An experimental approach to reduce the parametric dimensionality for rainfall–runoff models

Jie Chen, Richard Arsenault and François P. Brissette

ABSTRACT

Sobol' sensitivity analysis has been used successfully in the past to reduce the parametric dimensionality for hydrological models. However, the effects of its limitation, in that it assumes an independence of parameters, need to be investigated. This study proposes an experimental approach to assess the commonly used Sobol' analysis for reducing the parameter dimensionality of hydrological models. In this approach, the number of model parameters is directly pitted against an efficiency criterion within a multi-objective genetic algorithm (MOGA), thus allowing both the identification of key model parameters and the optimal number of parameters to be used within the same analysis. The proposed approach was tested and compared with the Sobol' method based on a conceptual lumped hydrological model (HSAMI) with 23 free parameters. The results show that both methods performed very similarly, and allowed 11 out of 23 HSAMI parameters to be reduced with little loss in model performance. Based on this comparison, Sobol' appears to be an effective and robust method despite its limitations. On the other hand, the MOGA algorithm outperformed Sobol' analysis for further reduction of the parametric space and found optimal solutions with as few as eight parameters with minimal performance loss in validation.

Key words | hydrological model, multi-objective genetic algorithm, parameter reduction, Sobol' sensitivity analysis

Jie Chen
State Key Laboratory of Water Resources and
    Hydropower Engineering Science,
Wuhan University,
Wuhan 430072,
China

Jie Chen (corresponding author)
Richard Arsenault
François P. Brissette
École de technologie supérieure,
Université du Québec,
1100, Rue Notre-Dame Ouest,
Montréal,
QC,
Canada H3C 1K3
E-mail: jie.chen@etsmtl.ca

INTRODUCTION

Hydrological models have been used in a wide range of water resources management activities, such as watershed streamflow quantification, reservoir system operations, groundwater protection, water distribution systems, water use (Wurbs 1998; Pechlivanidis et al. 2011), and climate change impacts assessment (e.g., Wilby & Harris 2006; Kay et al. 2009; Chen et al. 2011a, 2011b, 2012). The successful use of hydrological models largely depends on how well they are calibrated, and the complexity of model calibration depends, among others, on the number of parameters that must be optimized (van Werkhoven et al. 2009). Model performance over the calibration period is generally improved when introducing additional parameters due to the added degrees of freedom; however, this may lead to over-fitting and poorer performance when using the model over

different time periods (Myung & Pitt 2002). The number of 'free parameters' is generally too large for most commonly used hydrological models, and the problem of equifinality is ubiquitous (Beven & Binley 1992; Wagener et al. 2001; Tonkin & Doherty 2005). Owing to the equifinality problem, the uncertainty associated with the choice of an optimal parameter set can be large, since a single optimal parameter set for a hydrologic model may not be found during the calibration process (Klepper et al. 1991; van Straten & Keesman 1991; Beven & Binley 1992; Yapo et al. 1996). This is particularly true for hydrological models with a large parametric dimensionality, since both over-parameterization and parameter interactions can cause model parameters to be not uniquely identifiable (Zhan et al. 2013; Gan et al. 2014; Song et al. 2015). Additionally, the

use of different efficiency metrics for model optimization may result in different optimal parameter sets (van Werkhoven *et al.* 2009).

Several studies (e.g., Bastidas *et al.* 1999; Huang & Liang 2006; Cox *et al.* 2006; Wagener & Kollat 2007; Pechlivanidis *et al.* 2010) suggest that the most appropriate approach in dealing with equifinality consists in reducing the number of free parameters by fixing the less important parameters to constant values. The reduction in model parameters is also an efficient way to reduce the impacts of parameter uncertainty on hydrological simulations (Song *et al.* 2012). It also has the advantage of simplifying the optimization problem linked to model calibration. Additionally, if a reliable parameter identification technique is developed over regions where data are available, the reduced parameter set can be much more easily regionalized for predicting flows in ungauged regions (Lee *et al.* 2005; Pechlivanidis *et al.* 2010; Arsenault & Brissette 2014).

To reduce the number of parameters during the model calibration process, it is necessary to quantitatively evaluate the influence of each parameter on model performance. This has been done in several studies. For example, Huang & Liang (2006) introduced an alternative subsurface flow parameterization into a hydrological model (three-layer variable infiltration capacity). Two out of three parameters were reduced in the calibration process, and the results showed that the performance of the hydrological model with the one-parameter subsurface flow formulation was comparable to the model with the three-parameter version. This study further indicated that the reduction of model parameters is an effective way to reduce the parameter uncertainty for hydrological simulations. More recently, Pechlivanidis *et al.* (2010) overcame problems in distributed modeling associated with the lack of parameter identifiability through reduction of parameter dimensionality. The semi-distributed probability distributed moisture model was calibrated based on parameter regionalization, and a good balance in the model complexity was achieved assuming that some parameters directly correspond to their physical characteristics while other insensitive parameters were held constant. Cox *et al.* (2006) also showed that parametrically reduced models have lower prediction residual sums of squares (the sum of squared differences between the observed and predicted values) than the original model, strongly suggesting that the original model was over-fitted.

Sensitivity analysis (local and global sensitivity analysis) is one of the efficient ways to identify the influence of each parameter on the model performance. Thus, it can be used to reduce the parameter dimensionality for hydrological models by fixing insensitive parameters during the calibration process (Wilson *et al.* 1987a, 1987b; Pitman 1994; Gao *et al.* 1996; Bastidas *et al.* 1999). There are different approaches to assess parameter sensitivity. For example, Song *et al.* (2015) compared screening, regression, variance-based, meta-modeling, regionalized sensitivity analysis, and entropy-based methods. All have strengths and weaknesses which must be well understood by the end-user. For example, entropy-based methods require expert input to estimate parameter distributions, which can be difficult to perform correctly. Among all these sensitivity analysis methods (Gan *et al.* 2014), variance-based methods have been used extensively in the hydrological modeling community (Song *et al.* 2015). In particular, the Sobol' method (Sobol' 1993) is one of the most widely used global sensitivity analysis methods (e.g., Pappenberger *et al.* 2008; van Werkhoven *et al.* 2008; Nossent *et al.* 2011; Zhang *et al.* 2013; Arsenault *et al.* 2015; Zhang & Ross 2015). Current Sobol' analysis methods are capable of assessing the effect of each parameter and its interactions with other parameters on the model output, and have been used successfully in several studies (e.g., Tang *et al.* 2007; Van Werkhoven *et al.* 2009; Nossent *et al.* 2011; Baroni & Tarantola 2014). Arsenault *et al.* (2015) used the Sobol' variance-based sensitivity analysis to reduce parametric dimensionality of a 23-parameter model for use in regionalization. They found that up to 11 parameters could be fixed without significantly hurting the model performance, while at the same time increasing parameter identifiability. The Sobol' method assesses the sensitivity of each parameter and parameter interactions, based on their contributions to the total model variance. These contributions are called Sobol' sensitivity indices (SI), which are expressed by the ratio of the partial variance to the total variance. The first-order SI measures the variance contribution of the individual parameter to the total model variance. The second-order SI measures the variance contribution of the two-parameter interaction to the total model variance, and so on. The total SI measures the sensitivity due to the combined effect of one parameter and its interactions with all other parameters. When using the Sobol' method to reduce the parametric dimensionality for a hydrological model, parameters that do not make

important contributions to the total variance can be fixed with constant values and ignored during the optimization process (Sobol' 1993; Pappenberger *et al.* 2008; van Werkhoven *et al.* 2008; Nossent *et al.* 2011; Zhang *et al.* 2013).

Tang *et al.* (2007) compared the Sobol' method with three other sensitivity analysis tools (parameter estimation software, regional sensitivity analysis, and analysis of variance) and found that the Sobol' method yielded more robust sensitivity rankings than the other methods. Van Werkhoven *et al.* (2009) used Sobol' analysis as a screening tool to reduce the parametric dimensionality of hydrological models, and found that parameters explaining at least 20% of the model output variance should be included in the calibration process. This threshold generally reduced the number of model parameters by at least 30% for the Sacramento soil moisture accounting model, while maintaining good performance of predictions. This study further indicated that the reduced parameter sets changed across different hydroclimatic gradients, and that multiple metrics may be necessary for the selection of optimized parameters. More recently, Nossent *et al.* (2011) used the Sobol' method to analyze the parameter sensitivity of the Soil and Water Assessment Tool for flow simulations and found that no more than nine parameters out of 26 are needed to adequately represent the model output variability. However, the Sobol' method assumes no correlation between parameters, whereas strong inter-dependence of parameters is usually found in hydrological models

(Pechlivanidis *et al.* 2010). It is thus necessary to investigate the effects of this independence hypothesis on Sobol' performance in the reduction of hydrological model parameters by comparing it to other methods that do not depend on this hypothesis. While research is ongoing in attempting to address the independence issue (Chastaing *et al.* 2014), the effects of it have not yet been measured for hydrologic applications.

Accordingly, this work proposes a multi-objective genetic algorithm (MOGA)-based approach to assess the most commonly used Sobol' sensitivity analysis in the reduction of the hydrological model parametric dimensionality. The proposed approach directly uses the number of model parameters as one of the objectives, thus directly quantifying the value of using additional free parameters as well as identifying the best combination of model parameters. The proposed method was tested using a lumped conceptual rainfall–runoff model over two Canadian watersheds in the Province of Quebec, and then compared with the Sobol' method for one watershed.

STUDIED WATERSHEDS

Two Canadian watersheds (Peribonka and Yamaska; Figure 1), located in the Province of Quebec, were selected to test the proposed method and evaluate its applicability under different watershed characteristics. Both the
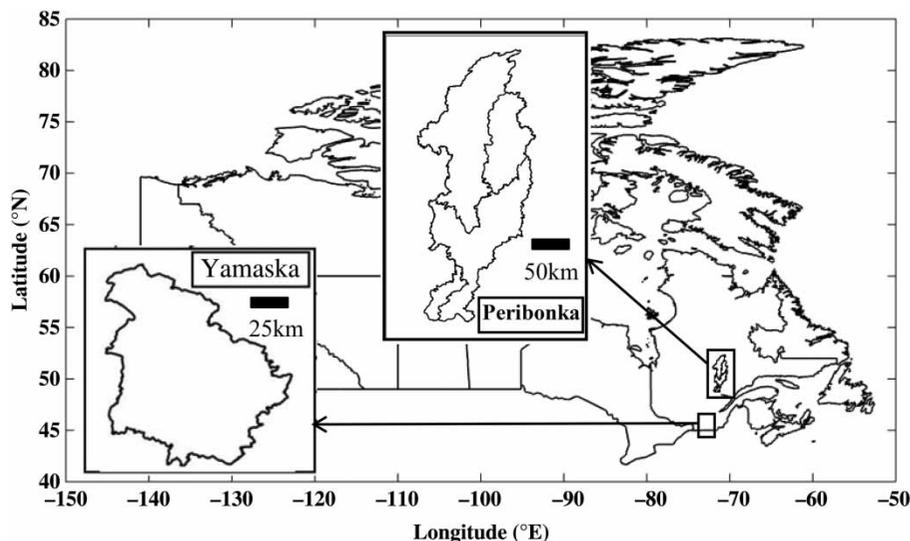

Figure 1 | Location map of the two watersheds.

Peribonka and Yamaska watersheds are composed of several tributaries draining basins of approximately 27,000 km^2 and 4,843 km^2 in southeastern and southern Quebec, respectively. The southern parts of the Peribonka and Yamaska watersheds, respectively, named as Chute-du-Diable (CDD) and Cowansville (COW) watersheds, are used in this study. The details on both watersheds are presented below.

CDD

The CDD watershed is located in the central part of the province of Quebec, and covers 9,700 km^2 of mostly forested areas with sparse population. The basin is part of the northern Quebec subarctic region, characterized by wide daily and annual temperature ranges, heavy wintertime snowfall, and pronounced rainfall and/or snowmelt peaks in the spring. The average annual precipitation in the area is 962 mm, of which about 36% is snowfall. The average annual maximum and minimum temperatures between 1,979 and 2003 were 5.49 °C and −5.85 °C, respectively. The CDD watershed contains a large hydropower reservoir managed by Rio Tinto Alcan for hydroelectric power generation. River flows are regulated by two upstream reservoirs. Snow plays a crucial role in the watershed management, with 35% of the total yearly discharge occurring during the spring flood. The mean annual discharge of the CDD watershed is 211.4 m^3/s. Snowmelt peak discharge usually occurs in May and averages about 1,220 m^3/s. Natural inflow series have been reconstructed by Rio Tinto Alcan using operation data from the upstream reservoir. Hydrologic model performance (to be detailed later) testifies to the quality of the reconstructed inflow data.

COW

COW is a non-regulated, 210 km^2 watershed located in southern Quebec. The average annual rainfall in the COW watershed is 1,267 mm with about 22% of snow. The average annual maximum and minimum temperatures were 11.59 °C and 1.14 °C, respectively, over the last 30 years. As opposed to the CDD watershed, the COW watershed's southern location and much smaller size results in hydrographs that are less dominated by snowmelt. In fact, annual maximum peak discharge often occurs during the summer season, a feature never seen in the CDD watershed. The mean annual discharge of the COW watershed is 4.5 m^3/s. Snowmelt peak discharge usually occurs in April and averages about 70 m^3/s.

METHODOLOGY

Hydrological modeling

HSAMI is a 23-parameter, lumped, conceptual, rainfall–runoff model developed by Hydro-Québec, and which has been used to forecast natural inflows for over 20 years (Fortin 2000). HSAMI is used by Hydro-Québec for daily forecasting of natural inflows on nearly 100 watersheds with drainage areas ranging from 160 to 69,195 km^2. HSAMI was also used in several flow forecasting and climate change impact studies (e.g., Minville *et al.* 2008, 2009; Chen *et al.* 2011a, 2011b, 2012; Poulin *et al.* 2011; Arsenault *et al.* 2013). Of HSAMI's 23 parameters, two account for evapotranspiration, six for snow accumulation/melting, ten for vertical water movement, and five for horizontal water movement (see Table 1). Vertical flows are simulated with four interconnected linear reservoirs (snow on the ground, surface water, unsaturated and saturated zones). Horizontal flows are routed through two unit hydrographs and one linear reservoir. In addition, the model takes into account snow accumulation, snowmelt, soil freezing/thawing, and evapotranspiration. Model calibration was done automatically using the covariance matrix adaptation evolution strategy (CMAES) (Hansen & Ostermeier 1996, 2001), following the conclusions of Arsenault *et al.* (2014).

The basin-averaged daily input data required for HSAMI are liquid and solid precipitation, as well as maximum and minimum temperatures. Cloud cover fraction and snow water equivalent can also be used as input, if available. A natural inflow or discharge time series is also needed for calibration/validation.

Two different metrics were used to evaluate the model's adequacy in representing the high and low flows. The first efficiency metric is the commonly used Nash–Sutcliffe efficiency (NSE) (Nash & Sutcliffe 1970), which is a normalized statistic that determines the relative magnitude

Table 1 | HSAMI 23 free parameters

Sub-model	ID	Acronym	Physical meaning	Unit	Parameter range
Evapo-transpiration	P1	FAETP	Factor multiplying potential evapotranspiration (PET) for the estimation of summer real evapotranspiration (RET)	–	[0.6 3]
	P2	POUR	Factor multiplying PET for estimating the RET in winter	–	[0 0.3]
Snowmelt	P3	FKM	Snow melting rate during daytime. ΔT in Celsius is calculated as the difference between T_{max} and parameter of T_{max} threshold for snowmelt (TB)	cm/Δ°C/day	[0.05 0.4]
	P4	FKYN	Snow melting rate during night-time. ΔT in Celsius is calculated as the difference between parameter TB and T_{min}	cm/Δ°C/day	[0.05 0.5]
	P5	TB	T_{max} threshold for snowmelt	°C	[−6 7]
	P6	TBN	T_{min} threshold for accelerated snowmelt	°C	[−6 6]
	P7	TACAL	Reference temperature for calculating the heat supplied by the rain to the snow cover	°C	[−6 4]
	P8	CC	Empirical parameter used to connect the state variables describing snow cover and cumulated snowmelt to the proportion of the basin covered by snow	–	[0.8 5]
Surface runoff	P9	CGEL	Empirical parameter used to connect the state variables describing soil freezing and thawing to the proportion of snowmelt water flowing on the surface	–	[0.8 15]
	P10	CIMAX	24-hour rainfall amount needed to generate 50% runoff with completely dry soil	cm	[10 45]
	P11	CSAP	24-hour rainfall amount needed to generate 50% runoff with completely saturated soil	cm	[1 8]
Vertical water movement	P12	RMIN	Water amount in the unsaturated zone that cannot drain by gravity	cm	[0 7]
	P13	RMAX	Maximum water amount that can be contained in the unsaturated soil zone	cm	[3 25]
	P14	SAPMAX	Maximum water amount that can be contained in the aquifer before generating surface runoff	cm	[4 30]
	P15	D	Proportion of surface water flowing through the intermediate hydrograph instead of moving through the soil column	–	[0.15 0.7]
	P16	X	Proportion of soil water that is directed to the intermediate hydrograph when the unsaturated zone overflows	–	[0.3 1]
	P17	PER	Emptying rate of the unsaturated zone to the groundwater reservoir	24h^{-1}	[0.09 0.07]
	P18	Y	Emptying rate of the groundwater reservoir (base flow)	24h^{-1}	[0.006 0.018]
Horizontal water movement	P19	ALP	Emptying rate of the intermediate reservoir, through the intermediate hydrograph	24h^{-1}	[0.6 1.2]
	P20	TPOINT	Time to peak for the surface unit hydrograph	day	[0.3 5]
	P21	BETA	Shape parameter of the surface hydrograph (using a gamma distribution function)	–	[0.4 5]
	P22	BET/GAMMA	Time to peak for the intermediate unit hydrograph	day	[1.5 13]
	P23	GAMMA	Shape parameter of the intermediate hydrograph (using a gamma distribution function)	–	[0.15 1.5]

T_{max}, maximum temperature; T_{min}, minimum temperature.

of the residual variance compared to the observed data variance. The second efficiency metric is a Box–Cox transformed version of the root mean square error (TRMSE) as used in the study of van Werkhoven *et al.* (2009). The NSE emphasizes high flows, whereas TRMSE puts a much greater emphasis on low flows and should weigh more on model parameters linked to vertical water movement. The effects of both efficiency metrics were

taken into account when determining the importance of each parameter of HSAMI. To ensure that conflicting objectives exist, in the case of the NSE, the objective to minimize will be (1-NSE). No such problem exists with TRMSE, whose value naturally decreases with a better fit.

In this study, 10 years of daily data (precipitation, temperature, and streamflow) were used for model calibration, and another 10 years of daily data were used for model validation, as presented in Table 2. The first year of the calibration and validation periods were used to warm up the hydrological model and the other 9 years were used for simulation. The observed daily precipitation and temperature were taken from the gridded observations of the Hutchinson *et al.* (2009) data set, which was obtained by interpolating station data to a 10-km grid using a thin plate smoothing spline surface fitting method. All the grid points within the watershed were averaged to obtain a single precipitation and temperature time series for each basin. The daily discharge at the watershed outlet was obtained from Hydro-Québec. HSAMI, with all 23 free parameters, was independently calibrated and validated 1,000 times within the same parametric space, but with different random seeds. A large number of calibrations were used to obtain a robust distribution for each parameter to quantify the parameter equifinality. The median of NSE and TRMSE for these 1,000 calibrations showed good performances of HSAMI for both watersheds. The results for the CDD watershed were consistently better than those of the COW watershed. This is because the input data quality of the former is considered better than that of the latter, due to there being more weather stations, and because the daily time step used in this study was less suited to the smaller watershed. In addition, snowmelt-dominated basins are also usually easier to model, because the winter streamflow

is not sensitive to the precipitation input. Since TRMSE is discharge dependent, its values are smaller for the COW watershed, even though HSAMI was better optimized for the CDD watershed based on NSE values.

MOGA

Multi-objective optimization has been quite commonly used in hydrology over the past decade (Sawaragi *et al.* 1985; Steuer 1986; Yapo *et al.* 1998; Vrugt *et al.* 2003; Shafii & De Smedt 2009; Efstratiadis & Koutsoyiannis 2010; Reed *et al.* 2013). The most likely use of a multi-objective approach in model calibration is to increase model robustness over different efficiency metrics (best-compromise solution) as well as to reduce parameter uncertainty. However, most hydrology studies have focused on finding an optimal parameter set (usually using an automated procedure), and not on the more fundamental aspect of reducing the parameter uncertainty. Since multi-objective optimization is the process of simultaneously optimizing two or more conflicting objectives, it allows the correlation and trade-offs between different efficiency metrics to be evaluated.

In this work, the relationship between the number of parameters that must be optimized during the calibration process and the model performance is viewed as a multi-objective problem. As such, the two conflicting objectives are as follows:

Objective   1 = minimize   the   number   of   calibration parameters
Objective 2 = minimize TRMSE (or minimize 1-NSE)

It is evident that reducing the number of free parameters will have a detrimental effect on model performance due to the loss of degrees of freedom. Therefore minimizing both criteria requires a compromise. Contrarily to univariate

Table 2 | Median, maximum (Max) and minimum (Min) values of the NSE and transformed root mean squared error (TRMSE) for 1,000 calibrations and validations of HSAMI with all 23 free parameters for two watersheds. Optimum NSE and TRMSE values are 1 and 0, respectively

Watershed	Source	Period	NSE			TRMSE		
			Median	Min	Max	Median	Min	Max
CDD	Calibration	1979–1988	0.89	0.87	0.90	1.52	1.49	2.42
	Validation	1989–1998	0.88	0.79	0.90	1.78	1.70	2.75
COW	Calibration	1990–1999	0.75	0.63	0.76	0.65	0.64	0.98
	Validation	2000–2009	0.72	0.59	0.74	0.78	0.75	1.11

optimization problems, the solution to a multi-objective optimization problem is a set of compromised trade-offs. For example, a large number of free parameters will improve model performance, whereas a small number of free parameters will have a detrimental effect on model performance. No single point minimizes both the number of parameters and model error as this would mean that the objectives are not conflicting. The multi-objective optimization problem will be solved using MOGA, which is a popular meta-heuristic multi-objective optimizer (Deb 2001; Konak *et al.* 2006), following these five steps:

1. Initial population of candidate sets. Each candidate set is a binary vector, where a 1 (0) at position j means that parameter j is calibrated (not calibrated). In this paper, 200 initial candidate sets were used as the initial population. One of the sets was a vector composed entirely of ones to ensure that the uppermost solution would be evaluated.

2. Perform the first evaluation. Each evaluation takes the candidate set (23-bit vector) and replaces zeros with a predefined fixed value. The fixed value is constant for the entire MOGA run. For example, if three bits are set to zero, then three parameters will be fixed and the model will calibrate the 20 remaining parameters. One evaluation in the MOGA run implies calibrating the HSAMI model with the remaining parameters.

3. Evaluate the two conflicting objective functions. The number of parameters is computed as the number of bits set to one in the candidate vector. The model performance is evaluated using either TRMSE or NSE during the model calibration.

4. If the following conditions are true:

   (1) For a given number of free parameters, the model performance is better than the best solution found to date.

   (2) For a given model performance, the number of parameters is lower than the current number of parameters for the given solution.

   Then the solution is retained as being part of the non-dominated solution while the newly replaced parameters are removed.

5. After the 200 candidates have been evaluated, the best candidates are slightly modified using parameter splicing and mutations to generate 200 more candidates. The

process is repeated for 20 generations, after which the final solution is produced. The final solution, called a Pareto front, is generated along which all solutions can be considered optimal. In other words, the Pareto front should indicate the best model performance in calibration when using an increasing number of model parameters. Conversely, minimization of the number of model parameters will only be achieved through a reduction in model performance due to the reduction in degrees of freedom. The Pareto front, and its translation into the validation data set, should allow for a quantitative evaluation of the true benefits of increasing the number of model parameters, as well as the identification of the most important model parameters, all in the same step.

In this study, a controlled elitist genetic algorithm was used. It consists of a variant of the non-dominated sorted genetic algorithm (Deb 2001; Deb *et al.* 2002). The algorithm's hyperparameters were set in such a way as to emphasize exploration of the search space because of the large amount of possible combinations. As such, a population of 200 individuals was selected. To compensate for the long processing time of generating 200 candidates per iteration (where each candidate requires a complete calibration of HSAMI), only 20 generations were performed.

When the algorithm chooses less than 23 parameters, an important issue arises with the value that must be assigned to all non-chosen parameters. If one decides not to calibrate all of HSAMI's 23 parameters, forced values have to be set for the remaining parameters. In this case, fixed parameter values were randomly assigned within a reduced search space defined for each parameter. The reduced search space is a subset of the total parameter space (Table 1) used by the optimization process. The total parameter space is chosen, while maintaining a physical sense, to minimize the chances that parameters are constrained by the search boundaries. The reduced search space was evaluated using the range of plausible values for each parameter following 1,000 different calibrations using all 23 parameters. This procedure eliminated the possibility of choosing an impossible fixed value for several parameters. Multi-objective optimization was performed eight times for each efficiency metric, with each iteration using different parameter sets for fixed parameters.

The eight different parameter sets were selected randomly from the reduced search space. The random assignments partly eliminated biases linked to a deliberate choice in fixed parameter values. Overall, 32 multi-objective optimization runs were conducted, since each of the eight random assignments was run for the two objective functions and both basins.

Sobol' sensitivity analysis

In this study, the Sobol' implementation was as follows. First, 100,000 parameter sets were sampled following a Sobol' sequence in the available parameter space. The Sobol' indices (SI) were then computed for the first and total orders from the selected parameter set using the Saltelli algorithm (Saltelli 2002). For this work, only the total order parameters are of importance since the idea is to fix the least important parameters to the total variance, thus their interactions must be considered.

The parameters were then sorted by order of importance, and the HSAMI model was calibrated with the least important parameter fixed to a random value within the predetermined boundaries. This operation was repeated 100 times to eliminate (or at least greatly reduce) the stochastic effects on the results. The results were noted in calibration and in validation, then the next least-important parameter was fixed and the process was launched again. However, the fixed parameters are kept at the same values throughout the study, i.e., the first parameter to be fixed keeps its value for the duration of the study, and thus the parameter space is kept intact for each calibration. This means that for each supplementary fixed parameter, the parameter space shrinks by one dimension but the parameter space is a subset of the previous parameter space. This methodology is continued until 22 parameters are fixed and a single one is left to calibrate. The idea to keep the previously fixed parameters to a constant value allows the effect of fixing the next parameter in a stable environment to be compared, instead of a constantly changing one (due to random selection of the parameters at each step).

RESULTS

This section consists of four sub-sections. The first sub-section involves analyzing the parametric equifinality for the hydrological model HSAMI, as a first step prior to the parameter reduction analysis. The second sub-section includes two steps to reduce HSAMI parameters using the multi-objective approach. The multi-objective optimization is initially conducted to obtain a Pareto front. Through the Pareto front, a judgment can be made about the number of parameters that can be fixed while keeping an adequate model performance. The third step involves ordering parameters based on their importance with an experimental approach. Finally, the proposed method is compared to the Sobol' sensitivity analysis method in terms of their performances in reducing HSAMI parameters.

Parametric equifinality analysis

Figure 2 presents the cumulative distribution function (CDF) of all parameter values obtained after performing 1,000 automatic calibrations with the CMAES algorithm. Model performance was reasonable for both watersheds, as indicated by the NSE and TRMSE values in Table 2. However, the problem associated with the determination of an optimal set of parameters is quite clear from the results presented in Figure 2. Most of the parameters had calibrated values covering the entire parameter space. In fact, the boundary values delimiting the parameter space were made larger than the operational range used by Hydro-Québec to make sure that parameters were not constrained over the possible range of physical parameter values. For parameters having a physical sense, boundaries were set as large as possible to encompass the full range of realistic possible values. For empirical parameters, boundaries were fixed using good judgment and user experience with the model. Figure 2 shows that despite being liberal in the boundary delineation process, most of the parameters would at one point very likely migrate outside the physically reasonable range of values if given the chance. This reflects on the problems of parameter identifiability and uniqueness that are the root cause of equifinality (Ebel & Loague 2006). In other words, Figure 2 shows that the optimization process found many local minima over 1,000 calibrations, whereas differences between the largest and smallest NSE were 0.03 for the CDD watershed and 0.13 for the COW watershed. Differences increased to 0.1 and 0.15, respectively, for the validation period (Table 2).
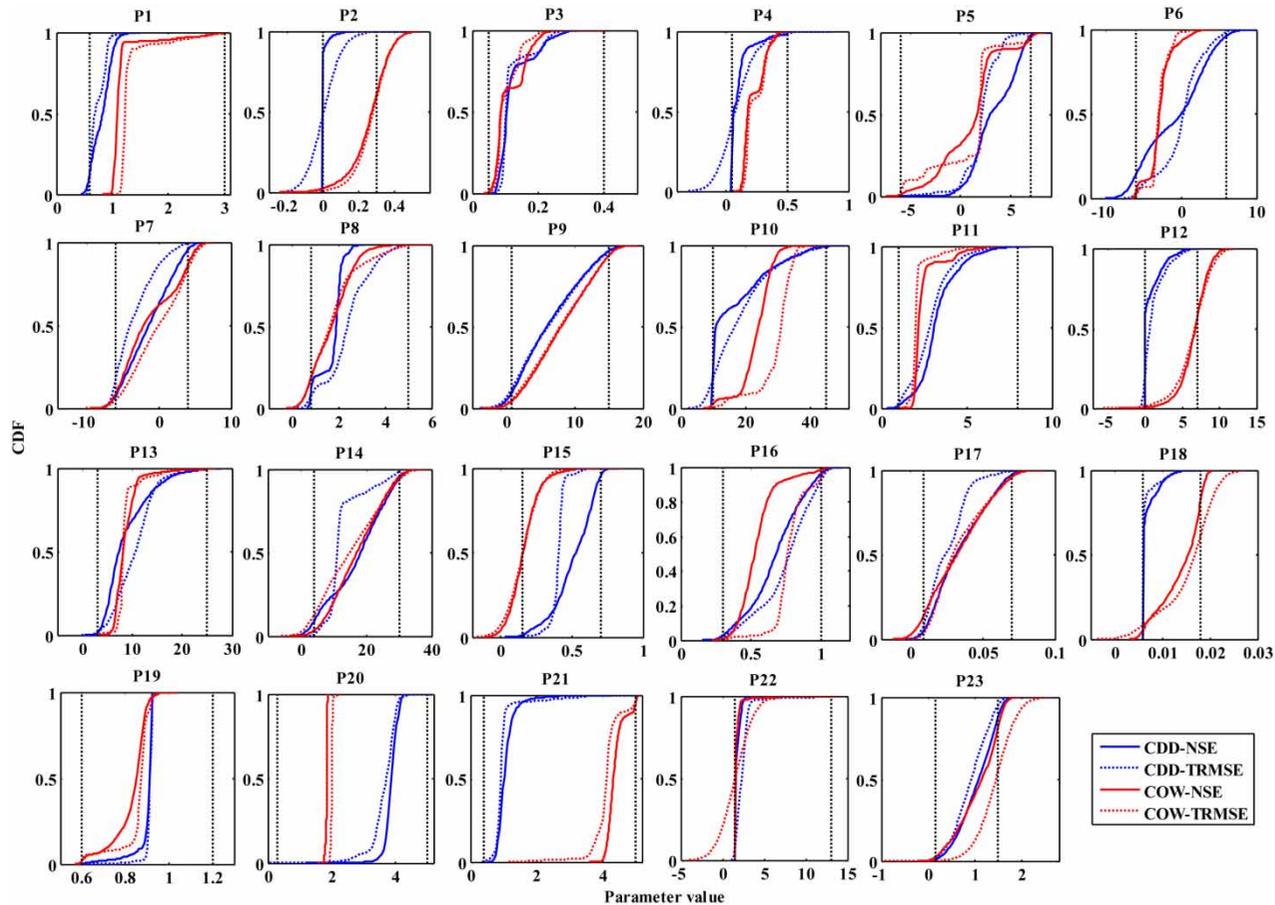
Figure 2 │ CDF of parameter values over 1,000 calibrations (23 optimized parameters) using NSE and transformed root mean square error (TRMSE) as efficiency metrics for both the CDD and COW watersheds. The *x*-axis covers the full boundary range, while the dashed lines show the upper and lower boundaries of the reduced search space for each parameter.

Multi-objective approach

Multi-objective optimization

Figure 3 presents the mean of all eight Pareto fronts obtained from the eight multi-objective optimizations for both watersheds and using both efficiency metrics. The circle data points in this graph define the Pareto front, and each circle corresponds to the best obtained calibration result using 'n' model parameters. The shape of the Pareto front indicates that adding additional free parameters to the model will result in an improvement in model performance. If adding a given parameter was to result in the same model performance, it would not appear on the Pareto front. This could explain why there is no point on the Pareto front with more than 19 parameters for the CDD watershed using

NSE as the efficiency metric and for the COW watershed using both efficiency metrics. However, in this case, the most likely reason for the lack of points using more than 19 parameters is the relatively low number of model iterations during the multi-objective optimization coupled with uncertainty associated with each calibration. Displaying more points (up to 23) on the Pareto front can be achieved, by increasing the iteration number. However, this may not be necessary, as the 'return on investment' of adding additional parameters rapidly diminishes (circles in Figure 3). In this case, the use of more than ten parameters results in only very modest performance improvements. The key question is whether or not these improvements are real and not simply the result of over-fitting. This can be very easily verified by using each combination of parameters along the Pareto front over the validation period. These
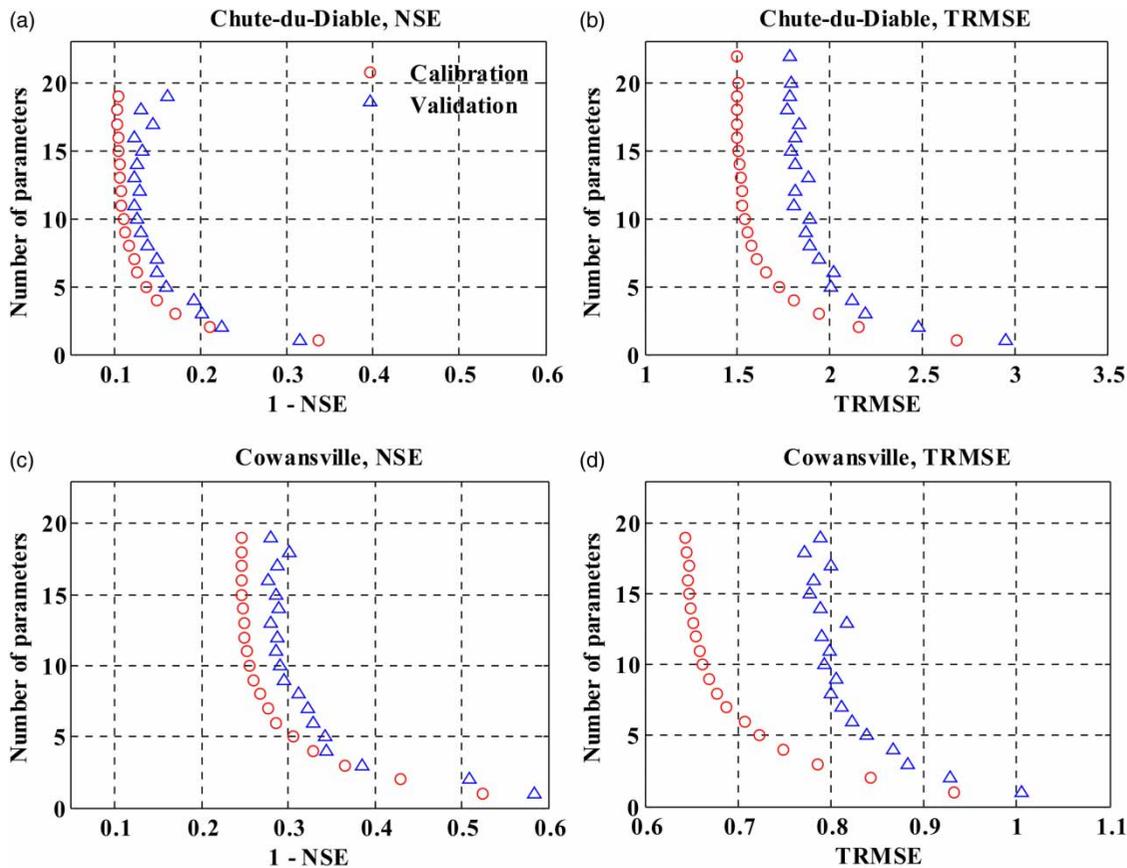
Figure 3 │ Mean of all eight Pareto fronts deriving from the eight multi-objective optimizations using the NSE and TRMSE as efficiency metrics for calibration and validation periods over both CDD and COW watersheds.

results are shown with the triangles in Figure 3. A similar pattern is observed for the validation, with performance improvements for up to ten parameters. However, above ten parameters, model performance is not improved any more. This demonstrates quite clearly that any improvement observed during calibration with more than ten parameters is due to over-fitting and not to a better representation of physical processes within the hydrological model. Any additional parameter will only increase parameter uncertainty with no additional benefit, and, in some cases, will be detrimental to model performance (see upper-left panel in Figure 3).

Parameters' importance ordering

Examination of Figure 3 indicates that at least 11 parameters out of 23 do not contribute much to model performance. The next step involves ranking parameters based on their relative importance and contribution to total variance in the model output. However, one problem arises in which the results of a choice of best parameters are not perfectly consistent from one multi-objective optimization to the other (results not shown). In other words, the best ten parameters resulting from one multi-objective optimization will not necessarily be the same ten parameters resulting from another. As discussed earlier, this is the consequence of the random assignation of fixed parameter values for each multi-objective optimization. To circumvent this problem, the results from all multi-objective optimizations were summed up to identify the leading parameters.

Table 3 presents the number of times that each parameter was optimized for the CDD watershed using the NSE, when moving from 1 to 19 parameters for the eight multi-objective optimizations. The vertical axis presents the 19 combinations of parameters from the Pareto front (as discussed earlier, the combinations of 20, 21, 22, and

Table 3 | Number of parameters selected as being important for the eight MOGA optimizations using the NSE for the CDD watershed

PN	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	RN
	Number of times that each parameter is selected during optimization for all eight multi-objective optimization runs																							
1	0	0	0	0	2	0	0	1	0	0	0	0	0	2	0	0	0	0	1	0	0	2	0	8
2	0	0	1	0	2	3	0	1	0	0	0	0	0	3	1	0	0	0	3	0	0	2	0	8
3	2	0	1	0	3	2	0	2	0	0	0	0	1	1	4	0	0	1	3	0	0	4	0	8
4	3	0	3	0	2	4	0	3	0	0	0	0	2	2	3	0	0	0	3	0	2	4	1	8
5	3	0	4	2	3	4	0	3	0	1	0	0	2	3	3	0	0	1	5	1	2	3	0	8
6	4	0	5	4	3	4	0	2	0	1	0	0	4	3	2	1	0	1	4	5	1	4	0	8
7	5	0	5	5	2	5	0	2	0	1	1	1	3	3	3	1	0	1	5	5	2	6	0	8
8	7	0	4	5	5	5	0	5	0	0	1	1	6	2	1	1	1	1	5	6	2	6	0	8
9	8	0	5	6	4	6	0	4	0	0	3	2	5	2	1	1	0	2	6	8	2	6	1	8
10	7	3	5	6	4	6	1	6	0	0	3	3	5	2	1	2	1	1	7	8	2	6	1	8
11	8	1	7	6	4	6	1	7	0	0	5	4	6	1	1	2	1	3	7	8	3	6	1	8
12	8	4	7	6	4	7	1	7	0	0	5	2	6	1	3	2	2	3	7	8	5	6	2	8
13	8	5	8	7	6	6	1	7	0	0	5	3	6	2	4	2	1	6	8	8	3	6	2	8
14	6	5	6	5	4	5	1	5	0	2	4	3	5	1	3	1	1	3	6	6	5	5	2	6
15	7	6	7	6	5	6	2	7	2	3	4	3	6	1	4	2	1	3	7	7	5	6	5	7
16	4	2	4	3	2	3	2	4	1	3	4	3	3	1	3	2	0	3	4	4	2	4	3	4
17	4	2	4	3	3	4	1	4	1	1	4	3	3	2	3	2	2	3	4	4	3	4	4	4
18	5	4	5	4	4	3	2	5	2	5	5	5	3	3	3	4	3	4	5	5	3	5	3	5
19	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1
Sum	90	33	82	69	63	80	12	76	6	18	45	34	67	35	44	24	14	37	91	84	43	86	25	131
Ratio	0.69	0.25	0.63	0.53	0.48	0.61	0.09	0.58	0.05	0.14	0.34	0.26	0.51	0.27	0.34	0.18	0.11	0.28	0.70	0.64	0.33	0.66	0.19	

Note: Roman, italic, bold, and underline denote the most important, important, average, and least important, respectively.

PN, the number of free parameters; RN, number of runs using PN; Sum, summation of the number of times the parameter was present in a Pareto-optimal set.

23 parameters were not optimal, as these combinations were no better than the one with a 19-parameter model). The horizontal axis represents the number of times each of the 23 parameters was selected for each combination defining the Pareto front. For example, the first row indicates that when one parameter was selected to define the Pareto front, parameters 5, 14, and 22 were chosen twice while parameters 8 and 19 were selected once for a total of eight corresponding to the eight performed optimizations. The last column of each row is the summation of all chosen parameters divided by the number of parameters and should be equal to eight. If the number is less than eight, it simply indicates that one of the eight multi-objective optimizations did not retain this number of parameters as optimal, thus indicating better performance with a lesser number of parameters.

The construction of Table 3 followed these four steps:

1. The total number of selections for each parameter was first summed in the vertical direction. For example, parameter 1 was selected 90 times (never when the total number of parameters was one or two, twice when three parameters were retained, etc.).
2. The total number of points selected on the eight Pareto fronts was summed up to 131 (sum of last column in Table 3). If one parameter was always selected for all eight multi-objective optimizations at every combination of parameters, its summation in step-(1) should be equal to 131.
3. The selection ratio was calculated for each parameter as the sum obtained in step-(1) divided by 131.
4. Three thresholds were used to qualitatively classify the parameters into four classes of importance according to the ratios calculated in step-(3). These thresholds are estimated based on our best professional judgment to insure that the number of parameters for each importance level is consistent with the Pareto front. The classes were defined as *most important* (greater than or equal to 0.4), *important* (less than 0.4, while greater than or equal to 0.25), *average* (less than 0.25, while greater than or equal to 0.15), and *least important* (less than 0.15).

It is important to restate that the uncertainty in parameter identification presented in Table 3 is the result of the random assignation of a fixed parameter value and of equifinality itself. For example, if two parameters are strongly correlated, a better assigned fixed value may result in the other one being selected. Technically speaking, if the most significant parameters in the hydrological model are randomly given optimal fixed values, they will not be retained by the algorithm. This is why multiple trials have to be done, and also why no parameter was always selected across the board. This approach did not maximize the differences between parameters because, clearly, a parameter being selected in the lower part of the Pareto front (few parameters) should have more weight than a selection in the steepest part of the front. Weighting schemes were tested but they had little effect on the identification of dominant parameters.

Table 3 only presents the parameter importance for the CDD watershed using the NSE. However, all the above procedures were applied to both watersheds and for both efficiency metrics. Table 4 presents the selection ratios and classes of importance for the four considered combinations. A first examination of Table 4 reveals that five parameters (P1, P3, P5, P13, and P22) were consistently identified as the *most important* in all four cases. At the other end of the spectrum, parameters P7 and P9 were the *least important* for all cases. The importance of each parameter can be ordered according to the summation of selection ratios for both watersheds and efficiency metrics (the last row in Table 4).

Sobol' sensitivity analysis

The Sobol' sensitivity analysis method was also used to reduce the parametric dimensionality of the HSAMI model. A case study was conducted for the CDD watershed, using the NSE as the efficiency metric. The total-order SI reflects the full impact of each parameter on the model output and therefore those that are most relevant in model calibration (van Werkhoven *et al.* 2009). Thus, the total-order SI is used to determine the parameter importance.

Table 5 presents the total SI and its cumulative sum for all 23 parameters. If contributing 90% of the total output variance is used as a threshold to identify the important parameters, ten parameters have to be kept for model optimization. Comparing these ten parameters (the last

Table 4 │ Selection ratio for parameters being important for eight MOGA optimizations using the NSE and TRMSE for both watersheds

Watershed	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23
CDD	NSE	0.69	0.25	0.63	0.53	0.48	0.61	0.09	0.58	0.05	0.14	0.34	0.26	0.51	0.27	0.34	**0.18**	0.11	0.28	0.69	0.64	0.33	0.66	**0.19**
	TRMSE	0.71	0.34	0.73	0.67	0.54	0.31	0.13	0.39	0.03	0.19	**0.19**	0.35	0.41	0.51	0.38	0.12	0.48	0.49	0.59	0.26	0.50	0.62	0.09
Cowansville	NSE	0.72	**0.19**	0.55	0.70	0.73	0.31	0.08	0.52	0.06	0.41	0.69	0.52	0.60	0.03	0.59	0.30	0.10	**0.18**	0.31	**0.19**	0.09	0.53	0.07
	TRMSE	0.57	0.31	0.63	0.32	0.51	0.40	0.04	0.38	0.05	0.29	0.59	0.67	0.44	0.06	0.79	0.06	0.04	0.42	0.37	0.44	0.33	0.44	0.26
Sum		2.68	1.09	2.53	2.22	2.26	1.64	0.34	1.87	0.18	1.03	1.80	1.97	0.87	2.10	0.67	0.73	1.37	1.96	1.53	1.25	2.25	0.61	
Order		16	2	5	3	12	22	9	23	17	10	11	7	18	6	20	19	14	8	13	15	4	21	

Note: Roman, italic, bold, and underline denote the most important, important, average, and least important, respectively.

row of Table 5) from the Sobol' method with the ten important parameters from the MOGA (the last row of Table 4), it can be found that seven parameters are identified to be commonly important by both methods, even though the importance order is not exactly the same. Especially, eight parameters are identified to be commonly important for the same watershed and metric (row 2 of Table 4 versus the last row of Table 5).

To further understand the ability of Sobol' sensitivity analysis with respect to reducing the parametric dimensionality, the trade-off between the number of parameters and the NSE is presented for the CDD watershed. Figure 4 presents the mean value of 1-NSE over 100 calibrations with the number of calibrated parameters ranging from 22 to 1. Values of 1-NSE for the validation period are also presented. Similarly to the Pareto front presented in Figure 3, Figure 4 shows that the increase in the number of fixed parameters generally results in a reduction in model performance for the calibration period. In particular, the model performance degenerates considerably for both calibration and validation periods when the number of free parameters is less than 12. In other words, 11 parameters can be reduced with little loss in model performance. When keeping 12 parameters, MOGA and Sobol' methods show even more similarity with respect to identifying the commonly important parameters. Specifically, 11 out of 12 parameters are identified to be commonly important for the same watershed and metrics, as well as for different watersheds and metrics.

DISCUSSION AND CONCLUSION

This study presented an approach for assessing the performance of sensitivity analysis in reducing the parametric dimensionality during the hydrological model calibration process. A method was proposed based on MOGA and tested using the conceptual lumped rainfall–runoff model HSAMI with 23 free parameters over two Canadian watersheds in the Province of Quebec. The model was calibrated using two efficiency metrics: NSE and TRMSE. Results indicated that at least 11 parameters can be reduced with little degeneration in model performance for both watersheds.

Table 5 | Total order sensitivity index (SI) and its cumulative sum of Sobol' sensitivity analysis for 23 parameters of HSAMI using the NSE for the CDD watershed

Parameter	P9	P7	P18	P23	P16	P14	P10	P22	P2	P17	P21	P1	P15	P20	P12	P19	P8	P4	P11	P5	P3	P6	P13
Total order SI	0.000	0.000	0.001	0.001	0.004	0.005	0.006	0.007	0.009	0.013	0.013	0.014	0.020	0.022	0.032	0.040	0.057	0.078	0.087	0.097	0.150	0.160	0.185
Cumulative sum	0.000	0.000	0.001	0.002	0.005	0.010	0.016	0.022	0.031	0.044	0.057	0.071	0.091	0.113	0.145	0.184	0.242	0.320	0.408	0.505	0.655	0.815	1.000
Order	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

The proposed method was used to validate the Sobol' sensitivity analysis method in reducing the HSAMI parameter dimensionality. Both methods indicated that a nearly equivalent model performance could be preserved when the number of HSAMI parameters was reduced from 23 to 12. In particular, 11 out of 12 parameters are identified to be commonly important, even though the relative importance of some parameters differs between methods. However, the proposed method allowed even further reduction of the number of parameters with minimal performance loss. In particular, there was little degradation in model performance for the validation period when reducing the number of parameters to eight. The Sobol' method was significantly less successful at finding appropriate combinations of smaller numbers of parameters as indicated by a sharp decrease in model performance below 12 parameters. For small parameter sets, the MOGA algorithm has the advantage of being able to find independent optimal combinations at each step in the parameter reduction process, whereas Sobol' analysis has to keep the parameters fixed in previous steps. In other words, the MOGA method is free to drop parameters fixed at an earlier step if advantageous. For example, a given parameter may be considered important for the 12-parameter version of the model, only to be dropped and replaced by two other parameters in the 11-parameter version.

Although the proposed method displayed significant benefits for a large reduction of the parameter-space, the advantages of Sobol' are numerous, such as execution time, ability to understand interactions and quantification of the variance explanation for each parameter. Of course, there are some drawbacks with the Sobol' analysis, such as the definition of the boundary space. For example, if a parameter has an *a priori* unknown value, setting the boundaries too narrow or too wide will result in an under (over)-estimation of the parameter's importance. Therefore care must be taken in setting the parameter boundaries. It is important to note that these drawbacks are effectively present in all parameter reduction methods. Also, in a case where two parameters are strongly or perfectly correlated, the analysis method will set both parameters to equal levels of importance. However, in reality, one of the parameters could be fixed without harming the model's performance. Fundamental research is ongoing to address
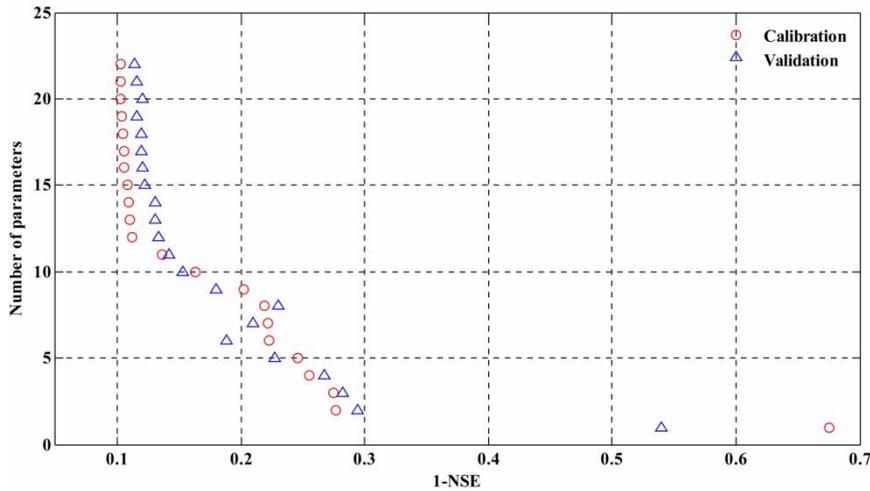
Figure 4 | Mean value of 1-NSE for 100 calibrations and validations with free parameters decreasing from 22 to 1. The non-calibrated parameters were fixed with random numbers. The fixed parameters were determined based on their contributions to total output variance (from the least to the most) according to Sobol' sensitivity analysis.

these issues (Chastaing *et al.* 2014), but this work has shown that the Sobol' method's limitations should not be problematic for hydrological model parameter dimensionality reduction in the interim.

One commonly raised concern in parameter fixing studies is that it could be possible to fix a parameter that turns out to be important for some processes but does not have a large effect on the NSE or TRMSE. While it is possible for this to happen, if the parameter has a significant impact on NSE then it will be considered influential and will be fixed at a later stage. On the other hand, the aim of the project is to remove parameters that are least important, therefore there is the implicit agreement that there will be a loss of performance with parameter fixing. And while the performance is measured through an aggregate measure such as NSE, which implies a loss of qualitative information on hydrological processes, it is important to note that the conceptual model itself is a powerful filter between simulated and observed processes. Parameter reduction has its drawbacks, one of which is a loss in process simulation and analysis.

It should be pointed out that the computational cost is the major disadvantage of the multi-objective-based method (Zhan *et al.* 2013). The lumped hydrological model with 23 free parameters used in this study was highly optimized, and one run of the model for a 10-year period took about one-tenth of a second on a single core of a modest processor. The multi-objective optimization algorithm demands thousands of calibrations to be performed, each demanding in turn several thousand model evaluations. Almost 1 week of computational time (one-core equivalent on a modest processor) was required for one multi-objective optimization. Consequently, the computational burden would be very high for models that take a longer time to run. A model that takes one second per evaluation over a 10-year period would require about 70 computational days on a single core. A shorter time period and using a smaller subset of model parameters could be used to counter this problem. Nonetheless, the computational burden would be high for complex models. From this point of view, the proposed approach would not be efficient for complex distributed hydrological models without access to parallel computing power. However, access to eight or 12 cores is relatively simple nowadays on a single desktop machine, and the proposed approach remains feasible with many hydrological models. Moreover, multi-objective optimization algorithms are very efficient at using parallel computing, and access to high-performance computing is getting relatively easier, especially with the advent of graphics processing unit computing and its accompanying breakthroughs. As such, this should not be viewed as an insurmountable problem.

The need to run several multi-objective optimizations to overcome the problem of choosing values for the parameters when not used in the calibration also adds to the computational burden. A possible option to circumvent this

problem is to deliberately use 'bad' fixed values for all parameters. By using known 'bad' values for each parameter, it is possible to avoid the pitfall of having one parameter deemed unimportant (i.e., its fixed *a priori* value was too good to begin with). The choice is obvious for several parameters, but not as clear for parameters whose uncertainty covers the entire search space. While this approach could have allowed only one multi-objective calibration to be done (per basin and per optimization metric), the underlying need to run hundreds of calibrations to determine the 'bad' parameter zones within the search space partly negates the advantage of this option. It was felt more appropriate to use a random assignation to remove the possible bias in parameter value selection.

A few important considerations must be taken with respect to the methodology used in this paper. First, the MOGA optimization algorithm was set up using hyperparameters defined by users' experiences with the algorithm and the model. Other hydrological models (depending on their complexity) could require different population/generation ratios. In this case, the 10:1 ratio was considered adequate to obtain a good exploration of the space while permitting some exploitation for refining the results. The number of generations and population size were computed according to the maximum allotted time for performing the computations. If more computing power was available, more model evaluations would have been performed. The final point to consider is the stochastic nature of model calibration. Uncertainty is also present due to initial conditions that the calibration algorithm uses. The initial seed value was left variable to ensure this uncertainty is present when calibrating the models. This gives more confidence, since there was no bias in 'forcing' the model into a specific region due to a particularly advantageous fixed seed.

Although the approach presented herein is based on two-objective metrics, other similar yet more targeted approaches could be introduced (i.e., exploring a three-objective space). Both the NSE and TRMSE could be minimized as well as the number of parameters. This would be a more objective way of defining the importance of each parameter (Rosolem *et al.* 2012, 2013). This would require more computing power, but the uncertainty would also be reduced when using the parameters in the validation space, as is the case when working on a single objective. Other approaches could use constraints to bind certain parameters together if they are known to be independent. This would reduce the search space while maintaining the same objectives.

ACKNOWLEDGEMENTS

REFERENCES

Arsenault, R. & Brissette, F. P. 2014 Continuous streamflow prediction in ungauged basins: the effects of equifinality and parameter set selection on uncertainty in regionalization approaches. *Water Resour. Res.* **50** (7), 6135–6153.

Arsenault, R., Malo, J. S., Brissette, F., Minville, M. & Leconte, R. 2013 Structural and non-structural climate change adaptation strategies for the Péribonka water resource system. *Water Resour. Manage.* **27** (7), 2075–2087.

Arsenault, R., Poulin, A., Côté, P. & Brissette, F. 2014 A comparison of stochastic optimization algorithms in hydrological model calibration. *J. Hydrol. Eng.* **19** (7), 1374–1384.

Arsenault, R., Poissant, D. & Brissette, F. 2015 Parameter dimensionality reduction of a conceptual model for streamflow prediction in Canadian, snowmelt dominated ungauged basins. *Adv. Water Resour.* **85**, 27–44.

Baroni, G. & Tarantola, S. 2014 A general probabilistic framework for uncertainty and global sensitivity analysis of deterministic models: a hydrological case study. *Environ. Modell. Software* **51**, 26–34.

Bastidas, L. A., Gupta, H., Sorooshian, S., Suttleworth, W. J. & Yang, Z. L. 1999 Sensitivity analysis of a land surface scheme using multicriteria methods. *J. Geophys. Res.* **104** (D16), 19481–19490.

Beven, K. & Binley, A. 1992 The future of distributed models–model calibration and uncertainty prediction. *Hydrol. Process.* **6** (3), 279–298.

Chastaing, G., Prieur, C. & Gamboa, F. 2014 Generalized Sobol sensitivity indices for dependent variables: numerical methods. *J. Statist. Comput. Simulation* **85** (7), 1306–1333.

Chen, J., Brissette, F. P., Poulin, A. & Leconte, R. 2011a Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resour. Res.* **47** (12), 1–16.

Chen, J., Brissette, F. P. & Leconte, R. 2011b Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *J. Hydrol.* **401**, 190–202.

Chen, J., Brissette, F. P. & Leconte, R. 2012 Downscaling of weather generator parameters to quantify the hydrological impacts of climate change. *Clim. Res.* **51**, 185–200.

Cox, G. M., Gibbons, J. M., Wood, A. T. A., Craigon, J., Ramsden, S. J. & Crout, N. M. J. 2006 Towards the systematic simplification of mechanistic models. *Ecol. Model.* **198** (1–2), 240–246.

Deb, K. 2001 *Multi-Objective Optimization using Evolutionary Algorithms*, 1st edn. Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons, Chichester, UK, p. 497.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evolut. Comput. **6** (2), 182–197.

Ebel, B. & Loague, K. 2006 Physics-based hydrologic-response simulation: seeing through the fog of equifinality. *Hydrol. Process.* **20**, 2887–2900.

Efstratiadis, A. & Koutsoyiannis, D. 2010 One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrol. Sci. J.* **55** (1), 58–78.

Fortin, V. 2000 *Le modèle météo-apport HSAMI: historique, théorie et application*. Institut de Recherche d'Hydro-Québec, Varennes, Canada, p. 68.

Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C. & Di, Z. 2014 A comprehensive evaluation of various sensitivity analysis methods: a case study with a hydrological model. *Environ. Modell. Software* **51**, 269–285.

Gao, X., Sorooshian, S. & Gupta, H. V. 1996 Sensitivity analysis of the biosphere-atmosphere transfer scheme. *J. Geophys. Res.* **101**, 7279–7289.

Hansen, N. & Ostermeier, A. 1996 Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In: *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, Nagoya, Japan, pp. 312–317.

Hansen, N. & Ostermeier, A. 2001 Completely derandomized self-adaptation in evolution strategies. *Evolut. Comput.* **9** (2), 159–195.

Huang, M. Y. & Liang, X. 2006 On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged basins. *J. Hydrol.* **320** (1–2), 37–61.

Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E. & Papadopol, P. 2009 Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961–2003. *J. Appl. Meteorol. Climatol.* **48**, 725–741.

Kay, A. L., Davies, H. N., Bell, V. A. & Jones, R. G. 2009 Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Clim. Change* **92**, 41–63.

Klepper, O., Scholten, H. & van de Kamer, J. P. G. 1991 Prediction uncertainty in an ecological model of the Oosterschelde Estuary. *J. Forecasting* **10**, 191–209.

Konak, A., Coit, D. W. & Smith, A. E. 2006 Multi-objective optimization using genetic algorithms: a tutorial. *Reliability Eng. Syst. Safety* **91**, 992–1007.

Lee, H., McIntyre, N., Wheater, H. & Young, A. 2005 Selection of conceptual models for regionalisation of the rainfall-runoff relationship. *J. Hydrol.* **312** (1–4), 125–147.

Minville, M., Brissette, F. & Leconte, R. 2008 Uncertainty of the impact of climate change on the hydrology of a Nordic watershed. *J. Hydrol.* **358**, 70–83.

Minville, M., Brissette, F., Krau, S. & Leconte, R. 2009 Adaptation to climate change in the management of a Canadian water-resources system exploited for hydropower. *Water Resour. Manage.* **23** (14), 2965–2986.

Myung, J. & Pitt, M. A. 2002 When a good fit can be bad. *Trends Cogn. Sci.* **6**, 421–425.

Nash, J. E. & Sutcliffe, W. H. 1970 River flow forecasting through conceptual models: Part 1. A discussion of principles. *J. Hydrol.* **10**, 282–290.

Nossent, J., Elsen, P. & Bauwens, W. 2011 Sobol' sensitivity analysis of a complex environmental model. *Environ. Modell. Software* **26** (12), 1515–1525.

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

Pechlivanidis, I. G., McIntyre, N. & Wheater, H. S. 2010 Calibration of the semi-distributed PDM rainfall-runoff model in the Upper Lee catchment, UK. *J. Hydrol.* **386** (1–4), 198–209.

Pechlivanidis, I. G., Jackson, B., McIntyre, N. & Wheater, H. S. 2011 Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST J.* **13** (3), 193–214.

Pitman, A. J. 1994 Assessing the sensitivity of a land-surface scheme to the parameter values using a single column model. *J. Clim.* **7** (12), 1856–1869.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R. & Malo, J. S. 2011 Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *J. Hydrol.* **409** (3–4), 626–636.

Reed, P. M., Hadka, D., Heman, J. D., Kasprzyk, J. R. & Kollat, J. B. 2013 Evolutionary multiobjective optimization in water resources: the past, present, and future. *Adv. Water Resour.* **51**, 438–456.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X. & de Gonçalves, L. G. G. 2012 A fully multiple-criteria

implementation of the Sobol' method for parameter sensitivity analysis. *J. Geophys. Res.* **117** (D7), D07103.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Gonçalves, L. G. G. & Zeng, X. 2013 Towards a comprehensive approach to parameter estimation in land surface parameterization schemes. *Hydrol. Process.* **27** (14), 2075–2097.

Saltelli, A. 2002 Sensitivity analysis for importance assessment. *Risk Anal.* **22**, 579–590.

Sawaragi, Y., Nakayama, H. & Tanino, T. 1985 Theory of Multiobjective Optimization *(vol. 176 of Mathematics in Science and Engineering)*. Academic Press, Orlando, FL, USA.

Shafii, M. & De Smedt, F. 2009 Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm. *Hydrol. Earth Syst. Sci.* **13**, 2137–2149.

Sobol', I. M. 1993 Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.* **1**, 407–417.

Song, X. M., Zhan, C. S. & Xia, J. 2012 Integration of a statistical emulator approach with the SCE-UA method in parameter optimization of a hydrological model. *Chin. Sci. Bull.* **57** (26), 3397–3403.

Song, X. M., Zhang, J. Y., Zhan, C. S., Xu, Y. Q., Ye, M. & Xu, C. G. 2015 Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework and applications. *J. Hydrol.* **523**, 739–757.

Steuer, R. E. 1986 *Multiple Criteria Optimization: Theory, Computations, and Application*. John Wiley & Sons, New York, USA.

Tang, Y., Reed, P., Wagener, T. & van Werkhoven, K. 2007 Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrol. Earth Syst. Sci.* **11** (2), 793–817.

Tonkin, M. J. & Doherty, J. 2005 A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour. Res.* **41**, W10412.

van Straten, G. & Keesman, K. J. 1991 Uncertainty propagation and speculation in projective forecasts of environmental change: a lake-eutrophication example. *J. Forecasting* **10**, 163–190.

van Werkhoven, K., Wagener, T., Reed, P. & Tang, Y. 2008 Characterization of watershed model behavior across a hydroclimatic gradient. *Water Resour. Res.* **44**, W01429.

van Werkhoven, K., Wagener, T., Reed, P. & Tang, Y. 2009 Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Adv. Water Resour.* **32**, 1154–1169.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. & Sorooshian, S. 2003 Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.* **39** (8), 1214.

Wagener, T. & Kollat, J. 2007 Numerical and visual evaluation of hydrologic and environmental models using the Monte Carlo Analysis Toolbox (MCAT). *Environ. Modell. Software* **22**, 1021–1033.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. & Sorooshian, S. 2001 A framework for development and application of hydrological models. *Hydrol. Earth Syst. Sci.* **5** (1), 13–26.

Wilby, R. L. & Harris, I. 2006 A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK. *Water Resour. Res.* **42**, W02419.

Wilson, M. F., Henderson-Sellers, A., Dickinson, R. E. & Kennedy, P. J. 1987a Sensitivity of the biosphere-atmosphere transfer scheme (BATS) to the inclusion of variable soil characteristics. *J. Clim. Appl. Meteorol.* **26**, 341–362.

Wilson, M. F., Henderson-Sellers, A., Dickinson, R. E. & Kennedy, P. J. 1987b Investigation of the sensitivity of the land surface parameterization of the NCAR Community Climate model in regions of tundra vegetation. *Int. J. Climatol.* **7**, 319–343.

Wurbs, R. A. 1998 Dissemination of generalized water resources models in the United States. *Water Int.* **23**, 190–198.

Yapo, P. O., Gupta, H. V. & Sorooshian, S. 1996 Calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *J. Hydrol.* **181**, 23–48.

Yapo, P. O., Gupta, H. V. & Sorooshian, S. 1998 Multi-objective global optimization for hydrologic models. *J. Hydrol.* **204**, 83–97.

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. Software* **41**, 39–52.

Zhang, C., Chu, J. G. & Fu, G. T. 2013 Sobol's sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. *J. Hydrol.* **480**, 58–68.

Zhang, J. & Ross, M. 2015 Comparison of IHM and MIKE SHE model performance for modeling hydrologic dynamics in shallow water table settings. *Vadose Zone J.* **14** (7).