Environmental Modelling & Software 26 (2011) 1515–1525

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

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

Sobol' sensitivity analysis of a complex environmental model

Jiri Nossent*, Pieter Elsen, Willy Bauwens

Department of Hydrology and Hydraulic Engineering, Earth System Sciences Group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

A R T I C L E I N F O

Article history:
Received 7 May 2011
Received in revised form
14 August 2011
Accepted 17 August 2011
Available online 16 September 2011

Keywords:
Sensitivity analysis
Process identification
Sobol' method
Environmental modeling
Model parameters
Interaction effects

A B S T R A C T

Complex environmental models are controlled by a high number of parameters. Accurately estimating the values of all these parameters is almost impossible. Sensitivity analysis (SA) results enable the selection of the parameters to include in a calibration procedure, but can also assist in the identification of the model processes. Additionally, a sensitivity analysis can yield crucial information on the use and meaning of the model parameters.

This paper presents a Sobol' sensitivity analysis for flow simulations by a SWAT model of the river Kleine Nete, with the objective to assess the first order, second order and total sensitivity effects. Confidence intervals for the resulting sensitivity indices are inferred by applying bootstrapping. The results indicate that the curve number value (CN2) is the most important parameter of the model and that no more than 9 parameters (out of 26) are needed to have an adequate representation of the model variability. The convergence of the parameter ranking for total sensitivity effects is relatively fast, which is promising for factor fixing purposes. It is also shown that the Sobol' sensitivity analysis enhances the understanding of the model, by e.g. pointing out 3 significant pairwise interactions.

In general, it can be concluded that the Sobol' sensitivity analysis can be successfully applied for factor fixing and factor prioritization with respect to the input parameters of a SWAT model, even with a limited number of model evaluations. The analysis also supports the identification of model processes, parameter values and parameter interaction effects.

1. Introduction

Developments over the past decade(s) in environmental modeling have led to an increasing complexity of the available tools. Despite the high computational demands, these models have become very popular in many applications, such as land use and climate change scenario analysis, flood prediction, water quality modeling, etc.

The high number of parameters in these complex models however constitute a major problem in their application, as the parameter estimation becomes a high-dimensional, multi-modal and mostly non-linear problem. To resolve this problem, a wide range of optimization algorithms have been developed (Beven and Binley, 1992; Doherty, 2003; Duan et al., 1992; Vrugt et al., 2003). Often, however, it is not feasible, nor necessary to include all model parameters in the calibration process to obtain an efficient optimization. A sensitivity analysis (SA) can determine the most influential parameters of a model, allowing a reduction of the number of parameters incorporated in the optimization (Saltelli, 2000). This can be achieved by factor fixing (FF), where non-influential parameters are set to a fixed value, or by factor prioritization (FP), where the modeler focuses on the parameters that have the potential to maximally reduce the output uncertainty (Saltelli et al., 2004; Sobol', 1990). Besides, an SA also provides information e.g. on the use of the parameters, on the influence of specific parameter values and the associated model outcome and on the model processes themselves, facilitating the understanding and interpretation of models (Saltelli, 2000). In recent years, this broad range of possibilities made sensitivity analysis a main topic in environmental modeling literature and, particularly, in this journal (Estrada and Diaz, 2010; Ravalico et al., 2010; Saltelli and Annoni, 2010; Yang, 2011).

A particular example of a complex, over-parameterized environmental model is the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1993). In this simulator, a large number of parameters rule the calculations for different processes like surface runoff and groundwater flow, but also pesticide and nutrient conversion and transport processes. To determine the most influential parameters of a SWAT model, the Latin-hypercube – One-factor-at-a-time (LH-OAT) algorithm is often applied, as this method is incorporated in the simulation tool (van Griensven et al., 2006). The LH-OAT method is a so called screening method (Campolongo et al.,

* Corresponding author. Tel.: +32 2 649 58 77; fax: +32 2 629 30 22.
E-mail addresses: jnossent@vub.ac.be (J. Nossent), Pieter.Elsen@vub.ac.be (P. Elsen), wbauwens@vub.ac.be (W. Bauwens).

1364-8152/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2011.08.010

2000), that provides a rough estimation of the parameter ranking according to their influence on the model output, with only a limited number of model evaluations (<1000). However, the results of such a screening method are qualitative and not quantitative. Where an SA normally can provide extra information on the model processes or makes it possible to verify underlying model assumptions, these possibilities are limited for screening methods and, in particular, for the LH-OAT results (Campolongo et al., 2000; Morris, 1991). For example, appointing the influence of parameter interactions on the model output is not possible, nor is it possible to quantify the ratio between the influences of the different parameters to account for the relative importance of a parameter as compared to the other parameters. The latter hinders a rational selection of the parameters that should be included in the optimization process. In general, a screening method is only suited for factor fixing and not for factor prioritization (Saltelli et al., 2008).

A sensitivity analysis method that is very popular in many fields, is the variance-based Sobol' method (Sobol', 1990). In general, variance-based sensitivity analysis methods aim to quantify the amount of variance that each parameter contributes to the unconditional variance of the model output. For the Sobol' sensitivity analysis method, these amounts, caused either by a single parameter or by the interaction of two or more parameters, are expressed as (Sobol') sensitivity indices (SI's). These indices represent fractions of the unconditional model output variance and can be employed for both FF and FP. Despite its high computational demands, this powerful SA technique has recently become more popular in environmental modeling (Pappenberger et al., 2008; Tang et al., 2007; van Werkhoven et al., 2008; Yang, 2011), because of its ability to incorporate parameter interactions and the relatively straightforward interpretation. Recently the Sobol' sensitivity analysis technique has also been applied for SWAT (Cibin et al., 2009).

Due to its advantageous properties and the limitations of the qualitative results of the LH-OAT method, a first objective of this research is to test the Sobol' method as an alternative method to perform a sensitivity analysis on a SWAT model. The first order and total sensitivity indices for 26 water quantity and water quality parameters will be calculated, considering the flow as the model output. Also, the evolution of the SI's with increasing Monte Carlo sample size will be investigated, as a measure for convergence. Bootstrapping (Efron and Tibshirani, 1993) will be applied to assess 95% confidence intervals (CI's) for the indices and the outcome of the importance analysis (i.e. the quantitative parameter ranking) will be evaluated with an adapted version of the SA repeatability test (Andres, 1997; Tang et al., 2007).

Besides the determination of highly influential or non-influential parameters, the Sobol' sensitivity analysis results can also provide background information on the parameters and the model processes. To enhance the extraction of such information, another objective is to calculate the second order sensitivity indices for the most important parameters, in order to estimate the most significant interactions between pairs of parameters (Saltelli et al., 2006). These interactions can e.g. highlight processes that are highly related.

2. Materials and methods

2.1. The study area

A SWAT model with a daily time step of the Kleine Nete catchment (Belgium) was used as a case study for this research. This relative small (580 km^2) subcatchment of the river Scheldt basin is located in the North-Eastern part of Belgium. It is a flat basin, dominated by sandy soils (>95%) and agricultural land use (57%). In this model, 14 subbasins are defined, taking into account the locations of gauging stations and discharge points of pollutants to the river. This leads to short river reaches (<25 km) and very small subbasins.

Belgium has a temperate maritime climate, with significant precipitation during the whole year, ranging from 700 mm/y to 1085 mm/y for the period under study (1997–2007). Records of daily precipitation, temperature, wind speed, solar radiation and relative humidity are uniformly applied over the catchment and were obtained as the mean of 5 different measuring stations. Stream discharge data at the outlet of the catchment are available on a daily basis for the period 1997–2007.

2.2. The soil and water assessment tool (SWAT)

SWAT (Arnold et al., 1993) is a semi-distributed, physically based, environmental model that is able to calculate water quantity and water quality variables on a daily or sub-daily time step. For these calculations, the basin is divided into subbasins and further into HRU's (Hydrological Response Units). The hydrological component of SWAT takes the following processes into account: interception, surface runoff, percolation, lateral subsurface flow, groundwater return flow, evapotranspiration and channel transmission losses. The behavior of nutrients and pesticides can also be simulated by the model. More details on the Soil and Water Assessment Tool can be found in (Arnold et al., 1993; Gassman et al., 2007).

2.3. The parameters considered for the sensitivity analysis

Based on a previous study on the sensitivity of the parameters of the SWAT model for the Kleine Nete (Nossent and Bauwens, in press), 26 parameters are selected for the Sobol' sensitivity analysis of the model for flow simulations. The set includes 21 water quantity parameters and 5 parameters dealing with water quality processes (Table 1).

As no prior information is available on the parameters, the input parameter values for the SA are sampled from a uniform distribution. The different parameter ranges are scaled between 0 and 1 with a linear transformation.

The sensitivity of the parameters will be assessed by varying the parameter values for the whole basin, resulting in one value of the SI's per parameter. Because the land use and soil type often have a big influence on the curve number value (CN2) and this parameter is traditionally one of the most important parameters in a SWAT model (Cibin et al., 2009; Green and van Griensven, 2008; Holvoet et al., 2005; van Griensven et al., 2006; White and Chaubey, 2005), the first order and total sensitivity indices for this factor will additionally be calculated at the level of the HRU's. In this way, fractions of the sensitivity indices for CN2 can be appointed to the different HRU types, enhancing the understanding of the model.

2.4. The Sobol' sensitivity analysis

The method of Sobol' (Sobol', 1990) is a global and model independent sensitivity analysis method that is based on variance decomposition. It can handle non-linear and non-monotonic functions and models.

Be the model represented by a function

$$Y = f(\boldsymbol{X}) = f(X_1, \ldots, X_p) \tag{1}$$

where Y is the model output (or objective function) and $\boldsymbol{X} = (X_1, \ldots, X_p)$ is the parameter set. Sobol' suggested to decompose the function f into summands of increasing dimensionality:

$$f(X_1, \ldots, X_p) = f_0 + \sum_{i=1}^{p} f_i(X_i) + \sum_{i=1}^{p} \sum_{j=i+1}^{p} f_{ij}(X_i, X_j) + \cdots + f_{1,\ldots,p}(X_1, \ldots, X_p) \tag{2}$$

If the input factors are independent and each term in this equation is chosen with zero average and is square integrable, then f_0 is a constant, equal to the expectation value of the output and the summands are mutually orthogonal. Additionally, this decomposition is unique.

The total unconditional variance can be defined as

$$V(Y) = \int_{\Omega^p} f^2(\boldsymbol{X}) \mathrm{d}\boldsymbol{X} - f_0^2 \tag{3}$$

with Ω^p representing the p-dimensional unit hyperspace (i.e. parameter ranges are scaled between 0 and 1). The partial variances, which are the components of the total variance decomposition, are computed from each of the terms in Equation (2) as

$$V_{i_1 \ldots i_s} = \int_0^1 \cdots \int_0^1 f_{i_1 \ldots i_s}^2 (X_{i_1}, \cdots, X_{i_s}) \mathrm{d}X_{i_1} \cdots \mathrm{d}X_{i_s} \tag{4}$$

where $1 \le i_1 \le \cdots \le i_s \le p$ and $s = 1,\ldots,$p. With the assumption that the parameters are mutually orthogonal, this results in Equation (5) for the variance decomposition.

$$V(Y) = \sum_{i=1}^{p} V_i + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} V_{ij} + \cdots + V_{1,\ldots,p} \tag{5}$$

Table 1
Parameters studied for the Sobol' sensitivity analysis.

Parameter	Definition	Process	Level
	Flow parameters		
ALPHA_BF	Baseflow recession factor (days)	Groundwater	HRU
BIOMIX	Biological mixing efficiency	Nitrogen cycle Phosphorus cycle	HRU
CANMX	Maximum canopy index	Evapotranspiration	HRU
CH_K2	Effective hydraulic conductivity in main channel alluvium (mm/hr)	Routing	Subbasin
CH_N	Manning coefficient for channel	Routing	Subbasin
CN2	SCS runoff curve number for moisture condition II	Surface runoff	HRU
EPCO	Plant evaporation compensation factor	Evapotranspiration	HRU
ESCO	Soil evaporation compensation factor	Evapotranspiration	HRU
GW_DELAY	Groundwater delay (days)	Groundwater	HRU
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	Groundwater	HRU
GW_REVAP	Groundwater 'revap' coefficient	Groundwater	HRU
RCHRG_DP	Groundwater recharge to deep aquifer (fraction)	Groundwater	HRU
REVAPMN	Threshold depth of water in the shallow aquifer for 'revap' (mm)	Groundwater	HRU
SLOPE	Average slope steepness (m/m)	Lateral flow sediment erosion	HRU
SLSUBBSN	Average slope length (m)	Concentration time sediment erosion	HRU
SMTMP	Snow melt base temperature (°C)	Snow	Subbasin
SOL_ALB	Soil albedo	Evapotranspiration	HRU
SOL_AWC	Available water capacity of the soil layer (mm/mm soil)	Soil water	HRU
SOL_K	Soil conductivity (mm/h)	Soil water	HRU
SURLAG	Surface runoff lag coefficient	Surface runoff	Subbasin
TIMP	Snow pack temperature lag factor	Snow	Subbasin
	Water quality parameters		
NPERCO	Nitrogen percolation coefficient	Nitrogen cycle	Subbasin
SOL_LABP	Initial labile (soluble) P concentration in surface soil layer (kg/ha)	Phosphorus cycle	HRU
SOL_NO3	Initial NO3 concentration (mg/kg) in the soil layer	Nitrogen cycle	HRU
SOL_ORGN	Initial organic N concentration in surface soil layer (kg/ha)	Nitrogen cycle	HRU
SOL_ORGP	Initial organic P concentration in surface soil layer (kg/ha)	Phosphorus cycle	HRU

In this way, the variance contributions to the total output variance of individual parameters and parameter interactions can be determined. These contributions are characterized by the ratio of the partial variance to the total variance, the Sobol' sensitivity indices:

$$\text{First order SI} \quad S_i = \frac{V_i}{V} \tag{6}$$

$$\text{Second order SI} \quad S_{ij} = \frac{V_{ij}}{V} \tag{7}$$

$$\text{Total SI} \quad S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + \cdots \tag{8}$$

The first order index, S_i, is a measure for the variance contribution of the individual parameter X_i to the total model variance. The partial variance V_i in (6) is given by the variance of the conditional expectation $V_i = V[E(Y|X_i)]$ and is also called the 'main effect' of X_i on Y. It can be described as the fraction of the model output variance that would disappear on average when X_i would be fixed to a value in its range (because $V(Y) = E[V(Y|X_i)] + V[E(Y|X_i)]$). The impact on the model output variance of the interaction between parameters X_i and X_j is given by S_{ij} and S_{Ti} is the result of the main effect of X_i and all its interactions with the other parameters (up to the p^{th} order).

The calculation of S_{Ti} can be based on the variance $V_{\sim i}$ that results from the variation of all parameters, except X_i (Homma and Saltelli, 1996).

$$S_{Ti} = 1 - \frac{V_{\sim i}}{V} \tag{9}$$

For additive models and under the assumption of orthogonal input factors, S_{Ti} and S_i are equal and the sum of all S_i (and thus all S_{Ti}) is 1. For non-additive models interactions exist: S_{Ti} is greater than S_i and the sum of all S_i is less than 1. On the other hand, the sum of all S_{Ti} is greater than 1. By analyzing the difference between S_{Ti} and S_i, one can determine the impact of the interactions between parameter X_i and the other parameters.

To compute the variances in order to obtain the sensitivity measures, Sobol' proposed a shortcut in the calculations, based on the assumption of mutually orthogonal summands in the decomposition. The shortcut is attained by transforming the double-loop integral of Equation (4) into an integral of the product of $f(X_{j_1}, \ldots, X_{j_{k-s}}, X_{i_1}, \ldots, X_{i_s})$ and $f(X'_{j_1}, \ldots, X'_{j_{k-s}}, X_{i_1}, \ldots, X_{i_s})$. The practical calculation will be discussed in Section 2.4.1.

2.4.1. Monte Carlo integrals and Sobol' quasi-random sampling

Because environmental models are mostly complex and non-linear, it is almost impossible to calculate the variances using analytical integrals. Hence, Monte Carlo integrals are applied in this research. The estimation of e.g. the total unconditional variance defined in Equation (3) therefore becomes:

$$\hat{V}(Y) = \frac{1}{(n-1)} \sum_{m=1}^{n} f^2(X_m) - \hat{f}_0^2 \tag{10}$$

where $\hat{\cdot}$ stands for the estimate, X_m is a sampled set of input parameters and n is the number of samples. Obviously, the estimation of the integral becomes more accurate for higher values of n.

Saltelli et al. (2010) suggest to use two independent input sample $n \times p$ -matrices (the "sample" matrix M_1 and the "resample" matrix M_2) to compute the Monte Carlo integrals, where n is the sample size and p is the number of parameters. Every row in M_1 and M_2 represents a possible parameter combination for the model. The parameter ranges in the matrices are scaled between 0 and 1. The two matrices are necessary to compile the sample sets for $f(X_{j_1}, \ldots, X_{j_{k-s}}, X_{i_1}, \ldots, X_{i_s})$ and $f(X'_{j_1}, \ldots, X'_{j_{k-s}}, X_{i_1}, \ldots, X_{i_s})$.

Basically, the calculation of the first order and total Sobol' sensitivity indices requires $n \cdot (2p+1)$ model evaluations. Saltelli (2002) introduced a method to almost halve the number of model runs required to $n \cdot (p+2)$. Additionally, he also developed a method to estimate S_i, S_{Ti} and S_{ij} using only $n \cdot (2p+2)$ model evaluations. In this study, we will apply the first simplified method (using $n \cdot (2p+2)$ model runs to calculate S_i and S_{Ti}) and afterwards add $n \cdot p'$ model evaluations to calculate S_{ij} for the p' most significant parameters.

Homma and Saltelli (1996) suggested to use Equation (11) to estimate the square of the expectation value for the calculation of the first order index.

$$\hat{f}_0^2 = \frac{1}{n} \sum_{m=1}^{n} f\left(X_m^{M_1}\right) * f\left(X_m^{M_2}\right) \tag{11}$$

where $X_m^{M_1}$ and $X_m^{M_2}$ are samples of M_1 and M_2. In this research, this equation is also successfully applied for the calculation of the other indices (S_{ij} and S_{Ti}).

The estimation of the unconditional total model output variance $\hat{V}$ was given by Equation (10). As matrices M_1 and M_2 are independently sampled, model evaluations based on all parameter sets of both matrices (instead of only M_1 or M_2) can be used to improve the accuracy and stability of the Monte Carlo integral:

$$\hat{V}(Y) = \frac{1}{2(n-1)} \sum_{m=1}^{2n} f^2\left(X_m^{M_1,M_2}\right) - \hat{f}_0^2 \tag{12}$$

where "M_1, M_2" is the $2n \times p$ -matrix obtained by combining M_1 and M_2 and $X_m^{M_1,M_2}$ is a sample of this matrix.

The partial variance estimate $\hat{V}_i$, which results from the main effect of parameter X_i, is given by:

$$\hat{V}_i = \frac{1}{(n-1)} \sum_{m=1}^{n} f\left(X_{\sim im}^{M_1}, X_{im}^{M_1}\right) * f\left(X_{\sim im}^{M_2}, X_{im}^{M_1}\right) - \hat{f}_0^2 \tag{13}$$

where $X_{\sim im}^{M_1}$ and $X_{\sim im}^{M_2}$ are the full sets of samples of M_1, resp. M_2 for all parameters except X_i. Oppositely, $X_{im}^{M_1}$ is the ith column of matrix M_1. In this way, $X_{\sim im}^{M_1}, X_{im}^{M_1}$, represents M_1 and $X_{\sim im}^{M_2}, X_{im}^{M_1}$ represents M_2, with the ith column replaced by the ith column of M_1. Finally, the estimates of the partial variances for the second order effect ($\hat{V}_{ij}$) and the effect of all parameters except X_i ($\hat{V}_{\sim i}$) is given by

$$\hat{V}_{ij} = \frac{1}{n-1}\sum_{m=1}^{n} f\left(X_{\sim im}^{M_2}, X_{im}^{M_1}\right) * f\left(X_{\sim jm}^{M_1}, X_{jm}^{M_2}\right) - \hat{f}_0^2 - \hat{V}_i - \hat{V}_j \tag{14}$$

$$\hat{V}_{\sim i} = \frac{1}{n-1}\sum_{m=1}^{n} f\left(X_{\sim im}^{M_2}, X_{im}^{M_2}\right) * f\left(X_{\sim im}^{M_2}, X_{im}^{M_1}\right) - \hat{f}_0^2 \tag{15}$$

Similar to $X_{im}^{M_1}$, $X_{jm}^{M_2}$ represents the jth column of matrix M_2. For the calculation of $\hat{V}_{ij}$, it is necessary to subtract the estimates for $\hat{V}_i$ and $\hat{V}_j$ from the Monte Carlo integral in Equation (14), as this integral actually determines an estimate of the partial variance that is closed within the subset of input parameters (X_i, X_j), i.e. $V_{ij}^c = V_i + V_j + V_{ij}$.

To sample M_1 and M_2, the Sobol' quasi-random sampling technique (Sobol', 1967; 1976) is used. In general, a quasi-random sampling algorithm adds samples to the sequence away from the earlier sampled points, to avoid clustering and to fill the parameter space as uniformly as possible. Therefore, the use of these sequences enhances the convergence of the Monte Carlo integrals. Where the Monte Carlo integration and thus the Sobol' sensitivity analysis normally converges at a rate of $1/\sqrt{n}$, the Sobol' quasi-random sampling enhances this to almost $1/n$ (Kucherenko et al., 2009).

2.4.2. Bootstrapping

As it would be impossible to repeat the $n\cdot(p+2)$ model runs several times, in order to assess confidence intervals for the first order and total Sobol' sensitivity index, we used bootstrapping with resampling instead (Efron and Tibshirani, 1993). The n samples used for the model evaluations, were 1000 times sampled with replacement, whereby for each resampling the SI's are calculated. In this way, distributions for the S_i's and S_{Ti}'s are obtained and the 95% confidence intervals are constructed by using the percentile method and the moment method (Archer et al., 1997). The percentile method is very simple, but a higher number of resamples are necessary than for the moment method to achieve a reliable estimate of the percentiles. On the other hand, the moment method can result in a poorly estimated confidence interval if the bootstrap distribution is skewed (Archer et al., 1997).

2.4.3. The evaluation of the sensitivity analysis

To support and evaluate the identification of the important parameters for the model output, we used an adapted version of the evaluation test proposed by Tang et al. (2007), which originated from the SA repeatability test (Andres, 1997). Three parameter sets, based on 1000 independent samples of the parameters are constructed for this evaluation. Set 1 consists of the 1000 samples as such, where in Set 2 the T most sensitive parameters have the values of Set 1 and the other, 'non-sensitive' parameters are fixed to an a priori value. Oppositely, in Set 3 the T parameters with the highest total sensitivity index are fixed to an a priori value and the other parameters have the values of Set 1. The value of T was varied from 4 parameters to 10 parameters to evaluate the identification of influential and non-influential parameters.

If a correct classification of important and non-important parameters was used, the correlation coefficient between the model outputs of Set 2 and Set 1 should theoretically become 1, representing a line with slope 1 on a plot of Set 2 against Set 1. Under the same conditions, the correlation coefficient between Set 3 and Set 1 should become 0, as the parameters that are varied in Set 3 should have no influence on the model outcome. The plot of the output results of Set 3 against Set 1 will show a horizontal trend.

2.5. The objective function

In this study, the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) is used for evaluating the daily model output:

$$\mathrm{NSE} = 1 - \frac{\sum_i \left(y_{o,i} - y_{s,i}\right)^2}{\sum_i \left(y_{o,i} - y_{o,\mathrm{mean}}\right)^2} \tag{16}$$

In this equation, $y_{o,i}$ is the observed value on day i, $y_{o,\mathrm{mean}}$ is the average of the observations and $y_{s,i}$ is the simulated value on day i. The accuracy of the variance estimation may decrease when the estimate of the square of the expectation value (11) becomes too large (Sobol', 2001). Therefore, the NSE was chosen as an objective function since e.g. the commonly used Sum of Squared Residuals (SSR) can yield high values of $\hat{f}_0^2$.

2.6. Graphical representation

A visualization of the sensitivity of the model output to changes in parameter values can enhance the interpretation of information contained in the results of the SA. In this study, the marginal influence of the parameters on the model output is represented by scatterplots. These scatterplots are constructed by performing model evaluations for a quasi-random sample sequence of 24000 parameter sets. The resulting model efficiencies (NSE values) are then projected on a (parameter specific) plane and yield a cloud of points. In this way, p graphs of 1-dimensional slices of the response surface are constructed, representing the global sensitivity of the model to a specific parameter. If the points in the plot are randomly spread over the parameter range, this can indicate that the parameter does not influence the model output. When, oppositely, a clear pattern is observed in the scatterplot, this suggests that the parameter influences the model output and/or interactions between parameters exist. To enhance and simplify the interpretation of the scatterplots, the parameter range can be subdivided into smaller ranges. For each subrange, the mean of the NSE values corresponding with a parameter value within this subrange can be calculated. Compared to the scatterplots, the plots of these mean NSE values often show clearer trends.

3. Results and discussion

For this research, a sample size (n) of 12,000 was used, resulting in 336,000 ($= n\cdot(p+2)$) model evaluations to calculate the first order and total sensitivity index. Some preliminary results showed that this sample size was necessary to assure convergence of the SI values. 72,000 ($n\cdot p'$) runs were added to calculate the second order sensitivity indices of the 6 (p') most significant parameters. For the bootstrapping analysis, the results were resampled 1000 times.

In the following sections, the Sobol' sensitivity analysis results are presented and discussed. In Section 3.1, the first order and total sensitivity index of the Sobol' analysis are handled. Also the evolution of the parameter ranking and the convergence of the results are presented in this section. In Section 3.2 the evaluation of the results, based on the adapted repeatability test (Section 2.4.3), are discussed. Section 3.3 presents the additional first order and

Table 2
First order and Total sensitivity index with bootstrap confidence intervals.

S_i		Sobol' value	Bootstrap average	95% CI percentile meth.		S_{Ti}		Sobol' value	Bootstrap average	95% CI percentile meth.	
1	CN2	0.252	0.251	0.205	0.297	1	CN2	0.652	0.653	0.605	0.701
2	CH_N	0.069	0.069	0.035	0.106	2	CH_N	0.498	0.499	0.445	0.555
3	GWQMN	0.067	0.066	0.047	0.085	3	GWQMN	0.178	0.178	0.110	0.247
4	RCHRG_DP	0.025	0.026	0.012	0.040	4	ALPHA_BF	0.160	0.160	0.092	0.230
5	SLSUBBSN	0.019	0.019	0.010	0.028	5	CH_K2	0.097	0.097	0.036	0.163
6	SOL_K	0.014	0.014	0.006	0.021	6	RCHRG_DP	0.088	0.088	0.022	0.155
7	CH_K2	0.011	0.011	−0.007	0.033	7	SURLAG	0.060	0.060	−0.004	0.125
8	SURLAG	0.010	0.010	−0.002	0.022	8	SLSUBBSN	0.052	0.052	−0.011	0.119
9	SOL_AWC	0.007	0.007	−0.002	0.016	9	SOL_AWC	0.039	0.039	−0.026	0.105
10	GW_REVAP	0.007	0.007	0.003	0.011	10	SOL_K	0.034	0.034	−0.029	0.101
…	…	…	…	…	…	…	…	…	…	…	…
13	SMTMP	0.001	0.001	0.000	0.003	17	GW_REVAP	0.012	0.012	−0.052	0.079
…	…	…	…	…	…	…	…	…	…	…	…
26	ALPHA_BF	0.000[a]	0.000[a]	−0.026	0.017	26	SMTMP	0.010	0.010	−0.055	0.076
	SUM	0.482					SUM	2.074			

[a] The values of S_i for 5 parameters were slightly smaller than 0, because of numerical integration. The values were therefore reset to 0.

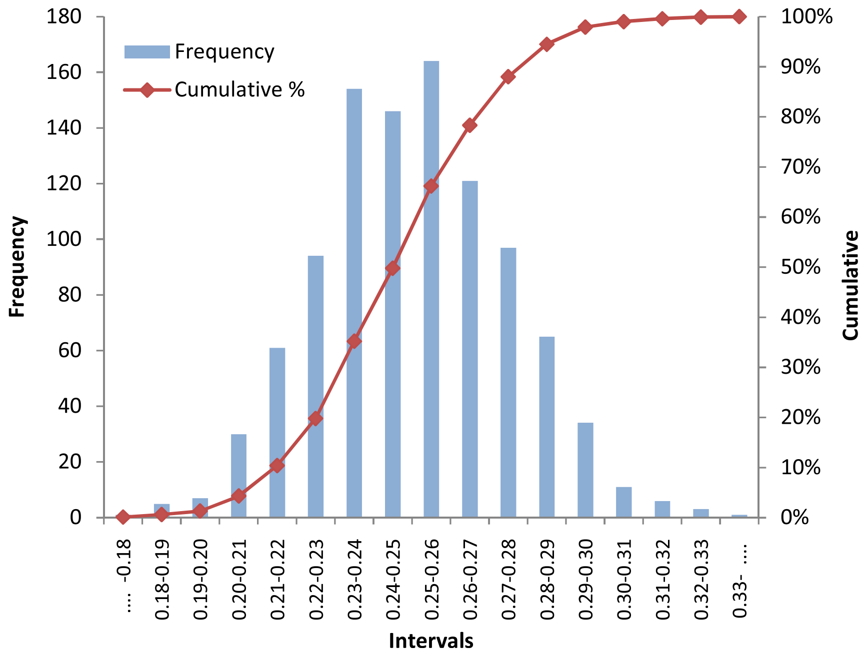

Fig. 1. Bootstrapping distribution of the first order sensitivity index of the parameter CN2.

total sensitivity results calculated at HRU level, followed by the results of the second order sensitivity analysis (in Section 3.4) and the graphical representations (in Section 3.5).

3.1. The first order and total sensitivity index

As mentioned in Section 2.4.1, the basic outcome of the Sobol' SA are the first order and total sensitivity index. Table 2 presents these two indices together with the 95% bootstrap confidence intervals, as calculated with the percentile method. Only the most sensitive parameters, together with a few non-sensitive ones are considered in this table.

The left part of the table (columns 1 until 6) presents the results of the first order sensitivity index S_i, while the right half (columns 7 until 12) shows the total sensitivity index S_{Ti}. The first and seventh column provide the ranking numbers of the parameters. Column 3 (9) shows the Sobol' sensitivity index, as calculated with the 12,000 samples. Column 4 (10) shows the averages of the 1000 bootstrap resampling results. It may be observed that the pairwise results of columns 3 and 4 (9 and 10) are almost equal, which indicates that our bootstrapping is unbiased. The confidence intervals on the sensitivity indices, as based on the percentile method are shown in columns 5 and 6 (11 and 12). With regard to these confidence intervals, it can also be stated that the results of the percentile method and the moment method were very similar, indicating an accurate estimation of the confidence intervals and a symmetric, average centered distribution. This is also confirmed by Fig. 1, which shows the bootstrapping distribution of the first order sensitivity index of the parameter CN2. Similarly, the bootstrapping distribution of the other parameters approaches a normal distribution.

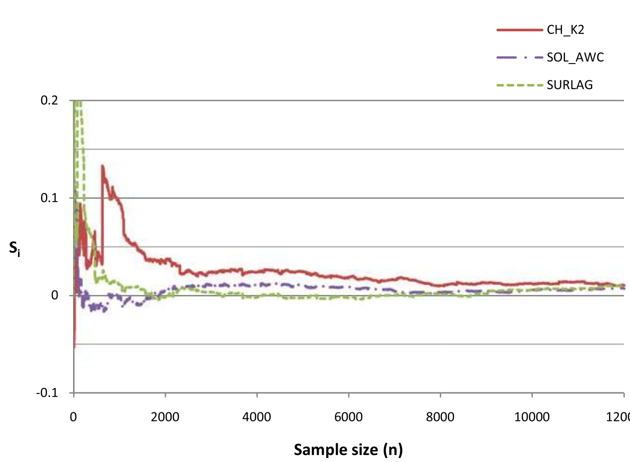

Fig. 2. Evolution of the first order sensitivity index for CH_K2, SOL_AWC and SURLAG (detail).

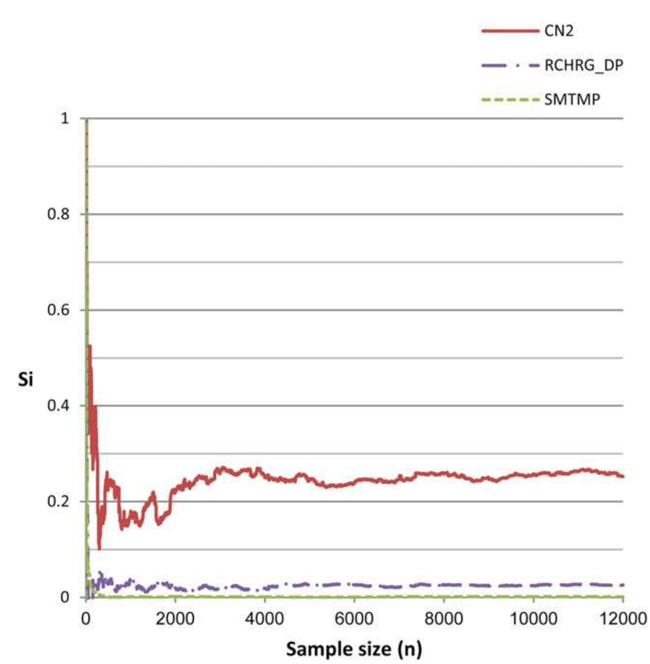

Fig. 3. Evolution of the first order sensitivity index for CN2, RCHRG_DP and SMTMP.

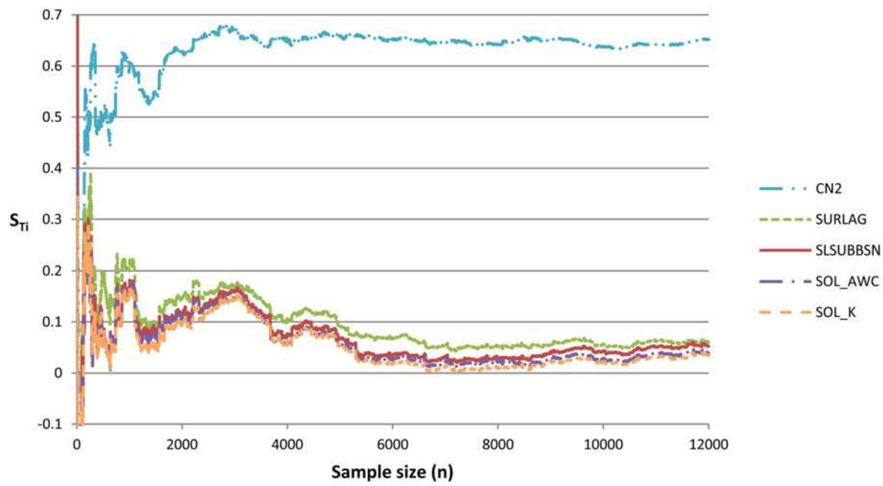

Fig. 4. Evolution of the total sensitivity index for CN2, SURLAG, SLSUBBSN, SOL_AWC and SOL_K.

From the left part of Table 2 (S_i) it can be clearly noticed that the curve number (CN2), is by far the most important single parameter. It has 4 times more influence on the total variance than the second and third ranked parameters, CH_N (Manning coefficient of the channel) and GWQMN (threshold depth of water in the shallow aquifer required for return flow to occur). Only 4 other parameters (RCHRG_DP, SLSUBBSN, SOL_K and GW_REVAP; see Table 1 for the definition of these parameters) have a significant main effect, as the lower limit of their confidence interval is larger than 0. Remarkably however, GW_REVAP has a lower nominal S_i value than the parameters CH_K2, SURLAG and SOL_AWC, while only the S_i value for GW_REVAP is statistically significant (which is not the case for the other pre-mentioned parameters). The explanation for this can be derived from the width of the confidence intervals of the parameters at hand (Table 2). The width of the CI of GW_REVAP is only 0.008, where this is 0.040 for CH_K2. This indicates that the values of S_i for a single sample are very stable for GW_REVAP while they are more fluctuating for CH_K2. The latter can also be noticed from Fig. 2, showing e.g. strong variations for the values of S_i for CH_K2 at a sample size below 2000. On the other hand, Fig. 2 also shows that the value of S_i for the parameters CH_K2, SOL_AWC and SURLAG is (almost) always above 0, independently of the sample size, and is converging to a positive value. This can be interpreted as an indication of a stable positive value of S_i, suggesting a significant main effect of these parameters, although this was not observed in the CI.

In general, the values of S_i for the different parameters have converged within the range of the 12000 samples. For most parameters, less than 5000 samples are sufficient to reach a stable value. From Fig. 3, showing the evolution of S_i for increasing sample size for parameters CN2, RCHRG_DP and SMTMP, it can be seen that for some parameters even less than 1000 samples are enough to converge to their final S_i value. Especially the parameters with a first order sensitivity index of about 0 (like SMTMP on Fig. 3) reach their final value very quickly.

It is also noted from the left part of Table 2 that the sum of all first order indices is less than 1, which means the model is non-additive, as could be expected.

The right part of Table 2 (S_{Ti}) shows that 65% of the variations in the simulated flow are caused by variations of CN2, either by the variation of the parameter itself (25%) or by interactions with other parameters. Together with CH_N (50%), it is by far the most influential parameter for the simulated flow. Also GWQMN and ALPHA_BF have a high total sensitivity index, although their S_{Ti} value is 3–4 times smaller than for CN2. 9 out of the top 10 parameters for the first order sensitivity index are also in the top 10 for the total sensitivity index, but in total, only 6 parameters are significantly sensitive according to the 95% confidence interval. Remarkably, SLSUBBSN, SOL_K and GW_REVAP are significant for the main effect, but have now a lower limit of their confidence interval for the total sensitivity index that is lower than 0. This

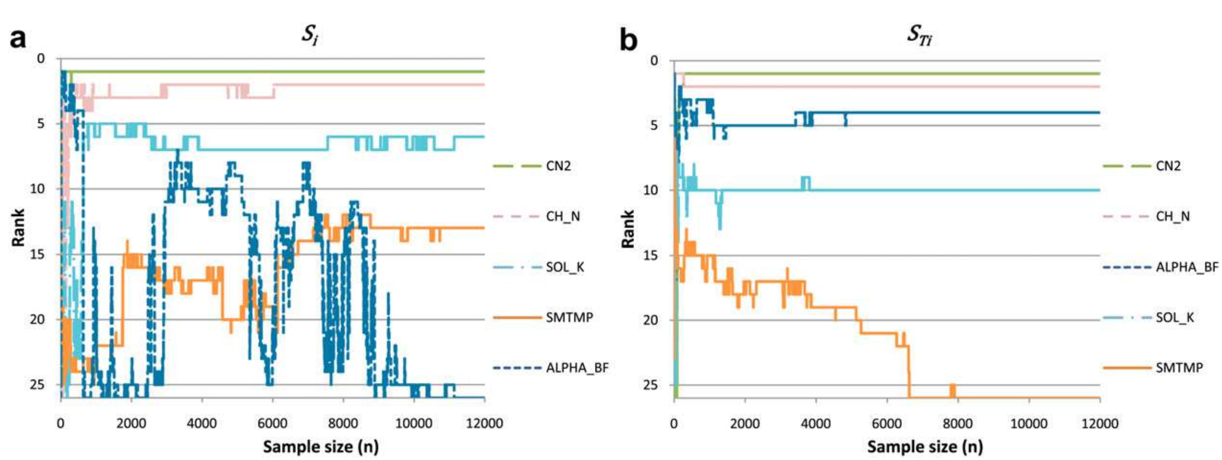

Fig. 5. Evolution of the parameter ranking based on S_i (a) and S_{Ti} (b) for increasing sample size.

Table 3
Correlation coefficients of the adapted repeatability test for a varying number of sensitive parameters.

Parameters	Correlation coefficient	
	Set 2 vs Set 1	Set 3 vs Set 1
4	0.699	0.254
6	0.822	0.191
7	0.922	0.195
8	0.938	0.134
9	0.963	0.106
10	0.972	0.092

observation makes the significance of the first order effect of these parameters questionable. On the other hand, on the detailed presentation of the evolution of S_{Ti} for SURLAG, SLSUBBSN, SOL_AWC and SOL_K for increasing sample size (Fig. 4), it can be noticed that these parameters almost always have a positive value.

A possible reason for SLSUBBSN, SOL_K and GW_REVAP becoming insignificant (according to the lower limit of their CI) is the slower convergence of S_{Ti} (Fig. 4). For most parameters, a quasi stable value of S_{Ti} is only reached for a sample size of 9000 or more, but even then smaller fluctuations can be observed. This all results in much wider 95% confidence intervals for these parameters for S_{Ti} than for S_i and, similar as for CH_K2, SURLAG and SOL_AWC in the case of S_i, to a negative lower limit of the CI. Based on this evidence and their S_i and S_{Ti} values, it may be assumed that the flow simulations are also sensitive to the variation of SURLAG, SLSUBBSN, SOL_AWC and SOL_K.

Another striking result is the value and ranking of parameter ALPHA_BF (Baseflow recession factor). One would expect this parameter to be sensitive, as in practice it is often classified as an important parameter (Holvoet et al., 2005; van Griensven et al., 2006). As a parameter on itself however, it has no influence on the total variance ($S_i = 0$), although it has a much wider 95% confidence interval than other non-influential parameters. Oppositely, ALPHA_BF is one of the most important parameters for the total sensitivity index (rank 4) (Table 2). This indicates that all of the influence of ALPHA_BF is caused by interactions with other parameters.

In general it is noticed that all total sensitivity indices are higher than the first order sensitivity indices. This is of course theoretically necessary, as S_i is a part of S_{Ti}, but as the numerical integration only yields an estimate of the indices (whose accuracy highly depend on the sample size), this needs to be checked. The difference between S_{Ti} and S_i is a measure for the influence of interactions between parameters on the model output. For the calculation of the second order sensitivity indices (discussed in §3.4), the 6 highest ranked parameters for S_{Ti} are used, because they also have the highest interaction value ($S_{Ti}-S_i$).

Very often, a sensitivity analysis is performed to achieve a parameter ranking. Fig. 5 presents the evolution of the parameter ranking for increasing sample size for S_i (a) and S_{Ti} (b) for 5 parameters. It is noticed that for the parameters with the highest sensitivity indices, the final ranking is reached very quickly, whereas for less important parameters, more fluctuations are observed. Remarkably, the fluctuations for the S_i ranking are more important than for S_{Ti}, where the values of the sensitivity indices reach their final value much faster for the first order sensitivity index (Fig. 3). Both these observations can be explained by the nominal values of the indices: small values (like for S_i and especially for the insensitive parameters) are more prone to the minor fluctuations that can appear with increasing sample size, as the differences between the values are small.

The results of the ranking evolution, based on the total sensitivity index, reveal an additional opportunity for the model of the Kleine Nete catchment: if the objective of the SA is only to calculate a parameter ranking prior to a calibration of the most important parameters, then the Sobol' analysis can be applied with a sample size of less than 2000 to have a reliable ranking result. As shown in Fig. 5(b), the S_{Ti} based ranking for the most important parameters remains indeed very stable for 2000 samples and more. This result looks very promising for future research, although uncertainties and specific characteristics of other catchments may lead to a higher number of evaluations required to achieve a stable sensitivity ranking.

3.2. The evaluation with the adapted repeatability test

When applying the repeatability test, a certain number of parameters are categorized as 'sensitive' and the remaining are said to be 'non-sensitive'. In our adapted version of the test, we vary this number of 'sensitive' parameters (see Section 2.4.3). In this way, we can infer how many parameters are necessary to represent the model variability in an acceptable way and assess the loss in variability when the remaining parameters are put on a fixed value. This information is very important when the SA is performed in order to indicate a limited number of parameters that should be included in the parameter estimation process (Factor Fixing/Factor Prioritization).

Based on the ranking for the total sensitivity index (Table 2), the number of 'sensitive' parameters was varied from 4 to 10, leading to increasing correlation coefficients when comparing the output of Set 2 with Set 1, and decreasing correlation coefficients when comparing the output of Set 3 with Set 1 (Table 3).

The six parameters that were significant (based on the CI's) for the total sensitivity index only yield a correlation coefficient of 0.82 for Set 2 vs. Set 1 (Fig. 6(a)) and 0.19 for Set 3 vs. Set 1. Although, removing the outlier visible in Fig. 6(a) increases the correlation coefficient for Set 2 vs. Set 1 in the case of 6 parameters to 0.89. An even better result is achieved when also SURLAG (7th ranked parameter) is added to the category of sensitive parameters. Adding SLSUBBSN and SOL_AWC brings the value of R higher than 0.95 for Set 2 vs. Set 1 (Fig. 6(b)) and to about 0.1 for Set 3 vs. Set 1 (Fig. 6(c)).

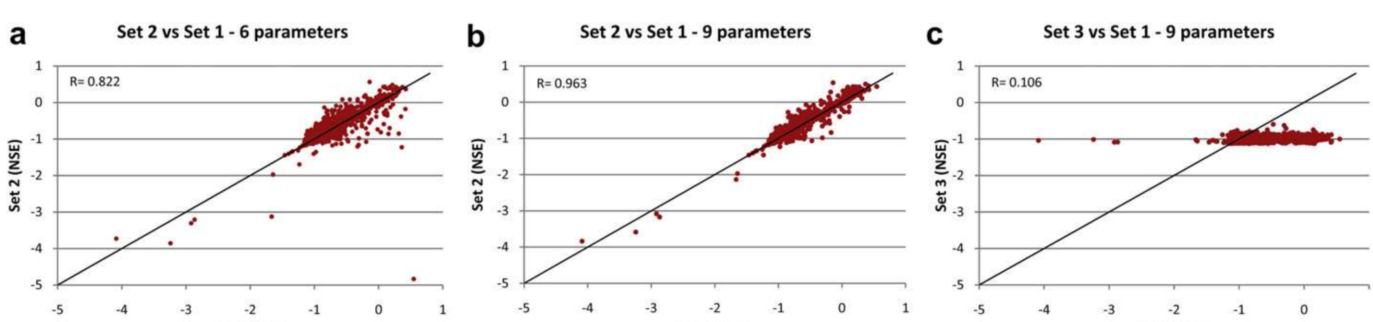

Fig. 6. Results of the adapted repeatability test: (a) Set 2 vs Set 1 for 6 parameters, (b) Set 2 vs Set 1 for 9 parameters, (c) Set 3 vs Set 1 for 9 parameters.

Table 4
First order and total sensitivity index for the curve number value for different land use and soil type combinations.

S_i	Area fraction	Sobol' value		95% CI percentile meth.		S_{Ti}	Area fraction	Sobol' value		95% CI percentile meth.	
Sand-Crop	37.4%	0.515	76.6%	0.422	0.606	Sand-Crop	37.4%	0.679	51.7%	0.593	0.770
Sand-Forest	25.8%	0.134	19.9%	0.045	0.212	Sand-Forest	25.8%	0.370	28.1%	0.277	0.472
Sand-Pasture	19.4%	0.020	2.95%	−0.052	0.084	Sand-Pasture	19.4%	0.234	17.8%	0.139	0.332
Sand-Urban	15.2%	0.004	0.55%	−0.022	0.029	Sand-Urban	15.2%	0.031	2.40%	−0.064	0.125
Loam	2.3%	0.000[a]	0.00%	−0.011	0.003	Loam	2.3%	0.000[a]	0.00%	−0.095	0.089
SUM	100%	0.672	100%			SUM	100%	1.313	100%		

[a] The values for loam were slightly smaller than 0, because of numerical integration. The values were therefore reset to 0.

The latter configuration with 9 sensitive parameters can be considered as a good matching of the output variability. The model retains its ability to forecast the flow when the remaining 17 (non-sensitive) parameters are fixed, as they have almost no influence on the total variance of the model output.

The results of the adapted repeatability test back up our belief that besides the parameters with a positive lower limit of the CI, also other parameters are important. Categorizing SURLAG, and also SLSUBBSN and SOL_AWC, as sensitive is necessary to sufficiently retain the model variability (R of the repeatability test higher than 95%).

In general, it can be noticed that the parameters that directly influence the evapotranspiration, like CANMX, EPCO, ESCO and SOL_ALB, but also the nutrient parameters (that influence the plant growth) are not significantly important for this model of the Kleine Nete catchment. As could be expected, also the snow related parameters (SMTMP and TIMP) do not have a significant influence on the model output variability.

3.3. Sensitivity analysis for the curve number value at HRU level

It was noted in Section 3.1 that the curve number value (CN2) is the most influential parameter of our model for flow predictions (Table 2). As this parameter value is strongly related to the land use and soil type, it is often estimated at HRU level in a model optimization, because aggregating it in one value reduces the model variability. For a particular analysis performed for this study, the CN2 parameter was also split up into several factors based on the land use and soil type, to infer the factors' influence on the model output. For sandy soils (more than 95% of the catchment area), four types of land use are distinguished for this analysis: urban area, forest, cropland and pastureland. The CN2 values for the second soil type (loam) for different land uses are combined in one value.

As for Table 2, the left part of Table 4 (columns 1 until 6) presents the results of this SA for the first order sensitivity index S_i and the right half (columns 7 until 12) shows the total sensitivity index S_{Ti}. Column 2 (8) provides the percentage of the total area for the different combinations of land use and soil type considered in this SA. In column 3 (9) the Sobol' sensitivity index is shown, accompanied by its relative value compared to the sum of the index values (column 4 (10)) and the 95% confidence intervals (column 5 (11) and 6 (12)).

Table 5
Second order sensitivity indices with the associated 95% confidence intervals for 3 parameter pairs.

S_{ij}	Interaction effect	95% CI percentile meth.	
CN2-CH_N	0.242	0.165	0.317
CN2-ALPHA_BF	0.055	0.002	0.110
GWQMN-RCHRG_DP	0.037	0.013	0.059

Based on the lower limits of the CI's, it can be seen that only for the combinations of sand with cropland and forest the CN2 value has a significant main effect on the simulated flow. For the total sensitivity index, besides the former two, also the curve number value for "sand—pasture" has a significant influence on the model output.

The curve number value associated with cropland has a very large influence on the model output, mainly caused by the first order effect. Moreover, the main effect is almost 4 times higher than the main effect of the second ranked combination and also the total effect is almost 2 times higher. As could be expected, a relation between the land use and soil type's fraction of the total catchment area and the sensitivity index of the related factor can be noticed, particularly for S_{Ti}. However, where it could be expected that the CN2 value of the urban area, is more influential on the model output as compared to the other land uses associated fraction of the area (because of its higher potential to generate surface runoff), this is not the case. On the contrary, it has a very low first order and total sensitivity index. The explanation for this irregularity is found in the model processes: the real curve number for urban areas in SWAT is a combination of 60% impervious area (with an associated CN2 value of 98) and 40% area with the curve number parameter that was incorporated in this SA. In this way, the influence of this input parameter is muted with regard to the model output variance.

3.4. The second order sensitivity index

The computation of the second order sensitivity index S_{ij} was performed for the 6 parameters with the largest second or higher order influences (S_{Ti}–S_i): CN2, CH_N, GWQMN, ALPHA_BF, CH_K2 and RCHRG_DP. They are each evaluated for combinations with the other 25 parameters.

It is noticed that for this model only 3 pairwise interactions have a significant influence on the model output variability: CN2—CH_N, CN2-ALPHA_BF and GWQMN-RCHRG_DP. The S_{ij} values and 95% CI's associated with these interactions are shown in Table 5. The other interactions had a S_{ij} value of 0 or a lower limit of their CI below 0. By using the values of the first order and total sensitivity indices (Table 2) and the now presented second order sensitivity indices, it would also be possible to calculate the remaining effect of higher order interactions ($S_{Ti} - S_i - S_{ij}$) for the 6 most important parameters.

The pairwise interactions that were highlighted by this SA elucidate some important model processes and particularly how these processes influence one another. The CN2-ALPHA_BF interaction could be expected, as these parameters have a large influence on the definition of the fast and slow runoff response of the system. The more water is diverted to the fast runoff, the less is available for the slow runoff and vice versa. This also leads to a trade-off between the parameter values. The relation between RCHRG_DP and GWQMN gives more insight on how the groundwater flow is regulated in the SWAT model and how both

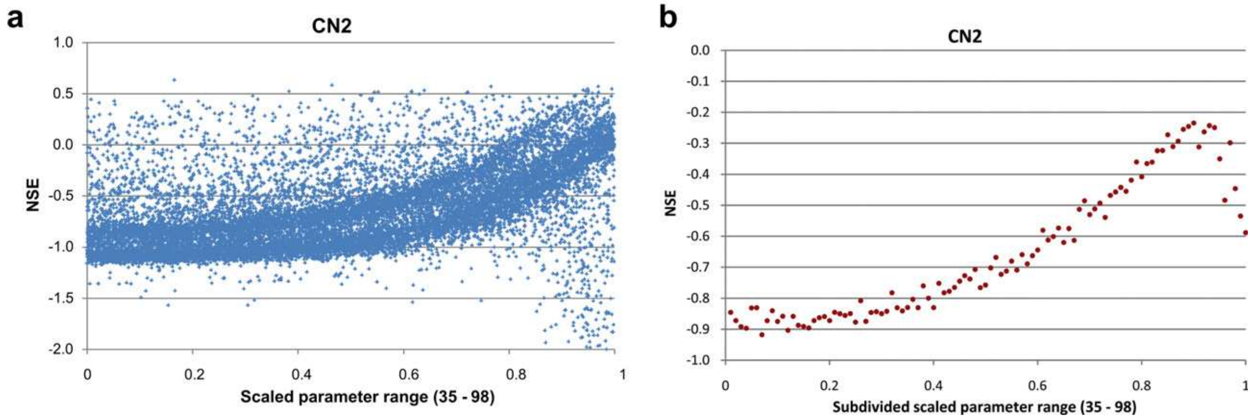

Fig. 7. Scatterplot (a) and mean NSE value plot (b) for CN2.

parameters contribute to the simulated outflow: RCHRG_DP defines the fraction of the recharge that goes to the deep aquifer. The remaining part goes to the shallow aquifer and partly determines the amount of water in this storage. If this amount of water is higher than the GWQMN value, return flow occurs and contributes to the total outflow. In this way, RCHRG_DP and GWQMN have an interaction influence on the simulated flow, as RCHRG_DP has an impact on the storage in the shallow aquifer and thus on GWQMN.

3.5. Graphical representation of the sensitivity

Figs. 7, 8 and 9 present the marginal response surface of 3 parameters (CN2 (Fig. 7), CH_N (Fig. 8) and SMTMP (Fig. 9) for flow simulations, by showing the variation of the NSE value over the full range of possible parameter values. Each figure consists of two graphs: a regular scatterplot (a) and a mean NSE value plot (b). The parameter ranges in Figs. 7, 8 and 9 are scaled (0–1) to allow a straightforward comparison.

The scatterplot of CN2 (Fig. 7(a)) shows a clear trend of increasing NSE values with increasing parameter values. However, for CN2 values higher than 88, a larger spread of the NSE values towards lower efficiencies is observed, which might indicate a decrease of the efficiency for very high CN2 values. The latter is confirmed by the plot of the mean NSE values (Fig. 7(b)). The decrease is somehow logical, because for high CN2 values almost all available water will become surface runoff, leading to a faster response of the simulated flow compared to the measured flow. On the other hand, the larger spread of the NSE values on the scatterplot can point out interactions of high CN2 values with other parameters (e.g. CH_N), as also high efficiencies are observed for these high curve number values.

Although a trend is still visible in the scatterplot of CH_N (Fig. 8(a)), it is less clear than for the scatterplot of the curve number value (Fig. 7(a)). Also the other parameters do not show a trend that is as strong as for CN2, confirming that the curve number value has the largest influence on the flow simulations of the SWAT model of the Kleine Nete catchment.

Nevertheless, the chart of the mean NSE values for CH_N (Fig. 8(b)), calculated by subdividing the parameter range into 100 equal subranges, confirms that CH_N does have an important impact on the flow simulations. Furthermore, the plot reveals a discontinuity for CH_N values higher than 0.047. The manning coefficient of the main channel (CH_N) rules the velocity of the flow in the reaches. If the CH_N value gets higher, the velocity is reduced and the simulated flow will become lagged compared to the measured flow. As this SWAT model of the Kleine Nete catchment has reaches of a similar length and the output is calculated with a daily time step, a low velocity will direct the majority of the flow to the next day, inducing a drop in the NSE value, instead of a continuous decrease.

The response surface for the snow melt base temperature parameter (SMTMP) (Fig. 9(a) & (b)) provides an illustration of a parameter that has (almost) no impact on the flow simulations. In general, the scatterplots for non-influential parameters show no clear pattern and the plot of the mean values is an almost horizontal line.

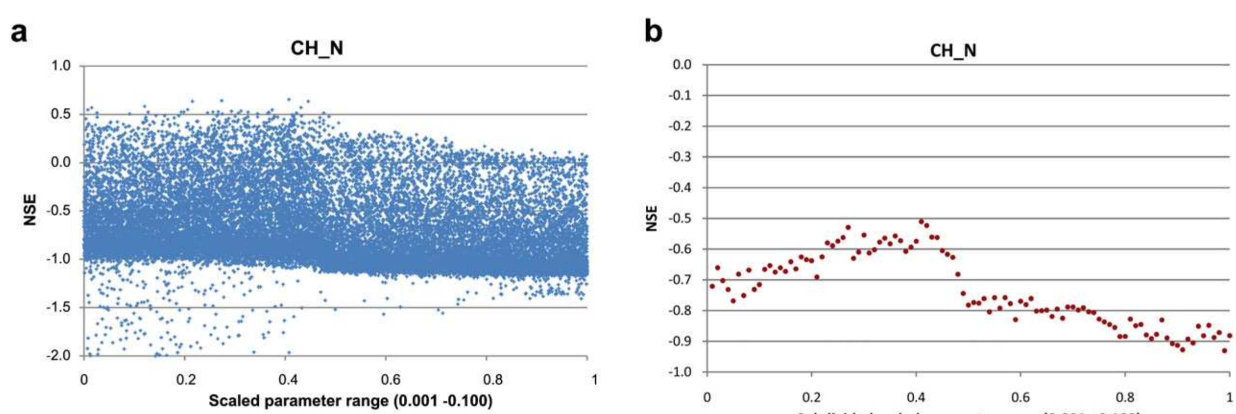

Fig. 8. Scatterplot (a) and mean NSE value plot (b) for CH_N.

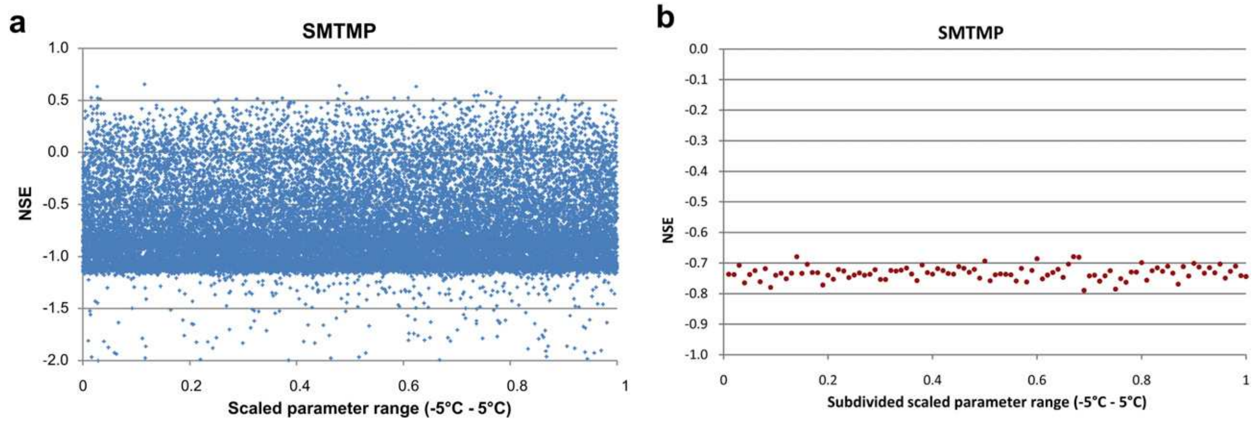

Fig. 9. Scatterplot (a) and mean NSE value plot (b) for SMTMP.

From the observations discussed in this section it is clear that graphical representations of the sensitivity of the model output to changes in parameter values can support the interpretation of information provided by the sensitivity analysis.

4. Conclusions

This paper presents the results of a Sobol' sensitivity analysis for flow simulations of a SWAT model of the Kleine Nete catchment, Belgium. The model independent Sobol' method was applied in order to overcome some of the major disadvantages of screening sensitivity analysis methods, which are often applied for SWAT. The Sobol' sensitivity analysis technique is based on variance decomposition and is able to handle non-linear and non-monotonic models.

A sample size of 12,000 was used to assess the first order and total sensitivity effect of twenty-six input parameters that affect stream flow and water quality simulations, leading to 336000 model evaluations. In general, the parameter rankings and quantitative results for these two effects can be used for both factor fixing and factor prioritization. Additionally, 72,000 model simulations yielded the second order sensitivity effects for six of the twenty-six parameters. These second order effects represent the interactions between pairs of parameters. Bootstrapping was applied to infer the 95% confidence intervals of the different sensitivity indices and the results were also evaluated with an adapted version of the SA repeatability test. Also the convergence of the sensitivity indices for increasing sample size was analyzed.

As the sum of all first order sensitivity indices is less than one, the model is non-additive, as could have been expected. Parameters CN2, CH_N and GWQMN appeared to have the most important first order and total effect on the flow simulations. Variations of the curve number (CN2) even contribute to 65% of the total unconditional variance of the model output, of which 25% without interactions with other parameters. The curve number for the HRU's with sandy soils and cropland has the highest influence on the stream flow. This is somehow logical because they represent the largest area in the basin. On the other hand, the CN2 value for urban area has a very small influence on the stream flow simulations compared to the corresponding area, although the opposite could have been expected. This is related to the use of an impervious fraction of the urban area with fixed curve number value, which is not included in the user defined CN2 value for this HRU. The baseflow recession factor (ALPHA_BF) has one of the highest total sensitivity indices (rank 4), but the parameter on itself has no influence on the total variance. This indicates that all influence of ALPHA_BF is caused by interactions with other parameters. The remaining parameters in the top 10 parameters for the total sensitivity index are also in the top 10 for the first order sensitivity index.

The results of the adapted SA repeatability test show that 9 parameters (CN2, CH_N, GWGMN, ALPHA_BF, CH_K2, RCHRG_DP, SURLAG, SLSUBBSN and SOL_AWC) should be considered to adequately capture the total model variance ($R = 0.963$). However, the confidence intervals inferred with the bootstrapping, which is unbiased and has a symmetric, average centered distribution, indicate that only 6 parameters have a significant total effect on the flow simulations. The latter could be related to more fluctuating values of the total sensitivity index for a single sample set, leading to wider confidence intervals.

The results for the first order sensitivity indices were converging quickly. After 5000 of the 12,000 samples, most of the parameters had reached their final value. Especially the non-sensitive parameters converged very fast. On the other hand, the total sensitivity indices were converging relative slow. For most of the parameters, it took about 9000 model evaluations to reach the final value and even then, small fluctuations were still present. Oppositely, the ranking of the parameters for the total sensitivity effect converged very quickly, especially for the most influential factors. A sample size of 2000 was needed for the significant parameters to attain their final rank. This is promising for the use of Sobol' sensitivity analysis as a preliminary step for parameter estimation, as the number of model evaluations could be limited in this way.

The difference between the total and first order sensitivity indices is a measure for the amount of interaction effects induced by a given parameter. It was noticed that for this model only 3 pairwise interactions have a significant influence on the model output variability: CN2-CH_N, CN2-ALPHA_BF and GWQMN-RCHRG_DP. Also some higher order interaction effects have an influence on the flow simulations.

The representations of the marginal response surface of the different parameters with scatterplots and mean NSE value plots can give additional insight on the influence of the input parameters on the model output. Clear trends are observed for (highly) influential parameters. A non-continuous evolution is noted for CH_N. This can be attributed to the decreasing channel velocity with higher values of CH_N, introducing a lag in the daily model output.

In general, it can be concluded that the Sobol' sensitivity analysis can be successfully applied for factor fixing and factor prioritization with respect to the input parameters of a SWAT model, even with a limited number of model evaluations. In addition, the sensitivity analysis can also support the identification of model processes, parameter values and parameter interaction effects. Therefore, the Sobol' method can be considered as a powerful and robust sensitivity analysis method, even though it requires more model evaluations than common screening methods.

Acknowledgement

The authors would like to thank GIS Vlaanderen, the Royal Meteorological Institute, the Flemish Environment Agency, the Flemish Land Agency, the Flemish Manure Bank and the Hydrologic Information Centre of the Flemish Region for providing all the necessary data.

The authors also highly appreciated the remarks and suggestions of the four anonymous reviewers.

References

Andres, T.H., 1997. Sampling methods and sensitivity analysis for large parameter sets. Journal of Statistical Computation and Simulation 57 (1–4), 77–110.

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 (2), 99–120.

Arnold, J.G., Allen, P.M., Bernhardt, G., 1993. A comprehensive surface-groundwater flow model. Journal of Hydrology 142 (1–4), 47–69.

Beven, K., Binley, A., 1992. The future of distributed models - model calibration and uncertainty prediction. Hydrological Processes 6 (3), 279–298.

Campolongo, F., Kleijnen, J., Andres, T., 2000. Screening methods. In: Saltelli, A., Chan, K., Scott, M. (Eds.), Sensitivity Analysis. Wiley, Chichester, pp. 65–80.

Cibin, R., Sudheer, K.P., Chaubey, I., 2009. Sensitivity and identifiability of stream flow generation parameters of the SWAT model. Hydrological Processes 24 (9), 1133–1148.

Doherty, J., 2003. Ground water model calibration using pilot points and regularization. Ground Water 41 (2), 170–177.

Duan, Q.Y., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research 28 (4), 1015–1031.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman Hall, New York, USA.

Estrada, V., Diaz, M.S., 2010. Global sensitivity analysis in the development of first principle-based eutrophication models. Environmental Modelling and Software 25 (12), 1539–1551.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. Transactions of the Asabe 50 (4), 1211–1250.

Green, C.H., van Griensven, A., 2008. Autocalibration in hydrologic modeling: using SWAT2005 in small-scale watersheds. Environmental Modelling and Software 23 (4), 422–434.

Holvoet, K., van Griensven, A., Seuntjens, P., Vanrolleghem, P.A., 2005. Sensitivity analysis for hydrology and pesticide supply towards the river in SWAT. Physics and Chemistry of the Earth 30 (8–10), 518–526.

Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering and System Safety 52 (1), 1–17.

Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C., Shah, N., 2009. Monte Carlo evaluation of derivative-based global sensitivity measures. Reliability Engineering and System Safety 94 (7), 1135–1148.

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

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – a discussion of principles. Journal of Hydrology 10 (3), 282–290.

Nossent, J., Bauwens, W., Multi-variable sensitivity analysis for a complex environmental model in view of integrated water quantity and water quality modeling. Water Science and Technology (in press).

Pappenberger, F., Beven, K.J., Ratto, M., Matgen, P., 2008. Multi-method global sensitivity analysis of flood inundation models. Advances in Water Resources 31 (1), 1–14.

Ravalico, J.K., Dandy, G.C., Maier, H.R., 2010. Management option rank equivalence (MORE) - a new method of sensitivity analysis for decision-making. Environmental Modelling and Software 25 (2), 171–181.

Saltelli, A., 2000. What is sensitivity analysis? In: Saltelli, A., Chan, K., Scott, M. (Eds.), Sensitivity Analysis. Wiley, Chichester, pp. 3–14.

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Computer Physics Communications 145 (2), 280–297.

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling and Software 25 (12), 1508–1517.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley & Sons, Ltd, Chichester.

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., European, C., Joint Res Ctr, I, 2006. Sensitivity analysis practices: strategies for model-based inference. Reliability Engineering and System Safety 91 (10–11), 1109–1125.

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

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications 181 (2), 259–270.

Sobol', I.M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. USSR Computational Mathematics and Mathematical Physics 7, 86–112.

Sobol', I.M., 1976. Uniformly distributed sequences with an additional uniform property. USSR Computational Mathematics and Mathematical Physics 16, 236–242.

Sobol', I.M., 1990. On sensitivity estimation for nonlinear mathematical models. Matematicheskoe Modelirovanie, 112–118.

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., Reed, P., Wagener, T., van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrology and Earth System Sciences 11 (2), 793–817.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, A., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. Journal of Hydrology 324 (1–4), 10–23.

van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2008. Characterization of watershed model behavior across a hydroclimatic gradient. Water Resources Research 44 (1).

Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resources Research 39 (8).

White, K.L., Chaubey, I., 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. Journal of the American Water Resources Association 41 (5), 1077–1089.

Yang, J., 2011. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. Environmental Modelling and Software 26 (4), 444–457.