Global sensitivity analysis for urban water quality modelling: Terminology, convergence and comparison of different methods

Peter A. Vanrolleghem [a], Giorgio Mannina [b,*], Alida Cosenza [b], Marc B. Neumann [a,c,d]

[a] modelEAU, Département de génie civil et de génie des eaux, Université Laval, 1065 av. de la Médecine, Québec, QC G1V 0A6, Canada
[b] Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy
[c] Basque Centre for Climate Change, Alameda Urquijo, 4 – 4°, 48008 Bilbao, Spain
[d] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

ARTICLE INFO

ABSTRACT

Sensitivity analysis represents an important step in improving the understanding and use of environmental models. Indeed, by means of global sensitivity analysis (GSA), modellers may identify both important (*factor prioritisation*) and non-influential (*factor fixing*) model factors. No general rule has yet been defined for verifying the convergence of the GSA methods. In order to fill this gap this paper presents a convergence analysis of three widely used GSA methods (SRC, Extended FAST and Morris screening) for an urban drainage stormwater quality–quantity model. After the convergence was achieved the results of each method were compared. In particular, a discussion on peculiarities, applicability, and reliability of the three methods is presented. Moreover, a graphical Venn diagram based classification scheme and a precise terminology for better identifying important, interacting and non-influential factors for each method is proposed. In terms of convergence, it was shown that sensitivity indices related to factors of the quantity model achieve convergence faster. Results for the Morris screening method deviated considerably from the other methods. Factors related to the quality model require a much higher number of simulations than the number suggested in literature for achieving convergence with this method. In fact, the results have shown that the term "screening" is improperly used as the method may exclude important factors from further analysis. Moreover, for the presented application the convergence analysis shows more stable sensitivity coefficients for the Extended-FAST method compared to SRC and Morris screening. Substantial agreement in terms of factor fixing was found between the Morris screening and Extended FAST methods. In general, the water quality related factors exhibited more important interactions than factors related to water quantity. Furthermore, in contrast to water quantity model outputs, water quality model outputs were found to be characterised by high non-linearity.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The evaluation of urban water quality represents a key issue in the urban drainage field in view of implementing environmental protection for receiving water bodies (Novotny et al., 1985). In this context mathematical models able to predict both water quantity and quality characteristics may provide useful support. Despite the fact that several water quality models are available in the urban drainage field, several aspects still limit their applicability, e.g. the extreme spatial and temporal variability of the water quality–quantity characteristics or the lack of distributed field data,

which consequently forces modellers to impose a considerable number of assumptions. Due to these assumptions their predictions are characterised by high uncertainty (Beck, 1987; Ashley et al., 2005; Deletic et al., 2012; Dotto et al., 2012; Mannina and Viviani, 2010). One may ask whether and how these model assumptions influence the output of the model. In this context, sensitivity analysis represents a very powerful tool to provide answers, as it is able to determine how uncertain factors affect the model outputs (Saltelli et al., 2004). The term "factors" includes all the input variables and the model parameters that are varied during the sensitivity analysis.

Several sensitivity analysis methods have been proposed in literature, mainly divided into two groups: local sensitivity analysis methods and global sensitivity analysis methods (Saltelli, 2000). The local methods provide a measure of the local effect on the model output of a given model factor by evaluating the change

* Corresponding author. Tel.: +39 091 23896556; fax: +39 0916657749.
E-mail addresses: Peter.Vanrolleghem@gci.ulaval.ca (P.A. Vanrolleghem), giorgio.mannina@unipa.it (G. Mannina), alida.cosenza@unipa.it (A. Cosenza), marc.neumann@bc3research.org (M.B. Neumann).

in model outputs under small changes of the model factors. Global sensitivity analysis (GSA) methods assess how the model outputs are influenced by the variation of the model factors over their entire variation range (Homma and Saltelli, 1996; Saltelli et al., 2004). GSA may help modellers in selecting important factors (factor prioritisation), non-influential factors (factor fixing) as well as identifying interactions among factors and performing factors identifiability. More specifically, by means of "factor prioritisation" the model factors that have the largest effect on model outputs are identified. Conversely, the "factor fixing" setting leads to the identification of factors that may be fixed at any given value over their range without changing the output (Saltelli et al., 2004).

In Saltelli (2000) the GSA methods are classified into: (i) global screening methods, e.g. Morris screening method (Morris, 1991; Campolongo et al., 2007); (ii) variance decomposition methods such as Extended Fourier Amplitude Sensitivity Testing (Extended-FAST) (Saltelli et al., 1999); (iii) regression/correlation-based methods such as the standardised regression coefficients (SRCs) method (Saltelli et al., 2008). Due to the high complexity of environmental models, the spread of the GSA applications has been limited due to their high computational cost (Campolongo et al., 2007; Yang, 2011). Therefore, modellers have often been reluctant to use GSA methods instead of local methods (Saltelli and Annoni, 2010). Despite such problems in applying GSA, during the last years modellers have spent considerable time in understanding the potentialities of each GSA method applied to complex models, especially in some areas such as hydrology (e.g., Massmann and Holzmann, 2012; Herman et al., 2013; Zhan et al., 2013).

1.1. Comparison of GSA methods

In Table 1 all the relevant studies on GSA found in literature are summarised. They are discussed below.

Some authors have compared the different GSA methods in order to highlight the potential of each method and the differences of the results.

Tang et al. (2007) compared four sensitivity analysis methods for a watershed model with 18 factors. In particular they applied one local sensitivity analysis method (Parameter Estimation Software, PEST) and three global sensitivity analysis methods (Regional Sensitivity Analysis – RSA; Analysis of Variance – ANOVA and Sobol's method). Convergence for the RSA and Sobol' methods was tested on the basis of the values of the sensitivity indices and on the reproducibility of the results. For the PEST method the authors imposed a maximum value of 30 iterations for each model factor. They found that, in terms of sensitivity classification, the results of the PEST method were significantly different compared to the other three methods due to the local nature of the PEST application. Moreover, among the three global methods, the Sobol' method was considered the most robust in terms of sensitivity rankings, detailed in terms of variance decomposition and easy to interpret.

Yang (2011) compared five different sensitivity analysis methods (Sobol', Morris screening, Linear Regression, RSA and SDP non-parametric regression/smoothing approach) for a hydrological model of the Leaf River watershed with five model factors. On the basis of the Central Limit Theorem the author established the achievement of convergence for the sample size for which no significant change in the coefficient of variance occurred. The author found similar results in terms of parameter ranking for the Sobol', Morris screening and Linear Regression methods and for the SDP non-parametric regression/smoothing approach. The different results obtained with the RSA were attributed to the choice of the filtering criteria.

Sun et al. (2012) compared three sensitivity analysis methods of a hydrological water quality model with 6 model factors: the local method, the Morris screening method and RSA. They concluded that the compared methods should be considered as complementary and not as mutually exclusive alternatives. The peculiar features of each method can assist the modeller in characterising the behaviour of the model studied. In case of a model with a large number of factors, Sun et al. (2012) suggested to use a two-step procedure including first a factors screening step (by using a local method) followed by a global sensitivity analysis step of the important factors identified during the first step.

Neumann (2012) presented a comparison among five sensitivity analysis methods (derivative-based local sensitivity analysis, Morris screening, Standardised Regression Coefficients, Extended-FAST and an entropy-based method) for a model predicting micropollutant degradation in drinking water treatment with 10 model factors. Although the author found the same parameter ranking for the different methods he underlined the poor approximation of 1st order effect indices obtained with the local methods or regression-based methods. Thus, when model non-linearity increases the factors classification can significantly differ when local methods or regression-based methods are applied.

Recently, Cosenza et al. (2013) compared three global sensitivity analysis methods (SRC, Morris screening and Extended FAST) to assess the most relevant processes occurring in membrane bioreactor wastewater treatment systems by using the numerical settings as suggested in literature. Morris screening and Extended-FAST showed low similarity in terms of both the number and type of influential/non-influential factors. The differences were attributed to convergence problems for the Morris screening results. Further, very similar results were obtained between the Extended-FAST and SRC methods despite the fact that SRC was applied outside its range of applicability ($R^2 < 0.7$). Thus, Cosenza et al. (2013) suggested, for the case studied, to use the SRC method (less computationally demanding compared to the Extended FAST method) in case the modeller is only interested in *factor prioritisation*.

1.2. Convergence analysis

Despite the aforementioned reports on the convergence issues for the GSA results, only few studies regarding the assessment of convergence exist in literature (e.g., Benedetti et al., 2011; Yang, 2011; Wang et al., 2013). Benedetti et al. (2011) proposed a method to minimise the computational cost of Monte Carlo based GSA methods in the wastewater treatment modelling field. They focused their attention on two criteria (the model output variability and the stability of the composition of the important factor set as the number of iterations increases) for selecting the minimum number of simulations to be performed. However, they found that by using different criteria the results of the convergence analysis are quite different highlighting that the achievement of convergence is strongly dependent on the model output considered during GSA application. Benedetti et al. (2011) suggested that the number of simulations required to reach convergence is between 40 and 150 times the number of uncertain model factors. Such a result is not in line with the findings of previous works where the maximum value of 50 times (the number of uncertain model factors) was suggested (Benedetti et al., 2010).

Yang (2011) proposed a method to investigate the convergence of the results of different Monte Carlo based GSA methods by using two techniques for monitoring the convergence: the Central Limit Theorem (CLT) and the bootstrap technique. Yang (2011) found, for a simple model characterised by 5 factors, that for each GSA method the bootstrap technique leads to a lower number of simulations required than the CLT. Further, Nossent and Bauwens

Table 1
Summary of relevant studies on GSA found in literature.

Method	Convergence analysis	Number of factors	Number of simulation runs	References	Research field
Sobol'	Yes		36,000	Yang (2011)	
Morris screening		5	12,000		
Linear regression			3000		
RSA			5000		
SDP			512		Hyd
PEST	Yes	18	–	Tang et al. (2007)	
ANOVA			1000		
RSA			10,000		
Sobol'			8192		
OAT	No	6	85	Sun et al. (2012)	
RSA			10,000		
Morris screening			10,000		
Derivative-based LSA	No	10	11	Neumann (2012)	
Morris screening			110		
SRC			1000		
Extended-FAST			10,240		WQ
Kullback–Leibler entropy			11,000		
SRC	Yes	15	600–2250	Benedetti et al. (2011)	
RSA	No	20	10,000	Mannina and Viviani (2010)	
SRC	No	79	800	Cosenza et al. (2013), Mannina and Cosenza (2015)	WWT
Morris screening			800		
Extended-FAST			39,500		
RSA	No	6	10,000	Thorndahl et al. (2008)	
RSA	Yes	6	30,000	Dotto et al. (2012)	
Sobol'	No	9	10,000	Vezzaro and Mikkelsen (2012)	
SRC	No	13	1000	Gamerith et al. (2013)	UD
Morris screening			300		
RSA	No	17	10,000	Mannina and Viviani (2010)	
RSA	No	28	10,000	Freni et al. (2009a,b)	
Extended-FAST	No	45	11,295	Donckels et al. (2014)	
Morris screening	No	90	24,500	Vanuytrecht et al. (2014)	
Extended-FAST	No	39	46,917		
Morris screening	No	25	3700	Moreau et al. (2013)	O
ANOVA	No	6	1875		
Extended-FAST	Yes	47	2049	Wang et al. (2013)	

Where: RSA = Regionalized Sensitivity Analysis; SDP = State-Dependent Parameter; PEST = Parameter Estimation Software; ANOVA = ANalysis Of VAriance; OAT = One-at-A-Time; SRC = Standardised Regression Coefficient; Extended-FAST = Extended Fourier Amplitude Sensitivity Testing; LSA = Local Sensitivity Analysis; Hyd = Hydrology; WQ = Water Quality; WWT = Wastewater Treatment; UD = Urban Drainage; O = Other.

(2012) explored the influence on the convergence of Sobol' sensitivity indices of the use of different expressions for computing the total variance of the model output influences.

Very recently Wang et al. (2013) investigated the influence of the factor sample size on the convergence of sensitivity measures for a crop growth model by using the E-FAST method. Wang et al. (2013) found that the factor sample size has little influence on the factor rankings. Thus the authors suggest using a small sample size in case the modeller is only interested in the factor rankings. On the other hand Wang et al. (2013) found that the value of the sensitivity measure is strongly influenced by the factor variation range, especially for the factors with high influence.

Despite the fact that researchers have tried in several ways to evaluate the convergence of GSA methods, it is still an open issue. Indeed, although all the approaches proposed in literature are promising and useful, no pragmatic rule has yet been defined to a priori suggest and/or determine the number of model simulations that should be run to achieve the convergence of the GSA results (Benedetti et al., 2011).

1.3. GSA in urban drainage

Although GSA offers many advantages compared to local methods, only few applications have been published in the urban drainage modelling field (Thorndahl et al., 2008 Gamerith et al., 2013). Thorndahl et al. (2008) applied the RSA method to a commercial urban drainage model (MOUSE) by using the General Likelihood Uncertainty Estimation (GLUE) methodology with the aim of quantifying the importance of each model factor in the model structure for different model set-ups. Recently, Gamerith et al. (2013) compared two GSA methods for a sewer flow and water quality model: the SRCs and the Morris screening method. In particular, by varying the model factors of the sewer model, Gamerith et al. (2013) demonstrated that both methods identified the same set of important factors. They also found important non-linear behaviour related to the sewer water quality model parameters, limiting the applicability of the SRC method. Vezzaro and Mikkelsen (2012) applied a variance decomposition method combined with the GLUE method in order to identify the major sources of uncertainty

in a stormwater quality model. They combined GSA and GLUE methods to identify the most relevant sources of stormwater model uncertainty. Donckels et al. (2014) applied the Extended-FAST method to an sewer process model for identifying the most influential input factors with regard to sulphide formation in rising mains and the concrete corrosion rate downstream of a rising main. They found that the maximum surface specific sulphide formation rate represented the most important parameter for predicting both sulphide-related processes and concrete corrosion rate.

Moreover, in the urban drainage modelling field, some applications of GSA aimed at uncertainty analysis have been found (Freni et al., 2009a,b; Mannina and Viviani, 2010; Dotto et al., 2012; Vezzaro and Mikkelsen, 2012). Dotto et al. (2012) compared different uncertainty analysis techniques in the urban drainage field using RSA with the GLUE approach. Freni et al. (2009a,b) analysed the parametric and structural uncertainty of an urban drainage integrated model by using RSA with the GLUE methodology. They found that the hydrological model factors have a relevant role in estimating both the quantity discharged in the river water body (RWB) and the quality of the RWB.

This paper presents a comparison of three GSA methods applied to the relatively complex urban drainage stormwater quality model of Mannina and Viviani (2010) in order to provide a discussion on peculiarities, applicability and reliability of the different methods. Attention has been focused on the different responses of the methods in terms of factor prioritisation and factor fixing. In particular, the SRC, Morris screening and Extended-FAST methods have been compared. Moreover, for each of the compared methods, a convergence analysis is carried out aimed at identifying the required number of simulations.

2. Materials and methods

2.1. GSA methods

2.1.1. Standardised regression coefficients – SRC

The SRC method consists of performing a multivariate linear regression between the model outputs and factors obtained by means of Monte Carlo (MC) simulations. For each ith factor (x_i) and for each model output (y) of interest the regression slope (b_i) is standardised according to Eq. (1) and the sensitivity coefficient is evaluated as:

$$SRC(x_i) = \beta_i = b_i \cdot \sigma_{x_i}/\sigma_y \tag{1}$$

where σ_{xi} and σ_y represent, respectively, the ith factor and the model output standard deviation. The β_i value represents a valid measure of sensitivity for the factor x_i when the linear regression coefficient R^2 is greater than 0.7 (Saltelli et al., 2004). The absolute value of β_i represents the order of magnitude of the influence of the ith factor. The sign of β_i represents the positive or negative effect that an increase of the ith factor has on the model output. β_i^2 approximates the variance contribution of the ith model factor to the total variance of the model output. For linear models: $\Sigma\beta_i^2 = 1$.

In order to select model factors the cut-off threshold for the SRC method (CT_{SRC}) has to be defined. The factors with $|\beta_i| > CT_{SRC}$ are considered important factors, while those with $|\beta_i| < CT_{SRC}$ are non-important factors.

The important factors represent those model factors that contribute most to the variance of the model output. Conversely, the non-important factors are those model factors that determine model output variance to a lesser degree.

The SRC method generally requires a number of MC simulations in the order of 500–1000 in the case of random sampling (Neumann, 2012). For Latin Hypercube Sampling (LHS) the

required number of simulations is typically 40–150 times the number of factors (NF) (Benedetti et al., 2011).

2.1.2. Morris screening

The Morris screening method is based on a one-at-a-time (OAT) perturbation of the model factors under investigation (Morris, 1991). The OAT analysis is repeated r times, at randomly sampled locations in factor space, resulting in r Elementary Effects (EEs) for the model output. More specifically, in the Morris screening method the Morris sampling design is applied. Such sampling design relies on a discrete partitioning of the factor space into p levels and a randomized sampling (repeated r times for each factor) inside the factor space (Morris, 1991). For each sample a perturbation (Δ) of the factor value is considered as a multiple of $1/(p-1)$. Regarding the sampling trajectories according to Pujol (2009) a simplex-based design is considered where the trajectories are assumed to be simplexes. The EEs related to each factor and each perturbation are computed by means of the relative difference of the model output with and without perturbation for that factor. According to Campolongo et al. (2007), for each factor the measure of sensitivity is summarised by the absolute mean (μ^*) and the standard deviation (σ) of the cumulative distribution function of the r EEs. More specifically, in order to avoid the problem of the effects of opposite signs of the EEs, μ^* is computed considering the cumulative distribution of the absolute value of the r EEs. In particular, μ^* and σ represent, respectively, the measure of the importance of the factor and to which degree it caused non-linearity or interacts with other factors. More specifically, for the ith factor a high value of μ^* shows that model output variation is due to the variation of this factor. Further, a high value of σ for the ith factor means that the model output variation is influenced by non-linearity or interactions. The line corresponding to $\mu_i^* = 2*SEMi$, where SEMi represents the standard error of the mean, is used for establishing the type of effect of factors (Morris, 1991; Ruano et al., 2011). SEMi is equal to $\sigma_i*r^{-1/2}$, where r (number of repetitions) is typically between 10 and 50 (Campolongo et al., 2007). Factors which lie outside the wedge formed by the line corresponding to the established CT_{MORRIS} for μ^* and the line $\mu_i^* = 2*SEM$ have a linear effect on the model outputs. Conversely, the factors which lie inside the area formed by the CT_{MORRIS} for μ^* and the line $\mu_i^* = 2*SEM$, have a non-linear effect. In order to select model factors the cut-off threshold for the Morris-screening (CT_{MORRIS}) method has to be defined. Thus, according to the Morris screening method the important factors are factors for which $\mu^* > CT_{MORRIS}$, the interacting factors have $\mu^* > CT_{MORRIS}$ and $\sigma > \mu^**\sqrt{r}/2$ while the non-influential factors have $\mu* < CT_{MORRIS}$.

In contrast to the SRC method, the Morris screening method also allows identification of interacting and non-influential factors.

Regarding the number of model runs, according to Morris (1991), $r*(NF+1)$ model simulations are required (Campolongo et al., 2007). Typical numbers are $r = 10$–20 (Campolongo et al., 2007). Recently, Ruano et al. (2011, 2012) found that the optimal number of r is between 60 and 70.

2.1.3. Extended-FAST

The E-FAST method belongs to the variance decomposition methods. The application of this method provides, for each factor, two sensitivity indices: the first-order effect index (S_i) and the total effect index (S_{Ti}). S_i is defined as follows:

$$S_i = \frac{Var_{x_i}(E_{x_{-i}}(Y|x_i))}{Var(Y)} \tag{2}$$

where E represents the expectancy operator and Var is the variance operator, the subscripts indicate that the operation is either applied "over the ith factor" $x_{i,}$ or "over all factors except the i-th

factor" x_{-i}. S_i measures how the ith factor contributes to the total variance of the model output, without taking into account the interactions among factors. Thus, the higher the S_i is, the higher is the influence of the factor in terms of factor prioritisation. The total effect index S_{Ti} is used to determine factor interactions: the difference between S_{Ti} and S_i represents the degree to which the ith factor is involved in interactions. A low S_{Ti} value indicates that the ith factor may be fixed anywhere within its range of uncertainty without reducing the variance of the model outputs. In order to select model factors the cut-off thresholds for the Extended-FAST method ($CT_{E-FAST1}$ and $CT_{E-FAST2}$) have to be defined. For the E-FAST method important factors are characterised by a $S_i >$ $CT_{E-FAST1}$, the interacting factors by $S_{Ti} - S_i > CT_{E-FAST2}$, the influential factors by $S_i > CT_{E-FAST1}$ or $S_{Ti} - S_i > CT_{E-FAST2}$ and, finally, the non-influential factors require $S_i < CT_{E-FAST1}$ and $S_{Ti} - S_i < CT_{E-FAST2}$.

Regarding the number of simulations, NF $*$ MC simulations are required for the E-FAST method application, where NF represents the number of the model factors and MC is between 500 and 1000 (Saltelli et al., 2005).

3. Terminology

In this section a definition of the sensitivity classification employed in the three analysed sensitivity methods is provided. Indeed, in the literature, as far as the authors are aware, a complete, clear and generally accepted definition is lacking. The main reason is likely due to the fact that the sensitivity methods have been developed in different periods and disciplines, and the authors of each method generally do not refer to the other methods. Further, since a comprehensive comparison is lacking so far, the terminology used in the different methods has not been standardised.

The objective of this section is to suggest a common terminology on the basis of the definitions drawn from the literature (among others, Saltelli, 2000; Campolongo et al., 2007; Pujol, 2009).

The cut-off threshold for the SRC method (CT_{SRC}) classifies factors in the following way (Fig. 1a):

1. Important factors: if sensitivity > CT_{SRC}.
2. Non-important factors: if sensitivity < CT_{SRC}.

The important factors represent those model factors that have a high sensitivity coefficient and where, therefore, the modeller should pay more attention. Conversely, the non-important factors are those model factors characterised by a low sensitivity coefficient. In the case of linear models, they can be fixed anywhere in their variation range. In the case of non-linear models ($R^2 < 0.7$), however, some of the non-important factors cannot be fixed due to interactions with other factors (see below) that cannot be quantified by using the SRC method.

Morris screening allows distinguishing between three different types of factors with respect to the mean and the standard deviation of the sensitivity (Fig. 1b):

1. Important factors: if mean sensitivity > CT_{MORRIS}.
2. Interacting factors: if mean sensitivity > CT_{MORRIS} and the standard deviation of the sensitivity is above a specified cone line.
3. Non-influential factors: if mean sensitivity < CT_{MORRIS}.

In particular, the Morris screening method (Morris, 1991) as modified by Campolongo et al. (2007) basically defines a cone whose edges are set by a CT_{MORRIS} and an oblique line that is a statistical function of the mean and standard deviation of the sensitivity (Fig. 1b) (quantitative characteristics are given below).

The Extended-FAST distinguishes three classes of factors on the basis of two CT ($CT_{E-FAST1}$ and $CT_{E-FAST2}$) (Fig. 1c):

1. Important factors: if sensitivity > $CT_{E-FAST1}$.
2. Interacting factors: if interaction > $CT_{E-FAST2}$.
3. Non-influential factors: if sensitivity < $CT_{E-FAST1}$ and interaction < $CT_{E-FAST2}$.

Non-influential factors which can be identified by both the Morris-screening as well as the Extended-FAST method can be fixed anywhere within their variation range without changing the model output variance. With respect to this, it is important to stress once more that non-important does not imply non-influential.

4. Convergence analysis

In order to better understand the potential in terms of computational cost of each of the three GSA methods applied, a convergence analysis has been performed for each method. Convergence analysis was performed by analysing the results of the GSA methods application using different k number of MC simulations (n_{MC}). Convergence was first analysed visually by examining the stability of the sensitivity index value with increasing n_{MC}. Further, a quantitative convergence analysis was performed by evaluating, for each model output j, the S_{SCj} as the sum of the sensitivity indices (SC_i) of all factors normalised with respect to the number of factors (NF) (Eq. (3)) and analysing its variability with increasing n_{MC}.

$$S_{SCj} = \frac{\sum_{i=1}^{NF} SC_{i,j}}{NF} \tag{3}$$

For the jth model output the variability of S_{SCj} was expressed as the percentage of change of S_{SCj} from n_{MCk-1} to n_{MCk} (Eq. (4))

$$\text{Variability } y = \left[\frac{\left(\sum_{i=1}^{NF} SC_{i,j} \right)_{n_{MCk-1}} - \left(\sum_{i=1}^{NF} SC_{i,j} \right)_{n_{MCk}}}{NF} \right] \cdot 100 \tag{4}$$

The convergence was considered achieved with n_{MCk} if the variability stays within the band of ± the precision threshold selected, with $n_{MCk}, n_{MCk+1}, \ldots, n_{MCn}$ for all model outputs.

5. Case study

5.1. Model description

The urban stormwater quality model used in this study is able to simulate the main phenomena that take place both in the catchment and in the sewer network during both dry and wet weather periods (Mannina and Viviani, 2010). It is divided into two connected modules: a flow module that calculates the hydrographs at the inlet (surface runoff) and at the outlet (sewer flow) of the sewer network, and a solids transport module, that calculates the pollutographs at the outlet of the sewer network for different pollutants (TSS, BOD and COD). The flow module consists of a hydrological and a hydraulic component. It evaluates the net rainfall by applying a loss function (initial and continuous) to the measured rain intensity. From the net rainfall, the model simulates the rainfall–runoff process and the flow propagation with a cascade of two reservoirs in series and a linear channel.

The solids transport module reproduces the accumulation and propagation of solids in the catchment and in the sewer network. The main simulated phenomena are build-up and wash-off of pollutants from catchment surfaces and sedimentation and re-suspension of pollutants in sewers (Bertrand-Krajewski et al., 1993). To simulate the build-up of pollutants on the catchment
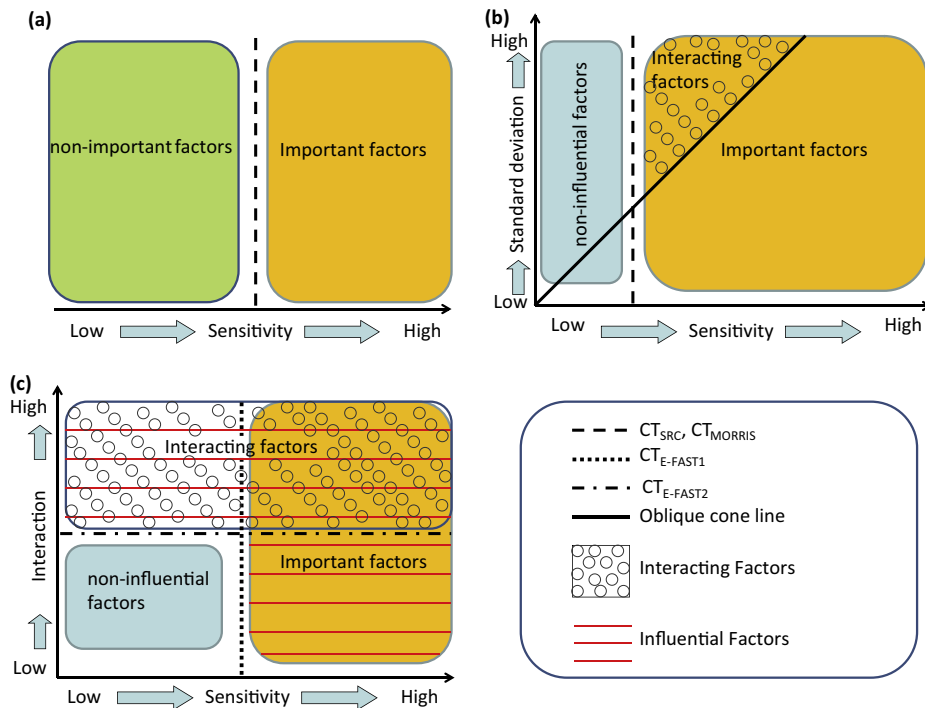
Fig. 1. Schematic overview of the suggested terminology for differentiating factors according different GSA methods: SRC (a), Morris screening (b) and Extended-FAST (c).

surfaces an exponential function was adopted (Alley and Smith, 1981). The solids wash-off caused by overland flow during a storm event was simulated with the formulation proposed by Jewell and Adrian (1978). The solids deposition in the sewers during dry weather is evaluated by adopting an exponential law. Two classes of particles are considered: fine particles and coarse particles. The fine particles are mainly transported as suspended load whereas the coarse particles are mainly transported as bed load (sediment transport) and are in suspension only at high flows. Particular care has been taken with regard to sediment transformation in sewers, considering their cohesive-like behaviour due to organic substances and to the physical–chemical changes during sewer transport (Crabtree, 1989; Ristenpart, 1995). In particular, the transport equation proposed by Parchure and Mehta (1985) is coupled to the bed sediment structure hypothesised by Skipworth et al. (1999) to simulate the sediment erosion rate. The pollutographs at the outlet of the sewer system have been evaluated by assuming the complex catchment sewer network to act as a reservoir and by considering an adapted version of Wiuff's model (Bertrand-Krajewski et al., 1993). The quality model focuses on describing TSS. BOD and COD dynamics are evaluated as a ratio of the TSS concentration. More specifically, during wet weather, a linear relationship between TSS and the COD and BOD concentration is assumed. The Montelepre experimental catchment is located near Palermo in the north-western part of Sicily, Italy. The total drained area is 70 ha with an impermeable area of 40 ha. The buildings in the area are mainly for residential use and minor service sector businesses; the number of inhabitant-equivalents is about 7000. The Montelepre sewer pipes are circular and egg-shaped with maximum dimensions of 100×150 cm. The sewer system is characterised by an average dry weather flow equal to 12.5 l/s (water supply: 195 l/capita/d), and an average dry weather BOD concentration of 225 mg/l.

Discharge has been estimated from water depth measured by an ultrasonic probe placed in the main channel. A refrigerated automatic sampler with 24 bottles, each with one litre volume,

was used for sampling of BOD, COD and TSS. The field campaign was carried out by DICA Palermo University (Candela et al., 2012).

For further details about the model and the case study the reader is referred to literature (Candela et al., 2012; Mannina and Viviani, 2009a, 2010; Freni et al., 2010c).

5.2. GSA methods application

The model is run with a long input time series to simulate both dry and wet weather periods. The simulation covered a period of 1 year during which 36 events were recorded, the rainfall depth was 802 mm and the average rainfall intensity was 8.54 mm/h. Seven model outputs have been considered for the whole simulated period: the maximum sewer flow rate (Q_{MAX}), the total sewer flow volume (V_{TOT}), the maximum sewer TSS concentration ($C_{MAX,TSS}$), the maximum BOD concentration ($C_{MAX,BOD}$), the sewer TSS load ($L_{TOT,TSS}$), the average sewer TSS concentration ($C_{AVERAGE,TSS}$) and the average sewer BOD concentration ($C_{AVERAGE,BOD}$). The seventeen model factors reported in Table 2 have been considered. The model factor ranges have been established by considering previous model applications to different case studies (Freni et al., 2010c; Mannina and Viviani, 2009a, 2010; Mannina et al., 2012). Quantity and quality model factors were changed simultaneously for each MC run. It is important to stress that the model structure is such that changing the quality model factors (Nos. 6–17, Table 2) has no effect on the quantity model outputs (Q_{MAX} and V_{TOT}), i.e. the quantity model outputs are insensitive to these factors. For each model factor reported in Table 2 a uniform distribution has been considered. Such a choice was driven by the fact that the prior information on the factors' behaviour was insufficient. As pointed out by Freni and Mannina (2010a; Mannina & Viviani, 2009b), a uniform distribution of model factors is preferred whenever relevant prior factor information is unavailable, as assuming a non-uniform shape may lead to wrong estimations of uncertainty in modelling results. For the SRC the sampling is carried out according to the LHS method.

Table 2
Model factors number, symbol, definition, units and variation range.

No.	Symbol	Definition	Unit	Min	Max
1	λ	Channel constant	min	0.04	6
2	W_0	Initial hydrological losses	mm	0.22	1.5
3	Φ	Catchment runoff coefficient	–	0.25	0.57
4	K_1	Catchment reservoir constant	min	2	7
5	K_2	Sewer reservoir constant	min	2	7
6	Accu	Build-up coefficient	kg ha^{-1} d^{-1}	0.01	40
7	Disp	Decay coefficient	d^{-1}	0.01	0.5
8	Arra	Wash-off coefficient	mm^{-Wh} h$^{(Wh-1)}$	0.01	2
9	W_h	Wash-off factor	–	0.1	3
10	K_{dep}	Sewer sediment accumulation coefficient	h^{-1}	0.001	2
11	h_{max}	Maximum sewer sediment height	m	0.01	0.1
12	d'	Depth of the weak layer	mm	0.01	0.4
13	b	Erosional resistance exponent	min	0.001	1
14	τ_{cu}	Yield strength at uniform layer	N m^{-2}	1.1	10
15	M	Erosion coefficient	g h^{-1}	1	200,000
16	K_{susp}	Sewer suspension delay	h	0.001	0.9
17	K_{bed}	Sewer bed transport delay	h	0.001	0.9

The GSA methods have been applied by using the package "sensitivity" developed by Pujol (2007) in the R environment (R Development Core Team, 2007).

5.3. GSA thresholds and criteria for comparison

For the SRC and Morris screening methods a CT (CT$_{SRC}$ and CT$_{MORRIS}$) equal to 0.1 has been established while a CT$_{E-FAST1}$ equal to 0.01 was used for E-FAST. This latter threshold has been established considering the fact that for linear models β_i^2 is equal to S_i (Saltelli, 2000) and therefore the value of CT$_{SRC}$ of 0.1 corresponds to a CT$_{E-FAST1}$ value of 0.01 for S_i in E-FAST.

For the Extended-FAST method a CT$_{E-FAST2}$ of 0.1 for the value of the interaction (i.e. $S_{Ti} - S_i$) was chosen. For each method a rank of importance has been determined for each model factor according to the factors prioritisation setting. For factors prioritisation the comparison between methods has been performed by making a comparison between the following indices (Campolongo et al., 2007; Saltelli et al., 2008):

- β_i^2 and S_i for the comparison between SRC and E-FAST method results.
- β_i^2 and μ^* for the comparison between SRC and Morris-screening method results.
- μ^* and S_i for the comparison between Morris-screening and E-FAST method results.
- Importance ranking obtained by applying each method.
- For factor fixing the following indices have been considered (Campolongo et al., 2007; Saltelli et al., 2008).
- μ^* versus S_{Ti} and σ versus S_{Ti} for the comparison between Morris screening and E-FAST.

The comparison among methods has been graphically presented by using the Venn diagram. This diagram is presented as a circle diagram for each method, the overlapping area among the circles contains all equal results among the methods.

For convergence analysis several precision thresholds of the variability have been established. Specifically, the values ranging from ±0.5% to ±3.5% have been selected.

6. Results and discussion

Results of the convergence analysis are presented in Figs. 2–4. Results for each method and the water quantity model outputs (Q_{MAX} and V_{TOT}) are summarised in Table 1A, while in Table 2A the results are reported for the quality model outputs ($L_{TOT,TSS}$,

$C_{MAX,TSS}$, $C_{MAX,BOD}$, $C_{AVERAGE,TSS}$ and $C_{AVERAGE,BOD}$). In the following sections the results of each method are presented and discussed in detail. For sake of conciseness, this discussion was limited to two of the seven investigated model outputs: Q_{MAX} (as water quantity model output) and $C_{MAX,BOD}$ (as water quality model output).

6.1. Convergence analysis

The convergence analysis has been performed by increasing the number of simulations until there was no significant change of the sensitivity measure. More precisely, the SRC method has been applied by increasing the number of simulations up to 10,000; for the Morris screening method the repetition number, r, has been increased up to 2500 resulting in 45,000 simulations; for the Extended-FAST method the number of MC has been increased up to 7000 with a maximum value of the simulation number of 119,000.

By analysing Fig. 2 one can observe that the sensitivity of factors λ, W_0, Φ, K_1 and K_2 (Fig. 2a and b), all related to the water quantity model outputs, achieve convergence with a reduced number of runs than water quality model outputs. Such a result is due to the fact that the quantity model has less factors than the quality model and it is characterised by a linear and additive behaviour, as discussed before.

For the water quality model outputs the sensitivity measure of the factors shows a high variability at low number of simulations, especially for the factors with high influence (Fig. 2c–g). Indeed, for example at a number of simulations below 3000 a high variability of the sensitivity measure of Accu has been found for the $C_{AVERAGE,BOD}$.

As shown in Fig. 3a and b the convergence analysis of the Morris screening method provides different results compared with the SRC method. Indeed, despite the above discussed simplicity of the quantity model, the application of the Morris screening method leads to a greater variability of the sensitivity measure (μ^*) with increasing number of simulations. Such instability of the Morris screening results, even at a high number of simulations (36,000), has also been observed by Yang (2011) and Cosenza et al. (2013). Indeed, Yang (2011) found that the instability of the Morris sensitivity measure for a simple model with 5 factors lasted until 2000 runs. This repetition number is quite different to the value of 10–50 typically chosen in literature (Campolongo et al., 2007). The same higher variability of the Morris sensitivity measure has also been found for the quality model outputs (Fig. 3c–g).

As shown in Fig. 4 very similar results to the SRC method, both in terms of quantity and quality model factors, have been obtained
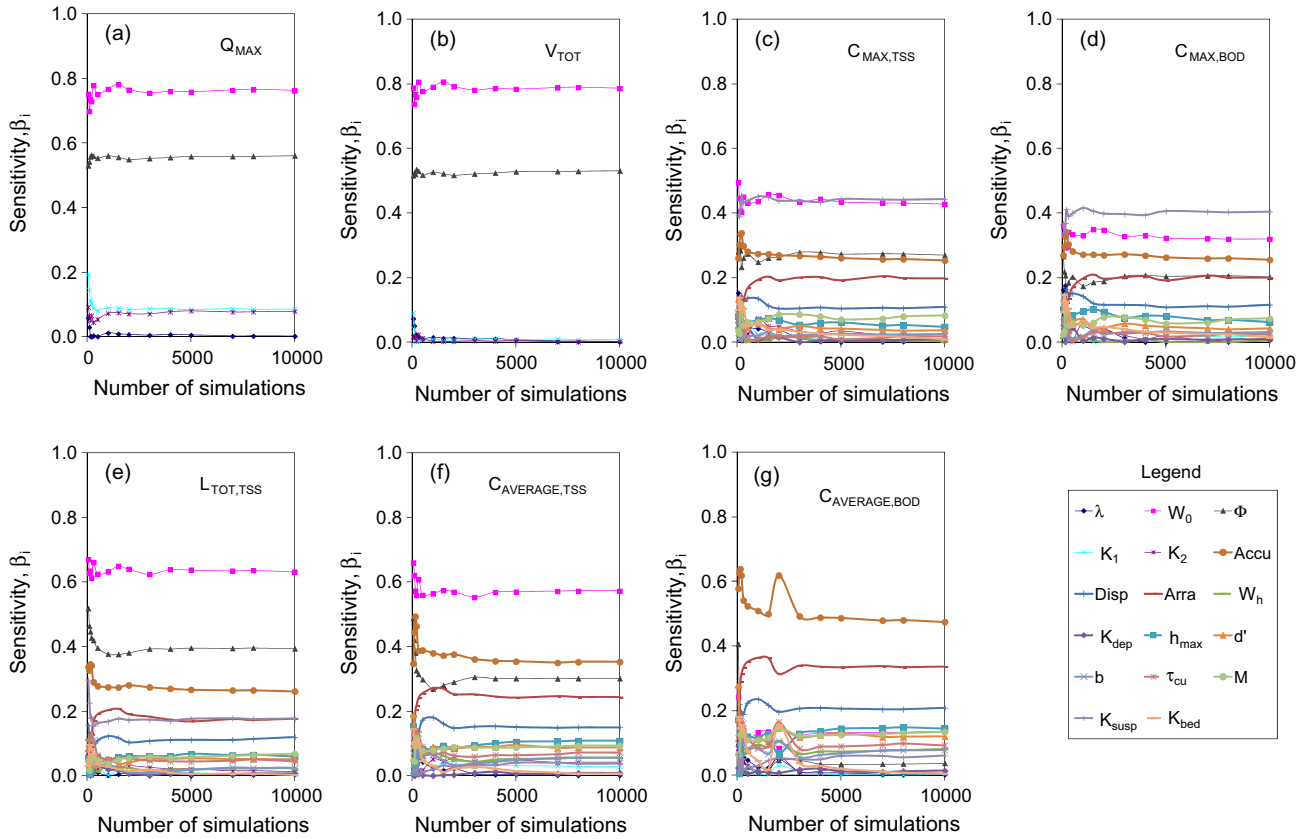
Fig. 2. Convergence analysis of the SRC method for each model output and factor with increasing sample size (expressed as number of simulations).
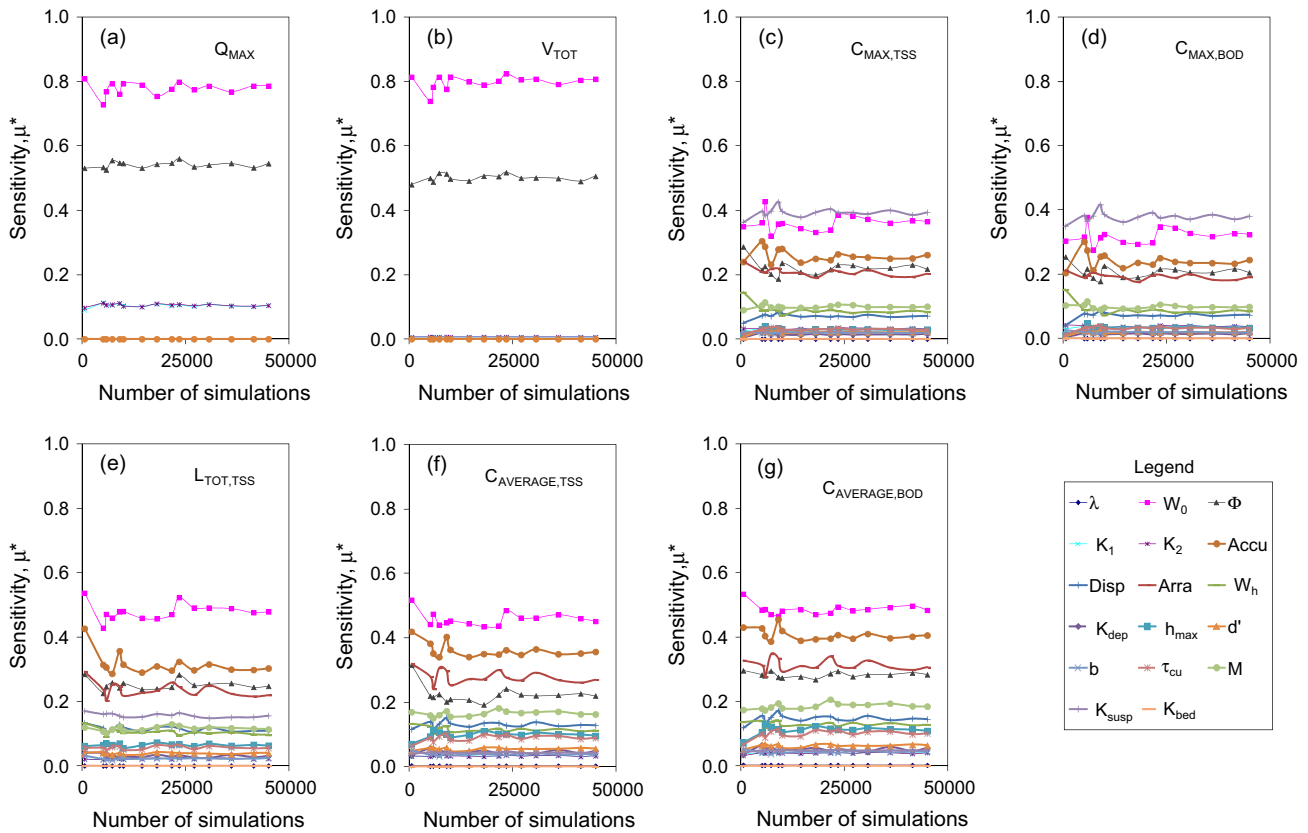

Fig. 3. Convergence analysis of the Morris screening method for each model output and factor with increasing sample size (expressed as number of simulations).
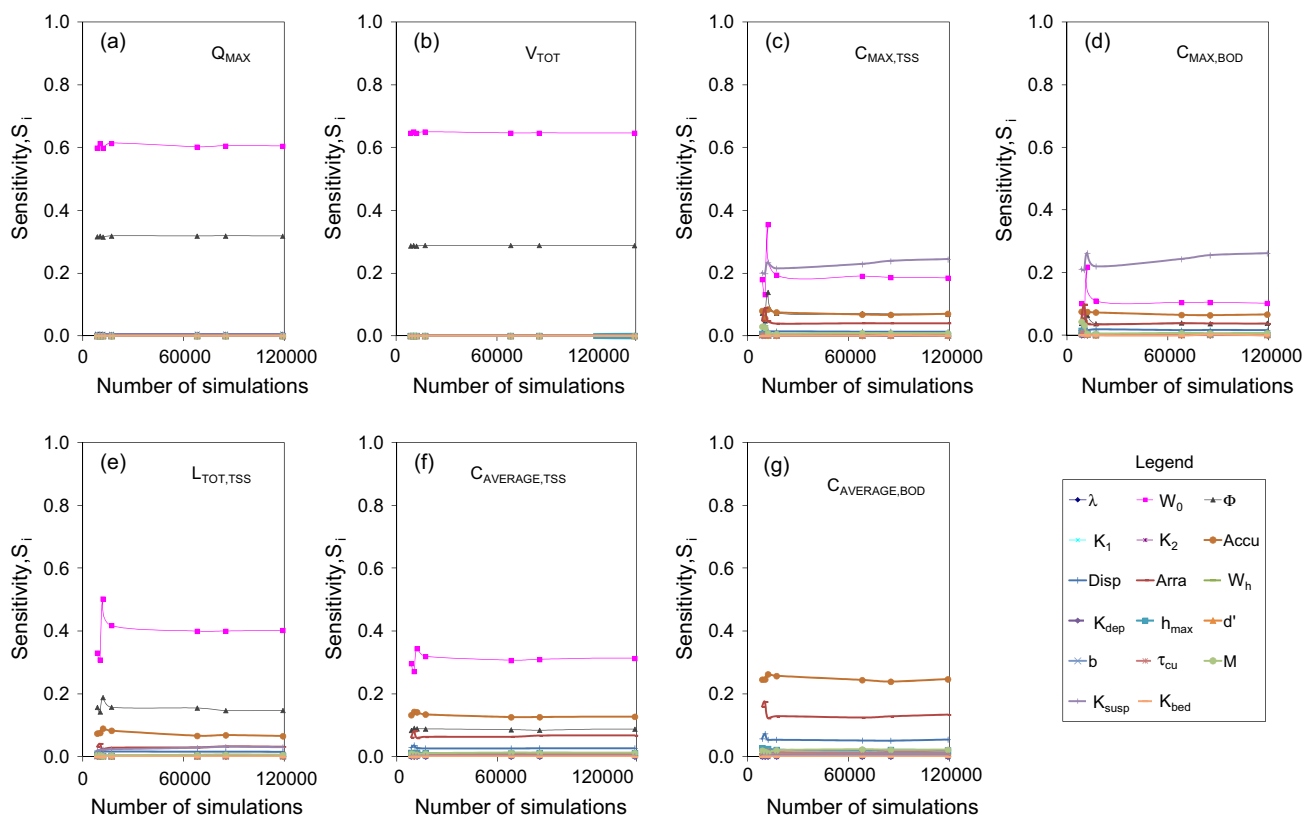
Fig. 4. Convergence analysis of the Extended-FAST method for each model output and factor with increasing sample size (expressed as number of simulations).

by applying the Extended-FAST method. Noteworthy, the factors with the highest influence showed the lowest convergence rate of their sensitivity indices. Wang et al. (2013) found similar results by applying the Extended-FAST method to a crop growth model. Thus, in case the modeller's aim in performing sensitivity analysis is the factor fixing, he/she can reduce the number of required simulations to test the convergence. Indeed, even if the convergence of model factors with the highest influence has not been achieved, no errors would be performed in fixing non influential model factors value.

In Fig. 5 results of the quantitative convergence analysis are reported. In particular, for each method and model output the variability, computed according to Eq. (4), is reported in Fig. 5 *versus* the number of simulations. The variability precision threshold value of ±3.5 has been reported in Fig. 5. By analysing Fig. 5a one can observe that the number of simulations required for achieving convergence depends on the model output and on the established precision threshold value. For example with a precision threshold value of ±1 convergence for Q_{MAX} and $C_{MAX,BOD}$ is achieved at 4000 and 8000 number of simulations, respectively. Obviously, by increasing the threshold value from ±1% up to ±3.5% the number of simulations required for convergence decreases.

The results shown in Fig. 5b confirm the findings of the Morris screening application discussed above. In terms of analysing the SC a higher variability was found for the Morris screening application than for SRC. Indeed, for the number of Monte Carlo simulations run, only starting from ±3% the variability trend of all model outputs stays within the band of the precision threshold. Thus, convergence of all model outputs is only achieved within 45,000 simulations performed with the Morris screening method, when a precision threshold of at least ±3% is specified. For Extended-FAST (Fig. 5c) the precision threshold of ±3.5% is required before acceptable convergence of all model outputs can be reached, within the overall number of Monte Carlo simulations run.

For each method, the results related to the model output having the lowest convergence rate were considered to select the number of simulations required for convergence. Moreover, the lowest value of the precision threshold for which the variability of all model outputs stays within the precision band for all methods was considered (namely, ±3.5%). Results of the quantitative convergence analysis for each method and precision threshold value are summarised in Fig. 6. One can observe that the convergence was achieved at 4000, 36,000 and 85,000 number of simulations respectively for SRC, Morris screening and Extended-FAST.

The proposed convergence method allows gaining insight on the method (or combination of methods) that can give convergence to a stable assessment of the influential and non-influential factors using the smallest number of model simulations.

6.2. SRC results

For the SRC method the results related to 4000 simulations, required to achieve convergence with the ±3.5% precision band, were considered. The R^2 values obtained by applying the SRC method were larger than 0.7 for Q_{MAX}, V_{TOT} and $L_{TOT,TSS}$ and smaller than 0.7 for $C_{MAX,TSS}$, $C_{MAX,BOD}$, $C_{AVERAGE,TSS}$ and $C_{AVERAGE,BOD}$ (see Tables 1A–2A). The low R^2 of these model outputs is probably caused by non-linearity due to the high complexity of the quality model (Freni et al., 2009a,b; Dotto et al., 2010). Indeed, several processes of the quality model control the TSS concentration (i.e., solids build-up and wash-off, sewer sediment accumulation, erosion and transport) with a variety of aspects (e.g. climate variables, land use or surface features) (Freni et al., 2009a,b; Dotto et al., 2010; Mannina and Viviani, 2010). The low R^2 for BOD follows from the functional relationship that exists between TSS and BOD in the model (see above). Consequently, the BOD concentration follows the same non-linear behaviour as TSS.
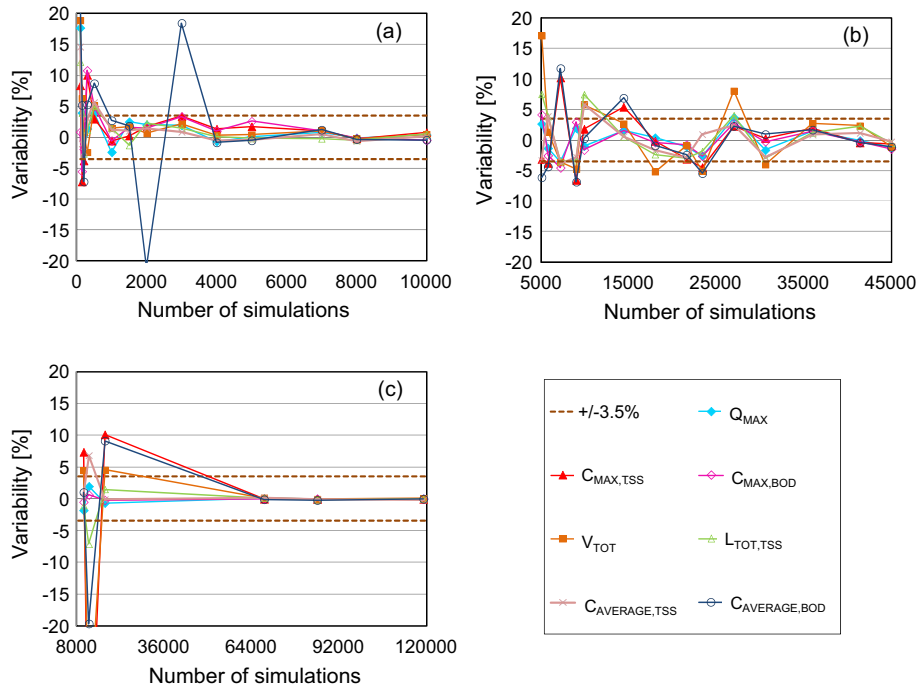
Fig. 5. Quantitative convergence analysis with increasing sample size (expressed as number of simulations) of different GSA methods: SRC (a), Morris screening (b) and Extended-FAST(c).
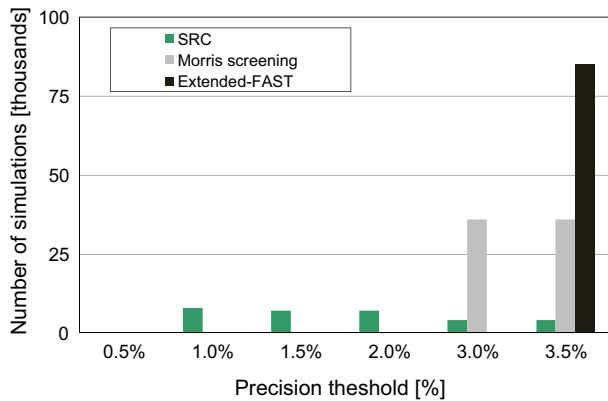

Fig. 6. Results of the quantitative convergence analysis for different precision thresholds (%).

Overall, the model factors λ, K_1, K_2, W_h, K_{dep}, b, τ_{cu} and K_{bed} were classified as being non-important for all model outputs (see Tables 1A–2A). For the quantity model outputs for which a high linearity was found these non-important model factors may be fixed anywhere within their range of uncertainty.

In Fig. 7 results related to Q_{MAX} (Fig. 7a) and $C_{MAX,BOD}$ (Fig. 7b) are presented. The effect of the quality factors on quantity output variables does not require investigation because they cannot have any effect given the model structure. By analysing Fig. 7a it is evident that only two model factors are important for Q_{MAX}: W_0 and Φ. This result highlights the strong importance of the hydrological losses for the quantity model. Indeed, W_0 and Φ account respectively for the losses in small ponds and in infiltration. In particular, a negative and a positive effect on Q_{MAX} was found respectively for W_0 ($\beta_i = -0.767$) and Φ ($\beta_i = +0.560$) (see Table 1A). Such results are in agreement with the physical meaning of these two factors: (i) increasing W_0 leads to a global reduction of Q_{MAX}, (ii) the Q_{MAX} value is directly proportional to Φ. For $C_{MAX,BOD}$ (Fig. 7b) six of the seventeen model factors were classified as being important. Among these six factors, W_0 and K_{susp} had the highest influence on $C_{MAX,BOD}$. More specifically, both W_0 and K_{susp} showed a negative effect on $C_{MAX,BOD}$ (see Table 2A). These results are in agreement with the physical meaning of the model factors. Indeed,
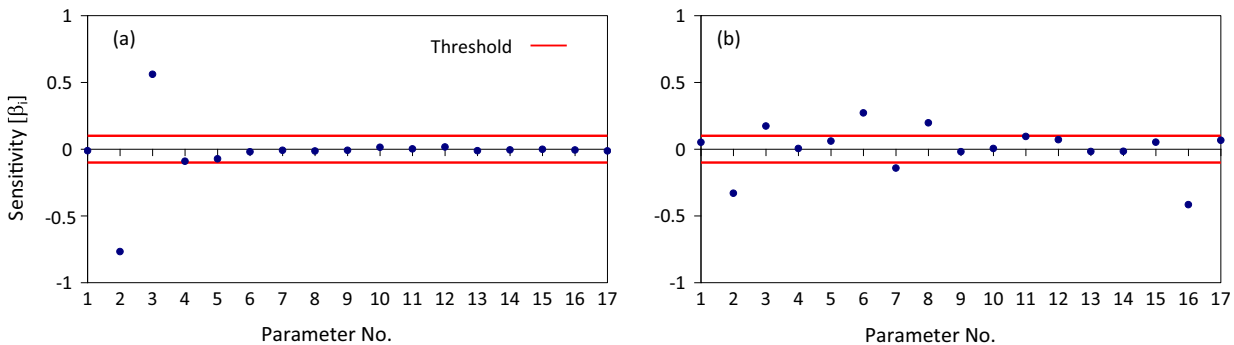

Fig. 7. Results of the SRC application related to Q_{MAX} (a) and $C_{MAX,BOD}$ (b); the lines represent the threshold value selected for β_i. Parameter numbers can be found in Table 2.

when increasing W_0 the wash-off effect decreases, thus reducing $C_{MAX,BOD}$. Moreover, increasing W_0 leads to a reduction of the erosion effect of sewer sediments, thus reducing $C_{MAX,BOD}$. An increase of K_{susp} increases the sewer flow storage and thus the $C_{MAX,BOD}$ decreases due to a dilution effect. The high influence of the model factor Accu confirms that the quality model outputs are strongly influenced by the solids accumulation during the dry weather period. This emphasizes the importance of having field data and detailed information on the catchment's land use because it influences the quantity of the solids that accumulates in the catchment. Despite this, the model factors h_{max} and d' were found to be nearly as important with their $|\beta_i|$ value close to 0.1. This confirms (as also found for all model outputs, see Table 2A) the importance of the sewer sediments for the pollutant load assessment (see, among others, Ashley et al., 2000; Banasiak et al., 2005). Moreover, the result related to the model factors h_{max} and d' confirms the need for accurate modelling of the sediments erosion process by considering the cohesive-like behaviour of sewer sediments (Skipworth et al., 1999).

6.3. Morris screening results

For the Morris screening the results related to $r = 2000$ random replicates (36,000 model simulations), required to achieve convergence with the ±3.5% precision band, were considered. Globally, the model factors λ, K_1, K_2, K_{dep}, h_{max}, d', b, τ_{cu} and K_{bed} were classified as non-influential for each model output according to the value of μ^* and σ. All other model factors had high μ^* value and were considered important (see Tables 1A–2A) for all model outputs. Overall, among the model factors with μ^* greater than 0.1 a higher non-linear effect was found for the quality model outputs than for the quantity model outputs (see Tables 1A–2A).

Thus, by applying Morris screening more than 50% of the model factors were found to be non-influential in terms of μ^* value and σ value (factor fixing).

In Fig. 8 results related to Q_{MAX} (Fig. 8a) and $C_{MAX,BOD}$ (Fig. 8b) are presented. For Q_{MAX} (and for all quantity model outputs, see Tables 1A–2A) the effect of the quality model factors (from 6 to 17) is nil due to the model structure. For Q_{MAX} the model factors W_0 and Φ are, as for SRC, the most important, in terms of μ^* value. Both W_0 and Φ, having a low value of σ, have a linear effect on Q_{MAX} (see Table 1A). Indeed, they both lie outside the wedge formed between the threshold line and the line $\mu^*_i = 2 * SEM_i$.

Regarding $C_{MAX,BOD}$ (Fig. 8b) a higher number of model factors were found to be influential than for Q_{MAX} because the model factors numbered 6–17 (quality model factors) do not have any effect on Q_{MAX}. For $C_{MAX,BOD}$ (Fig. 8b) model factors W_0, Φ, Accu, Arra, W_h, M and K_{susp} were influential. While these influential model factors all lie outside the wedge formed between the threshold line and

the line $\mu^*_i = 2 * SEM_i$, a higher non-linear effect is shown than for Q_{MAX} (see σ values on Table 2A). The physical interpretation of the important model factors is the same as discussed in the previous paragraph.

6.4. E-FAST results

For the E-FAST method the results related to the 85,000 model runs (corresponding to 5000 simulations for each model factor), required to achieve convergence with the ±3.5% precision band, were considered. By applying E-FAST the model factors W_0, Φ, Accu, Disp, Arra, h_{max}, d', τ_{cu}, M and K_{susp} were found to be important at least for one model output in terms of S_i value. Among these important model factors the parameters W_0, Φ, Accu, Disp and Arra were also found to be interacting in terms of normalised interaction, computed with respect to the maximum interaction value for each output variable. Moreover, factor W_h was found to be interacting for $C_{MAX,TSS}$, $C_{MAX,BOD}$ and $C_{AVERAGE,BOD}$ (see Tables 1A–2A). Model factors λ, K_1, K_2, K_{dep}, b and K_{bed} were found to be non-influential for each model output according to the S_i and the normalised $S_{Ti} - S_i$ values (see Tables 1A–2A). Overall 11 of the 17 model factors were found to be influential.

In Fig. 9 results related to Q_{MAX} (Fig. 9a) and $C_{MAX,BOD}$ (Fig. 9b) are shown. Again, the most important model factors for Q_{MAX} (Fig. 9a) are W_0 and Φ which account for 60% and 32% of the variance (see the S_i value for these model factors in Table 1A). As shown by the dark grey bars on Fig. 9a, the interaction among model factors is negligible for Q_{MAX}. Indeed, the model under study is characterised by an additive behaviour for Q_{MAX}. This characteristic is also demonstrated by the fact that the sum of S_i is close to 1 (see Table 1A). Consequently, the S_{Ti} values do not differ significantly from the S_i values (see Table 1A).

For $C_{MAX,BOD}$ (Fig. 9b) the number of important model factors is higher than for Q_{MAX} due to the fact that the quality model factors do not have any effect on the quantity model outputs. The model factors W_0, Φ, Accu, Disp, Arra, h_{max}, M and K_{susp} were found to be important in terms of S_i values. However, a high interaction is found as demonstrated by the sum of the S_{Ti} values and by the higher difference between S_{Ti} and S_i (see Table 2A). The highest interaction contribution to the total variance was found for model factors W_0 and K_{susp}. The interaction of model factors W_0 and K_{susp} with all model factors contributes to the variance of $C_{MAX,BOD}$ by respectively 23% and 24% (see Table 2A). Indeed, the most influential model factor in terms of S_i value accounts for only 21% of the total variance of $C_{MAX,BOD}$ demonstrating that the highest contribution is provided by the interaction among model factors. The higher interactions for the quality model factors is likely due to two aspects: the higher variation that generally comes with the quality processes compared to the quantity processes (see Freni
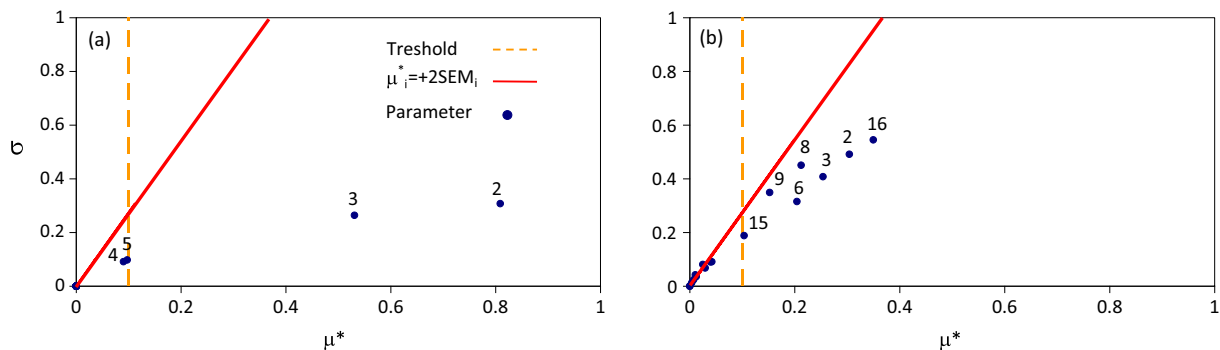

Fig. 8. Results of the Morris screening application related to Q_{MAX} (a) and $C_{MAX,BOD}$ (b); the labels indicate the number of the model factor according to Table 2. Parameter numbers can be found in Table 2.
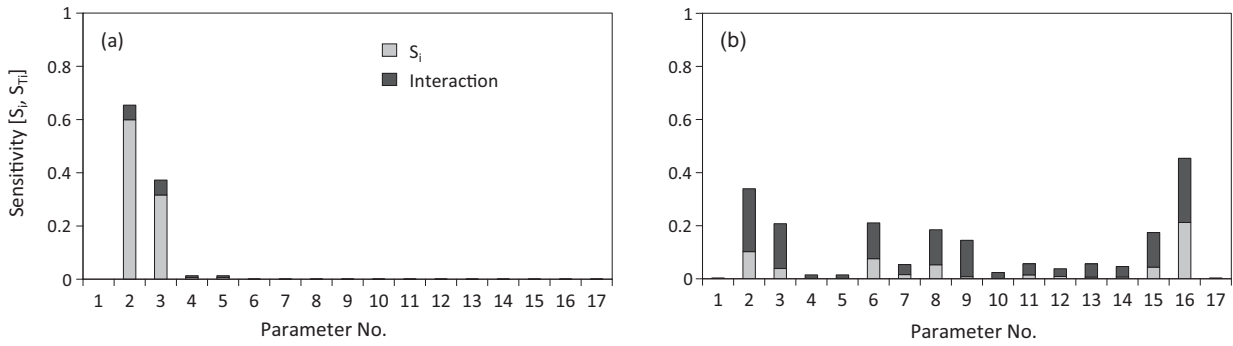
Fig. 9. Results of Extended-FAST application related to Q_{MAX} (a) and $C_{MAX,BOD}$ (b). Parameter numbers can be found in Table 2.

and Mannina, 2010b) and the higher number of model factors considered compared to the number of factors considered for the quantity output variables (namely, 5 and 17 model factors for the quantity and quality modelling, respectively). The factor W_h was found to be interacting a lot, contributing with 14% to the total variance of $C_{MAX,BOD}$ via the interaction with the other model factors.

Important to stress is that only by means of the Extended-FAST method a numerical quantification of the interactions among model factors is possible.

6.5. Comparison of the methods

In Table 3 the results related to the comparison among the three methods are summarised for the two model outputs discussed here (Q_{MAX} and $C_{MAX,BOD}$). Similar results were obtained for the other quantity and quality model outputs analysed (see Tables 1A and 2A, respectively). Conveniently, the three methods are able to provide the same result in a qualitative and quantitative way for Q_{MAX} and $C_{MAX,BOD}$.

Indeed, for Q_{MAX} a high linearity has been found in the application of the SRC method as demonstrated by the R^2 value close to 1. Such result was also confirmed by the low value of the sum of σ_i in the Morris screening application. The low value of the sum of σ_i means that globally the model factors have a linear effect. This high linearity for Q_{MAX} has also been confirmed for E-FAST, i.e. the sum

of S_i and the sum of S_{Ti} are very similar. This confirms the low interactions for this model output.

For $C_{MAX,BOD}$ a similar agreement in the results was found. However, the R^2 of SRC application is only 0.46 (outside the range of applicability of SRC) showing a substantial non-linearity for this model output. The same GSA result is found for the application of the Morris screening and E-FAST methods. Indeed, for $C_{MAX,BOD}$ the sum of σ_i is quite high (3.19) demonstrating a higher interaction among model factors than for Q_{MAX}. Moreover, the sums of S_i and S_{Ti} related to the E-FAST application differ considerably (Table 3).

In terms of factors prioritisation the comparison between β_i^2 and μ^* allows concluding that the SRC and Morris screening methods agree quite well. Excellent agreement among the results has been obtained when comparing β_i^2 and S_i for factor prioritisation, and this for both Q_{MAX} and $C_{MAX,BOD}$ (Table 3). Thus, the SRC and E-FAST methods lead to very similar results in terms of factors prioritisation.

In terms of factor fixing corresponding results have been obtained with the Morris screening and E-FAST methods, i.e. comparing μ^* with S_{Ti} and σ with S_{Ti} (Table 3).

A new way to compare classification between methods is the Venn diagram of Figs. 10 and 11. The overlapping area between SRC, Morris screening and Extended-FAST contains those model factors that can be considered important or non-influential for each method. Fig. 10 shows a Venn diagram related to the comparison of SRC, Morris screening and E-FAST in terms of important model

Table 3
Results obtained by applying SRC, Morris screening and E-FAST methods for Q_{MAX} and $C_{MAX,BOD}$.

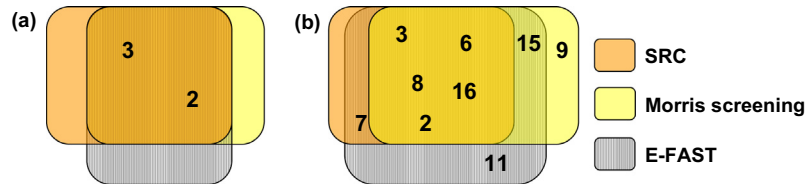| Method →	SRC		Morris screening			E-FAST			SRC		Morris screening			E-FAST		
Variable →				Q_{MAX}								$C_{MAX,BOD}$					
	R^2 0.92			$\Sigma\sigma_i$ 0.76		ΣS_i 0.93	ΣS_{Ti} 1.06		R^2 0.46			$\Sigma\sigma_i$ 3.19		ΣS_i 0.59	ΣS_{Ti} 2.02	
No. factor order	β_i	Rank	μ^*	σ	Rank	S_i	S_{Ti}	Rank	β_i	Rank	μ^*	σ	Rank	S_i	S_{Ti}	Rank
1	−0.012	10	0.000	0.000	5	0.000	0.000	17	0.051	12	0.000	0.000	16	0.000	0.003	17
2	−0.767	1	0.809	0.307	1	0.599	0.654	1	−0.330	2	0.304	0.491	2	0.102	0.339	2
3	0.560	2	0.531	0.264	2	0.316	0.372	2	0.173	5	0.254	0.408	3	0.038	0.207	6
4	−0.091	3	0.090	0.091	4	0.006	0.013	3	0.005	16	0.029	0.068	10	0.001	0.015	13
5	−0.073	4	0.097	0.098	3	0.006	0.013	4	0.061	10	0.042	0.091	8	0.001	0.015	14
6	−0.020	5	0.000	0.000	6	0.000	0.001	5	0.272	3	0.204	0.315	5	0.076	0.210	3
7	−0.009	12	0.000	0.000	7	0.000	0.001	6	−0.142	6	0.040	0.090	9	0.016	0.054	7
8	−0.014	7	0.000	0.000	8	0.000	0.001	7	0.197	4	0.212	0.451	4	0.053	0.185	4
9	−0.009	13	0.000	0.000	9	0.000	0.001	8	−0.019	13	0.152	0.349	6	0.008	0.145	10
10	0.014	8	0.000	0.000	10	0.000	0.001	9	0.005	17	0.004	0.012	15	0.001	0.023	15
11	0.002	16	0.000	0.000	11	0.000	0.001	10	0.095	7	0.012	0.036	12	0.015	0.057	8
12	0.016	6	0.000	0.000	12	0.000	0.001	11	0.071	8	0.007	0.024	14	0.009	0.038	9
13	−0.011	11	0.000	0.000	13	0.000	0.001	12	−0.017	14	0.025	0.081	11	0.007	0.056	12
14	−0.005	15	0.000	0.000	14	0.000	0.001	13	−0.016	15	0.011	0.042	13	0.007	0.046	11
15	−0.001	17	0.000	0.000	15	0.000	0.001	14	0.052	11	0.104	0.188	7	0.044	0.174	5
16	−0.006	14	0.000	0.000	16	0.000	0.001	15	−0.416	1	0.350	0.545	1	0.212	0.454	1
17	−0.013	9	0.000	0.000	17	0.000	0.001	16	0.066	9	0.000	0.000	17	0.000	0.003	16

Fig. 10. Venn diagram for important model factors (factors prioritisation).
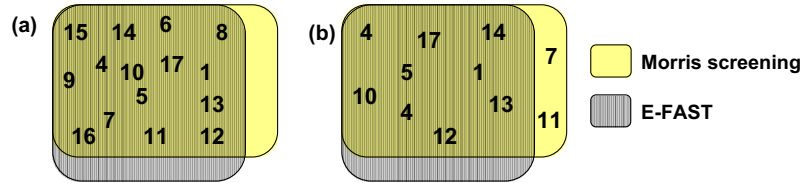

Fig. 11. Venn diagram for non-influential model factors (factor fixing) obtained by applying Morris screening and Extended-FAST methods related to Q_{MAX} (a) and $C_{MAX,BOD}$ (b); numbers refer to the factors according to Table 2.

factors for Q_{MAX} (Fig. 10a) and for $C_{MAX,BOD}$ (Fig. 10b). From Fig. 10 it is evident that for Q_{MAX} the three methods provide the same results. For $C_{MAX,BOD}$ (Fig. 10b) the three methods provide similar results for those model factors having the highest influence (W_0, Φ, Accu, Arra and K_{susp}). Further, the model factor W_h, selected as interacting by means E-FAST, was selected as important for $C_{MAX,BOD}$ (Fig. 10b) by using the Morris screening method. Such result shows a good agreement between the results obtained by means of E-FAST and Morris screening methods. It is worth mentioning though that model factor h_{max} is only important for $C_{MAX,BOD}$ for the E-FAST method (Fig. 7b). Such a result may indicate that both SRC and Morris screening may eliminate factors that may turn out to be important. However, it should be noted that for factor h_{max} the sensitivity indices are close to their respective thresholds.

Fig. 11 shows a Venn diagram related to the non-influential factors of Q_{MAX} (Fig. 11a) and $C_{MAX,BOD}$ (Fig. 11b). This diagram can only compare the Morris screening and E-FAST methods (SRC doesn't provide such information). From Fig. 11a one may observe that Morris screening and Extended-FAST provide exactly the same results in terms of non-influential model factors for Q_{MAX}. However, for $C_{MAX,BOD}$, the Morris screening method overestimates the number of non-influential model factors compared to the E-FAST method. Factors Disp and h_{max} are identified as being non-influential with the Morris screening. This is problematic as the same factors are considered important with Extended-FAST (Fig. 10). It seems that the Morris screening is not conservative enough, despite the name ("screening") given to the method.

7. Conclusions

- A comparison between three GSA methods (SRC, Morris screening and E-FAST) was performed in order to identify important, non-influential and interacting model factors of an urban drainage water quality model; seventeen model factors and seven model outputs (quality/quantity) were investigated.
- The compared methods give the same results for water quantity selected model outputs and slightly different results for water quality selected model outputs.
- For this study the method that provides a better compromise between the computational burden and the results reliability (both in terms of factor fixing and factor prioritisation) was the Extended-FAST method.
- A common terminology related to GSA method applications was proposed in this paper with the aim to fill literature gaps and to facilitate the comprehensive comparison among GSA methods.

- A Venn diagram based classification scheme to group important, interacting and non-influential factors was proposed.
- A complete convergence analysis exposed the danger of obtaining unreliable sensitivity index values for SRC and Morris when using standard literature sample sizes. Thus, convergence analysis is strictly required when these two methods are applied.
- The results of the Morris screening method application have shown that despite its name ("screening") this method may not be conservative enough and may exclude factors that turn out to be important after all.
- The robustness of GSA is significantly increased by using multiple methods and multiple objectives and testing convergence because depending on the objective and/or model complexity different numbers of simulations can be required.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2014.12.056.

References

Alley, W.M., Smith, P.E., 1981. Estimation of accumulation parameters for urban runoff quality modelling. Water Resour. Res. 17 (4), 1657–1664.

Ashley, R.M., Fraser, A., Burrows, R., Blanksby, J., 2000. The management of sediment in combined sewers. Urban Water 2 (4), 263–275.

Ashley, R.M., Bertrand-Krajewski, J.L., Hvitved-Jacobsen, T., 2005. Sewer solids: 20 years of investigation. Water Sci. Technol. 52 (3), 73–84.

Banasiak, R., Verhoeven, R., De Sutter, R., Tait, S., 2005. The erosion behaviour of biologically active sewer sediment deposits: observations from a laboratory study. Water Res. 39 (20), 5221–5231.

Beck, M.B., 1987. Water quality modelling: a review of the analysis of uncertainty. Water Resour. Res. 23 (8), 1393–1442.

Benedetti, L., De Baets, B., Nopens, I., Vanrolleghem, P.A., 2010. Multi-criteria analysis of wastewater treatment plant design and control scenarios with the Benchmark simulation model No. 2 under uncertainty. Environ. Modell. Softw. 25 (5), 616–621.

Benedetti, L., Claeys, F., Nopens, I., Vanrolleghem, P.A., 2011. Assessing the convergence of LHS Monte Carlo simulations of wastewater treatment models. Water Sci. Technol. 63 (10), 2219–2224.

Bertrand-Krajewski, J.L., Briat, P., Scrivener, O., 1993. Sewer sediment production and transport modeling: a literature review. J. Hydraul. Res. 31 (4), 435–460.

Campolongo, F., Carboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Modell. Softw. 22 (10), 1509–1518.

Candela, A., Freni, G., Mannina, G., Viviani, G., 2012. Receiving water body quality assessment: an integrated mathematical approach applied to an Italian case study. J. Hydroinform. 14 (1), 30–47.

Cosenza, A., Mannina, G., Vanrolleghem, P.A., Neumann, M.B., 2013. Global sensitivity analysis in wastewater applications: a comprehensive comparison of different methods. Environ. Modell. Softw. 49, 40–52.

Crabtree, R.W., 1989. Sediment in sewers. J. Inst. Water Env. Man. 3 (6), 569–578.

Deletic, A., Dotto, C.B.S., McCarthy, D.T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T.D., Rauch, W., Bertrand-Krajewski, J.L., Tait, S., 2012. Assessing uncertainties in urban drainage models. Phys. Chem. Earth 42–44, 3–10.

Donckels, B.M.R., Kroll, S., Van Dorpe, M., Weemaes, M., 2014. Global sensitivity analysis of an in-sewer process model for the study of sulfide-induced corrosion of concrete. Water Sci. Technol. 69 (3), 647–655.

Dotto, C.B.S., Kleidorfer, M., Deletic, A., Fletcher, T.D., McCarthy, D.T., Rauch, W., 2010. Storm water quality models: performance and sensitivity analysis. Water Sci. Technol. 62 (4), 837–843.

Dotto, C.B.S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D.T., Freni, G., Rauch, W., Deletic, A., 2012. Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. Water Res. 46 (8), 2545–2558.

Freni, G., Mannina, G., 2010a. Bayesian approach for uncertainty quantification in water quality modelling: the influence of prior distribution. J. Hydrol. 392, 31–39.

Freni, G., Mannina, G., 2010b. Uncertainty in water quality modelling: the applicability of variance decomposition approach. J. Hydrol. 394, 324–333.

Freni, G., Mannina, G., Viviani, G., 2009a. Uncertainty assessment of an integrated urban drainage model. J. Hydrol. 373, 392–404.

Freni, G., Mannina, G., Viviani, G., 2009b. Urban runoff modelling uncertainty: comparison among Bayesian and pseudo-Bayesian methods. Environ. Modell. Softw. 25, 1100–1111.

Freni, G., Mannina, G., Viviani, G., 2010c. Urban water quality modelling: a parsimonious holistic approach for a complex real case study. Water Sci. Technol. 61 (2), 521–536.

Gamerith, V., Neumann, M.B., Muschalla, D., 2013. Applying global sensitivity analysis to the modelling of flow and water quality in sewers. Water Res. 47 (13), 4600–4611.

Herman, J.D., Kollat, J.B., Reed, P.M., Wagener, T., 2013. Technical note: method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. Hydrol. Earth Syst. Sci. Discuss. 10, 4275–4299.

Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliab. Eng. Syst. Saf. 52, 1–17.

Jewell, T.K., Adrian, D.D., 1978. SWMM storm water pollutant washoff function. J. Environ. Eng. 104 (5), 1036–1040.

Mannina, G., Cosenza, A., 2015. Quantifying sensitivity and uncertainty analysis of a new mathematical model for the evaluation of greenhouse gas emissions from membrane bioreactors. J. Membr. Sci. 475 (1), 80–90.

Mannina, G., Viviani, G., 2009a. Separate and combined sewer systems: a long-term modelling approach. Water Sci. Technol. 60 (3), 555–565.

Mannina, G., Viviani, G., 2009b. Hybrid moving bed biofilm reactors: an effective solution for upgrading a large wastewater treatment plant. Water Sci. Technol. 60 (5), 1103–1116.

Mannina, G., Viviani, G., 2010. An urban drainage stormwater quality model: model development and uncertainty quantification. J. Hydrol. 381, 248–265.

Mannina, G., Schellart, A.N.A., Tait, S., Viviani, G., 2012. Uncertainty in sewer sediment deposit modelling: detailed vs simplified modelling approaches. Phys. Chem. Earth 42–44, 11–20.

Massmann, C., Holzmann, H., 2012. Analysis of the behavior of a rainfall–runoff model using three global sensitivity analysis methods evaluated at different temporal scales. J. Hydrol. 475, 97–110.

Moreau, P., Viaud, V., Parnaudeau, V., Salmon-Monviola, J., Durand, P., 2013. An approach for global sensitivity analysis of a complex environmental model to spatial inputs and parameters: a case study of an agro-hydrological model. Environ. Modell. Softw. 47, 74–87.

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

Neumann, M.B., 2012. Comparison of sensitivity analysis techniques for modelling micropollutant oxidation in water treatment. Sci. Total Environ. 433 (1), 530–537.

Nossent, J., Bauwens, W., 2012. Optimising the convergence of a Sobol' sensitivity analysis for an environmental model: application of an appropriate estimate for the square of the expectation value and the total variance. In: Proceedings of the International Environmental Modelling and Software Society conference (iEMSs) 2012, Leipzig, Germany, 1–5 July 2012.

Novotny, V., Sung, H.M., Bannerman, R., Baum, K., 1985. Estimating non point pollution from small urban watersheds. J. Water Pollut. Control Federat. 57, 339–348.

Parchure, T.M., Mehta, A.J., 1985. Erosion of soft cohesive sediment deposits. J. Hydraulic Eng. 111 (10), 1308–1326.

Pujol, G., 2007. Sensitivity: Sensitivity Analysis, R package version 1.3-0.

Pujol, G., 2009. Simplex-based screening designs for estimating metamodels. Reliab. Eng. Syst. Saf. 94, 1156–1160.

R Development Core Team. 2007. R Foundation for Statistical Computing; Vienna, Austria. R: A Language and Environment for Statistical Computing. Version 2.6.1, URL http://www.R-project.org/.

Ristenpart, E., 1995. Sediment properties and their changes in sewer. Water Sci. Technol. 25 (8), 1–12.

Ruano, M.V., Ribes, J., Ferrer, J., Sin, G., 2011. Application of the Morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants. Water Sci. Technol. 63 (10), 2199–2206.

Ruano, M.V., Ribes, J., Seco, A., Ferrer, J., 2012. An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. Environ. Modell. Softw. 37 (11), 103–109.

Saltelli, A., 2000. Sensitivity Analysis. John Wiley & Sons, Chichester, UK.

Saltelli, A., Annoni, P., 2010. How to avoid perfunctory sensitivity analysis. Environ. Modell. Softw. 25 (12), 1508–1517.

Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41 (1), 426, 39–56.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice. A guide to assessing scientific models. In: Probability and Statistics Series. John Wiley & Sons, Chichester, UK.

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., 2005. Sensitivity analysis for chemical models. Chem. Rev. 105, 2811–2827.

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., The Atrium, Southern Gate, Chichester, UK.

Skipworth, P.J., Tait, S.J., Saul, A.J., 1999. Erosion of beds in sewers: model development. J. Environ. Eng. 125 (6), 566–573.

Sun, X.Y., Newham, L.T.H., Croke, B.F.W., Norton, J.P., 2012. Three complementary methods for sensitivity analysis of a water quality model. Environ. Modell. Softw. 37, 19–29.

Tang, Y., Reed, P., Wahener, T., van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrol. Earth Syst. Sci. 11, 793–817.

Thorndahl, S., Beven, K.J., Jensen, J.B., Schaarup-Jensen, K., 2008. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. J. Hydrol. 357, 421–437.

Vanuytrecht, E., Raes, D., Willems, P., 2014. Global sensitivity analysis of yield output from the water productivity model. Environ. Modell. Softw. 51, 323–332.

Vezzaro, L., Mikkelsen, P.S., 2012. Application of global sensitivity analysis and uncertainty quantification in dynamic modelling of micropollutants in stormwater runoff. Environ. Modell. Softw. 27–28, 40–51.

Wang, J., Xin, Li., Ling, L., Feng, F., 2013. Parameter sensitivity analysis of crop growth models based on the extended Fourier amplitude sensitivity test method. Environ. Modell. Softw. 48, 171–182.

Yang, J., 2011. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. Environ. Modell. Softw. 26 (4), 444–457.

Zhan, C.-S., Song, X.-M., Xia, J., Tong, C., 2013. An efficient integrated approach for global sensitivity analysis of hydrological model parameters. Environ. Modell. Softw. 41, 39–52.