Vol. 19, No. 1 (2020) 423-444
Revista Mexicana de Ingeniería Química

PERFORMANCE PROFILES FOR BENCHMARKING OF GLOBAL SENSITIVITY ANALYSIS ALGORITHMS

PERFILES DE DESEMPEÑO PARA LA EVALUACIÓN COMPARATIVA DE ALGORITMOS DE ANÁLISIS DE SENSIBILIDAD GLOBAL

F.A. Lucay[1], T. Lopez-Arenas[2]*, M. Sales-Cruz[2], E.D. Gálvez[3], L.A. Cisternas[4]

[1]*Escuela de Ingeniería Química, Pontificia Universidad Católica de Valparaíso, Chile.*
[2]*Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana - Cuajimalpa, Ciudad de México, México.*
[3]*Departamento de Ingeniería de Metalurgia y Minas, Universidad Católica del Norte, Antofagasta, Chile.*
[4]*Departamento de Ingeniería Química y Procesos de Minerales, Universidad de Antofagasta, Antofagasta, Chile.*

Received: March 11, 2019; Accepted: July 9, 2019

Abstract

Nowadays, sensitivity analysis (SA) is a methodology commonly used to identify important parameters that determine the behavior of the model. The SA of a model allows to determine how uncertainties in the model responses (outputs) can be assigned to the values of the model parameters (input variables). The related literature indicates that there are several methods to perform SA. This work addresses the benchmarking of four widely used methods for Global SA (GSA): Sobol-Jansen, Sobol-Baudin, Sobol-Owen and Sobol 2007, based on the concept of performance profile introduced by Dolan and Moré (2002) and the extension by Mahajan *et al*. (2012). To evaluate these methods, a set of 21 models and their variations were considered, which correspond to various applications in chemical engineering (such as heap leaching, water distribution network, milling, flotation circuit, among others). These comparisons show that, although the four GSA methods based on the decomposition of the variance proved to be quite stable, the Sobol-Jansen method presented the best performance, since it is the first to perform GSA in 83% of the models considered and maintains a high performance up to 100%.

Keywords: Global sensitivity analysis, uncertainty, Sobol Method, parametric sensitivity, performance profile.

Resumen

Hoy en día, el análisis de sensibilidad (AS) es una metodología comúnmente utilizada para identificar parámetros importantes que determinan el comportamiento del modelo. El AS de un modelo permite determinar cómo las incertidumbres en las respuestas del modelo (salidas) se pueden asignar a los valores de los parámetros del modelo (variables de entrada). La literatura relacionada indica que hay varios métodos para realizar el AS. Este trabajo aborda la evaluación comparativa de cuatro métodos ampliamente utilizados para el AS global (ASG): Sobol-Jansen, Sobol-Baudin, Sobol-Owen y Sobol 2007, basados en el concepto de perfil de desempeño introducido por Dolan y Moré (2002) y la extensión hecha por Mahajan *et al*. (2012). Para evaluar estos métodos, se consideró un conjunto de 21 modelos y sus variaciones, los cuales corresponden a diversas aplicaciones en ingeniería química (tales como lixiviación, red de distribución de agua, molienda, circuito de flotación, entre otros). Estas comparaciones muestran que, aunque los cuatro métodos ASG basados en la descomposición de la varianza demostraron ser bastante estables, el método Sobol-Jansen presentó el mejor rendimiento, ya que es el primero en realizar ASG en el 83% de los modelos considerados y mantiene un alto rendimiento hasta el 100%.

Palabras clave: Análisis de sensibilidad global, incertidumbre, método de Sobol, sensibilidad paramétrica, índices de desempeño.

1 Introduction

Mathematical models have become indispensable in many fields of science and engineering, since they are used in a wide range of scientific research, such as environmental engineering, chemistry, biology, water resources, molecular design, life sciences, and economic analysis, among many others. The parameters associated with these models have a great influence on the model performance. The specification of the model parameters is not an easy task, some parameters can be measured, but there are many others that cannot be measured or inferred directly.

* *Corresponding author. E-mail*: mtlopez@correo.cua.uam.mx
Tel. (55) 5814 6500
https://doi.org/10.24275/rmiq/Sim547
issn-e: 2395-8472

To determine which parameters have the greatest impact on the model performance and to identify which are the most appropriate values for them, it is necessary to find a way to screen out the most sensitive parameters and evaluate qualitatively the influence of each parameter on the model performance. The sensitivity analysis (SA) is the study of the uncertainties in the responses (outputs) of a mathematical model, which can be distributed to different sources of uncertainties in the values of the model parameters (input variables). SA can identify the parameters of which a reduction in the uncertainty specification will have the most significant impact on improving the model performance. Thus, if some parameters have little or no influence, they can be identified and fixed reasonably at given values over their ranges, then, the computational cost may decrease without reducing the model performance. Therefore, interest in the application of SA has increased in recent years because it has proven to be an effective tool in the development and application of models that present uncertainty (Lucay *et al*., 2015b). These uncertainties can be related to the lack of knowledge (epistemic uncertainties) or randomness (stochastic uncertainties) (Jamett *et al*., 2015). The effect of these uncertainties must be examined to avoid poor design, safety problems, unreliable systems, among other problems. SA methods can help to analyze the effect of uncertainties. However, they can be computationally expensive for complex problems of chemical processes (Duong et al, 2016). Therefore, a way of comparing and discriminating them is needed in order to select the most appropriate method.

The related literature shows the existence of several methods to perform SA, in general, they can be classified into two main groups: local SA and global SA (GSA). The local SA explores changes in the model response by varying one parameter while keeping other parameters constant. In contrast, the GSA examines changes in the model response by varying all parameters at the same time (Gan *et al*., 2014). When SA is applied to process engineering, the local SA methods are the most common (i.e. Austin *et al*., 1975; Gonzalez-Cruz *et al*., 2012; Hernandez-Suarez *et al*., 2008; Soto-Cruz and Paez-Lerma, 2005). In general, the choice of local methods is due to time and computational cost reasons.

In a local SA, the impact of small input perturbations on the model output is studied, considering the assumptions of linearity and normality, as well as local variations. Where small perturbations occur around nominal values (for example, the mean of a random variable) (Ioosss and Lemaître, 2015). Therefore, GSA methods based on a statistical framework appear to overcome the limitations of the local SA methods. GSA considers the whole variation range of the inputs, and a wide set of GSA methods can be found within Saltelli *et al*. (2008). However, the GSA depends not only on the chosen method, but also in the clear specification of the objectives of a study before making a SA, which can be: identify and prioritize the most influential inputs, identify non-influential inputs to correct the nominal values, assign the output behavior depending on the inputs by focusing on a specific domain of the inputs, calibrate some model inputs using available information, and so on (Pappenberger *et al*., 2010; Ioosss and Lemaître, 2015).

The comparison of SA methods is not new, in general, different methods are compared considering the time they take to perform the SA or the similarity of the results obtained. In both cases, this comparison is usually based on specific cases. Some reported works are: regressions-based methods (e.g. Partial Correlation Coefficients, standardized regression coefficients) (Nguyen and Reiter, 2015), regressions-based methods using rank transformation techniques (e.g. Partial Rank Correlation Coefficient, standardized rank regression coefficients) (Nguyen and Reiter, 2015; Van Duc Long and Lee; 2012; Hopfe and Hensen, 2011), variance decomposition-based methods (Sobol and FAST) and screening-based method (Morris method) (Herman *et al*., 2013; Heiselberg *et al*., 2009). Gan *et al*. (2014) compared various SA methods using a simple conceptual hydrological model. While Sepulveda *et al*. (2013) used the Sobol-Jansen and Morris methods to improve mineral concentration plants. The model used by these authors was the global recovery of the plant, and the results indicated similarity in the results obtained with both methods. Brevault *et al*. (2013) compared Sobol, Morris, ANOVA methods and standardized regression coefficients to determine the most significant input variables in the optimal design of aerospace vehicles. Other works that compare GSA methods are the works of Cosenza *et al*. (2013) for applications in sewage water, Ikonen (2016) for fuels and, Galvez and Capuz-Rizo (2016) for project planning.

A similar situation was observed in the development of mathematical programming algorithms. The conventional way of comparing these mathematical programming algorithms has been by interpreting data obtained through benchmarking, i.e. by interpreting tables that show the performance

of each algorithm in a set of problems for a set of metrics such as the time required to solve the problem, number of iterations, estimation error, among other metrics. The literature indicates that there are many ways to interpret the data obtained, in some cases the average or the total accumulated for each metric in all problems is a way to evaluate the performance of an algorithm (Bongartz *et al*., 1997; Benson, 2000). While other authors developed a ranking of methods through a particular metric, such as Vanderbei (1999) and Nash and Nocedal (1991). Billups *et al*. (1995) proposed comparing methods through the relationship (runtime of a specific method) / (best runtime), then the methods are classified according to the percentage of problems for which their resolution time is "very competitive" or "competitive". Beiranvand *et al*. (2017) presented the best practices for comparing optimization algorithms, and provided suggestions for benchmarking those algorithms, such as tables and graphs, and using the concept of performance profiles proposed by Dolan and Moré (2002). These authors used the relationship (computing time of an algorithm) / (best time of all algorithms) as a performance metric. The methodology of Dolan and Moré (2002) is regularly used to compare solvers in mathematical programming with more than 800 citations of their work. However, one of the limitations of Dolan and Moré's proposal is that the performance profile only allows to determine the probability that a method has a performance at least τ times slower than the best performance solver, for a given factor $\tau > 1$. Therefore, Mahajan *et al*. (2012) extended the definition of performance profile to determine the probability that a method is faster than another method for a given factor $\tau < 1$.

In this paper, four GSA methods based on the variance decomposition are evaluated and compared based on the performance profiles and their extension indices proposed by Dolan and Moré (2002) and Mahajan *et al*. (2012). A set of 21 mathematical models and their variations are considered as case studies, corresponding to several applications in chemical engineering (such as heap leaching, water distribution network, milling, flotation circuit, among others). The four GSA methods are selected because, according to Saltelli *et al*. (2010), the methods based on the variance decomposition show great versatility and effectiveness. Among the aspects of versatility, these methods allow the uncertainties of the input variables to have different types of probability distribution functions, which is not possible with Morris method for instance.

2 Materials and methods

2.1 Sensitivity analysis

The sensitivity analysis consists of identifying the contribution of the uncertainties of the input variables in the uncertainties of the outputs of a mathematical model. There are two types of sensitivity analysis: local SA and GSA. The first quantifies the rate of change of the model output due to small variations in the uncertainties of the input variables. This quantification is based on the derivative of the model output, which is calculated at a specific point in the space of the input variables (Cariboni *et al*., 2007). One of the disadvantages of local sensitivity analysis is that the choice of the assessment point could sharply influence the results, especially when the sampling space of the input variables is affected by considerable uncertainty. While GSA considers the full range of uncertainties of the input variables, such uncertainties are characterized by probability distribution functions (PDF). GSA considers the following six steps: 1) determine the objective function, 2) select the input variables, 3) assign a range and type of distribution of the probabilities of the input variables, 4) applying a sampling design to generate samples, size N from the distributions of the input variables 5) assess the model for the generated samples, thus obtaining N values of the objective function, 6) implement the results of step 5 to perform uncertainty analysis and determine the importance of the input variables on the model output (Lilburne and Tarantola, 2009). As mentioned above, there are several methods to perform the sensitivity analysis, but among them, the methods that are based on the decomposition of the variance stand out, due to their versatility and efficiency (Saltelli *et al*., 2010). Therefore, in this study, we will implement such methods, specifically those based on the Sobol method. In the Sobol method, the variance of the model output is broken down in terms of increasing dimensions, called partial variances, which represent the contribution of the input variables to the model uncertainty, that is, the variance of the model output can be expressed as follows (Confalonieri *et al*., 2010):

$$V(Y) = \sum_{i=1}^{n} V_i + \sum_{i \le j \le n}^{n} V_{ij} + \ldots + V_{i,\ldots,n} \qquad (1)$$

or equivalently

$$1 = \sum_{i=1}^{n} \frac{V_i}{V(Y)} + \sum_{i \le j \le n}^{n} \frac{V_{i,j}}{V(Y)} + \ldots + \frac{V_{i,\ldots,n}}{V(Y)} = \sum_{i=1}^{n} S_i + \sum_{i \le j \le n}^{n} S_{i,j} + \ldots + S_{1,2,\ldots,n} \quad (2)$$

where $V(Y)$ is the model variance, $Y = f(x_1, x_2, \ldots, x_n)$ is a scalar, x_i is a model input, V_i represent the first order effects of each input variable $x_i (V_i = V[E(Y|x_i)]$ and $V_{ij} (V_{ij}[E[Y|x_i, x_j)] - V_i - V_j)$ to $V_{i,\ldots,n}$ are the interactions of the n factors. Since the calculation of the partial variances has a high computational cost, Homma and Saltelli (1996) introduced the concept of total sensitivity index. This index allows determining the average effect of an input variable, taking into account all its interactions with the other input variables. The methods considered in this work allow to determine such indexes, i.e., they allow to calculate the first order index (S_i):

$$S_i = \frac{V[E(Y|x_i)]}{V(Y)} \quad (3)$$

and the total sensitivity index (S_i^T)

$$S_i^T = \frac{E(V(Y|x_{\sim i}))}{V(Y)} = 1 - \frac{V[E(Y|x_{\sim i})]}{V(Y)} \quad (4)$$

for the input variable x_i, where $x_{\sim i}$ indicates the consideration of all the input variables and their combinations, except those involving the input variable x_i. The first order index is important when the goal is to determine which of the input variables is the most influential in the model output. While the total sensitivity index is important when the goal is to identify the input variables that do not affect the model output.

In the case that the input variable x_i does not interact with the other variables, we will have $S_i = S_i^T$, otherwise $S_i < S_i^T$. If $S_i^T = 0$, then we will have that the input variable x_i is not important in the model output and can be fixed at a given value. An important implication of the sensitivity indexes is that if the sum of the first order indexes is equal to $1\left(\sum_{i=1}^{n} S_i = 1\right)$ and the sum of the total indexes is also equal to $1\left(\sum_{i=1}^{n} S_i^T = 1\right)$, then the model is perfectly additive, otherwise the model is not additive (Saltelli *et al.*, 2008).

According to Lilburne and Tarantola (2009), a way to estimate the terms in the decomposition (Eq. (2)) is through the Sobol method (Sobol, 1993). This method involves: (1) choose an integer N; (2) generate a matrix, size $(N, 2k)$ of quasi random numbers from sampling the input variables from their respective distribution functions (here k is the number of variables of the model input); (3) divide the matrix into two sub-matrices A and B size (N, k); (4) form the matrix D_i from the columns of the matrix A, except the *ith* column, which is taken from the matrix B; and similarly, form the matrix C_i from the columns of matrix B, except the *ith* column, which is taken from matrix A; (5) assess the model output in the matrices A, B, C_i and D_i, obtaining $Y_A = f(A)$, $Y_B = f(B)$, $Y_{C_i} = f(C_i)$, $Y_{D_i} = f(D_i)$, with such vectors and using the following equation:

$$S_i = \frac{V[E(Y|x_i)]}{V(Y)} = \frac{Y_A T_{C_i} - f_0^2}{Y_A Y_A - f_0^2} = \frac{\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{C_i}^{(j)} - f_0^2}{\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_A^{(j)} - f_0^2}$$

$$with \quad f_0^2 = \left(\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)}\right)^2 \quad (5)$$

where the Sobol method can estimate the first order index for the input variable x_i; (6) finally, estimate the total sensitivity index for the input variable xi with the Sobol method, using the following expression:

$$S_i^T = \frac{E(V(Y|x_{\sim i}))}{V(Y)} = 1 - \frac{V[E(Y|x_{\sim i})]}{V(Y)} = 1 - \frac{Y_A Y_{D_i} - f_0^2}{Y_A Y_A - f_0^2} = 1 - \frac{\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{D_i}^{(j)} - f_0^2}{\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_A^{(j)} - f_0^2} \quad (6)$$

Over the years, many authors have proposed different modifications to equations (5) and (6) made by Sobol (1993). Table 1 shows some of these changes. A somewhat different proposal was given by Baudin *et al.* (2016), who considered the linear correlation coefficient to propose the following expressions to determine the first order and total order sensitivity indexes, respectively:

$$S_i = \rho(Y_B, Y_{D_i}) = \frac{\sum_{j=1}^{N}\left(Y_B^{(j)} - \overline{Y}_B\right)\left(Y_{B_i}^{(j)} - \overline{Y}_{D_i}\right)}{N S_B S_{D_i}} \quad (7)$$

Table 1. Formulas for calculating S_i and S_i^T.

Table 1. Formulas for calculating S_i and S_i^T.

$V\left(E\left(Y/x_{\sim i}\right)\right)$ for S_i	Reference
$\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{C_i}^{(j)} - \left(\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)}\right)^2$	Sobol (1993)
$\frac{1}{N}\sum_{j=1}^{N} Y_B^{(j)} Y_{D_i}^{(j)} - \left(\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)}\right)\left(\frac{1}{N}\sum_{j=1}^{N} Y_B^{(j)}\right)$	Saltelli (2002)
$\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{C_i}^{(j)} - \left(\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)}\right)\left(\frac{1}{N}\sum_{j=1}^{N} Y_B^{(j)}\right)$	Tarantola et al. (2006)
$\frac{1}{N}\sum_{j=1}^{N} Y_B^{(j)}\left(Y_{D_i}^{(j)} - Y_A^{(j)}\right)$	Saltelli et al. (2010)
$V(Y) - \frac{1}{2N}\sum_{j=1}^{N}\left(Y_B^{(j)} - Y_{D_i}^{(j)}\right)^2$	Jansen (1999)
$V\left(E\left(Y/x_{\sim i}\right)\right)$ for S_i^T	Reference
$\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{D_i}^{(j)} - \left(\frac{1}{N}\sum_{k=1}^{N} Y_A^{(k)}\right)^2$	Sobol (1993)
$V(Y) - \frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)} Y_{D_i}^{(j)} + f_0^2$	Homma and Saltelli (1996)
$\frac{1}{N}\sum_{j=1}^{N} Y_A^{(j)}\left(Y_A^{(j)} - Y_{D_i}^{(j)}\right)$	Sobol (2007)
$\frac{1}{2N}\sum_{j=1}^{N}\left(Y_A^{(j)} - Y_{D_i}^{(j)}\right)^2$	Jansen (1999)

$$S_i^T = 1 - \rho(Y_A, Y_{D_i}) = 1 - \frac{\sum_{j=1}^{N}\left(Y_A^{(j)} - \overline{Y}_A\right)\left(Y_{D_i}^{(j)} - \overline{Y}_{D_i}\right)}{NS_A S_{D_i}} \tag{8}$$

While Owen (2013) proposed using three sample matrices: A, B and C size (N,k), which form matrix D_i from the columns of matrix C, except the *ith* column, which is taken from matrix A. Similarly, the matrix E_i is formed from the columns of matrix A, except the *ith* column, which is taken from the matrix B. Then, his proposal is:

$$V(E(Y|x_{\sim i})) = \frac{1}{N}\sum_{j=i}^{N}\left(Y_A^{(j)} - Y_{D_i}^{(j)}\right)\left(Y_{E_i}^{(j)} - Y_B^{(j)}\right) \tag{9}$$

2.1.1 *Performance profiles*

Consider that we are interested in evaluating and comparing a set of GSA methods applied to an M set of models. The set contains n_{sa} GSA methods and the set M contains n_m models to evaluate. Benchmarking is obtained by applying the set of SA methods to the M set of models and recording relevant information, such as the standard deviation (SD) error of the total sensitivity index obtained when applying each method and the time it took each method to perform the sensitivity analysis. Where $t_{m,sa}$ is defined as the time required for the *sa* method to perform the GSA in the model m.

The slow performance radius ($r^s_{m,sa}$) of the *sa* method in the model m is defined as the calculation time of the *sa* method divided by the shortest calculation time of all GSA methods in performing the GSA in model m, i.e.

$$r^s_{m,sa} = \frac{t_{m,sa}}{min\{t_{m,i} : i \in SA\}} \tag{10}$$

To evaluate the performance of a method in all models of the M set, the slow performance profile ($\rho^s_{sa}(\tau)$) is used, which is defined by:

$$\rho^s_{sa}(\tau) = \frac{1}{n_m} size\{m \in M : r^s_{m,sa} \leq \tau\} \tag{11}$$

This last expression is the cumulative distribution function of the slow performance radius. The definition of the slow performance profile can be interpreted as the probability that the GSA *sa* method can perform the sensitivity analysis at least τ times slower than the best performance method in range $(1,\tau)$, in case that $\tau > 1$. These definitions are based on the work of Dolan and Moré (2002).

Now consider that we want to know the probability that a GSA *sa* method is faster than another GSA method for a factor $\tau < 1$. For this, similar to the work of Mahajan *et al.* (2012) for mathematical programming solvers, we propose a fast performance radius ($r^f_{m,sa}$) define as

$$r^f_{m,sa} = \frac{t_{m,sa}}{min\{t_{m,i} : i \in SA, i \neq sa} \tag{12}$$

this is, the fast performance radius is defined as the calculation time of the *sa* method divided by the shortest calculation time of all GSA methods excluding the *sa* method to perform GSA in the model m. Moreover, the fast performance profile ($\rho^f_{sa}(\tau)$) is defined as:

$$\rho^f_{sa}(\tau) = \frac{1}{n_m} size\{m \in M : r^f_{m,sa} \leq \tau\} \tag{13}$$

Therefore, we have that

$$size\{m \in M : r^s_{m,sa} \leq \tau\} = size\{m \in M : r^f_{m,sa} \leq \tau\},$$

$$\tau \geq 1 \tag{14}$$

and, consequently, $\rho^s_{sa}(\tau) = \rho^f_{sa}(\tau)$ when $\tau \geq 1$ motivates the expression "extension of the slow

performance profile". In particular, $\rho^{s}_{sa}(1)$ represents the probability that the GSA *sa* method is faster than any other method for $\tau = 1$ and $\rho^{f}_{sa}(\tau)$ indicates the probability that the GSA *sa* method is at least $1/\tau$ times faster than another method in the range $(0, \tau)$.

Onwards, the goal is to graph the performance profile of the different methods and compare them.

3 Benchmarking data

The times used to obtain the performance profiles of the GSA methods that will be shown in Section 4 were generated when performing the sensitivity analysis with the methods of Sobol-Jansen (Jansen, 1999), Sobol-Baudin (Baudin *et al*., 2016), Sobol-Owen (Owen, 2013) and Sobol 2007 (Sobol, 2007) in the *M* set of models. Such set consisted of the following base models: heap leaching, Bateman equation, Legendre polynomial (P), Lotka-Volterra equation, semi-autogenous grinding (SAG) milling, transport model, pipe network design (ND), flotation circuit, Ishigami function (F), flotation circuit, Lorenz's attractor, reverse osmosis (RO) stage, RO membrane, linear model, heat equation, matrix system, polynomial (P) division, Linear function (F), trigonometric function (F), Sobol function, K function and B function. Moreover, some variations of the base models were considered, which can be seen in Table 2. Equations and some details of each model can be found in Appendix A. It is relevant to mention that the model parameters are given in the corresponding references of Appendix A, and for this work, the base values of the model parameters were the average of the range in the case of uniform distributions, and the mean in the case of the normal distributions.

Here it is worth mentioning that models such as K function, Sobol function, B function and Ishigami function were chosen because they commonly appear as test models in the sensitivity literature. While models such as SAG milling, flotation circuit, transport, heap leaching, pipe ND and RO models were chosen because they represent important mining processes. The rest of the models were chosen because they usually appear in chemical engineering problems, besides they add diversity within the set of models. Moreover, models as B function, linear function, K function, or pipe ND have a low GSA time, which hinders the development of the performance profiles and, consequently, hinders the analysis of the results as well.

In Table 2, the different variations performed on the base models can be seen. In some cases, the input variables of the models had distributions of the following types: uniform, normal or uniform-normal. While in other cases, the uncertainty interval of the input variables was varied by ±10% of the initial interval. In other cases, the uncertainty interval and the type of uncertainty were varied, while at times there was no changes in the base model. Besides, the same table shows in what type of industry, scientific area or process the models are applied.

Table 2. Description, type, variations and main variables of the models from M set.

No.	Model	Type distribution	Variation ranges of uncertainty	No. variables	Main variables	Implementation
1	Heat equations	U	base	3	t	Photovoltaic energy (Bien and Musikowski, 2008)
2	Ishigami F	U	base	3	x_1,x_2	
2a	Ishigami F	N	10% base	3	x_2	
2b	Ishigami F	U	10% base	3	x_1, x_2	
2c	Ishigami F	N	base	3	x_2	
2d	Ishigami F	N-U	base	3	x_2	
3	Flotation circuit	U	base	6	T^C, T^{SC}, T^{CS}	Mining (Lucay *et al*., 2015c), paper recycling (Nie *et al*., 1998)
3a	Flotation circuit	N	10% base	6	T^C, T^{SC}, T^{CS}	
3b	Flotation circuit	U	10% base	6	T^C, T^{SC}, T^{CS}	
3c	Flotation circuit	N	base	6	T^C, T^{SC}, T^{CS}	
3d	Flotation circuit	N-U	base	6	T^C,T^{SC}	
4	Flotation-	U	base	45	J_g, $k_{\max,mol,C}$,	

	El Salvador				$R_{\max,mol,R}$	
5	Legendre P	U	base	2	d,x	Dairy industry (Pool
5a	Legendre P	U	10% base	2	d,x	*et al*., 2000;
5b	Legendre P	N-U	base	2	d, x	Campolina *et al*.,
5c	Legendre P	N	base	2	d, x	2014)
6	Milling	U	base	7	F, J_b	Mining (Lucay *et al*., 2017), Cement treatment (Austin *et al*., 1975)
7	RO stage	U	base	6	n	Seawater desalination, sewage water treatment, ultra-pure water production (Garud *et al*., 2011)
8	RO membrane	U	base	4	T, P_f, Q_f, CC_f	
9	Lorenz equation	U	base	3	x_1, x_2, x_3	Climate change (Palmer *et al*., 2005)
10	Lotka-Volterra eq.	U	base	3	d	Population modeling in ecosystems (Zhue and Yin, 2009)
11	Linear F	U	base	6	x_4	
11a	Linear F	N	10% base	6	x_4	
11b	Linear F	U	10% base	6	x_2, x_4, x_5	
11c	Linear F	N	base	6	x_4	
11d	Linear F	N-U	Base	6	x_4	
12	Matrix system	U	base	4	x_1	
12a	Matrix system	U	10% base	4	x_1	
12b	Matrix system	N	base	4	x_1	
13	Transport model	U	base	5	x_2	
13a	Transport model	U	10% base	5	x_2	
13b	Transport model	N	base	5	x_2	
14	Pipe ND	U	base	13	M_3, M_4, M_5	Desalinated water
14a	Pipe ND	N	10% base	13	M_3, M_4, M_5	transport (Herrera
14b	Pipe ND	N	base	13	M_3, M_4, M_5	*et al*., 2015)
14c	Pipe ND	U	10% base	13	M_3, M_4, M_5	
14d	Pipe ND	N-U	base	13	M_3, M_5	
15	P division	U	base	5	x_2, x_3	
15a	P division	U	10% base	5	x_2, x_3	
15b	P division	N	base	5	x_2, x_3	
16	Trigonometry F	U	base	2	x_1	
16a	Trigonometry F	N	10% base	2	x_1	
16b	Trigonometry F	U	10% base	2	x_1	
16c	Trigonometry F	N	base	2	x_1	
16d	Trigonometry F	N-U	base	2	x_1	
17	Bateman eq.	U	base	8	l_4, c_1	Environmental impact
17a	Bateman eq.	N	base	8	l_4, c_1	of uranium mining
17b	Bateman eq.	N-U	base	8	l_4, c_1	(Cotta and Naveira,
17c	Bateman eq.	U	10% base	8	l_4, c_1	2007)
18	Heap leaching	U	base	11	R_∞	Mining (Mellado *et al*., 2009), agriculture (Hu *et al*., 2014)
19	Sobol function	U	base	8	X_1	

19a	Sobol function	N	base	8	X_1
20	K function	U	base	5	X_1, X_2, X_3, X_4, X_5
21	B function	N	base	10	X_1, X_2, X_3, X_4, X_5 w_1, w_2, w_3, w_4, w_5

U (uniform), N (normal)

Table 3. Standard deviation error of the total sensitivity index (error).

		error			
No.	Model	Sobol-Jansen	Sobol-Baudin	Sobol-Owen	Sobol 2007
1	Heat equations	0.0128	0.011	0.013	0.0123
2	Ishigami F	0.0137	0.0127	0.0136	0.0138
2a	Ishigami F	0.0165	0.0158	0.015	0.015
2b	Ishigami F	0.0152	0.0154	0.0142	0.0134
2c	Ishigami F	0.0141	0.0131	0.0134	0.0133
2d	Ishigami F	0.012	0.0128	0.0137	0.0128
3	Flotation circuit	0.0094	0.009	0.0137	0.015
3a	Flotation circuit	0.0109	0.01	0.013	0.017
3b	Flotation circuit	0.0099	0.009	0.0145	0.0187
3c	Flotation circuit	0.0102	0.0104	0.0137	0.019
3d	Flotation circuit	0.0093	0.0086	0.0131	0.017
4	Flotation-El Salvador	0.0106	0.0103	0.0177	0.018
5	Legendre P	0.0108	0.0108	0.0129	0.0107
5a	Legendre P	0.0108	0.00997	0.013	0.011
5b	Legendre P	0.0096	0.0096	0.0128	0.0101
5c	Legendre P	0.01	0.0098	0.015	0.0103
6	Milling	0.0123	0.012	0.018	0.017
7	RO stage	0.0143	0.0135	0.018	0.015
8	RO membrane	0.0145	0.0112	0.017	0.018
9	Lorenz equation	0.017	0.0163	0.016	0.018
10	Lotka-Volterra	0.0186	0.017	0.017	0.018
11	Linear F	0.01	0.01	0.0146	0.015
11a	Linear F	0.01	0.01	0.0172	0.0172
11b	Linear F	0.01	0.01	0.0138	0.0132
11c	Linear F	0.01	0.01	0.0163	0.017
11d	Linear F	0.01	0.01	0.0146	0.0164
12	Matrix system	0.01	0.0104	0.015	0.016
12a	Matrix system	0.0106	0.01	0.014	0.012
12b	Matrix system	0.0104	0.0098	0.017	0.018
13	Transport	0.0102	0.00995	0.0167	0.017
13a	Transport	0.0101	0.0102	0.0163	0.017
13b	Transport	0.0108	0.0102	0.017	0.018
14	Pipe ND	0.0085	0.0066	0.018	0.016
14a	Pipe ND	0.0116	0.0106	0.018	0.018
14b	Pipe ND	0.0083	0.0078	0.017	0.018
14c	Pipe ND	0.0104	0.0082	0.0172	0.0171
14d	Pipe ND	0.00784	0.00786	0.0178	0.017
15	P division	0.01	0.01	0.012	0.0123
15a	P division	0.01	0.0096	0.0144	0.01
15b	P division	0.0107	0.0098	0.013	0.015

16	Trigonometry F	0.011	0.0093	0.0135	0.0117
16a	Trigonometry F	0.0177	0.0099	0.019	0.017
16b	Trigonometry F	0.011	0.01	0.012	0.0116
16c	Trigonometry F	0.0098	0.01	0.017	0.019
16d	Trigonometry F	0.01	0.01	0.0148	0.0094
17	Bateman eq.	0.0103	0.0098	0.017	0.0152
17a	Bateman eq.	0.00959	0.00971	0.0145	0.0177
17b	Bateman eq.	0.00959	0.00994	0.0162	0.0157
17c	Bateman eq.	0.0097	0.0094	0.0135	0.0167
18	Heap leaching	0.01001	0.0106	0.015	0.018
19	Sobol function	0.0105	0.0104	0.013	0.012
19a	Sobol function	0.0105	0.0102	0.0182	0.0103
20	K function	0.0108	0.0107	0.014	0.0101
21	B function	0.0105	0.0105	0.016	0.0102

U (uniform), N (normal)

4 Results and discussion

4.1 Sensitivity analysis

The requirement of GSA methods was only to have 2 or 3 sample matrices, size (N,k), where N is the number of rows and k is the number of input variables and a semi-random sampling design, in our work the Monte Carlo method was used. Here it is important to mention that all the calculations were developed in the computational platform called Rstudio, which is an integrated development environment for the R programming language, dedicated to statistical computing and graphics. After performing the sensitivity analysis, the time taken by the GSA method to perform sensitivity analysis, the standard deviation (SD) error of the global sensitivity index and the most relevant variables in each model were recorded. The breakdown of the sensitivity analysis results is reported in Table 3, where it can be seen that all values of the SD error of the total sensitivity index are small, in the range [0.0066, 0.019]. When analyzing these values, it is not possible to define the method with greater efficiency to perform GSA, because sensitivity indices provided by GSA methods depend on the sample size used. Pianosi *et al*. (2016) reported that for each input variable of the model, close to 1000 data should be taken to guarantee reliable results. However, this criterion is not feasible in large models.

In particular, the GSA methods included in Rstudio (the programming language used in this work) provide the SD related to sensitivity indices. These SDs are calculated using the bootstrap method, which considers the resampling of the initial sample with a certain number of times. Then, the sensitivity indexes are calculated for each resampling.

Finally, the SDs are determined using all the resampling for each sensitivity index. Note that SD decreases as the sample size increases. When the maximum SD was approximately equal to 0.01 (error), the sample size was considered optimal. The error was chosen arbitrarily and the resampling was equal to 100 in all cases. Therefore, the reported SDs indicate that the used sample sizes provided reliable results, in other words, the effect of the sample sizes on the variation of the sensitivity indexes was minimal.

Table 4. Time required for the *sa* method to perform GSA ($t_{m,sa}$)

		$t_{m,sa}$ (s)			
No.	Model	Sobol-Jansen	Sobol-Baudin	Sobol-Owen	Sobol 2007
1	Heat equations	1551.9	594.9	2270.7	679.4
2	Ishigami F	0.58	0.98	104.74	0.8
2a	Ishigami F	0.83	1.34	101.6	1.52
2b	Ishigami F	0.74	1.09	103.75	1.19
2c	Ishigami F	0.61	0.84	97.59	1.14

2d	Ishigami F	0.58	0.85	96.59	1.06
3	Flotation circuit	2.55	4.01	227.14	1113.2
3a	Flotation circuit	1.85	3.18	191.62	902.75
3b	Flotation circuit	2.86	5.11	293.97	1085.52
3c	Flotation circuit	2.29	3.8	210	1249.48
3d	Flotation circuit	1.87	3.13	188.95	967.69
4	Flotation-El Salvador	66.69	201.03	1071.07	2452.716
5	Legendre P	4.11	2.92	87.83	2.39
5a	Legendre P	18.94	6.65	80.07	4.81
5b	Legendre P	4.17	2.85	86.94	2.28
5c	Legendre P	7.48	6.56	56.45	1.83
6	Milling	249.7	439.58	15763	49078
7	RO stage	777600	1041984	2083968	7293888
8	RO membrane	256801	343283.18	686566	3432831.8
9	Lorenz equation	19574.5	22152.3	79834.5	63800.7
10	Lotka-Volterra	1974.5	682	4536.3	3902
11	Linear F	0.12	0.31	117.33	8.31
11a	Linear F	0.14	0.33	120.58	32.62
11b	Linear F	0.09	0.22	111.33	8.56
11c	Linear F	0.12	0.3	114.12	30.38
11d	Linear F	0.19	0.44	117.28	19.58
12	Matrix system	3.84	4.91	92.83	51.23
12a	Matrix system	3.07	4.66	92.19	51.33
12b	Matrix system	3.25	5.11	86.85	91.69
13	Transport	22.89	35.7	131.73	1502.9
13a	Transport	22.37	32.17	167.85	1670.94
13b	Transport	23.54	30.87	198.92	1749.61
14	Pipe ND	9.83	16.49	215.47	726.7
14a	Pipe ND	6.83	12.97	201.89	791.41
14b	Pipe ND	8.05	19.25	295.97	968.2
14c	Pipe ND	8.32	15.7	238.89	924.44
14d	Pipe ND	7.95	14.07	232.89	755.11
15	P division	0.75	1.3	173.92	30.49
15a	P division	0.92	1.31	141.39	31.43
15b	P division	0.63	1.18	182.25	60.39
16	Trigonometry F	0.67	0.5	44.03	18.7
16a	Trigonometry F	0.58	0.58	46.78	34.38
16b	Trigonometry F	0.48	0.5	46.21	19.08
16c	Trigonometry F	1.56	0.59	54.879	21.63
16d	Trigonometry F	0.52	0.48	46.02	35.63
17	Bateman eq.	15.42	26.13	263.42	29.75
17a	Bateman eq.	13.95	21.74	257.01	33.67
17b	Bateman eq.	14.65	17.29	266.31	24.65
17c	Bateman eq.	11.9	20.2	266.64	20.11
18	Heap leaching	1.34	2.01	244.2	851.4
19	Sobol function	0.48	0.98	242.81	1.49
19a	Sobol function	1.08	1.15	247.16	1.97
20	K function	1.33	2.26	137.76	5.59
21	B function	2.13	3.99	433.8	3.36

Therefore, the time required for the *sa* method to perform GSA ($t_{m,sa}$) was calculated and the corresponding values are in Table 4. It is important to mention that the two models, the RO stage and the RO membrane, with the greatest time $t_{m,sa}$ are due to the high non-linearity exhibited by these models together with the use of a solver. Then, these time values were used to calculate the slow and fast performance profiles, which will help to discriminate the best GSA method.

4.2 *Performance profiles*

In this section, the performances of Sobol-Jansen, Sobol-Baudin, Sobol-Owen and Sobol 2007 methods in the *M* set are examined.

Figure 1a shows the slow performance profile from the four GSA methods, which establishes that the Sobol-Jansen method is the most efficient to perform GSA in the *M* set. The Sobol-Jansen method is higher in the $\rho^{s}_{Sobol-Jansen}(1) = 83\%$ of the cases considered, exceeded only in four and five cases by the Sobol 2007 and Sobol-Baudin methods, respectively. The Sobol-Baudin method reaches a performance of 40%, 60% and 80% at least 1.5, 1.7, 2.0 times slower, respectively, than the Sobol-Jansen method. While the Sobol-Owen and Sobol 2007 methods do not present a competitive performance within the interval considered for τ. On the other hand, Figure 1b shows the faster performance profile, this allows to deduce that the Sobol-Jansen method achieves a performance close to 10%, 40% and 80% at least 2.0, 1.66 and 1.03 times faster, respectively, than any other method of the SA set.

Also from Figure 1, we have that approximately $\tau = 2$ the Sobol-Baudin method achieves a competitive performance with respect to the Sobol-Jansen method. While performance of Sobol 2007 approaches 25% and Sobol-Owen achieves a performance of 0%, i.e., it is not capable of carrying out the global sensitivity analysis to any model of the SA set.

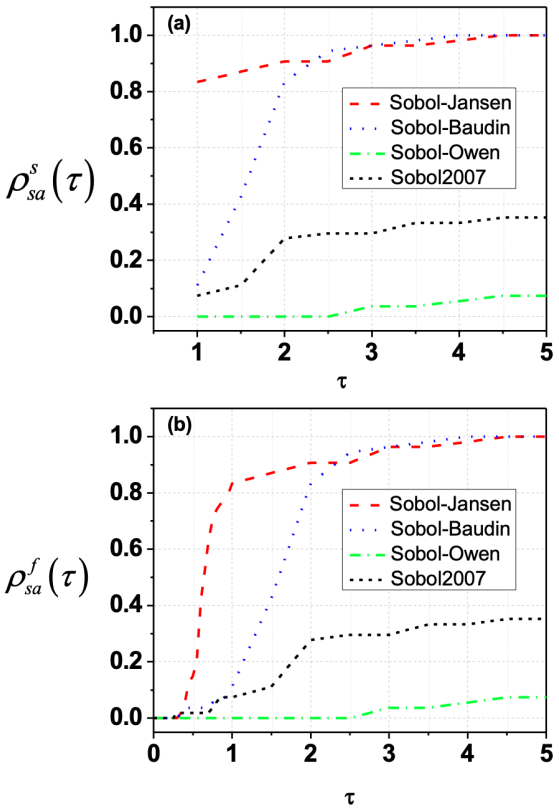

Fig. 1. Performance profiles in [0, 5]: (a) slow performance profile and (b) fast performance profile.

When $\tau = 5.0$ the Sobol-Baudin and Sobol-Jansen methods reach a 100% performance, while the method of Sobol 2007 achieves a performance of approximately 36% and the Sobol-Owen method has a performance of approximately 8%. In general, Figure 1 shows that the slow and fast performance profiles are functions that do not decrease and are piecewise constant. To see how follows the evolution of the methods performance profiles, the range of the interval of the values for τ are expanded, as shown in Figure 2.

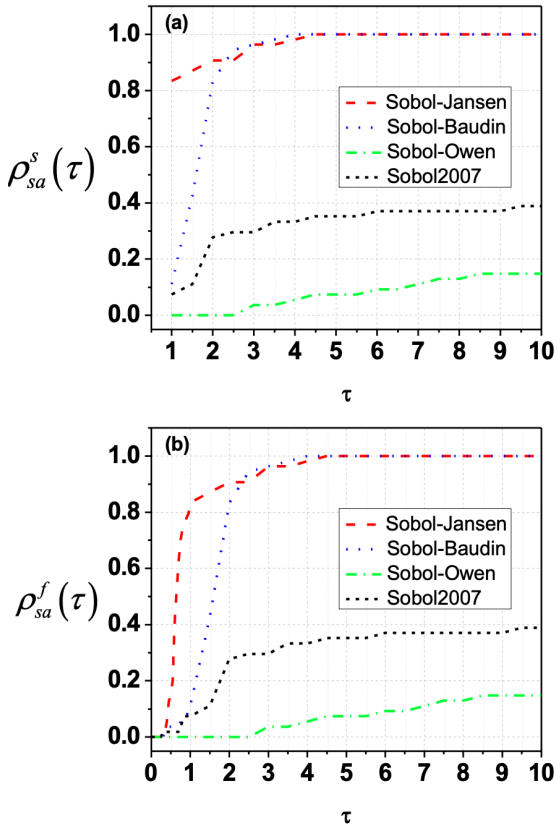

Fig. 2. Performance profiles in [0, 10]: (a) slow performance profile and (b) fast performance profile.

Figure 2a shows that the slow performance of the Sobol 2007 method achieves a performance of almost 20%, 30% and 38% at least 1.7, 3.3 and 9.2 times slower, respectively, than the Sobol-Jansen method. While the Sobol-Owen method reaches a slow performance, almost 10%, at least 7.1 times slower than the Sobol-Jansen method. For $\tau = 10$, the performance of the Sobol 2007 and Sobol-Owen methods is approximately 17% and 40%, respectively. Considering that is the fraction of the models that the sa sensitivity analysis method is not able to perform within a range $(1,\tau)$, we have that the Sobol-Owen method cannot perform the sensitivity analysis to almost 83% of the models of the SA set in the range of τ considered. According to Figures 1 and 2, the progress of the performance profile is not linear, so a logarithmic scale in base 2 will be used for the values of τ, in order to not to lose information, so only values of $\tau \geq 1$ are considered.

Figure 3 shows the slow performance profile for the four GSA methods. From this figure, the following order is established for the performance of the GSA methods: 1) Sobol-Jansen, 2) Sobol-Baudin, 3) Sobol 2007 and 4) Sobol-Owen. The fourth place for Sobol-Owen method could be because this method uses 3 sample matrices and not 2 as the other methods, which increases the number of calculations and, therefore, the time to perform the sensitivity analysis. Also, Owen (2013) indicates that this method is not efficient when the first order index that is calculated is large. The third place of the method of Sobol 2007 could be because its performance is competitive only in models such as Sobol-function, K-function, B-function, Bateman equation, polynomial Legendre and models that include trigonometric functions, i.e. strongly non-linear and non-monotonic models (see Appendix A). In the other models, their performance is poor because to obtain a reliable index, a greater number of iterations was necessary than with other methods to achieve a small error in the standard deviation of the total sensitivity index. This was especially observed in complex models and models that involve a large number of variables. The second place for the Sobol-Baudin method is that, although its performance is very competitive with respect to the Sobol-Jansen method from approximately $\log_2(2)$, the latter performs a sensitivity analysis first in almost 83% of the cases considered in this study. As mentioned above, the Sobol-Baudin method calculates the sensitivity indexes using the equation of linear correlation coefficients, which considering the results is quite efficient. In part, these results were presented by Saltelli *et al*. (2010), who compared the Sobol-Jansen and Sobol 2007 methods through three models. Their results indicate that the Sobol-Jansen method is computationally more efficient, which can also be derived from Sobol (2001), specifically from the Theorem 4 described in his work.

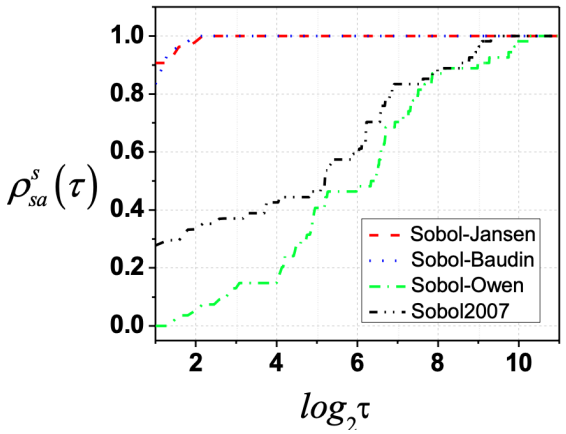

Fig. 3. Slow performance profile in $[1, \log_2(2000)]$.

On one hand, Figure 3 shows that all GSA methods are able to perform sensitivity analysis in 100% of the models considered in this study. On the other hand, these results are positive, since they indicate that the GSA methods based on variance decomposition are quite stable when performing this procedure. However, on the other hand, they are negative, since they question the way of selecting the different models considered. Such models were selected considering different applications and because they appeared in different publications related to GSA, others because they have interesting mathematical characteristics to be analyzed. As Dolan and Moré (2002) indicated that the way of selecting the elements of the test set remains an open problem.

One of the implications of the results of this work is the validation of the use of the Sobol-Jansen method in the methodology proposed by Lucay *et al*. (2015b) to determine critical variables to avoid unwanted responses in the process design. While the methodologies, as proposed by Lamoureux *et al*. (2014) for early validation of health indicators, could be improved since it involves the combination of the sensitivity analysis of Morris and Sobol. First, the input variables with low sensitivity index are determined with Morris, which are then fixed, and then with a reduced input variables model, the most relevant input variables are determined through the Sobol method. That is, in the work of Lamoureux *et al*. (2014), the Sobol method could be changed by the Sobol-Jansen method in the hope of reducing computational costs.

Conclusions

This work presents a procedure for comparing global sensitivity analysis methods based on performance profiles. A large set of models were considered, which were selected because some of them appeared in different works related to GSA, others because they have interesting mathematical characteristics to analyze in the chemical engineering area. The slow performance profile allows determining the probability that a GSA method can perform GSA at least τ times slower than the best performance method for $\tau \geq 1$. The fast performance profile allows determining the probability that a GSA method is faster than another method at least $1/\tau$ times, for $\tau < 1$. Although the four GSA methods based on the decomposition of the variance proved to be quite stable, the Sobol-Jansen method presented the best performance, since it is the first to perform GSA in 83% of the considered models and maintains a high performance up to 100%. Which additionally means that the Sobol-Jansen method uses less computational resources. And, finally, it is worth mentioning that the GSA methodology can be extended and applied to multiple response models, but analyzing one response at a time.

Acknowledgements

The financial support from CONICYT (Fondecyt 1171341) is gratefully acknowledged.

References

Archer, G. E. B., Saltelli, A., and Sobol, I. M. (1997). Sensitivity measures, anova-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation https://www.tandfonline.com/loi/gscs2058*, 99-120. DOI:10.1080/00949659708811825.

Austin, L. G. (1990) A mill power equation for SAG mills. *Minerals & Metallurgical Processing 7*, 57-62.

Austin, L. G., Luckie, P. T., and Wightman, D. (1975). Steady-state simulation of a cement-milling circuit. *International Journal of Mineral Processing 2*, 127-150. DOI:10.1016/0301-7516(75)90017-4.

Baudin, M., Boumhaout, K., Delage, T., Iooss, B., and Martinez, J.-M. (2016). Numerical stability of Sobol indices estimation formula. *Proceedings of the 8th International Conference on Sensitivity Analysis*. Model Output. University of Reunion Island, Le Tampon.

Beiranvand, V., Hare, W., and Lucet, Y. (2017). Best practices for comparing optimization algorithms. *Optimization Engineering 18*, 815-848. DOI:10.1007/s11081-017-9366-1.

Benson H. Y., Shanno, D. F., Vanderbei, R. J. (2000). Interior-point methods for nonconvex nonlinear programming: Jamming and comparative numerical testing. *Operations Research and Financial Engineering*. Princeton University, Princeton.

Billups, S., Dirkse, S., and Ferris, M.C., (1995). A comparison of algorithms for large scale mixed complementarity problems. Technical Report. *Computer Science Department*, University of Wisconsin, Madison.

Bongartz, I., Conn, A. R., Gould, N. I. M., and Sauders, M. A. (1997). A numerical comparison between the LANCELOT and MINOS packages for large-scale nonlinear optimization. *Technical Report*. Namur University, Belgium.

Brevault, L., M., Balesdent, N., Bérend, and Le Riche, R. (2013). Comparison of different global sensitivity analysis methods for aerospace vehicle optimal design. *Proc. 10th World Congress on Structural and Multidisciplinary Optimization*, Orlando Florida, 1-12.

Cariboni, J., Gatelli, D., Liska, R., and Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. *Ecology Modeling 203*, 167-182. doi.org/10.1016/j.ecolmodel.2005.10.045

Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., and Acutis, M. (2010). Comparison of sensitivity analysis techniques: A case study with the rice model WARM. *Ecology Modeling 221*, 1897-1906. doi.org/10.1016/j.ecolmodel.2010.04.021

Cosenza, A., Mannina, G., Vanrolleghem, P. A., and Neumann, M. B. (2013). Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods. *Environmental Models and Software 49*, 40-52. DOI:10.1016/j.envsoft.2013.07.009.

Dolan, E. D., and Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematics Programming Series B 91*, 201-213. DOI:10.1007/s101070100263.

Duong, P. L. T., Ali, W., Kwok, E., and Lee, M. (2016). Uncertainty quantification and global sensitivity analysis of complex chemical process using a generalized polynomial chaos approach. *Computers and Chemical Engineering 90*, 23-30. DOI:10.1016/j.compchemeng.2016.03.020.

Galvez, E. D., and Capuz-Rizo, S. F. (2016). Assessment of global sensitivity analysis methods for project scheduling. *Computers & Industrial Engineering 93*, 110-120. DOI:10.1016/j.cie.2015.12.010.

Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C., and Di, Z. (2014). A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model. *Environmental Modelling & Software 51*, 269-285. DOI:10.1016/j.envsoft.2013.09.031.

Gonzalez-Cruz, M.B., Pire-Rivas, S.F., and Lopez-Jimenez, P.A. (2012). Sensitivity of conservative pollutants estimation using ISCST3 to input parameters for a case of study. *Revista Mexicana de Ingeniería Química 11*, 287-298.

Heiselberg, P., Brohus, H., Hesselholt, A., Rasmussen, H., Seinre, E., and Thomas, S. (2009). Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy 34*, 2030-2036. DOI:10.1016/j.renene.2009.02.016.

Herman, J. D., Kollat, J. B., Reed, P. M., and Wagener, T. (2013). Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. *Hydrology and Earth System Sciences 17*, 2893-2903. DOI:10.5194/hess-17-2893-2013.

Hernandez-Suarez, R., Zamora-Mata, J.M., Murrieta-Guevara, F., and Lugo-Leyte, R. (2008). Analysis and optimal retrofit of effluent lines entering a multicontaminant treatment unit. *Revista Mexicana de Ingeniería Química 7*, 151-162.

Herrera, S., Cisternas, L. A., and Gálvez, E. D. (2015). Simultaneous design of desalination plants and distribution water network. *Computer Aided Chemical Engineering 37*, 1193-1198. doi.org/10.1016/B978-0-444-63577-8.50044-9

Homma, T., and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety 52*, 1-17. DOI:10.1016/0951-8320(96)00002-6.

Hopfe, C. J., and Hensen, J. L. M. (2011). Uncertainty analysis in building performance simulation for design support.

Energy and Buildings 43, 2798-2805. DOI:10.1016/j.enbuild.2011.06.034.

Ikonen, T. (2016). Comparison of global sensitivity analysis methods - Application to fuel behavior modeling. *Nuclear Engineering and Design 297*, 72-. DOI:10.1016/j.nucengdes.2015.11.025.

Iooss B., Lemaître P. (2015) A Review on Global Sensitivity Analysis Methods. In Dellino G., Meloni C. (Eds) *Uncertainty Management in Simulation-Optimization of Complex Systems. Operations Research/Computer Science Interfaces Series 59*. Springer, Boston, MA. DOI: 10.1007/978-1-4899-7547-8_5

Ishigami, T., Homma, T. (1990). An importance quantification technique in uncertainty analysis for computer models. *Proc. First International Symposium on Uncertainty Modeling and Analysis*, 398-403.

Jamett, N., Cisternas, L. A., and Vielma, J. P. (2015). Solution strategies to the stochastic design of mineral flotation plants. *Chemical Engineering Science 134*, 850-860. DOI:10.1016/j.ces.2015.06.010.

Jansen, M. J. W. (1999). Analysis of variance designs for model output. *Computer Physics Communications 117*, 35-43. DOI:10.1016/S0010-4655(98)00154-4.

Kaghazchi, T., Mehri, M., Takht, M., and Kargari, A. (2010). A mathematical modeling of two industrial seawater desalination plants in the Persian Gulf region. *Desalination 252*, 135-142. DOI:10.1016/j.desal.2009.10.012.

Kucherenko, S., Feil, B., Shah, N., and Mauntz, W. (2011). The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety 96*. DOI:10.1016/j.ress.2010.11.003.

Lamoureux, B., Mechbal, N., and Massé, J. R. (2014). A combined sensitivity analysis and kriging surrogate modeling for early validation of health indicators. *Reliability Engineering & System Safety 130*, 12-26. DOI:10.1016/j.ress.2014.03.007.

Lilburne, L., and Tarantola, S. (2009). Sensitivity analysis of spatial models. *International Journal of Geographical Information Science 23*, 151-168. DOI:10.1080/13658810802094995.

Lorenz, E. N. (1963). Deterministic non-periodic flows. *Journal of the Atmospheric Sciences 20*, 130-141.

Lucay, F., Cisternas, L. A., and Gálvez, E. D. (2015a). A new group contribution method for mineral concentration processes. *Computers and Chemical Engineering 74*, 28-33. DOI:10.1016/j.compche meng.2014.12.009.

Lucay, F., Cisternas, L. A., and Gálvez, E. D. (2015b). Global sensitivity analysis for identifying critical process design decisions. *Chemical Engineering Research and Design 103*, 74-83. DOI:10.1016/j.cherd.2015.06.015.

Lucay, F., Cisternas, L. A., and Gálvez, E. D. (2015c). Retrofitting of concentration plants using global sensitivity analysis. *Computers Aided Chemical Engineering 37*, 311-316. doi.org/10.1016/B978-0-444-63578-5.50047-5

Lucay, F.A., Gálvez, E.D., Sales-Cruz, M., Cisternas, L. A. (2017). Improving milling operation using uncertainty and global sensitivity analyses. *Proc. 6th International Computational Modelling Symposium*. MEI, Falmouth.

Mahajan, A., Leyffer, S., and Kirches, C. (2012). *Solving mixed-integer nonlinear programs by QP- diving*. Argonne National Laboratory, Argonne.

Mellado, M. E., Cisternas, L. A., and Gálvez, E. D. (2009). An analytical model approach to heap leaching. *Hydrometallurgy 95*, 33-38. DOI:10.1016/j.hydromet.2008.04.009.

Nash, S. G., and Nocedal, J. (1991). A numerical study of the limited memory BFGS method and the truncated-Newton method for large scale optimization. *SIAM Journal of Optimization 1*, 358-372. DOI:10.1137/0801023.

Nguyen, A., and Reiter, S. (2015). A performance comparison of sensitivity analysis methods for building energy models. *Buildings Simulation 8*, 651-664. doi.org/10.1007/s12273-015-0245-4

Owen, A. B. (2013). Better estimation of small Sobol sensitivity indices. ACM Trans. *Modeling and Computer Simulation 23*, 1-17. DOI:10.1145/2457459.2457460.

Pappenberger, F., Ratto, M., Vandenberghe V. (2010) Review of sensitivity analysis methods. In P.A. Vanrolleghem (Ed.), *Modelling Aspects of Water Framework Directive Implementation*, IWA Publishing, 191-265.

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T. (2016) Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software 79*, 214-232. DOI: 10.1016/j.envsoft.2016.02.008

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications 181*, 259-270. DOI:10.1016/j.cpc.2009.09.018.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis*. The Primer, John Wiley & Sons, Ltd: Chichester, UK.

Saltelli, A., Tarantola, S., and Campolongo, F. (2000). Sensitivity analysis as an ingredient of modeling.*Statistical Science 15*, 377-395. DOI:10.2307/2676831.

Sepulveda, F. D., Cisternas, L. A., and Gálvez, E. D. (2013). Global sensitivity analysis of a mineral processing flowsheet. *Computer Aided Chemical Engineering 32*, 913-918. https://doi.org/10.1016/B978-0-444-63234-0.50153-6

Sobol, I. M. (1993) Sensitivity analysis for non-linear mathematical models. *Mathematics and Computers in Simulation 1*, 407-414.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation 55*, 271-280. DOI:10.1016/S0378-4754(00)00270-6.

Sobol, I. M. (2007). Global sensitivity analysis indices for the investigation of nonlinear mathematical models. *Mathematical Modeling 19*, 23-24. DOI:10.18287/0134-2452-2015-39-4-459-461.

Soto-Cruz, O., and Paez-Lerma, J. (2005) fermentation process balances: consistency and metabolic flux analysis. *Revista Mexicana de Ingeniería Química 4*, 59-74.

Van Duc Long, N., and Lee, M. (2012). Dividing wall column structure design using response surface methodology. *Computers & Chemical Engineering 37*, 119-124. doi.org/10.1016/j.compchemeng.2011.07.006

Vanderbei, R. J. (1999). LOQO user's manual - version 3.10. *Optimization Methods and Software 11*, 485-514. DOI:10.1080/10556789908805760.

Yianatos, J., Bucarey, R., Larenas, J., Henríquez, F., and Torres, L. (2005). Collection zone kinetic model for industrial flotation columns. *Minerals Engineering 18*, 1373-1377. DOI:10.1016/j.mineng.2005.01.014.

Yianatos, J. B., and Henríquez, F. D. (2006). Short-cut method for flotation rates modelling of industrial flotation banks.*Minerals Engineering 19*, 1336-1340. DOI:10.1016/j.mineng.2005.12.010.

Appendix A. Models

A.1. Heat equations

The problem P1 describes the heat transfer through conduction in a thin wire, here $u(x,t)$ is the temperature:

$$P1 = \begin{cases} \alpha \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}, & \alpha > 0 \\ u(0,t) = u(L,t) = 0, & t > 0 \\ u(x,0) = f(x) & x \in (0,L) \end{cases} \tag{A.1}$$

with

$$\alpha = \frac{K}{\rho c} \tag{A.2}$$

where L is the wire length, ρ is the density of the metal wire, c is the specific heat and K is the thermal conductivity. Assuming that $f(x)$ can be expressed as an infinite series of $\sin(x)$, this is

$$f(x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{L}\right) \tag{A.3}$$

a general solution of the heat equation is given by

$$u(x,t) = \sum_{n=1}^{\infty} a_n \exp\left(-\alpha \frac{n^2\pi^2}{L^2} t\right) \sin\left(\frac{n\pi}{L} x\right) \tag{A.4}$$

taking

$$f(x) = x(1-x) \quad and \quad L = 1 \tag{A.5}$$

the coefficients a_n have the following expression

$$a_n = \frac{4(1-(-1)^n)}{n^3\pi^3} \tag{A.6}$$

Then, a particular solution of the problem P1 is given by

$$u(x,t) = \sum_{n=1}^{\infty} \frac{4(1-(-1)^n)}{n^3\pi^3} \exp(-\alpha n^2\pi^2 t) \sin(n\pi x) \tag{A.7}$$

The model to which the GSA was applied had t, α and x as input variables, while the model output was u.

A.2 Ishigami function

The Ishigami function was defined by Ishigami and Homma (1990) as:

$$f(x_1, x_2, x_3) = \sin(x_1) + a \cdot \sin^2(x_2) + b \cdot x_3^4 \sin(x_1) \tag{A.8}$$

This equation shows the property of being strongly non-linear and non-monotonic. The model considered to apply the GSA had x_1, x_2 and x_3 as input variables, while the model output was f.

A.3 Flotation circuit

Flotation circuits are used to separate the valuable ore from the gangue ore. Lucay *et al*. (2015a) presented a group contribution method to estimate the recovery of a large number of flotation circuits. This method was validated with 30 circuits among which is the plant shown in Figure A.1.

Taking T^R, T^{C1}, T^{S1}, T^{S2}, T^{SC11} and T^{CS21} as the recovery of valuable ore in step rougher (R), cleaner (C_1), scavenger 1 (S_1), scavenger 2 (S_2), scavenger-cleaner (SC_{11}) and cleaner-scavenger (CS_{21}), respectively, it is possible to determine by mass balances the global recovery circuit shown in Figure A.1, with the following equation:

$$R = \frac{f}{g} \tag{A.9}$$

where

$$\begin{aligned} F = {} & T^C\left(T^R - T^R T^{S_1} - T^R T^{S_2} + T^{CS} T^{S_1} T^{SC}\right. \\ & + T^{CS} T^{S_2} T^{SC} + T^R T^{S_1} T^{S_2} + T^R T^{S_1} T^{SC} \\ & + T^R T^{S_2} T^{SC} + T^{CS} T^{S_1} T^{SC} - T^{CS} T^R T^{S_1} T^{SC} \\ & - T^{CS} T^R T^{S_2} T^{SC} - T^{CS} T^{S_1} T^{S_2} T^{SC} \\ & \left. - T^R T^{S_1} T^{S_2} T^{SC} + T^{CS} T^R T^{S_1} T^{S_2} T^{SC}\right) \end{aligned} \tag{A.10}$$

$$\begin{aligned} g = {} & \left(T^C T^{CS} - T^{CS} + 1\right)\left(T^{S_1} T^{S_2} - T^{S_2}\right. \\ & \left. - T^{S_1} + T^{S_1} T^{SC} + T^{S_2} T^{SC} - T^{S_1} T^{S_2} T^{SC} + 1\right) \end{aligned} \tag{A.11}$$

The model considered to apply the GSA had the recovery in each stage as input variables and the global recovery as model output.

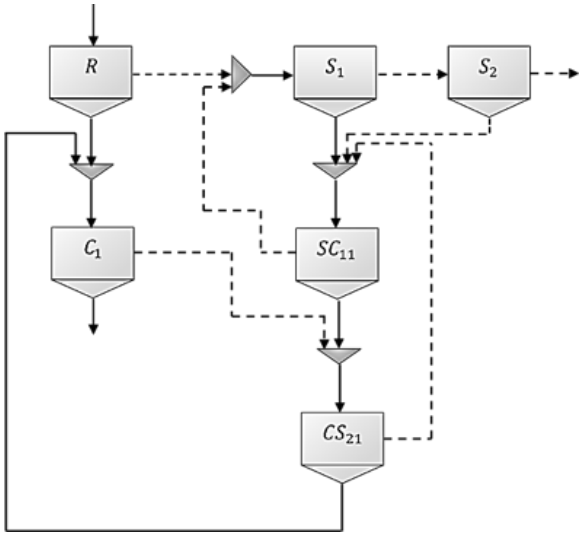

Fig. A.1. Flotation circuits used to validate the group contribution method.

A.4 El Salvador Flotation circuit

The flotation process is used to concentrate valuable minerals, in practice the number of variables in the process is very large, considering this, this time we will include a flotation model in each stage of the circuit. Through mass balances and the following equations (Yianatos *et al*., 2005; Yianatos and Henríquez, 2006):

$$R = R_{max}\left[1 - \frac{\left(1-(1+k_{max}\tau)^{1-N}\right)}{(N-1)k_{max}\tau}\right] \tag{A.12}$$

$$R_C = R_{max}\left[1 - \frac{4}{3k_{max}\tau}\left(\frac{12}{k_{max}\tau+12} - 1 + \frac{10}{9}\ln\left(\frac{10k_{max}\tau+12}{k_{max}\tau+12}\right)\right)\right] \tag{A.13}$$

$$R_f = 95\exp\left(-0.0144\frac{H_F(1+3J_w)}{J_g^3}\right) \tag{A.14}$$

$$R_G = \frac{R_C R_F}{1 - R_C + R_C R_F} \tag{A.15}$$

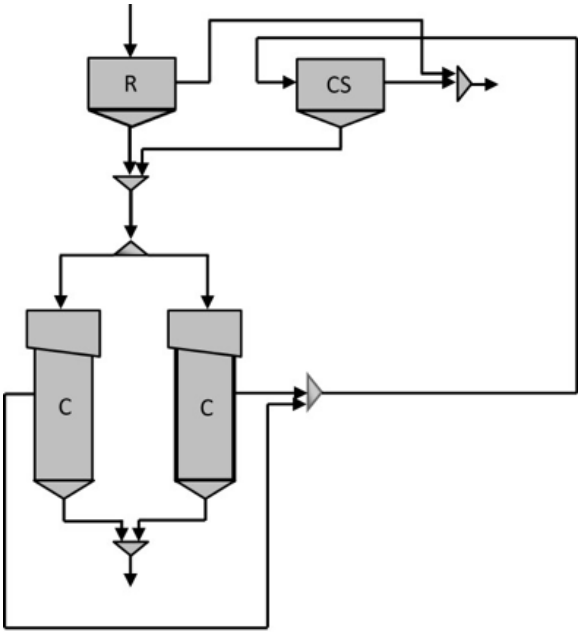

Fig. A.2. El Salvador flotation circuit.

It is possible to determine the global recovery of the circuit shown in Figure A.2, with the following equation

$$Rec(Cu) = \frac{\sum_{i=1}^{n} CF_i \cdot l(i)}{\sum_{i=1}^{n} CF_i} \tag{A.16}$$

where CF_i is the mass flow of the species i in the concentrate of the circuit, L_i is the grade of Cu in the species i present in the concentrate and F is the fresh feed flow to the circuit. The model considered to apply the GSA had 47 input variables and the model output was $Rec(Cu)$, for further details see Lucay *et al*. (2015c).

A.5 Legendre polynomial

The Legendre differential equation can be solved using the power series method, in the particular case d (polynomial order) is a non-negative integer. The solution forms a family of orthogonal polynomials, called Legendre polynomial, which has the following expression:

$$L_d(x) = \frac{1}{2^d}\sum_{m=0}^{d/2}(-1)^m\left[\frac{d!}{m!(d-m)!}\frac{(2d-2m)!}{d!(d-2m)!}x^{d-2m}\right] \tag{A.17}$$

where $x \in [-1,1]$, $d = \{1,2,3,4,5\}$ and $\lfloor v \rfloor$ indicates the greatest integer $\leq v$. The model considered to apply the GSA has x and d as input variables, and the model output was $L_d(x)$. For further details see Saltelli *et al*. (2000).

A.6 SAG Milling

Semi-Autogenous Grinding (SAG) is a process that reduces the size of minerals. Usually, for modeling, the mill is divided into two areas according to the process carried out in each of them (see Figure A.3): size reduction and classification. In general, it is assumed that the SAG mill behaves like a well-mixed reactor, with a mass W retained in the mill volume V, and first order kinetics (Austin, 1990).

It is possible to determine the retained mass W with the following system of equations (Lucay *et al*., 2017):

$$0 = f_i + \tau\sum_{\substack{j=1 \\ i>j}}^{i-1} b_{ij}K_j w_j - \left(K_i\tau + (1-c_i)(1+C^*)\right)w_i \qquad i = 1,2,\ldots,n \tag{A.18}$$

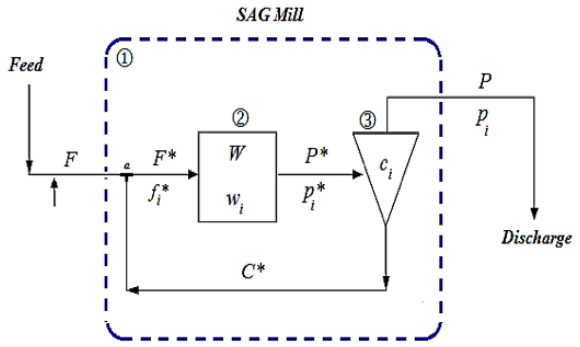

Fig. A.3. Conceptual representation for internal SAG mill operation.

$$0 = W - W\sum_{i=1}^{n} w_i \tag{A.19}$$

$$0 = \tau - \frac{W}{F} \tag{A.20}$$

where F is the feed to mill, F_i is the mass fraction of feed to mill, τ is the residence time, and K_i is the specific breakage rate of species i present in the feed to the SAG Mill, b_{ij} is the breakage distribution function, w_i is the mass fraction of W, $C*$ is the ratio of internal recirculation and ci is the classification efficiency of the internal grid mill.

Once W is determined, it is possible to calculate the specific energy consumption of the mill:

$$E = \sum_i \frac{M_p}{K_i W} \tag{A.21}$$

with

$$M_p = GD^{2.5}L(1-AJ)\left((1-\varepsilon_b)J\rho_s(1-w_c) + 0.6J_b(\rho-\rho_s)(1+w_c)\right)\phi_c\left(1-\frac{0.1}{2^{9-10\phi_c}}\right) \tag{A.22}$$

Where M_p is the mill power consumption. The model considered to apply the GSA had as input variables: the feed to the mill, the fraction of the species present in the feed, the mill volume occupied by the steel balls, the percentage of solids in the discharge mill, the operating speed equal to % of the critical speed. While the model output was the specific energy consumption of the mill, E.

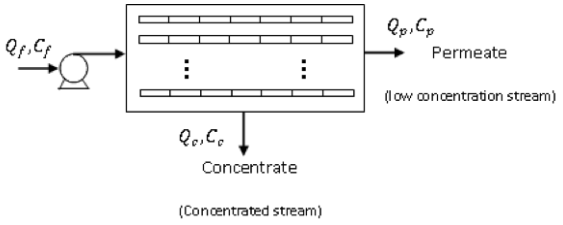

Fig. A.4. Stage of desalinization with RO membrane.

A.7 Reverse osmosis stage

Reverse osmosis (RO) is a water purification technology that uses membranes to remove particles, ions, and molecules. This process is performed in circuits or stages as shown in Figure A.4.

This procedure is modeled by the following equations

$$Q_p = J_w A \tag{A.23}$$

with

$$J_w = A_w\left(P_f - \Delta P_L - P_p - \Delta\pi\right) \tag{A.24}$$

$$\Delta\pi = 0.9524(C_b - C_p)\exp\left(\frac{J_w}{k}\right)\left((C_b - C_p)\exp\left(\frac{J_w}{k}\right) + 2C_p + 81633\right) \tag{A.25}$$

$$\Delta P_L = \frac{\rho u^2 L C_{td}}{2d_h} \tag{A.26}$$

where Q_p is the permeate flow, J_w is the water flow, P_f is the operating pressure of the process, $\Delta\pi$ is the difference of the osmotic pressure across the membrane, ΔP_L is the pressure drop across the feed channel of the membrane, P_P is the permeate pressure, A_w is the coefficient permeability of water, L is the length of one element of a membrane, k is the mass transport coefficient, C_b is the bulk concentrate, u is the axial velocity, A is the membrane area, ρ is the water density, C_{td} is the total drag force, and d_h is the hydraulic diameter of the feed spacer channel. While the concentration of salt in the permeate is determined by

$$C_p = C_b\frac{B_s\exp\left(\frac{J_w}{k}\right)}{\left(J_w + B_s\exp\left(\frac{J_w}{k}\right)\right)} \tag{A.27}$$

with

$$C_b(x) = C_f + \left(\frac{C_c - C_f}{L}\right)x \tag{A.28}$$

where L is the membrane length and B_s is the permeability coefficient of salt. For further information see Kaghazchi *et al*. (2010). The model considered to apply the GSA had as input variables: the operating pressure (P_f), the feed flow (Q_f), the salt concentration in the feed flow (C_f), the number of membranes (n), the number of pressurized tank (m) and the temperature (T). While the concentration in the permeate was selected as the model output.

A.8 Reverse osmosis membrane

The previous equations for the RO stage are also valid for the case of a single RO membrane, i.e., tanks in parallel or membranes in series are not considered. The model to which the GSA was applied had the operating pressure, the feed flow, the salt concentration in the feed flow and the temperature as input variables, while the salt concentration in the permeate generated by the membrane was the model output. The fact of not considering tanks or membranes in series is an important factor because the model output for the RO membrane has an additive behavior, while the model output of the RO stage is not additive.

A.9 Lorenz attractor

The Lorenz attractor is a three-dimensional deterministic nonlinear dynamic system (Lorenz, 1963). The system of equations is given by:

$$\frac{dx}{dt} = a(y - x) \tag{A.29}$$

$$\frac{dy}{dt} = x(b - z) - y \tag{A.30}$$

$$\frac{dz}{dt} = xy - cz \tag{A.31}$$

The model to which the GSA was applied had a, b and c as input variables. Once the system was determined, this was solved by Runge-Kutta 4^{th} order and the model output was the variable x.

A.10 Lotka-Volterra equation

This model is used to describe the dynamic behavior of two interacting species, one as preys and the other as predator (Cariboni *et al*., 2007). The system of differential equations is given by:

$$\frac{dX}{dy} = aX\left(1 - \frac{X}{K}\right) - bXY \tag{A.32}$$

$$\frac{dY}{dt} = cY - dXY \tag{A.33}$$

where X is the number of prey, Y is the number of predators, the constant a is the birth rate of prey, b is the rate of prey elimination by predators, c is the birth rate of predators, d is the predator increase rate as a result of the prey consumption, K is the maximum number of prey that the environment can endure. The model considered to apply the GSA had a, b, c, d and K as inputs, and the number of prey as model output.

A.11 Linear models

A classic model in many publications is the linear model. In this work, the following expression is considered:

$$f(X_1, X_2, X_3, X_4, X_5, X_6) = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 \tag{A.34}$$

The model considered to apply the GSA had X_i as input variables, while the model output was f.

A.12 Matrix system

Many engineering problems can be transformed into matrix systems:

$$AX = b \tag{A.35}$$

where A is a matrix of nxn, X is a vector of unknowns of dimension $nx1$, and b is a vector of dimension $nx1$. In our case, we have

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 2 & -1 & 2 & 1 \\ 3 & 0 & 1 & 1 \\ 2 & 2 & 2 & -1 \end{bmatrix}$$

$$and \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{A.36}$$

The model input variable is the vector b, and the model output is $\|X\|_2$.

A.13 Transport model

Consider that we want to determine the transportation schedule that minimizes costs and meets the market demands and materials factories shown in Table A.1. The total costs are defined as

$$total\ \ cost = \sum_i \sum_j c(i,j) \cdot X(i,j) \tag{A.37}$$

with

$$c(i,j) = F \cdot D(i,j) \tag{A.38}$$

where $c(i,j)$ are the transport cost, $X(i,j)$ are the transport quantities, F is the cost freight and $D(i,j)$ are the distances between plants and markets. The model considered to apply the GSA had the markets supply and the plants demand as input variables, while the total costs as model outputs.

Table A.1. Transport model parameters.

	Distances			
	Markets			
Plants	New York	Chicago	Topeka	Supply
Seattle	2.5	1.7	1.8	350
San Diego	2.5	1.8	1.4	600
Demand	325	300	275	

A.14 Pipes network design

Herrera *et al*. (2015) proposed a methodology for the design of a water distribution network, among the model equations are:

$$TC_{i,j} = \frac{\kappa_p \left(1 + \frac{H_{i,j}}{H_b}\right) L_{i,j} D_{i,j}^m}{PL} \tag{A.39}$$

$$TC_n = \left(\frac{\kappa_N (1+s_b)\rho g}{\eta PL} + \frac{8.76 F_D F_A E_C \rho g}{\eta}\right) \sum_{j \in output} Q_{n,j} H_{n,j} \tag{A.40}$$

$$TC_{SO} = PA \cdot UPC_{SO,C} Q_{SO} \tag{A.41}$$

where Eq. (A.39) is used to determine the cost of the water transport network, Eq. (A.40) is used to determine the operational cost of pumping stations, and Eq. (A.41) is used to determine the cost associated with the RO plants. Moreover, the head $H_{i,j}$ was determined by the following equation:

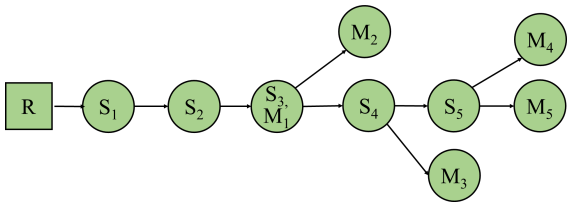

Fig. A.5. Optimum pipe network design.

$$H_{i,j} = \Delta Z_{i,j} + H_0 + \frac{8 f_i L_{i,j} Q_{i,j}^2}{\pi^2 g D_{i,j}^5} \tag{A.42}$$

While the function objective was to minimize the total annualized cost

$$TC = \sum_{so} TC_{so} + \sum_{i,j} TC_{i,j} + \sum_n TC_n \tag{A.43}$$

The methodology considers a network consisting of 5 mining operations, 3 RO plants and 14 pumping stations. Then, the best design was determined using the GAMS software. The result was a network that uses 5 pumping stations and 1 RO plant, the design is shown in Figure A.5.

The model considered applying the GSA had as input variables: the flows of desalinated water required by each mining operation and the diameters of the pipes of the water distribution network. The model output was the total cost.

A.15 Polynomial division

This model is given by:

$$P(x_1,x_2,x_3,x_4,x_5) = \frac{x_1 + x_2 + x_3 + x_4 x_5}{x_1 x_2 + x_3 x_4 x_5} \tag{A.44}$$

here the x_i is the input variable and P is the model output.

A.16 Trigonometric function

This model has the following expression:

$$F(x_1,x_2) = \sin(\pi x_1) + \cos\left(\frac{\pi x_2}{4}\right) + (x_1 x_2)^{0.5} \tag{A.45}$$

The model considered to apply the GSA had x_1 and x_2 as input variables, while the model output was F.

A.17 Bateman equation

This equation describes a simple chemical or radioactive chain, where the growth rate of each element is proportional to the initial concentration, and the decay rate of each element λ_i is proportional to the concentration C_i of the element itself. Bateman equation is given by the following system of differential equations:

$$\frac{dC_1}{dt} = -\lambda_1 C_1, \qquad C_1(0) = C_1^0 \tag{A.46}$$

$$\frac{dC_2}{dt} = -\lambda_2 C_2 + \lambda_1 C_1, \qquad C_2(0) = C_2^0 \tag{A.47}$$

$$\frac{dC_3}{dt} = -\lambda_3 C_3 + \lambda_2 C_2, \qquad C_3(0) = C_3^0 \tag{A.48}$$

$$\frac{dC_1}{dt} = -\lambda_1 C_1, \qquad C_1(0) = C_1^0 \tag{A.49}$$

$$\frac{dC_4}{dt} = -\lambda_4 C_4 + \lambda_3 C_3, \qquad C_4(0) = C_4^0 \tag{A.50}$$

The analytical solution of the system of differential equations is

$$C_i(t) = \sum_{m=1}^{i} C_m^0 \left[\prod_{\substack{r=m \\ m\neq i}}^{i-1} \lambda_r \sum_{n=m}^{i} \frac{\exp(-\lambda_n t)}{\prod_{\substack{l=m \\ m\neq i \\ l\neq n}}^{i} \lambda_l - \lambda_n} \right] \quad i = 1,2,3,4. \tag{A.51}$$

The model considered to implement the GSA had as λ_1, λ_2, λ_3, λ_4, C_1^0, C_2^0, C_3^0, C_4^0 as input variables. While the model output was . For further details see Saltelli *et al*. (2000).

A.18 Heap leaching

Leaching is a hydrometallurgical process that has been modeled by many authors, among them Mellado *et al*. (2009) proposed analytical models to describe the heap leaching. The model is given by the following equation

$$R(t) = R_\infty \left[1 - \alpha \exp\left(\frac{-k_\theta u_s}{Z\varepsilon_b} \left(t - \frac{\varepsilon_b Z w}{u_s} \right) \right) - (1-\alpha) \exp\left(\frac{-k_\tau D_{Ae}}{r^2 \varepsilon_0} \left(t - \frac{\varepsilon_b Z W}{u_s} \right) \right) \right] \tag{A.52}$$

where Z is the heap death, u_s is the superficial bulk flow velocity, D_{Ae} is the effective pore diffusivity of the reagent, r is the particle size, ε_b is the volume fraction of the bulk solution, R_∞ is the recovery in an infinite time, ε_0 is the porosity of the mineral, k_θ is the kinetic constant at the particle level, k_τ is the kinetic constant at the heap level and w is the delay time. While that parameters α, k_θ, k_τ and w must be adjusted with experimental data.

The model considered to apply the GSA had all the variables mentioned above, except the adjustment parameters, as input variables. While the model output was the full recovery $R(t)$.

A.19 Sobol function

This function is defined by Archer *et al*. (1997) as

$$G(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) = \prod_{i=1}^{8} \frac{|4X_i - 2| + a_i}{1 + a_i} \tag{A.53}$$

where $a = (0, 1, 4.5, 9, 99, 99, 99, 99)$. The input variables are $X_i \in U[0,1]$, and the model output is G.

A.20 K function

This function is defined by Kucherenko *et al*. (2011) as

$$K(X_1, X_2, X_3, X_4, X_5) = \sum_{i=1}^{5} (-1)^i \prod_{j=1}^{i} X_i \tag{A.54}$$

Here the input variables are $X_i \in U[0,1]$, and the model output is K.

A.21 B function

This function is defined by Saltelli *et al*. (2000) as

$$B(X_1, X_2, X_3, X_4, X_5, w_1, w_2, w_3, w_4, w_5) = \sum_{i=1}^{5} X_i w_i \tag{A.55}$$

Here the input variables are $X_i \in N(0, \sigma_{X_i}) and w_i \in N(0, \sigma_{w_i})$, and the model output is B.