the model output.  $\tau_y$  can be viewed as a measure of the change in the output that one would consider to be significant or representative.  $\tau_y$  is especially important when we combine multiple observations of different scales to create an integral measure of parameter influence or output sensitivity, such as the analysis by [Finsterle and Pruess \(1995\)](#) and the composite scaled sensitivity index in [Hill and Tiedeman \(2007\)](#). The local sensitivity method requires  $(k+1)$  simulations, i.e., a simulation for the reference parameter set plus  $k$  simulations for small-increment changes in the  $k$  parameters.

### 2.2. Morris sensitivity method

The Morris one-at-a-time (OAT) method ([Morris, 1991](#)) can be considered as an extension of the local sensitivity method. Each parameter range is scaled to the unit interval [0, 1] and partitioned into  $(p-1)$  equally-sized intervals. The reference value of each parameter is selected randomly from the set  $\{0, 1/(p-1), 2/(p-1), \dots, 1-\Delta\}$ . The fixed increment  $\Delta=p/(2(p-1))$  is added to each parameter in random order to compute the elementary effect (EE) of  $x_i$

$$EE_i = \frac{1}{\tau_y} \frac{f(x_1^*, \dots, x_i^* + \Delta, \dots, x_k^*) - f(x_1^*, \dots, x_k^*)}{\Delta} \quad (2)$$

where  $\{x_i^*\}$  is the randomly selected parameter set, and  $\tau_y$  is the output-scaling factor. To compute  $EE_i$  for  $k$  parameters, we need  $(k+1)$  simulations (called one “path”) in the same way as that of the local sensitivity method. By having multiple paths, we have an ensemble of EEs for each parameter. The total number of simulations is  $r(k+1)$ , where  $r$  is the number of paths.

We compute three statistics: the mean EE, standard deviation (STD) of EE, and mean of absolute EE (mean |EE|). Since the mean EE represents the average effect of each parameter over the parameter space, the mean EE can be regarded as a global sensitivity measure. As noted by Saltelli (2008), the mean |EE| is used to identify the non-influential factors, and the STD of EE is used to identify nonlinear and/or interaction effects. The standard error of mean (SEM) of EE, defined as  $SEM=STD/r^{0.5}$ , is used to calculate the confidence interval of mean EE ([Morris, 1991](#)).

### 2.3. Sobol'/Saltelli sensitivity method

While the local and Morris sensitivity methods are difference-based, the Sobol'/Saltelli method is variance-based ([Sobol', 2001](#); [Saltelli et al., 2008](#)). Here we define the random variable  $Y$  and the random vector  $\{X_i\}$  for the system response and the parameters, respectively. The sampled response and parameters are  $y$  and  $\{x_i\}$ . The first-order sensitivity index (i.e., Sobol' index) is defined by  $S_i = V[\mathbf{E}[Y|X_i]]/V[Y]$ , where  $\mathbf{E}[\bullet]$  and  $V[\bullet]$  represent mean and variance, respectively.  $S_i$  quantifies the first-order effect, i.e., the relative contribution of  $X_i$  to the uncertainty of  $Y$  excluding the interaction effect with other parameters. In addition, the total sensitivity index of  $X_i$  is defined by  $S_{ti} = 1 - V[\mathbf{E}[Y|X_{-i}]]/V[Y]$ , where  $\mathbf{E}[Y|X_{-i}]$  represents the mean of  $Y$  conditioned on all the parameters but  $X_i$ .  $S_{ti}$  accounts for the total effect of  $X_i$  including interaction effects, and is used to identify parameters with negligible effects and parameters that can be fixed.

To compute  $S_i$  and  $S_{ti}$ , we use an algorithm developed by Saltelli (2008) and modified by [Glen and Isaacs \(2012\)](#). It first generates two sets of sample matrices A and B, each of which is an  $n \times k$  matrix containing  $n$  sets of  $k$ -dimensional parameter vectors from Monte-Carlo (MC) sampling. From A and B, we create matrices  $C_i$  ( $i=1, 2, \dots, k$ ) such that the  $i$ -th column of  $C_i$  is the same as the  $i$ -th column of A (i.e.,  $C_{i,(m, i)} = A_{(m, i)}$  for  $m=1, 2, \dots, n$ ), and the other columns of  $C_i$  are the same as B (i.e.,  $C_{i,(m, j)} = B_{(m, j)}$  for  $m=1, 2, \dots, n$  and  $j \neq i$ ). The simulation results from the parameter sets A, B and  $C_i$  are  $n$ -dimensional vectors:  $\{a_m\}$ ,  $\{b_m\}$ , and  $\{c_{i,m}\}$  ( $m=1, \dots, n$ ), respectively. The number of required simulations is  $n(k+2)$ .  $S_i$  is computed as a correlation coefficient between  $\{a_m\}$  and  $\{c_{i,m}\}$

$$S_i = \frac{1}{\sigma_y^2} \frac{1}{n-1} \sum_{m=1}^n (a_m - \mu_y)(c_{i,m} - \mu_y), \quad (3)$$

where  $\mu_y$  and  $\sigma_y^2$  are the overall sample mean and variance of  $Y$ , respectively. We can use an analytical form of the confidence interval given for the correlation coefficient ([Fisher, 1921](#)) rather than computing it with the bootstrap method (e.g., [Archer et al., 1997](#); [Tang et al., 2007](#); [Saint-Geours et al., 2010](#)). The 95% confidence interval of  $S_i$  is given as  $\tanh\{\text{arctanh}(S_i) \pm 1.96SE\}$ , where SE is the standard error given by  $SE=(n-3)^{-0.5}$ . The width of confidence interval increases for small  $S_i$  due to the tanh transform, which implies that more simulations are required to rank minor parameters.

[Eq. \(3\)](#) offers an intuitive way to understand  $S_i$ . The parameter sets  $C_i$  and A share the same values only for  $X_i$ . If  $X_i$  is more influential,  $X_i$  determines the results so that  $\{a_m\}$  and  $\{c_{i,m}\}$  should be similar and hence have higher correlation. In [Eq. \(3\)](#), the covariance term can be re-written as a semivariogram ([Deutsch and Journel, 1992](#)) such that

$$S_i = 1 - \frac{1}{\sigma_y^2} \frac{1}{2(n-1)} \sum_{m=1}^n (a_m - c_{i,m})^2. \quad (4)$$

Note that the stationarity is warranted, since  $a_m$  and  $c_{i,m}$  ( $i=1, \dots, k$ ) are the same system response (the difference comes from random sampling), and the mean of them should be the same.  $S_i$  is high when the difference between  $\{a_m\}$  and  $\{c_{i,m}\}$  is small. Here, we interpret  $S_i$  as a difference-based measure between the original

$\{a_m\}$ ) and perturbed responses ( $\{c_{i,m}\}$ ) caused by a variation in all parameters except  $X_i$ .

Similar to  $S_i$ ,  $S_{ti}$  can be computed as

$$S_{ti} = 1 - \frac{1}{\sigma_y^2 n-1} \sum_{m=1}^n (c_{i,m} - \mu_Y)(b_m - \mu_Y). \quad (5)$$

Using the covariance–semivariogram relationship, we can re-write  $S_{ti}$  as

$$S_{ti} = \frac{1}{\sigma_y^2} \frac{1}{2(n-1)} \sum_{m=1}^n (c_{i,m} - b_m)^2. \quad (6)$$

Since  $C_i$  and  $B$  have the same values except for  $X_i$ ,  $(c_{i,m} - b_m)$  is equivalent to taking a difference in  $Y$  when perturbing  $X_i$  with the other parameters fixed. This procedure is the same as the Morris method, except that the output difference is not divided by the parameter difference ( $\Delta$ ). This similarity could be the reason why [Campolongo et al. \(2007\)](#) observed the mean IEEI being a good proxy for  $S_{ti}$ . We would note that [Sobol' \(2001\)](#) also included the same equation as Eq. (6) from a mathematical derivation. However, we believe that the progression from the correlation to the semivariogram provides us more intuitive insight as described above.

Eqs. (4) and (6) can also explain the difference between  $S_i$  and  $S_{ti}$ . When we perturb  $X_i$  with the other parameters fixed, the difference  $(c_{i,m} - b_m)$ , hence  $S_{ti}$ , accounts not only for the impact of  $X_i$  as a single factor, but also for the interaction effects with the other parameters.  $S_i$ , on the other hand, is computed by fixing  $X_i$  and changing the other parameters. Perturbing all the other parameters except for  $X_i$  includes the total effects involving all the parameters except for  $X_i$ . In other words, the second term in Eq. (4) presents the total effect involving all the parameters except for the first-order effect of  $X_i$ . Since the sum of the all the effects becomes 1, Eq. (4) shows that  $S_i$  represents the first-order effect of  $X_i$  excluding interaction effects.

#### 2.4. Alternative approximation method for the Sobol' index

We may recall the original definition of  $S_i$  as  $\mathbf{V}[\mathbf{E}[Y|X_i]]/\mathbf{V}[Y]$ . Since  $\mathbf{E}[Y|X_i]$  is the mean of  $Y$  as a function of  $X_i$ ,  $\mathbf{E}[Y|X_i]$  corresponds to the mean (or fitted) line of MC samples as a function of  $X_i$ . Therefore,  $S_i$  can be computed by the variance of the fitted line  $\mathbf{V}[\mathbf{E}[Y|X_i]]$  divided by  $\mathbf{V}[Y]$ . In other words,  $S_i$  can be calculated by determining  $\mathbf{E}[Y|X_i]$  in an one-dimensional space, which typically requires much less computational effort than fitting a response surface in a multi-dimensional parameter space (e.g., [Marrel et al., 2009](#); [Oladyshevkin et al., \(2012\)](#)). Although fitting introduces additional assumptions, one-dimensional fitting allows us to use less model-dependent approaches, such as semiparametric or non-parametric regression methods.

Another advantage of this approach is that the number of simulations is not directly dependent on  $k$ , since MC samples are projected onto each parameter axis independently. Although the number of simulations to achieve overall convergence still depends on  $k$ , the minor parameters do not affect the convergence significantly. The original approach requires additional  $n$  simulations for each additional parameter, whereas this approach does not increase the number of simulations for parameters with a minor impact. The disadvantage of this method is that it does not provide  $S_{ti}$ . However, this approach is cost effective to identify influential parameters, especially when MC simulations are performed anyway for other purposes (e.g., UA).

### 3. Demonstration problem setup

For demonstration, we use a synthetic example involving fluid injection and leakage in an idealized reservoir–aquitard–aquifer system of a lateral infinite extent. Three homogeneous isotropic

layers (i.e., reservoir, aquitard, and aquifer) are considered as shown in Fig. 1a. The top and bottom of the system are assumed impervious except for a constant volumetric rate of injection into the reservoir,  $Q=5700 \text{ m}^3 \text{ d}^{-1}$ , through the injection well. The leaky well, located at 2000 m away from the injection well, can bring resident brine from the reservoir to the aquifer through focused leakage. Diffuse leakage from the reservoir to the aquifer occurs through the aquitard. We also assume that the radius of injection and leaky wells is 0.15 m. We use the semi-analytical model developed by [Cihan et al. \(2011\)](#), which calculates both flow through aquitards (diffusive brine migration) and flow through leaky wells (focused leakage); see [Cihan et al. \(2011\)](#) for details.

In SA, we perturb seven parameters: hydraulic conductivity  $K$  and storativity  $S$  for the reservoir, aquitard, and aquifer, and well hydraulic conductivity. Table 1 shows the reference parameter values

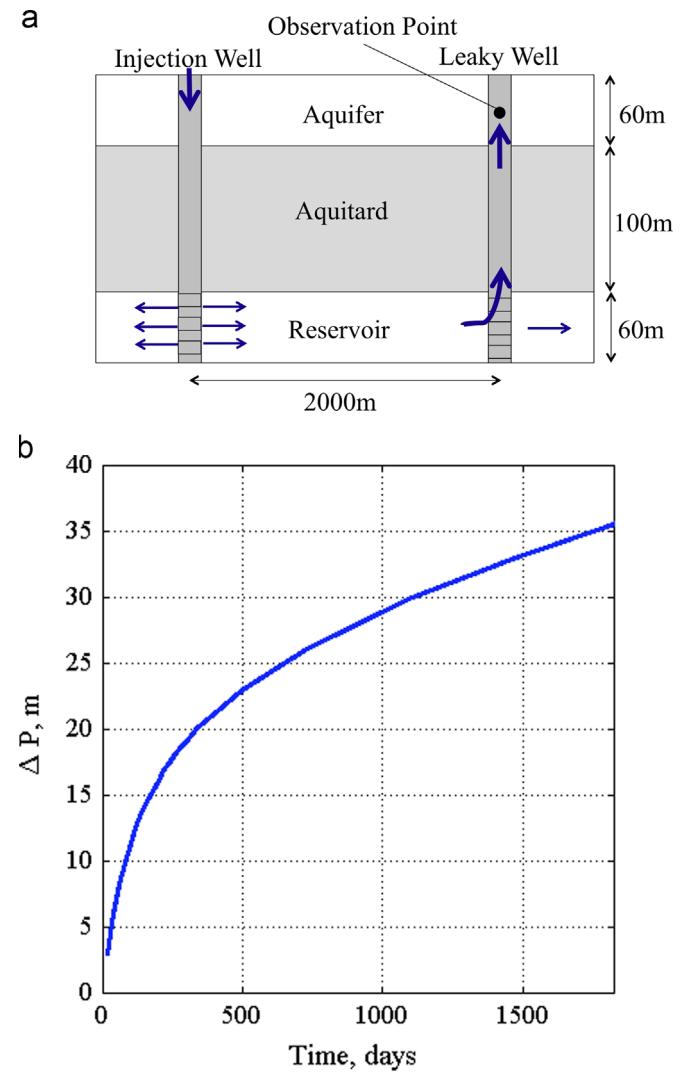


Fig. 1. (a) Conceptual model setup for the pressure leakage problem, and (b)  $\Delta P$  (m) at the observation point.

Table 1

Reference parameter values and uncertainty ranges: hydraulic conductivity  $K$  and specific storativity  $S$ .

	Aquifer	Aquitard	Reservoir	Well	Range
$K$ (m/d)	2.00E-1	2.00E-6	2.00E-1	2.00E+5	One order of magnitude
$S$ (1/m)	1.88E-6	1.47E-6	1.88E-6	N/A	Factor of five

and parameter ranges. We assume a uniform distribution within the range for each parameter. As a performance measure, we are interested in the pressure buildup  $\Delta P$  at the leaky well location in the shallow aquifer.  $\Delta P$  as a function of time in the reference case is shown in Fig. 1b.

## 4. Results and discussion

### 4.1. Comparison of sensitivity coefficients

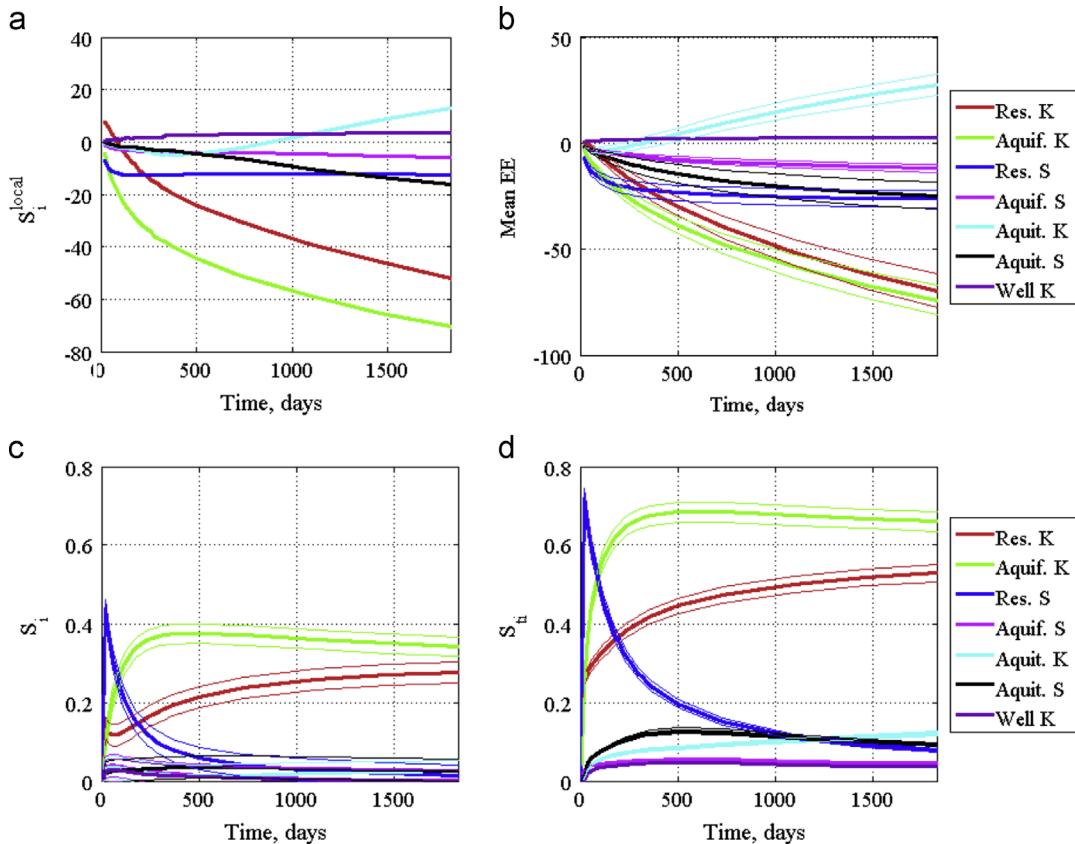
Fig. 2 shows the time evolution of the sensitivity indices from the three SA methods;  $S_i^{local}$ , mean EE,  $S_i$  and  $S_{ti}$ . The number of simulations is 8, 3200 ( $r=400$ ,  $p=6$ ), and 45,000 ( $n=5000$ ) for the local, Morris and Sobol'/Saltelli methods, respectively. For the GSA indices, the confidence intervals are shown so that we can evaluate the parameter importance with a given number of simulations. For the local and Morris methods,  $\tau_y$  is fixed to one with the same unit as  $y$ , which is acceptable because we do not combine multiple outputs. For easier comparison, Fig. 3 shows the rescaled sensitivity indices:  $|S_i^{local}|$ , mean EE,  $3\sigma_y\sqrt{S_i}$  and  $3\sigma_y\sqrt{S_{ti}}$ .  $S_i$  and  $S_{ti}$  are rescaled because they are the square of difference-based measures, according to Eqs. (4) and (6). The factor three comes from the fact that the mean difference of two parameters sampled from the scaled uniform distribution ( $\in [0, 1]$ ) is 1/3. Since all the values in Fig. 3 are positive, the rescaled indices represent the magnitude of parameter effects.

In Fig. 3, all the rescaled sensitivity indices provide a consistent pattern of parameter importance. The sensitivity indices increase monotonically as time increases, following the increase in  $\Delta P$  (Fig. 1b). The reservoir  $S$  is most influential at early time, whereas aquifer and reservoir  $K$  have a dominant effect afterward. The sensitivity to aquitard  $K$  and  $S$  increases later, as the brine diffuses

through the aquitard and reaches the shallow aquifer. There are differences among the four sensitivity indices. For  $|S_i^{local}|$  (Fig. 3a), the sensitivity to the aquitard  $K$  decreases between 500 and 1000 days, since  $S_i^{local}$  (Fig. 2a) changes from negative to positive.  $3\sigma_y\sqrt{S_i}$  (Fig. 3c) is smaller than the mean EE (Fig. 3b) and  $3\sigma_y\sqrt{S_{ti}}$  (Fig. 3d), since  $S_i$  does not include interaction effects. For all the methods, the leaky-well conductivity appears to be the least sensitive parameter. This may be because the very high reference well conductivity ( $2 \times 10^5$  m/d) and its range of one order of magnitude for GSA; in this range of well conductivity, the leakage flux is controlled by the capacity of brine supply from the reservoir (Jung et al., 2013; Fig. 7). There are some differences (e.g., the relative importance of aquitard  $S$  compared to the other parameters varies depending on the method used), requiring further analysis (see Sections 4.2 and 4.4).

Additional information can be gained from the sensitivity indices in Fig. 2. For example,  $S_i^{local}$  and mean EE provide the sign of the sensitivity, which helps understand the physics of the problem. Increasing  $K$  of leakage pathways (i.e., well  $K$  and aquitard  $K$  at later times) increases  $\Delta P$  (hence a positive effect), whereas increasing the aquifer and reservoir  $K$ s dissipates pressures more quickly and thus generally decreases  $\Delta P$  (hence a large negative effect). The reservoir  $S$  has a negative effect at the beginning, since the compressible matrix pores absorb the initial pressure increase. We see that the pressure perturbation due to diffusive brine migration may arrive between 500 and 1000 days, since the sensitivity to the aquitard  $K$  changes from negative to positive both in the  $S_i^{local}$  and mean EE.

In Fig. 2c and d,  $S_i$  and  $S_{ti}$  provide the relative contribution of each parameter to the uncertainty of the output. The ranking of parameter importance is more easily recognized than  $S_i^{local}$  and mean EE. The large effect of reservoir  $S$  at early times can be clearly seen. The patterns of the three most influential parameters are



**Fig. 2.** Time evolution of sensitivity index: (a)  $S_i^{local}$ ; (b) mean EE; (c)  $S_i$  and (d)  $S_{ti}$ . In (b)–(d), the thin lines represent the 95% confidence intervals.

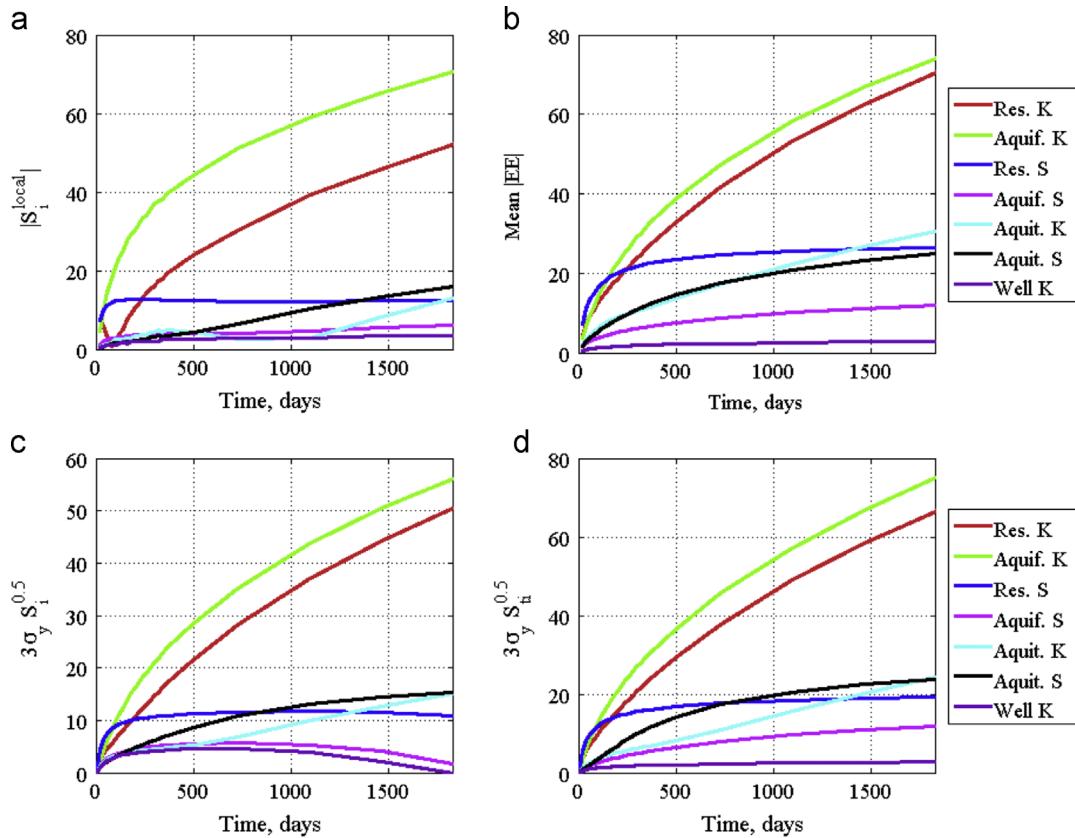


Fig. 3. Time evolution of re-scaled sensitivity index: (a)  $|S_i^{\text{local}}|$ ; (b) mean  $|EE|$ ; (c) and (d).

similar in Fig. 2c and d, suggesting that a change in any of these parameters would have a significant impact on  $\Delta P$  even without taking into account the interaction effects.

#### 4.2. Identification of nonlinear and interaction effects in GSA

Fig. 4 shows a cross-plot between the mean and STD of EE, following Morris (1991). Each curve corresponds to the time evolution for a parameter's sensitivity index. The two black lines represent Mean  $EE = \pm 2\text{SEM}$ . Due to the large number of paths ( $r=400$ ), the SEM is small, and all the parameters are below the black lines, indicating that their non-zero impact is statistically significant. All the parameters have a non-zero value of STD, indicating that they have nonlinearity and/or interaction effects. The ratio between the mean and STD of EE is larger for aquitard  $K$  and  $S$ , which have a large difference in the evaluated sensitivity indices among the three SA methods (Fig. 3).

Fig. 5 shows the difference between  $S_{ti}$  and  $S_i$  as a function of  $S_i$ , identifying the ratio between the first-order effect and the interaction effects. All the parameters show interaction effects, since  $(S_{ti}-S_i)$  is larger than zero. The aquitard  $K$  and  $S$  have a particularly large difference  $(S_{ti}-S_i)$  relative to  $S_i$ , suggesting that they have a large interaction effect compared to the first-order effect. Comparing Figs. 4 and 5 allows us to separate interaction from nonlinearity effects, since Morris's STD of EE includes both, but  $(S_{ti}-S_i)$  represents only the interaction effects. For example, the reservoir  $K$  and aquifer  $K$  would have a similar magnitude of the interaction effect (Fig. 5), but the nonlinear effect is higher for reservoir  $K$  (Fig. 4).

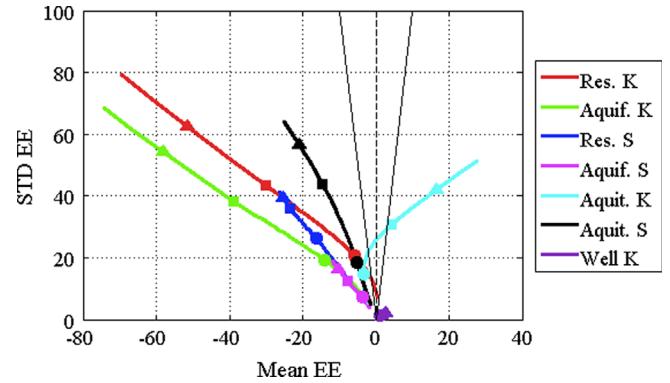


Fig. 4. Mean EE vs. STD from the Morris method. The circle, square and triangle on each curve represent 100, 500 and 1100 days, respectively. The end of each line corresponds to 1825 days.

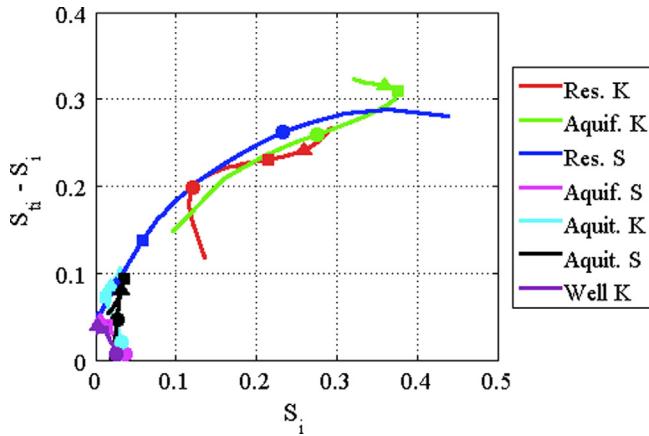
#### 4.3. Computational cost in GSA

In the Morris method, we need to determine the number of partitions ( $p$ ) as well as the number of paths ( $r$ ). The influence of these two parameters is investigated here. Fig. 6 shows the mean and STD of EE at 5 years for each parameter as a function of  $r$  with  $p=6$  and 50. The estimated mean and STD of EE appear stabilized after 200 and 300 paths, respectively. Running fewer paths, however, still provides a reasonable estimate of the relative importance of a parameter.

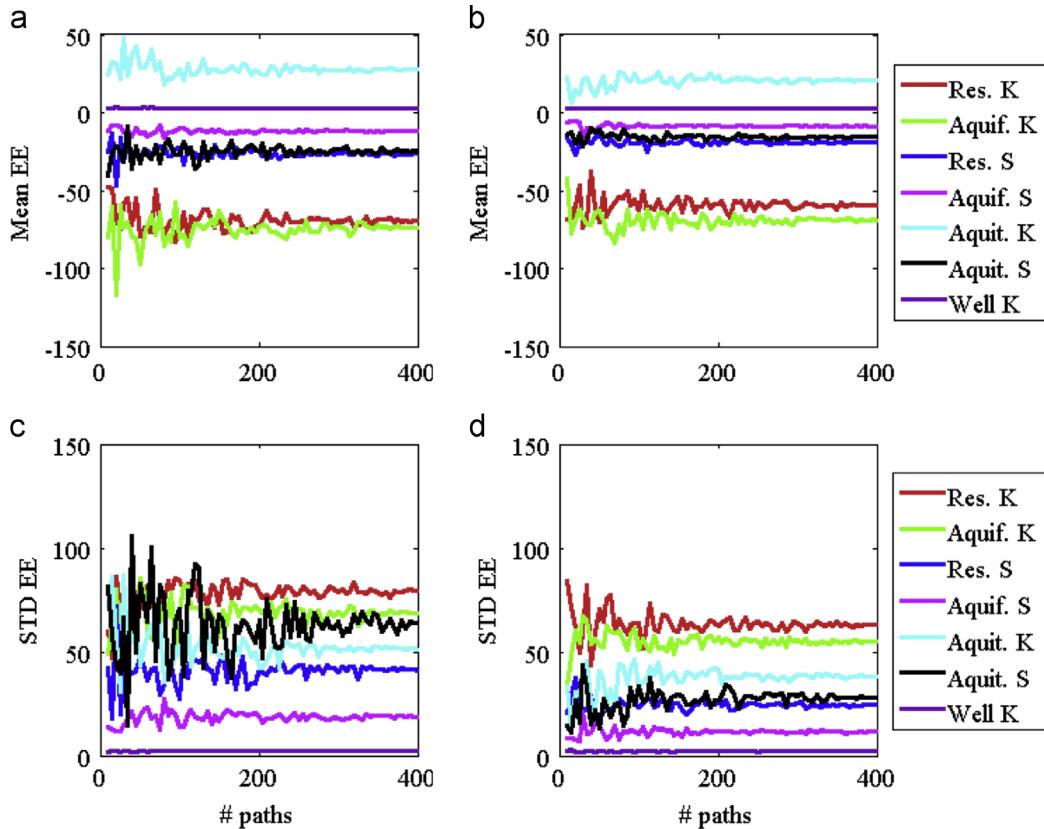
The number of partitions has a small impact on the convergence of both mean EE and STD of EE in this case. This is surprising since  $p$  determines the possible number of parameter combinations;  $p^k/2$

(Morris, 1991). This is because having fewer partitions leads to individual EE values varying significantly, which is seen as a large variability of curves with fewer than 100 paths. In addition, Fig. 6 shows that the converged values are slightly different for  $p=6$  and  $p=50$ . We found that increasing  $p$  tends to systematically reduce the mean EE (in magnitude) and STD of EE in this system. This is due to the presence of nonlinearity; the parameter difference  $\Delta$  becomes smaller as  $p$  increases, and the difference in outputs divided by  $\Delta$  (hence EE) changes depending on  $\Delta$ . The impact of the different number of partitions is further discussed in Section 4.4.1.

In the Sobol'/Saltelli method, we only need to determine  $n$ . Fig. 7 shows  $S_i$ ,  $S_{ti}$ , and mean |EE| (computed from B and  $C_i$ ) for each parameter at 5 years as a function of  $n$ . The mean |EE| (Fig. 7c) is computed based on EE by taking the difference ( $c_{i,m}-b_m$ ) and



**Fig. 5.** ( $S_{ti} - S_i$ ) as a function of  $S_i$ . The circle, square and triangle on each curve represent 100, 500 and 1100 days, respectively. The end of each line corresponds to 0 days and 1825 days.



**Fig. 6.** Mean EE and STD of EE at 5 years as a function of  $r$ : (a) mean EE with  $p=6$ ; (b) mean EE with  $p=50$ ; (c) STD of EE with  $p=6$  and (d) STD of EE with  $p=50$ .

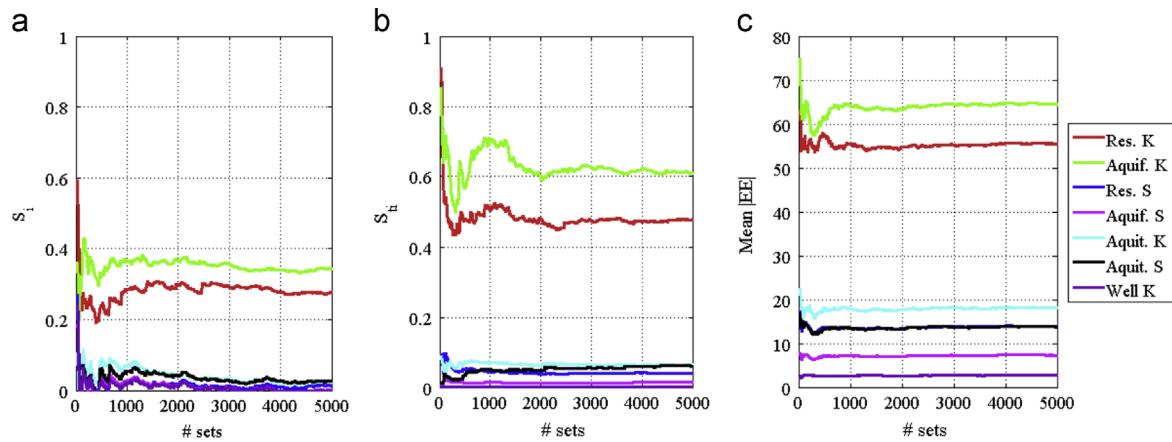
dividing it by the parameter difference, according to Eq. (2). In other words, this mean |EE| is equivalent to Morris's mean |EE| with a variable  $\Delta$  ( $\Delta = C_{i,(m,i)} - B_{(m,i)}$ ).  $S_i$  and  $S_{ti}$  (in Fig. 7a and b) require several thousand sets to stabilize, whereas the mean |EE| only requires several hundred simulations.

Fig. 7 shows that the large computational cost of the Sobol'/Saltelli method compared to the Morris method is attributed to the fact that the Sobol'/Saltelli method calculates a variance-based measure (i.e., a second-order statistics). Eq. (6) implies that Morris's mean |EE| quantifies the magnitude of each parameter effect including its interaction with others, similar to  $S_{ti}$ . The major difference is that Morris's mean |EE| uses a fixed parameter difference, and that the Morris mean |EE| is not normalized by the variance of the system response. If each observation is analyzed separately, the normalization does not affect the parameter importance ranking. Our analysis suggests that Morris's mean |EE| can be used instead of  $S_{ti}$  to rank the parameter effect including the interaction effect, considerably saving computational costs.

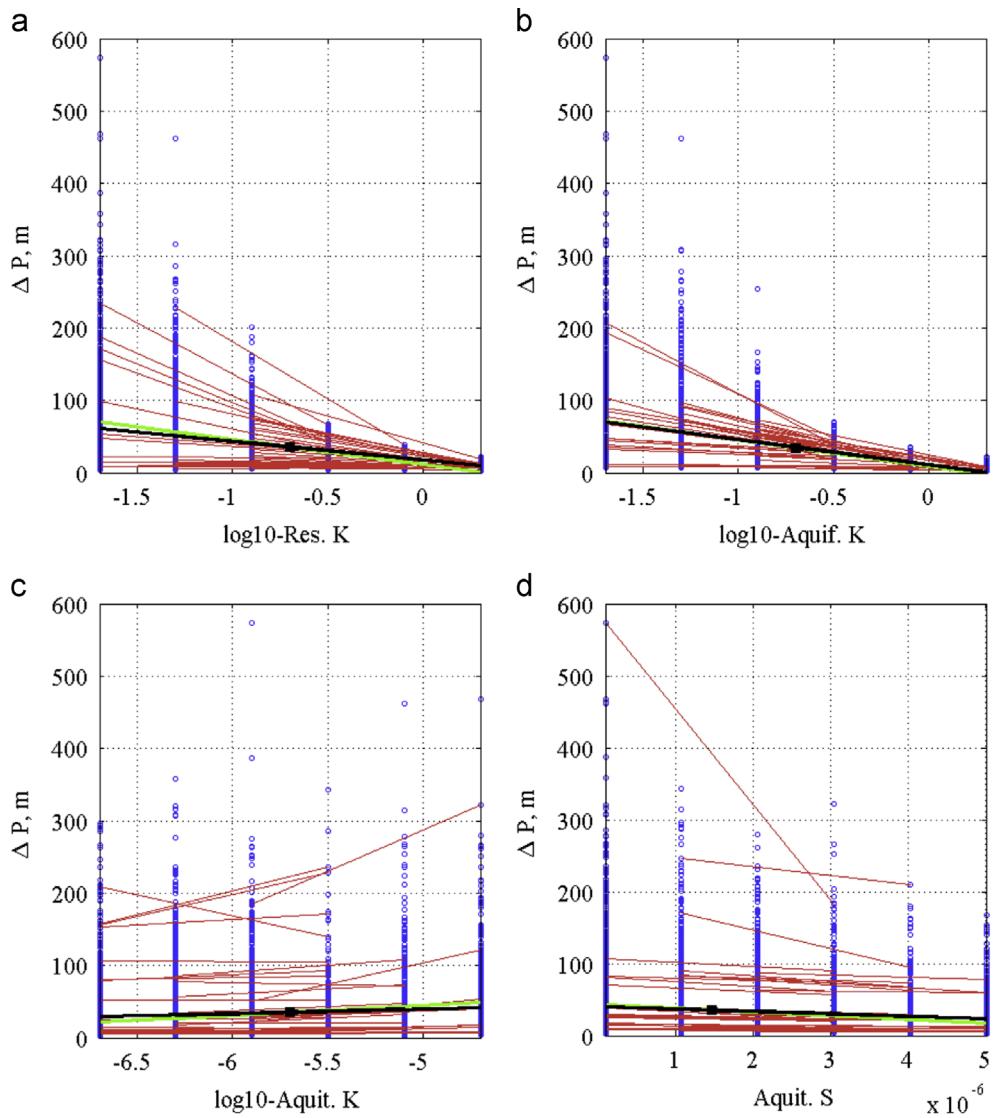
#### 4.4. Additional information from GSA

##### 4.4.1. Morris method

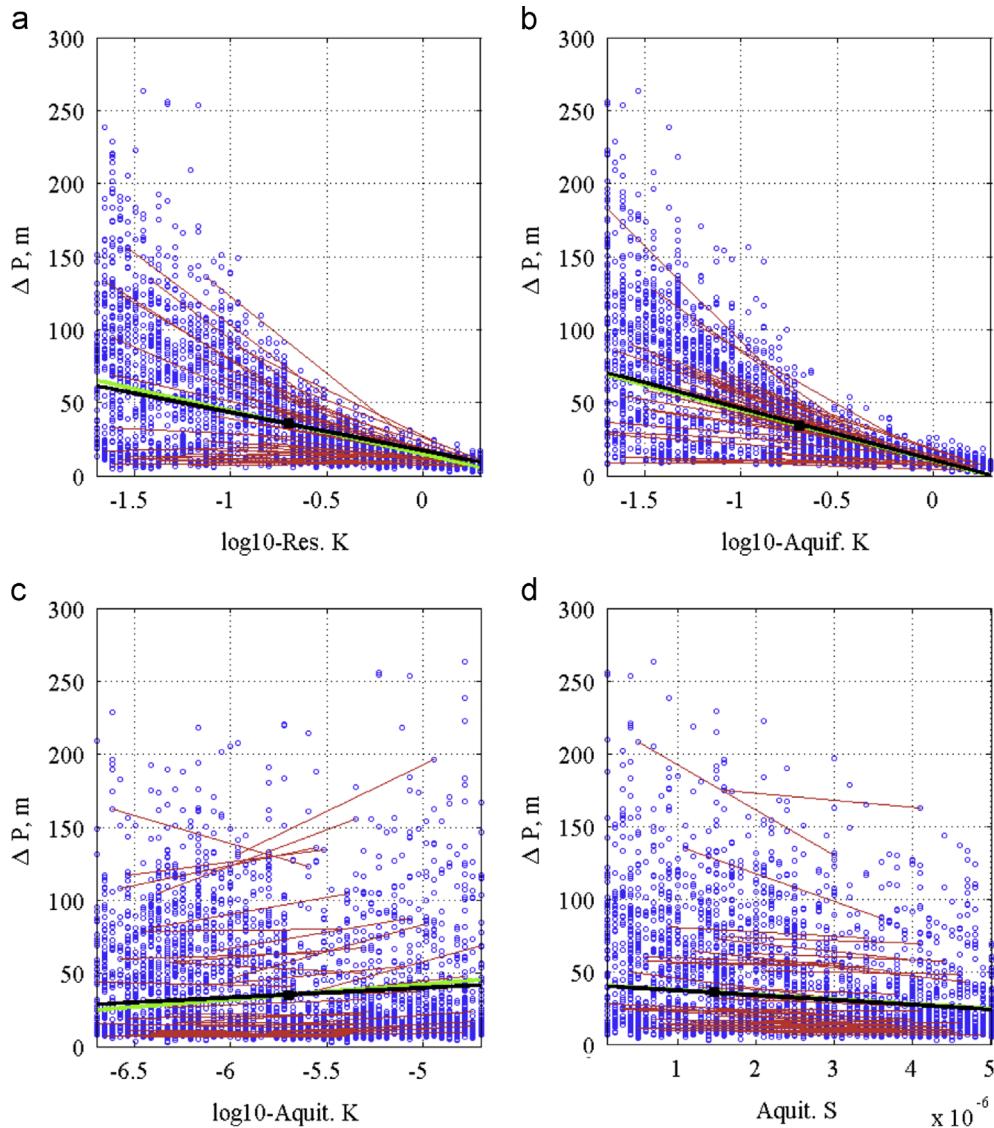
Figs. 8 and 9 show all the simulated  $\Delta P$  at 5 years from the Morris method as a function of four parameters: reservoir  $K$ , aquifer  $K$ , aquitard  $K$ , and aquitard  $S$ . The number of partitions is different:  $p=6$  in Fig. 8 and  $p=50$  in Fig. 9. Although the blue dots are not from random sampling (except for the first set of each path), they are distributed over the entire parameter space and show the variability of possible  $\Delta P$  depending on each parameter. The dependence of  $\Delta P$  on the reservoir  $K$  and aquifer  $K$  are obvious, although the large scatter in the low  $K$  regions suggests an interaction effect: when



**Fig. 7.** (a)  $S_1$ , (b)  $S_2$  and (c) mean |EE| at 5 years as a function of  $n$  from the Sobol'/Saltelli method.



**Fig. 8.** Scatter plot of  $\Delta P$  at 5 years from the Morris method ( $p=6$ ) as a function of: (a) reservoir  $K$ ; (b) aquifer  $K$ ; (c) aquitard  $K$  and (d) aquitard  $S$ . The red lines represent the change in  $\Delta P$  when the corresponding parameter is changed (1 in 10 paths is shown). The black square is the reference point in the local SA. The green and black line use the slope equal to  $S^{local}$  and mean EE, respectively, passing through the local SA reference point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 9.** Scatter plot of  $\Delta P$  at 5 years from the Morris method ( $p=50$ ) as a function of: (a) reservoir  $K$ ; (b) aquifer  $K$ ; (c) aquitard  $K$  and (d) aquitard  $S$ . The red lines represent the change in  $\Delta P$  when the corresponding parameter is changed (1 in 10 paths is shown). The black square is the reference point in the local SA. The green and black line use the slope equal to  $S_{\text{local}}$  and mean EE, respectively, passing through the local SA reference point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

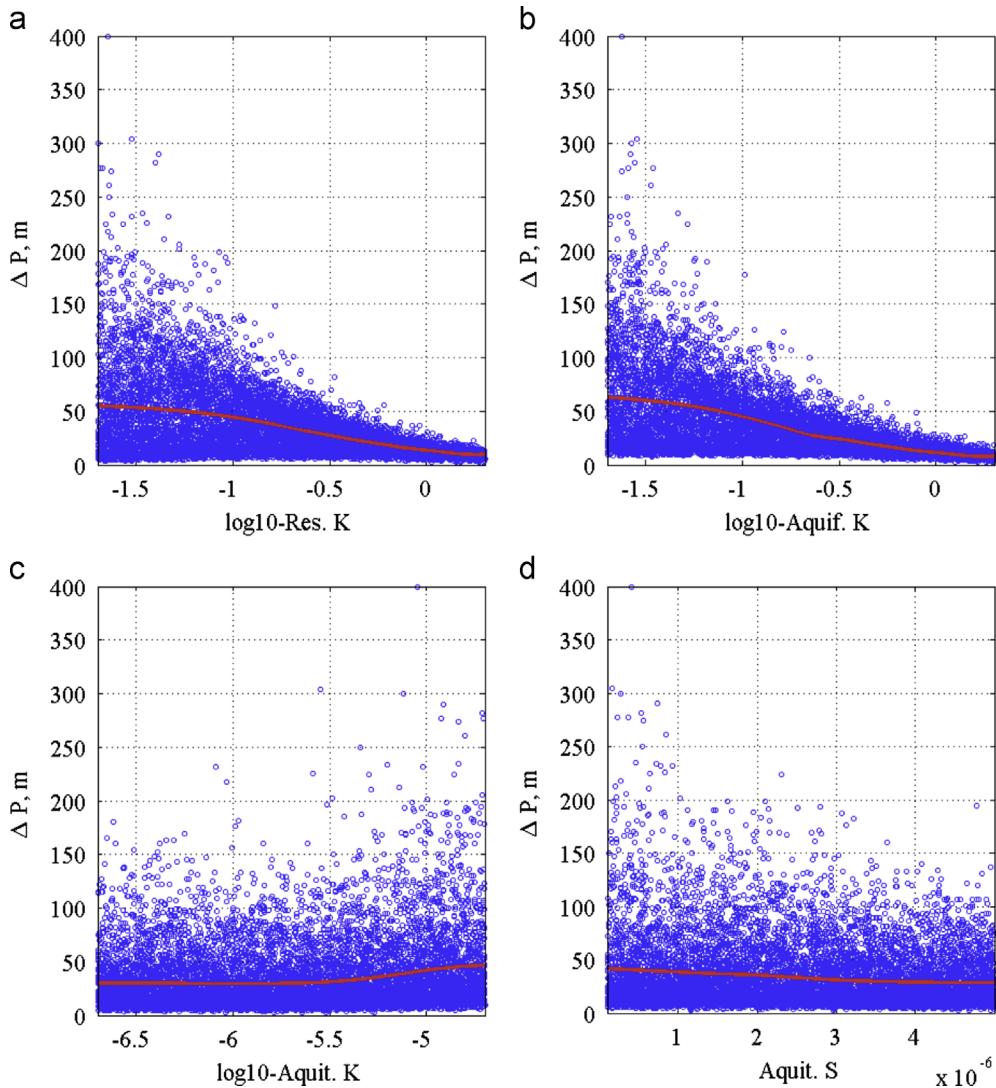
reservoir  $K$  or aquifer  $K$  is low,  $\Delta P$  varies significantly depending on other parameters.

The red lines in each figure represent the change in  $\Delta P$  when the corresponding parameter is changed.  $\Delta$  is represented by the two end points of each line in the  $x$ -axis. The slope of each red line represents individual EEs (Eq. (2)). The average slope of all the red lines (i.e., mean EE, shown by the green line) is similar to the slope of the local sensitivity (black line) calculated at the reference point (black square), which suggests a good match between the local and mean EE, although slight differences are visible for the reservoir  $K$  and aquitard  $K$ .

In addition, these red lines effectively visualize nonlinearity and/or the interaction among parameters. We may find nonlinearity when the slope is different depending on the starting parameter value. In Fig. 8c, for example, the slope is steeper when the aquitard  $K$  is higher. We may find an interaction effect when the starting parameter value ( $x$ -axis) is the same, but the slope depends on the starting output value ( $y$ -axis) due to the impact of other parameters. In Fig. 8c and d, for example, the red lines have a steeper slope

when the starting  $\Delta P$  is high. This means that the aquitard  $K$  and  $S$  are more influential when  $\Delta P$  is high, which is equivalent to the case where the reservoir  $K$  and/or aquifer  $K$  are low. Although we do not show here, the 2D contour plots (i.e., the system response as a function of two parameters) were also examined to confirm the observations obtained from the scatter plots.

The difference between  $p=6$  and  $p=50$  can be seen by comparing Figs. 8 and 9. The range of  $\Delta P$  variability is larger in Fig. 8, since more extreme  $\Delta P$  values are sampled when  $p$  is small.  $S_{\text{local}}$  The slope is generally larger in Fig. 8 due to nonlinearity, which creates a difference in the converged mean EE and STD values between  $p=6$  and  $p=50$  in Fig. 6. In addition, the six discrete points along each parameter axis (Fig. 8) seem to capture the variability of the pressure reasonably well in this case. From these figures (Figs. 6, 8 and 9), we may conclude that we should use a small  $p$  as long as the discrete points can capture the variability in the output, since using a small  $p$  can capture the extremes of output values easily and lead to a larger SEM (i.e., not underestimating the uncertainty of mean EE estimate).



**Fig. 10.**  $\Delta P$  at 5 years from MC sampling as a function of: (a) reservoir  $K$ ; (b) aquifer  $K$ ; (c) aquitard  $K$  and (d) aquitard  $S$ . Each red line represent  $E[Y|X_i]$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

#### 4.4.2. Sobol'/Saltelli method

Fig. 10 shows the scatter plots of  $\Delta P$  (5 years) as a function of the four parameters from the MC sampling used in the Sobol'/Saltelli method. It shows scatters similar to Fig. 9, suggesting that the samples from the Morris method successfully capture the output variability in this one-dimensional parameter space, even though the Morris sampling is not purely random. The red line in each figure represents  $E[Y|X_i]$  obtained by locally weighted scatterplot smoothing (LOESS) fitting (Hastie et al., 2001). The fitting is done locally using least-square fitting and a second-order polynomial function. The  $E[Y|X_i]$  lines show nonlinearity; aquitard  $K$ , for example, has an impact only when it is larger than  $-5.5$  in  $\log_{10}$ , since higher aquitard  $K$  allows brine to arrive diffusively in the shallow aquifer before 5 years.

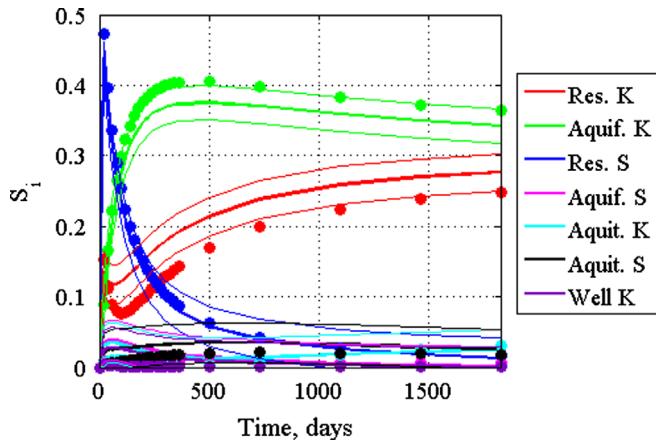
Fig. 10 (along with Figs. 8 and 9) visualizes the conceptual difference between  $S_i$  and mean EE, as well as between  $S_i$  and  $S_{ti}$ . Fig. 8c and d (and also Fig. 9c and d) show that the aquitard  $K$  and  $S$  sometimes have a large impact depending on the other parameters, since some of the red lines have a large slope when the starting point of the red line is high. Such an interaction effect is not represented in  $S_i$ , since the  $E[Y|X_i]$  line averages out such effects. The uncertainty of  $E[Y|X_i]$  would be higher for minor parameters, since the variability from the other parameters is

overwhelming. This again suggests that the Morris method and  $S_{ti}$ —determined by changing one parameter while fixing the others—can capture interaction effects and are necessary to identify the impact of minor parameters.  $S_{ti}$  has rarely been used in hydrogeological UQ studies, although  $S_i$  has commonly been evaluated. As Saltelli et al. (2008) suggested, both have to be combined in SA, especially when the objective is to find negligible parameters that can be fixed for PE and UA.

#### 4.5. Alternative approximation method for the Sobol' index

As is discussed in Section 2.3, the Sobol' index ( $S_i$ ) can be computed by fitting of  $E[Y|X_i]$  only in one dimension. Fig. 10 shows the fitted line  $E[Y|X_i]$  based on the MC samples in this system.  $S_i$  can be computed by taking the variance of this fitted line and dividing it by  $V[Y]$ .

Fig. 11 compares  $S_i$  determined by this approach with the original approach. The approximated sensitivity index (dots) fall in the or near confidence intervals computed by the original algorithm, suggesting that this approach is a valid method. The number of simulations is reduced to 1/9 (45,000 → 5000 simulations). This approach uses a standard MC sampling, which is likely to be available from other applications (e.g., UA). In other words, when we perform



**Fig. 11.**  $S_i$  from the original algorithm (thick lines) compared with that from the alternative method (dots). The thin lines represent the 95% confidence intervals computed by the original algorithm.

UA, we can compute  $S_i$  (i.e., Sobol' index) as a by-product without any additional computation.

## 5. Conclusions

In this study, we compared the interpretation and computational cost of local SA, and the Morris and Sobol'/Saltelli GSA methods. We re-interpreted the Sobol'/Saltelli sensitivity indices as difference-based measures, which enabled direct comparison to the Morris method. We also developed an alternative approximation approach to efficiently compute the first-order sensitivity index (i.e., Sobol' index). We demonstrated that the three methods could provide additional information to better understand the system behavior, in addition to the traditional use of ranking parameter importance.

Re-interpreting the Sobol'/Saltelli sensitivity indices as difference-based measures provided better understanding of the similarities and differences among the three methods, and also more intuitive understanding of the Sobol'/Saltelli sensitivity indices and the algorithm to compute these indices. It also suggested that Morris's mean |EE| has similar information to the total sensitivity index when we are interested in the parameter importance ranking (i.e., when the normalization by the total variance does not matter).

We demonstrated the comparison of the three methods using a pressure perturbation problem with fluid injection and leakage in a reservoir-aquitard-aquifer system. The demonstration results showed that the three sensitivity methods give similar interpretations and importance rankings. We found that the local sensitivity method is sufficient to identify the influential parameters in our case. The Morris method provides many uses (e.g., identifying positive/negative effects, influential parameters, and nonlinear and/or interaction effects) with relatively small computational burden. The Sobol'/Saltelli method gives a more quantitative sensitivity measure in the context of UQ, although it is computationally expensive. It is noted that a highly nonlinear system would create larger differences among the three methods. However, we believe that the insights offered in this study generally improve our understanding of SA.

We also explored the computational cost of GSA. In the Morris method, increasing the number of partitions did not change the convergence of the mean and STD of EE significantly, but changed the converged values due to nonlinearity. In the Sobol'/Saltelli method, the number of simulations is much higher than that for the Morris method. Our results showed that the difference comes from the fact that the Sobol'/Saltelli method uses variance-based measures, which are second-order statistics.

Examining the scatter plots of the output values as a function of each parameter from GSA provided much richer information, improving our understanding of the system. Morris sampling was especially useful to check the nonlinearity in the output and identify the individual interaction effects. In MC sampling, we can also check whether the number of partitions in the Morris method is enough to capture the variability in the response.

Based on our analysis, one may conclude that the Morris OAT method is sufficient. This study, however, does not imply that the Sobol'/Saltelli method is not necessary. Although the total sensitivity index could be substituted by the Morris mean |EE| for parameter importance ranking, the first-order sensitivity index (i.e., Sobol' index) is required when we are interested in the first-order effect. In addition, the Sobol'/Saltelli indices are necessary when we are interested in the relative contribution of each parameter to the uncertainty and/or variability of each output, in addition to the parameter importance ranking. However, in this study, we could offer some alternatives when the forward model is too computationally intensive to use the Sobol'/Saltelli method. For example, we can use the alternative method to compute the Sobol' index from UA results (if available), and the Morris mean |EE| can be used as a proxy for the total sensitivity index to choose non-influential parameters.

From this study, we make several recommendations for SA.

- A local sensitivity analysis should be done first. Even when the system is nonlinear, it still provides considerable insight into the system behavior.
- The number of partitions in the Morris method should be small as long as the discrete points capture the variability of the system response. A small set of MC simulations would be helpful to determine the minimum number of partitions.
- Examining the scatter plots of the model outputs from the Morris sampling visualizes nonlinear effects and interaction with the other parameters.
- The Sobol'/Saltelli method requires a large number of simulations to converge. It is important to compute the confidence interval.
- The Sobol' index does not differentiate the minor parameters well; the total sensitivity index is necessary (or alternatively the Morris mean |EE| can be used).
- When there is a set of MC simulations already available from UA, the proposed alternative method can be used to compute the Sobol' index as a by-product.

Although some of these recommendations were documented elsewhere (Saltelli et al., 2008), we emphasize them here for applications in hydrogeology.

## Acknowledgment

This work was funded by the Assistant Secretary for Fossil Energy, National Energy Technology Laboratory, National Risk Assessment Partnership, of the US Department of Energy under Contract no. DEAC02-05CH11231. The authors wish to thank George Pau of Lawrence Berkeley National Laboratory for his technical review, and three anonymous reviewers for their constructive comments.

## References

- Archer, G.E.B., Saltelli, A., Sobol', I.M., 1997. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation* 58, 99–120.  
 Cacuci, D.G., 2003. *Sensitivity and Uncertainty Analysis, Vol. 1: Theory*. Chapman and Hall/CRC Press, Boca Raton, FL.

- Campolongo, F., Cariboni, J., Saltelli, A., An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10), 1509–1518, 2007.
- Cihan, A., Zhou, Q., Birkholzer, J.T., 2011. Analytical solutions for pressure perturbation and fluid leakage through aquitards and wells in multilayered aquifer systems. *Water Resources Research* 47 (10), W10504.
- Deutsch, C.V., Journel, A.G., 1992. Gstat: Geostatistical Software Library and User's Guide. Oxford University Press, New York, USA.
- Fisher, R.A., On the probable error of a coefficient of correlation deduced from a small sample, *Metron* 1 (4): 3–32, 1921.
- Finsterle, S., Pruess, K., 1995. Solving the estimation–identification problem in two-phase flow modeling. *Water Resources Research* 31 (4), 913–924.
- Finsterle, S., Kowalsky, M.B., Pruess, K., 2012. TOUGH: model use, calibration and validation. *Transactions of the American Society of Agricultural and Biological Engineers* 55 (4), 1275–1290.
- Foglia, L., Hill, M.C., Mehl, S.W., Burlando, P., 2009. Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function. *Water Resources Research* 45, W06427, <http://dx.doi.org/10.1029/2008WR007255>.
- Glen, G., Isaacs, K., 2012. Estimating Sobol sensitivity indices using correlations. *Environmental Modelling & Software* 37, 157–166, <http://dx.doi.org/10.1016/j.envsoft.2012.03.014>.
- Hastie, T., Tibshirani, R., Friedman, J.H., 2001. *The elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, USA.
- Hill, M.C., Tiedeman, C.R., 2007. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*. John Wiley & Sons, New York.
- Jung, Y., Zhou, Q., Birkholzer, J.T., Early Detection of Brine and CO<sub>2</sub> Leakage through Abandoned Wells Using Pressure and Surface-Deformation Monitoring Data: Concept and Demonstration, *Advances in Water Resources*, in press, <http://dx.doi.org/10.1016/j.advwatres.2013.06.008>.
- Marrel, A., Iooss, B., Laurent, B., Roustant, O., 2009. Calculations of Sobol indices for the Gaussian process metamodel. *Reliability Engineering & System Safety* 94 (3), 742–751.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33 (2), 161–174.
- Oladyshkin, S., de Barros, F.P.J., Nowak, W., 2012. Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. *Advances in Water Resources* 37, 10–22, <http://dx.doi.org/10.1016/j.advwatres.2011.11.001>, ISSN: 0309-1708.
- Saint-Geours, N., Bailly, J.S., Lavergne, C., Grelot, F., 2010. Latin hypercube sampling of Gaussian random field for Sobol' global sensitivity analysis of models with spatial inputs and scalar output. In: Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Accuracy 2010.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. *Global Sensitivity Analysis: The Primer*. John Wiley and Sons, New York.
- Sobol, I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55 (1–3), 271–280.
- Tang, Y., P. Reed, K. van Werkhoven, and T. Wagener (2007), Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis, *Water Resour. Res.*, 43, W06415, doi:10.1029/2006WR005813.
- Van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology* 324 (1), 10–23.