Calibration of Water Distribution Hydraulic Models Using a Bayesian-Type Procedure

Zoran S. Kapelan[1]; Dragan A. Savic[2]; and Godfrey A. Walters[3]

Abstract: Estimating model parameters is a difficult, yet critical step in the use of water distribution system models. Most of the optimization-based approaches developed so far concentrate primarily on efficient and effective ways of obtaining optimal calibration parameter values. At the same time, very little effort has been made to determine the uncertainties (i.e., errors) associated with those values (and related model predictions). So far, this has typically been done using the first-order second moment (FOSM) method. Even though reasonably computationally efficient, the FOSM approach relies on several restrictive assumptions and requires computationally demanding calculation of derivatives. To overcome these limitations, the recently developed shuffled complex evolution metropolis (SCEM-UA) global optimization algorithm is linked to the Epanet2 hydraulic model and used to solve a least-squares-type calibration problem. The methodology is tested and verified on the Anytown literature case study. The main advantage of the SCEM-UA algorithm over existing approaches is that both calibration parameter values and associated uncertainties can be determined in a single optimization model run. In addition, no model linearity or parameter normality assumptions have to be made nor any derivatives calculated. The main drawback of the SCEM-UA methodology is that it could, potentially, be computationally demanding, although this is not envisaged as a major problem with current computers.

DOI: 10.1061/(ASCE)0733-9429(2007)133:8(927)

CE Database subject headings: Bayesian analysis; Calibration; Markov chains; Optimization; Pipe networks; Water distribution systems.

Introduction

Various hydraulic simulation models are widely used nowadays by planners, water utilities, consultancy companies, and many others involved in analysis, design, operation, or maintenance of water distribution systems. In order to make a hydraulic model useful, it is necessary to calibrate it first (Walski 1983). Calibration of pipe network models consists of determining the physical and operational characteristics of an existing system. This is achieved by determining various parameters that when input into a hydraulic simulation model will yield a reasonably good match between measured and predicted variables (Shamir and Howard 1968).

Initial water distribution system (WDS) calibration methodologies were based on various trial and error procedures (Bhave 1988; Rahal et al. 1980; Walski 1983). Shortly after that, more systematic, explicit-type calibration approaches were introduced (Boulos and Wood 1990; Ormsbee and Wood 1986). These approaches were soon replaced with "automatic," optimization-based calibration methodologies (Lansey and Basnet 1991; Ormsbee 1989). However, most (if not all) of the optimization-based WDS calibration approaches developed so far have focused primarily on the most computationally efficient and effective way of obtaining the optimal calibration parameter values. At the same time, very little effort has been made to determine the uncertainties (i.e., errors) associated with these values and related model predictions.

So far, uncertainty quantification has typically been done using the first-order second moment (FOSM) method (Bush and Uber 1998; Kapelan et al. 2003; Lansey et al. 2001) in a post-calibration-type procedure. Note that, even though reasonably computationally efficient, the FOSM approach relies on several underlying assumptions that are not necessarily true in the WDS modeling case. These include: (1) model linearity; (2) independence of calibration parameter values and measurement errors; and (3) normality of calibration parameter values and measurement errors. The methodology also requires calculation of derivatives of model-dependent variables with respect to calibration parameters that may be computationally demanding and prone to numerical errors.

The aim of this paper is to develop and present an alternative calibration approach to overcome some of the above deficiencies. This paper is organized as follows: after this introduction, the WDS model calibration problem is formulated. This is followed by a description of the shuffled complex evolution metropolis (SCEM-UA) optimization methodology. This methodology is then applied to both literature and real-life case studies. Finally, relevant conclusions are drawn.

[1]Senior Lecturer, Dept. of Engineering, School of Engineering, Computer Science and Mathematics, Univ. of Exeter, North Park Rd., Exeter, EX4 4QF, U.K. E-mail: Z.Kapelan@exeter.ac.uk

[2]Professor, Dept. of Engineering, School of Engineering, Computer Science and Mathematics, Univ. of Exeter, North Park Rd., Exeter, EX4 4QF, U.K. E-mail: D.Savic@exeter.ac.uk

[3]Professor, Dept. of Engineering, School of Engineering, Computer Science and Mathematics, Univ. of Exeter, North Park Rd., Exeter, EX4 4QF, U.K. E-mail: G.A.Walters@exeter.ac.uk

Note. Discussion open until January 1, 2008. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on July 5, 2005; approved on December 8, 2006. This paper is part of the ***Journal of Hydraulic Engineering***, Vol. 133, No. 8, August 1, 2007. ©ASCE, ISSN 0733-9429/2007/8-927–936/$25.00.

WDS Model Calibration

All WDS hydraulic models map model inputs (e.g., demands, network configuration data, etc.) into model outputs (nodal pressure/total heads and link flows). This mapping involves also a number of parameters (e.g., pipe roughness coefficients) whose values are usually determined in the calibration process. Let $\mathbf{x}$ denote a vector of all WDS hydraulic model inputs; $\mathbf{y}$, a vector of model outputs; and $\mathbf{a}$, a vector of N_a model parameters. A WDS hydraulic model can then be defined as $\mathbf{y}=f(\mathbf{x}|\mathbf{a})$ where f=function denoting the aforementioned mapping. Given vector $\mathbf{y}^*$ of N_o observed (i.e., measured) output values and the vector of errors (i.e., residuals) $\mathbf{e}$, the following can be written:

$$\mathbf{y}^* = f(\mathbf{x}|\mathbf{a}) + \mathbf{e} \tag{1}$$

The above equation represents the fundamental formulation of a nonlinear regression approach. Assuming that residuals $\mathbf{e}$ are mutually independent, each following an exponential power density function $Ex(\sigma,\gamma)$ (encapsulating simultaneously several error types, see below), the likelihood that a parameter vector $\mathbf{a}$ describes the observed data $\mathbf{y}^*$ can be computed as follows (Box and Tiao 1973):

$$p(a|y^*,\gamma) = \left(\frac{\omega(\gamma)}{\sigma}\right)^{N_o} \exp\left[-c(\gamma)\sum_{i=1}^{N}\left|\frac{e_i(a)}{\sigma}\right|^{2/1+\gamma}\right] \tag{2}$$

where

$$\omega(\gamma) = \frac{\{\Gamma[3(1+\gamma)/2)]^{1/2}}{(1+\gamma)\{\Gamma[(1+\gamma)/2)]^{3/2}}$$

$$\text{and} \quad c(\gamma) = \left\{\frac{\Gamma[3(1+\gamma)/2]}{\Gamma[(1+\gamma)/2]}\right\}^{1/1+\gamma}$$

where exp=exponential function; and Γ=Gamma function. Parameter γ is used to specify the error model type. The residuals $\mathbf{e}(\mathbf{a})$ are assumed to be normally distributed when $\gamma=0$, double exponentially distributed when $\gamma=1$, and tend to uniform distribution when $\gamma\to-1$. Assuming a noninformative prior PDF of the form $p(\mathbf{a})\sim\sigma^{-1}$, Box and Tiao (1973) showed that the influence of parameter σ can be integrated out, leading to the following form of the posterior PDF of $\mathbf{a}$:

$$p(a|y^*,\gamma) \propto \left[\sum_{i=1}^{N_o}|e_i(a)|^{2/(1+\gamma)}\right]^{-N_o(1+\gamma)/2} \tag{3}$$

In general, the objective of the calibration problem is to maximize the likelihood defined in Eq. (2) or (3). In the case of assumed Gaussian residuals ($\gamma=0$), the above likelihood maximization reduces to an equivalent minimization of the sum of squared errors, i.e., it reduces to the simple least-squares problem

$$\text{Minimize } E = \mathbf{e}^T\mathbf{e} = \sum_{i=1}^{N_o} e_i^2 \tag{4}$$

where E=scalar objective function denoting the total calibration error; and T is the vector transpose operator.

In the methodology presented here, the WDS calibration problem is formulated as an optimization problem as in Eq. (4), subject to the following sets of constraints: (1) all WDS hydraulic model equations (mass and energy balance) must be satisfied; and (2) values of all calibration parameters must be between corresponding, prespecified search bounds. In addition, a uniform prior PDF is assumed here for each calibration parameter considered. Bounds on the uniform PDF are equal to the parameter search bounds. This, however, is not essential for the application of the calibration methodology described here (any prior PDF can be used to describe beliefs in the calibration parameter values).

In all the case studies shown here pipe roughness coefficients are the calibration parameters and observed data consist of measured pressures heads only. This, however, does not mean that the calibration methodology presented here cannot be used to calibrate for other unknown model parameters (e.g., nodal demands, pipe diameters, etc.) or that it cannot use other measurements available (e.g., flows).

SCEM-UA Algorithm

The above calibration optimization problem is solved here using the SCEM-UA algorithm (Vrugt et al. 2003). The SCEM-UA is a Bayesian recursive algorithm where during the search, parameter PDF functions are constantly updated until convergence to the stationary PDFs is achieved. The algorithm builds upon the SCE-UA global optimization methodology (Duan et al. 1992) that determines the values of calibration parameters only (e.g., like genetic algorithms). SCE-UA has been used successfully for the deterministic calibration and sampling design of transient WDS hydraulic models (Vitkovsky et al. 2003). When compared to SCE-UA, the SCEM-UA algorithm incorporates three improved features (Vrugt et al. 2003): (1) the downhill simplex method (Press et al. 1990) is replaced by the Metropolis annealing algorithm (Metropolis et al. 1953); (2) complexes are not further subdivided into subcomplexes when generating off-spring (i.e., candidate points); and (3) a different replacement procedure is used. The modifications were introduced in the SCEM-UA algorithm to prevent the search from concentrating on a localized region of calibration parameter values, i.e., so that the SCEM-UA search algorithm becomes capable of identifying the whole posterior target PDF.

The SCEM-UA search process starts by generating N_s random samples of unknown calibration parameters $[\mathbf{a_1}^{(0)},\mathbf{a_2}^{(0)},\ldots,\mathbf{a_{Ns}}^{(0)}]^{\mathrm{T}}$ from the corresponding (assumed) marginal prior distributions followed by the computation of associated posterior densities $p(\mathbf{a}|\mathbf{y}^*,\gamma)$ using either Eq. (2) or (3). The sampled points are then ordered by decreasing posterior density and used to initialize N_q parallel sequences $\mathbf{S}_k(k=1,2,\ldots,N_q)$ where the kth ranked point is used as the starting point in the kth sequence. The set of all samples is also subdivided into N_q complexes $\mathbf{C}_k$ $(k=1,2,\ldots,N_q)$, each containing $N_m=N_s/N_q$ points such that kth complex contains every $N_q(i-1)+k$ previously ranked point $(k=1,2,\ldots,N_q;\ i=1,2,\ldots,N_m)$. Note that the use of complexes enables the collection of information gained about the search space by each parallelized sequence. Each parallel sequence represents a separate search thread (equivalent to a single Markov chain), thus enabling an independent exploration of the search space and hence reducing the chance of the search algorithm converging to a local (rather than the global) optimum.

Once all the sampled points are subdivided into complexes $\mathbf{C_k}$ and each parallel sequence $\mathbf{S_k}$ is initialized, the sequences are evolved using the sequence evolution metropolis (SEM) algorithm. The SEM algorithm is a custom made algorithm whose development is motivated by the well-known Metropolis annealing algorithm (Metropolis et al. 1953). The SEM algorithm generates new candidate points in each of the parallel sequences by generating draws from an adaptive multinormal proposal distribution. Parameters of this distribution (mean and the covariance

J. Hydraul. Eng. 2007.133:927-936.

Table 1. Recommended Values of the SCEM-UA Algorithm Parameters (Vrugt et al. 2003)

Number	Parameter	Description	Recommended value(s)
1	N_s	Number of samples	$\leq$100 for simple problems
2	N_q	Number of parallel sequences/complexes	$\geq$250 for complex problems $\leq$5 for simple problems $\geq$10 for complex problems
3	T	Threshold	10^6–10^7
4	c_n	Jump rate	1.0 or $2.4/\sqrt{N_a}$
5	N_L	Number of evolution steps in each sequence before reshuffling	$N_m/10$ to $N_m/5$

Note: Problems with independent or correlated Gaussian target distribution are problems.

matrix) are updated during the search process based on the information embedded in complexes. After generating the new candidate point, its posterior density is computed using either Eq. (2) or (3) and the Metropolis annealing criterion is used to decide whether the candidate point should be added to the sequence or not. When the candidate point is accepted it is added to the sequence and it replaces the best member of the associated complex. When the candidate point is rejected, it replaces the worst (i.e., N_mth) point in the associated complex providing that (1) the ratio of the posterior density of the first (best) to the last (worst) member of that complex is larger than some prespecified threshold value T; and (2) the posterior density of the candidate point is larger than the posterior density of the worst member of the complex. More details of the SEM algorithm can be found in Vrugt et al. (2003).

Once each parallel sequence is evolved by using the SEM algorithm, the complexes are reshuffled by unpacking, i.e. grouping them together and then reordering by decreasing posterior density. The associated parallel sequences are then reinitialized according to the previously described procedure. At this point, the algorithm convergence is checked by using a convergence criterion (see below). If this criterion is not satisfied then another iteration of the SCEM-UA algorithm is performed by reshuffling complexes and evolving parallel sequences. Otherwise, the search process is stopped.

The SCEM-UA algorithm convergence is checked by using the Gelman–Rubin (Gelman and Rubin 1992) convergence statistic and/or simply by checking whether the prespecified number of calibration function evaluations has been reached (note that neither criterion is an exact one). The Gelman–Rubin (Gelman and Rubin 1992) convergence statistic is defined as follows:

$$\sqrt{\mathrm{SR}} = \sqrt{\frac{N_g - 1}{N_g} + \frac{N_q + 1}{N_q N_g} \cdot \frac{B}{W}} \tag{5}$$

where N_g=number of iterations within each sequence $\mathbf{S_k}$ ($k=1,2,\ldots,N_q$); B=variance between the N_q sequence means; and W=average of the N_q within-sequence variances. Note that a separate $\sqrt{\mathrm{SR}}$ value is calculated for each calibration parameter. The SCEM-UA algorithm run is considered (approximately) converged when the value of the Gelman–Rubin convergence statistic is less than or equal to 1.2 for all calibration parameters (Vrugt et al. 2003).

The SCEM-UA algorithm has several parameters that need to be specified prior to its use: number of samples, i.e., population size N_s, number of sequences/complexes N_q, threshold likelihood ratio T, jump rate c_n and the number of evolution steps N_L taken by each sequence in the SEM algorithm before complexes are reshuffled. The values recommended by Vrugt et al. (2003) can be found in Table 1. According to the SCEM-UA authors, these values should be valid for a wide range of different calibration problems. This fact was further tested in the preliminary analyses of the case studies considered here by performing a number of trial SCEM-UA runs. The SCEM-UA parameter values selected in this process are shown in the relevant case studies sections (see below). As can be seen, these values are compatible with the suggestions made in Table 1.

The SCEM-UA is, fundamentally, a Markov chain Monte Carlo (MCMC) sampler. Therefore, the candidate point $\mathbf{a}^{(t+1)}$ is generated from point $\mathbf{a}^{(t)}$ only, using some proposal distribution (multinormal distribution here, see above). The most general and earliest MCMC sampler is the Metropolis–Hastings (MH) algorithm (Hastings 1970; Metropolis et al. 1953). When compared to the local search MH algorithm, the main advantage of the global search SCEM-UA algorithm is that it is more likely to identify the globally optimal posterior target PDFs of calibration parameters. In addition, it does not require good prior (i.e., close to target) estimates of calibration parameter PDFs to be able to converge. The above is the consequence of the better mixing of Markov chains in the SCEM-UA algorithm. These abilities, however, come at a price: the SCEM-UA algorithm is more computationally demanding than the MH algorithm.

Case Studies

Case 1: Literature Calibration Problem

Problem Description

The aim of the analysis presented here is to test the performance of the SCEM-UA algorithm on the "Anytown" network model

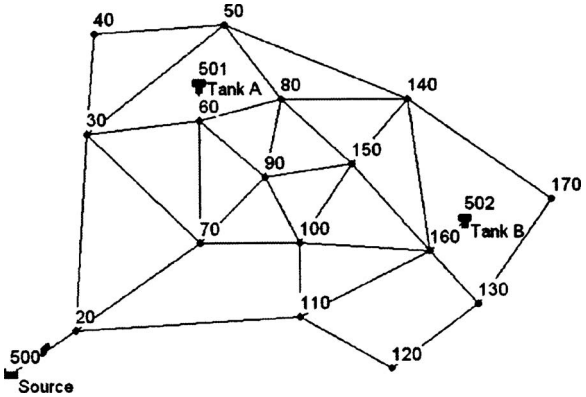

Fig. 1. Case 1: Anytown network layout

J. Hydraul. Eng. 2007.133:927-936.

Table 2. Case 1.1: Calibration Parameter Statistics

Parameter ID	True value	Search limits		SCEM-UA			
		Low	High	Mean	SD	2.5 (percentile)	97.5 (percentile)
P1	120	50	150	120.01	0.01	119.99	120.04
P2	70	50	150	70.01	0.01	69.99	70.02
P3	90	50	150	89.93	0.05	89.85	90.03
P4	130	50	150	130.08	0.06	129.97	130.16

Note: SD=standard deviation.

(Walski et al. 1987). This network (see Fig. 1) has been used previously to test different calibration models (Ferreri et al. 1994; Kapelan et al. 2005; Lansey and Basnet 1991; Ormsbee 1989). Even though the Anytown calibration problem is artificial it is considered to be close to reality in terms of the principal network layout (city center area with older pipes with low Hazen–Williams coefficients and boundary, residential areas with newer pipes with high Hazen–Williams coefficients) and the use of fire flow test data as calibration measurement data. The Anytown steady-state hydraulic model is calibrated for five loading conditions (average demand, peak demand, and three fire flow conditions). The model and network configuration details used here can be found in Ormsbee (1989).

A total of $N_a=4$ grouped Hazen–Williams pipe roughness coefficients are considered as calibration parameters. The true (i.e. assumed) parameter values are shown in Tables 2 and 3. Parameter grouping details can be found in Ormsbee (1989). All calibration parameters are assumed to follow the uniform prior PDF defined on interval [50,150]. The observed information consists of "measured" heads at four network nodes (40, 90, 120 and 140) for all five loading conditions (i.e., $N_o=20$). The Epanet2 hydraulic solver (Rossman 2000) is used to solve the steady-state hydraulics. The noninformative PDF defined in Eq. (3) and for $\gamma=0$ has been used here to drive the SCEM-UA algorithm search for the posterior target distribution. The least-squares calibration problem defined in Eq. (4) has been solved.

The following two cases are analyzed: (1) Case 1.1: perfect head measurement data; and (2) Case 1.2: imperfect head measurement data.

Case 1.1: Perfect Measurement Data

The objective of this case study is to test the performance of the SCEM-UA algorithm assuming perfect head "measurement" data. The perfect measurement data are created by running the Epanet2 hydraulic solver for each of the five aforementioned loading conditions using the "true" (i.e., assumed) calibration parameter values. The following SCEM-UA parameters are used: $N_s=100$, $N_q=5$, $T=10^6$, $c_n=1.0$, and $N_L=4$. The SCEM-UA run is stopped after 2,000 objective function evaluations.

The value of the Gelman–Rubin (Gelman and Rubin 1992) convergence statistic for all four calibration parameters is shown in Fig. 2. As can be seen, at the end of the SCEM-UA run, the value of this statistic is well below the target convergence level (1.2) for all calibration parameters, indicating successful algorithm convergence to a posterior PDF. Note also from Fig. 2 that stopping the SCEM-UA run immediately after the Gelman–Rubin statistic reaches the threshold value of 1.2 for all parameters considered (this happens after approximately 700 objective function evaluations in the example shown here) may not always be reliable due to potential oscillations that may (sometimes) occur later on in the search process (at around 1,200 evaluations in the example shown here). The additional runs performed (up to 2,000 evaluations) are used to make sure that this does not happen again. Note that additional evaluations performed should not be seen as a waste of computational resources as they can always be used as additional samples for the estimation of joint posterior PDF (not done here).

Optimal calibration parameter values (means), the associated uncertainty indicators (standard deviations), and the corresponding 95% confidence intervals are shown in Table 2. As can be seen, the mean parameter values obtained are almost identical to the assumed calibration parameter values. At the same time, standard deviations of all parameters are very low and the corresponding 95% confidence intervals are narrow. All of these findings indicate that the Bayesian recursive procedure implemented in the SCEM-UA algorithm successfully replicates the true calibration parameter values, even though the number of samples used is small (only 100).

Case 1.2: Imperfect Measurement Data

The objective of this case study is to further verify the performance of the SCEM-UA algorithm by calibrating the Anytown network model using imperfect head measurements and by comparing the SCEM-UA results to the results obtained using Monte Carlo simulation (MCS).

The perfect head measurements used in case 1.1 are first corrupted by introducing random normal noise with zero mean and standard deviation of 0.10 m. The Anytown calibration problem is

Table 3. Case 1.2: Calibration Parameter Statistics

Parameter ID	True value	Search Limits		SCEM-UA		MCS	
		Low	High	Mean	SD	Mean	SD
P1	120	50	150	119.8	0.39	119.8	0.48
P2	70	50	150	69.7	0.22	69.8	0.30
P3	90	50	150	89.4	1.24	89.3	1.54
P4	130	50	150	130.8	1.42	130.9	1.72

Note: SD=standard deviation.

J. Hydraul. Eng. 2007.133:927-936.

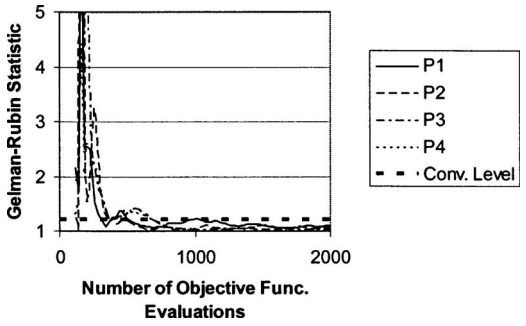

Fig. 2. Case 1.1: SCEM-UA algorithm convergence

then solved using both the SCEM-UA algorithm and the MCS analyses. In the MCS analyses, 10,000 random observed head sets are first created by adding normal noise with zero mean and 0.10 m standard deviation to the previously defined imperfect head measurements. The calibration parameter PDFs are then obtained by solving the simple least squares calibration problem for each random head set. This is done by applying repeatedly the improved version of the Levenberg-Marquardt search algorithm (Kapelan 2002). Once the 10,000 sets of calibration parameter values are obtained, the corresponding 10,000 pressure/total head prediction sets are obtained by repeatedly running the Epanet2 hydraulic solver.

The following SCEM-UA parameters are used in the above analysis: $N_s=1{,}000$, $N_q=5$, $T=10^6$, $c_n=1.0$, and $N_L=40$. The SCEM-UA run was stopped after 10,000 objective function evaluations (total computational time of approximately 12 s on a 3.0 GHz personal computer). When compared to Case 1.1, the larger number of samples (and objective function evaluations) used here is required to accurately compare the marginal target PDFs of calibration parameters obtained using the SCEM-UA algorithm to the corresponding PDFs obtained using the MCS approach.

The Gelman–Rubin (Gelman and Rubin 1992) convergence statistics obtained are equal to or less than 1.01 for all four calibration parameters, indicating an almost ideal convergence to a joint posterior PDF. The root-mean-square error (RMSE) obtained is equal to 0.08 m indicating an excellent model fit. The coefficient of determination of 0.9995 and Pearson's correlation coefficient of $R^2=0.9998$ further indicate an almost ideal calibration model fit, i.e., an excellent agreement between the corresponding model predictions and measurements.

The optimal calibration parameter values (means) and the associated uncertainty indicators (standard deviations) are shown in Table 3. The corresponding head predictions and their uncertainties (standard deviations) for a number of selected nodes are shown in Table 4. Tables 3 and 4 indicate an almost perfect agreement between the corresponding mean values (both parameters and predictions) and a very good agreement in terms of the standard deviations obtained (again, for both parameters and predictions).

The marginal cumulative density functions (CDFs) obtained for all calibration parameters are shown in Fig. 3. As can be seen, a good agreement has been obtained, especially bearing in mind that the SCEM-UA algorithm started the search with uniformly distributed calibration parameters over a wide interval of possible values (50,150) and that a noninformative prior PDF was used to drive the search process. The marginal CDF agreement is slightly better for Parameters P1 and P2, which was an expected outcome. The reason for this is that these two parameters get more information from the observed heads than the other two parameters (as indicated by lower corresponding uncertainties, i.e., standard deviations in Table 3).

Table 4. Case 1.2: Selected Head Prediction Statistics (m)

			SCEM-UA		MCS	
Node ID	LC index	Observed	Mean	SD	Mean	SD
20	0	—	91.47	0.01	91.47	0.02
20	1	—	91.90	0.01	91.91	0.01
20	2	—	91.97	0.01	91.98	0.01
20	3	—	92.07	0.01	92.07	0.01
20	4	—	92.02	0.01	92.02	0.01
40	0	74.58	74.64	0.02	74.64	0.03
40	1	68.48	68.42	0.07	68.42	0.09
40	2	77.25	77.31	0.02	77.31	0.03
40	3	78.81	78.72	0.02	78.72	0.02
40	4	76.84	76.96	0.02	76.95	0.03
70	0	—	78.44	0.03	78.44	0.03
70	1	—	80.89	0.02	80.89	0.02
70	2	—	80.81	0.02	80.81	0.02
70	3	—	82.04	0.02	82.04	0.02
70	4	—	81.48	0.02	81.48	0.02
90	0	70.83	70.71	0.01	70.71	0.02
90	1	74.26	74.23	0.02	74.23	0.03
90	2	65.15	65.12	0.08	65.12	0.10
90	3	75.71	75.90	0.01	75.90	0.02
90	4	73.85	73.88	0.02	73.88	0.03
110	0	—	84.14	0.11	84.15	0.14
110	1	—	86.21	0.08	86.21	0.10
110	2	—	86.20	0.08	86.21	0.10
110	3	—	85.68	0.09	85.69	0.11
110	4	—	86.20	0.09	86.21	0.11
140	0	70.97	70.99	0.01	70.99	0.01
140	1	73.94	74.03	0.01	74.03	0.04
140	2	74.14	74.05	0.01	74.05	0.04
140	3	76.13	76.16	0.01	76.16	0.02
140	4	70.56	70.56	0.07	70.56	0.11
170	0	—	69.73	0.02	69.73	0.05
170	1	—	73.85	0.02	73.86	0.05
170	2	—	73.86	0.02	73.86	0.05
170	3	—	72.59	0.04	72.59	0.10
170	4	—	71.00	0.07	71.01	0.12

Note: LC=loading condition; SD=standard deviation. LC index: 0=normal demand; 1=peak demand; and 2–4=fire flows.

The calibration parameter correlations obtained are shown in Fig. 4. This indicates that the joint parameter PDF obtained using SCEM-UA can be utilized to calculate the parameter correlation coefficients. Fig. 4 also demonstrates that parameter correlation is low for any pair of parameters except for Parameters P3 and P4 (correlation coefficient equal to −0.91).

Case 2: Real-life Calibration Problem

Problem Description

This case study aims to apply the SCEM-UA algorithm to calibrate a hydraulic model of a real-life water distribution system. The system analyzed is the WDS located in a small town in the United Kingdom (name withheld for confidentiality reasons).

J. Hydraul. Eng. 2007.133:927-936.

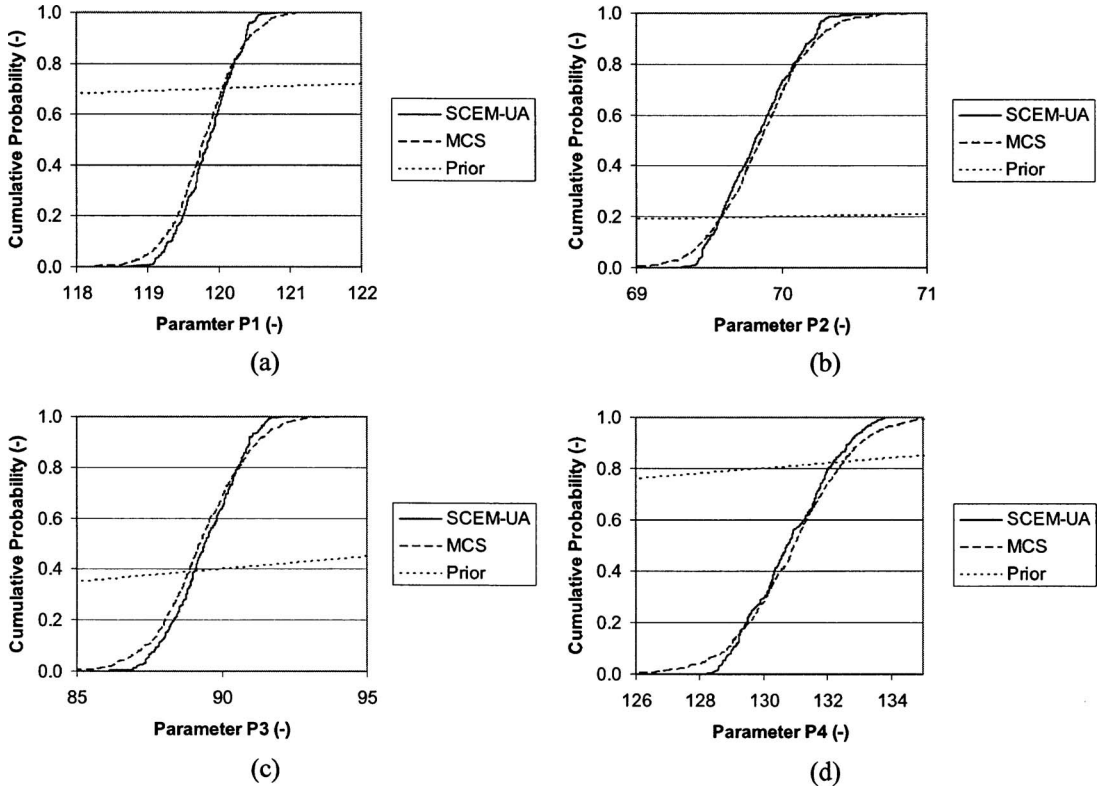

Fig. 3. Case 1.2: Marginal CDFs of calibration parameters

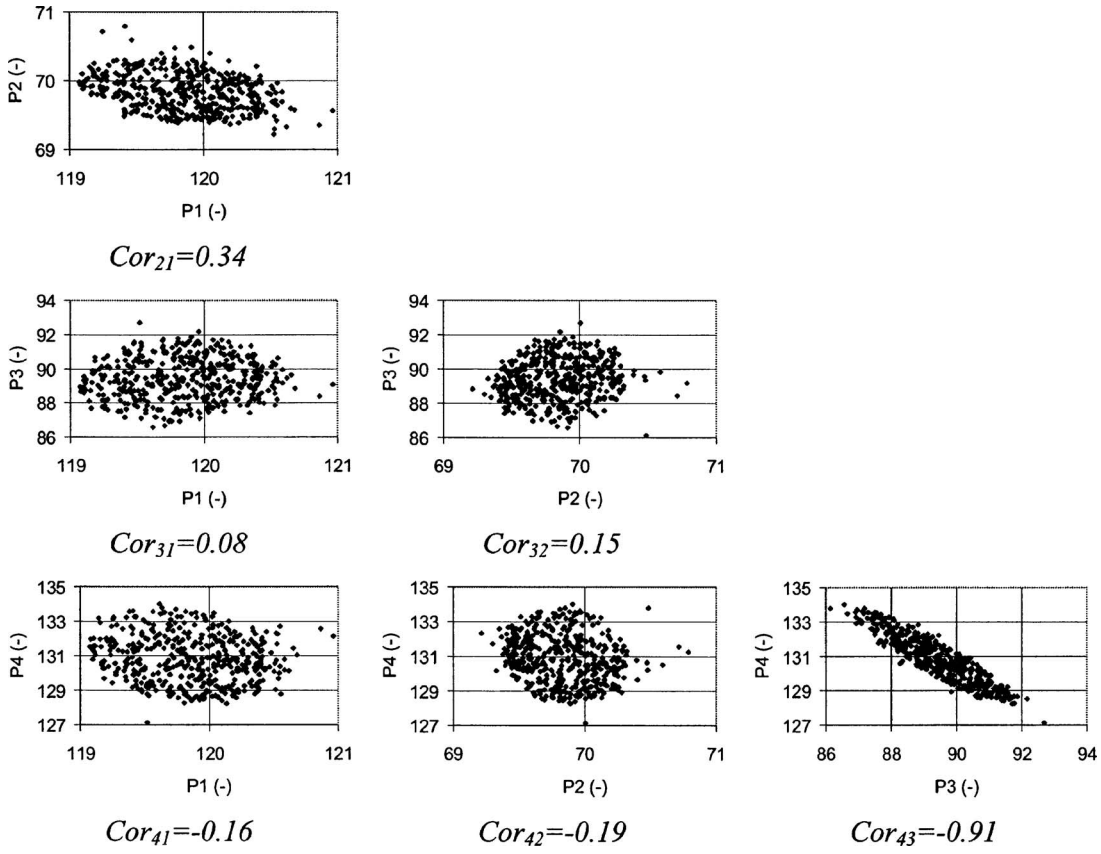

Fig. 4. Case 1.2: Lower triangular parameter correlation matrix

J. Hydraul. Eng. 2007.133:927-936.

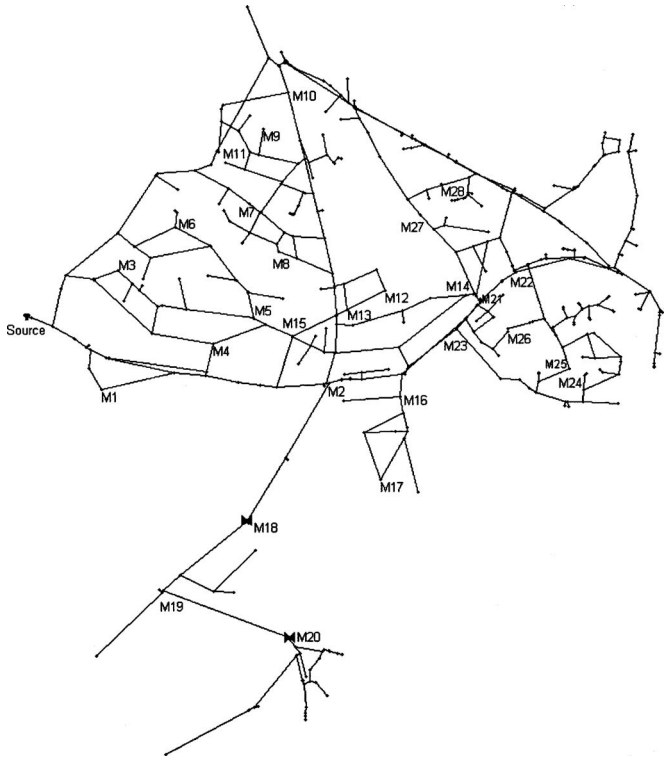

Fig. 5. Case 2: United Kingdom water distribution system layout

This system was studied previously (de Schaetzen et al. 2000; Kapelan et al. 2003). The general layout is presented in Fig. 5. The WDS covers approximately 6 km^2, with a population of around 4,500. Ground levels vary from 54 to 200 m above datum. Model demands are predominantly domestic with some commercial users to the East. The system is supplied by gravity from a service reservoir (see "Source" in Fig. 5) and includes two pressure reducing valves (PRVs) in the South. The majority of pipes are cast iron or ductile iron. A normal water use field test was carried out on June 29, 1994, with an estimated average demand of 14.4 L/s. Based on all available information, an Epanet2 hydraulic model was constructed containing one tank, 451 nodes, 497 pipes, and two PRVs.

A total of 32 loggers were initially used to observe pressures in the system. Evidence of obvious anomalies (e.g., nodal heads higher than the corresponding reservoir level) led to the exclusion of data collected by four loggers. Locations of the remaining 28 pressure loggers are depicted in Fig. 5 (e.g., M17). The network model is calibrated for 24 steady-state loading conditions. The 24 characteristic loading conditions represent conditions at 1 h time intervals during June 29, 1994. Therefore, the total number of observations is $N_o=28\times 24=672$.

The water network model is calibrated for 10 grouped Hazen–Williams (HW) pipe roughness coefficients. The 10 calibration parameters are obtained by grouping the network pipes based on pipe material/lining and diameter (see Table 5). Pipe age is not used here as an additional criterion simply because this information was not available. This however, is not envisaged as a major problem here since pipe age is expected to be highly correlated to the pipe material/lining. Note that all pipes located downstream of final measurement points were excluded from the calibration process (i.e., grouping) since there was no information on which to calibrate them. All calibration parameters are assumed to follow the uniform prior PDF defined on intervals specified in Table 5.

The noninformative PDF defined in Eq. (3) with $\gamma=0$ is used here to drive the SCEM-UA algorithm search for the posterior target distribution. The following SCEM-UA parameters are used: $N_s=1{,}000$, $N_q=5$, $T=10^6$, $c_n=1.0$, and $N_L=40$. The SCEM-UA run is stopped after 5,000 objective function evaluations (total computational time of approximately 2.5 min on a 3.0 GHz personal computer).

Results and Discussion

The Gelman–Rubin (Gelman and Rubin 1992) statistics obtained for all calibration parameters are shown in Fig. 6. Fig. 6 indicates an almost ideal convergence of the SCEM-UA algorithm to a joint posterior PDF. Note, however, that it takes only approximately 2,600 objective function evaluations to achieve the recommended Gelman–Rubin convergence of 1.2 for all 10 parameters. For comparison, the same calibration problem (without estimation of the parameter/prediction uncertainty) was solved by de Schaetzen (2000) using a genetic algorithm (Goldberg 1989) with approximately 20,000 calibration function evaluations.

The optimal calibration parameter values (means) and the associated uncertainty indicators (standard deviations) are shown

Table 5. Case 2: Calibration Parameter Data and Results

					SCEM-UA					
					Prior uniform PDF parameters		5,000 evaluations		2,600 evaluations	
Parameter ID	Material	Lining	Diameter (mm)	Number of pipes	Minimum	Maximum	Mean	SD	Mean	SD
P1	Cast iron	None	76	50	20	100	25	0.8	26	1.6
P2	Cast iron	None	102	34	20	100	48	1.4	50	2.3
P3	Cast iron	None	152	45	20	100	42	4.5	42	7.9
P4	Cast iron	None	254	37	20	100	66	1.8	65	2.7
P5	Ductile iron	Cement	100	22	80	130	113	8.9	108	12.3
P6	Ductile iron	Cement	150	15	80	130	100	12.9	108	13.3
P7	Ductile iron	Cement	250	1	80	130	104	13.5	112	10.2
P8	Cast iron	Epoxy	76	43	90	130	112	10.4	109	9.3
P9	MDPE	None	145	7	110	150	130	9.8	126	9.1
P10	PVC	None	152	2	110	150	130	10.9	133	8.9

Note: SD=standard deviation.

J. Hydraul. Eng. 2007.133:927-936.

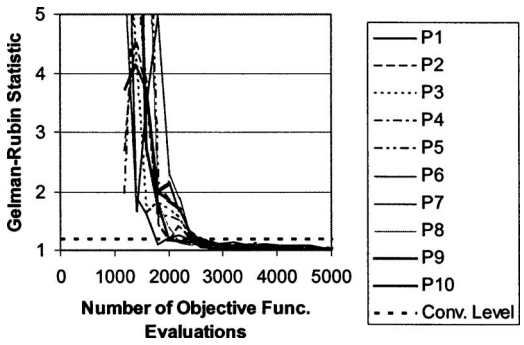

Fig. 6. Case 2: SCEM-UA algorithm convergence

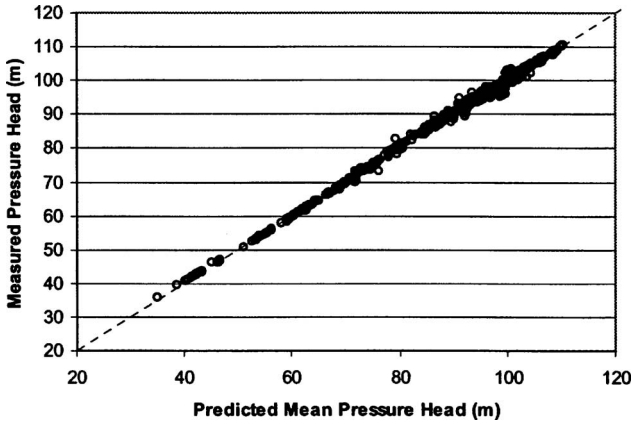

Fig. 8. Case 2: SCEM-UA predicted versus measured pressure head

in Table 5. As can be seen from Table 5, parameter uncertainty (as indicated by the corresponding parameter standard deviations) is lowest for the first four calibration parameters (P1–P4). The reason for this is that most of the pipes in the network are grouped in four groups corresponding to these parameters and therefore, the four parameters are getting the measurement information from most (if not all) pressure loggers. The same thing can be observed from the marginal PDF plots for Parameters P1 and P10 (see Fig. 7). Obviously, Parameter P10 is getting very little information from the pressure measurements which, in turn, results in a uniform-like posterior distribution of its value. From Table 5 it can also be noted that the calibration parameter values obtained after 2,600 objective function evaluations (i.e., at the point where the Gelman–Rubin statistic for all 10 calibration parameters is equal to or less than the recommended convergence value of 1.2) are close to the corresponding values obtained after 5,000 objective function evaluations. This is especially true for the first four parameters which, as already noted above, are getting most of the information from the available pressure measurements.

The RMSE obtained is equal to 0.79 m indicating a very good calibration model fit. The coefficient of determination of 0.998 and Pearson's correlation coefficient of $R^2=0.999$ both further indicate an almost ideal calibration model fit, i.e., an excellent overall agreement between the predicted and measured pressure heads. This point is also illustrated in a plot of predicted versus measured nodal heads in Fig. 8. Further examples of predicted versus measured pressure heads at several pressure logger locations are shown in Fig. 9. The average model prediction error (i.e., residual) is −0.08 m indicating a very low bias.

Finally, note that the estimated average pressure prediction uncertainty (i.e., standard deviation) is equal to 0.05 m (maximum equal to 0.24 m). Since these values are significantly smaller than the mean pressure values, the 95% prediction confidence intervals are not shown in Fig. 9.

Conclusions

A new calibration procedure based on a Bayesian recursive approach is shown here. The SCEM-UA algorithm presented is used to solve literature and real-life WDS model calibration problems. The results obtained in the case studies demonstrate that the SCEM-UA algorithm can be successfully used to calibrate WDS hydraulic models.

The main advantages of the Bayesian calibration approach over other existing calibration approaches are identified as follows: (1) Both calibration parameter values and associated uncertainties are determined in a single, optimization-type model run (in fact, the whole joint parameter probability density function is determined); (2) The SCEM-UA methodology does not require any FOSM-model-type assumptions (e.g., linearity and normality) to be made; (3) The SCEM-UA methodology does not require any derivatives to be calculated (avoiding possible numerical errors and discontinuity issues); and (4) The Bayesian recursive procedure encapsulated in the SCEM-UA methodology enables the specification of prior information on calibration parameters in a flexible, probabilistic rather than deterministic way (Greco and Del Guidice 1999; Kapelan et al. 2004).

The potential disadvantage of the SCEM-UA calibration ap-

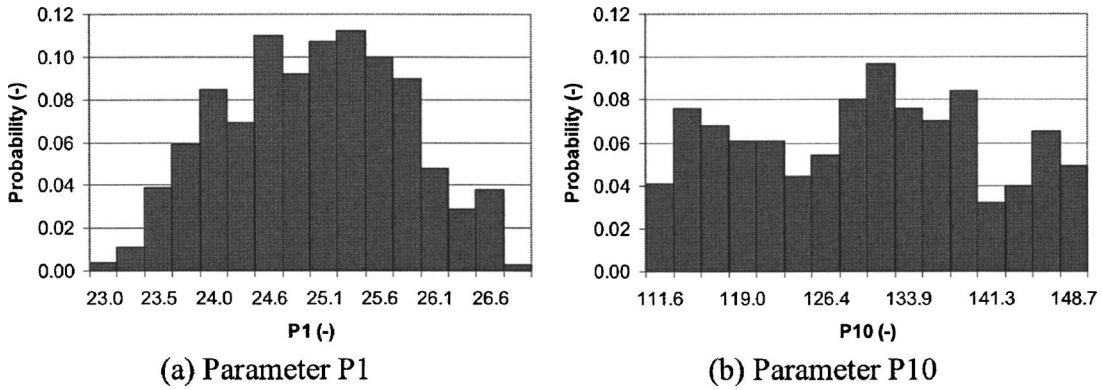

Fig. 7. Case 2: Marginal PDFs of selected calibration parameters

J. Hydraul. Eng. 2007.133:927-936.

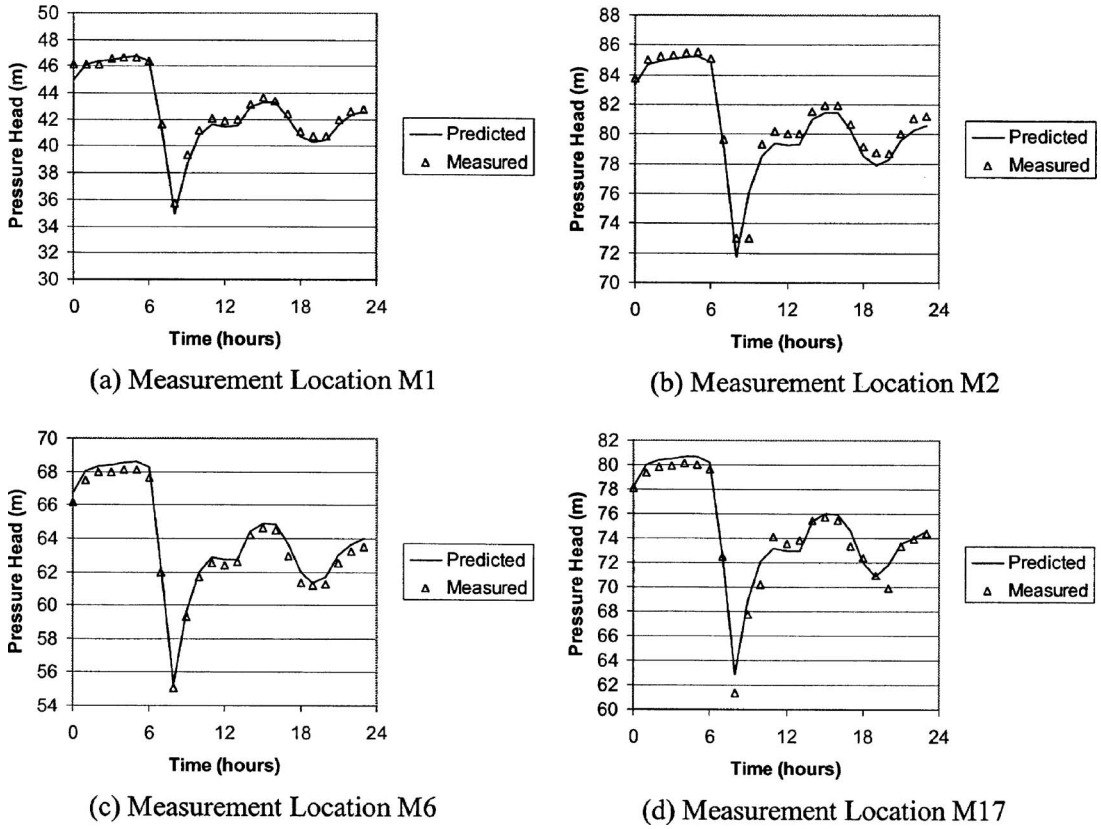

Fig. 9. Case 2: Examples of predicted versus measured pressure heads at several network locations

proach is the computational effort required to converge to a posterior, joint target distribution, especially if the number of calibration parameters is large (due to the matrix operations involved). This, however, is not envisaged as a major problem with modern computers, even for fairly large, real-life network calibration problems. Also, based on the second case study analysis, this computational effort is expected to be similar to (if not smaller than) the more classical approach where some global optimization method (e.g., genetic algorithms) is used to determine the (mean) calibration parameter values followed by the FOSM model application to calculate the associated uncertainties.

Acknowledgments

The research work presented here was partially supported by U.K. Engineering and Physical Sciences Research Council Platform Grant GR/T26054/01.

Notation

The following symbols are used in the paper:

$\mathbf{a}$ = vector of calibration parameters;
$\mathbf{a}^{(t)}$ = value of parameter set at tth iteration of the SCEM-UA algorithm;
B = variance between the N_q sequence means;
c = likelihood parameter;
$\mathbf{C}$ = matrix of complexes;
c_n = jump rate;
$\mathbf{e}$ = residual (i.e., error) vector;
e_i = ith element of vector $\mathbf{e}$;
E = least-squares objective;
Ex = exponential PDF;
exp = exponential function;
f = function denoting the mapping of $\mathbf{x}|\mathbf{a}$ into $\mathbf{y}$;
i = index (general);
k = index (general);
N_a = number of calibration parameters;
N_g = number of iterations within each sequence;
N_o = number of observations (i.e., measurements);
N_q = number of parallel sequences/complexes;
N_L = number of evolution steps taken by each sequence before complex reshuffle;
N_m = number of points in each sequence/complex;
N_s = number of samples;
p = prior/posterior probability or likelihood;
$\mathbf{S}$ = matrix of sequences;
SR = Gelan–Rubin statistic squared;
t = SCEM-UA algorithm iteration index;
T = threshold likelihood ratio;
W = average of the N_q within-sequence variances;
$\mathbf{x}$ = vector of WDS model inputs;
$\mathbf{y}$ = vector of WDS model predicted variables;
$\mathbf{y}^*$ = vector of measurements (i.e., observations);
γ = likelihood parameter denoting the PDF type; and
σ = standard deviation.

References

Bhave, P. R. (1988). "Calibrating water distribution network models." *J. Environ. Eng.*, 114(1), 120–136.

J. Hydraul. Eng. 2007.133:927-936.

Boulos, P. F., and Wood, D. J. (1990). "Explicit calculation of pipe-network parameters." *J. Hydraul. Eng.*, 116(11), 1329–1344.

Box, G. E. P., and Tiao, G. C. (1973). *Bayesian inference in statistical analyses*, Addison-Wesley-Longman, Reading, Mass.

Bush, C. A., and Uber, J. G. (1998). "Sampling design methods for water distribution model calibration." *J. Water Resour. Plann. Manage.*, 124(6), 334–344.

de Schaetzen, W. (2000). "Optimal calibration and sampling design for hydraulic network models." Ph.D. thesis, Univ. of Exeter, Exeter, U.K.

de Schaetzen, W., Walters, G. A., and Savic, D. A. (2000). "Optimal sampling design for model calibration using shortest path, genetic, and entropy algorithms." *Urban Water*, 2, 141–152.

Duan, Q., Sorooshian, S., and Gupta, V. K. (1992). "Effective and efficient global optimization for conceptual rainfall–runoff models." *Water Resour. Res.*, 28(4), 1015–1031.

Ferreri, G. B., Napoli, E., and Tumbiolo, A. (1994). "Calibration of roughness in water distribution networks." *Proc., 2nd Int. Conf. on Water Pipeline Systems*, Edinburgh, U.K., 379–396.

Gelman, A., and Rubin, D. B. (1992). "Inference from iterative simulation using multiple sequences." *Stat. Sci.*, 7, 457–472.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, Reading, Mass.

Greco, M., and Del Guidice, G. (1999). "New approach to water distribution network calibration." *J. Hydraul. Eng.*, 125(8), 849–854.

Hastings, W. K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." *Biometrika*, 57, 97–109.

Kapelan, Z. (2002). "Calibration of WDS hydraulic models," Ph.D. thesis, Univ. of Exeter, Exeter, U.K.

Kapelan, Z., Savic, D. A., and Walters, G. A. (2003). "Multiobjective sampling design for water distribution model calibration." *J. Water Resour. Plann. Manage.*, 129(6), 466–479.

Kapelan, Z., Savic, D. A., and Walters, G. A. (2004). "Incorporation of prior information on parameters in inverse transient analysis for leak detection and roughness calibration." *Urban Water*, 1(2), 129–143.

Kapelan, Z., Savic, D. A., and Walters, G. A. (2005). "Optimal dampling design methodologies for water distribution model calibration." *J. Hydraul. Eng.*, 131(3), 190–200.

Lansey, K. E., and Basnet, C. (1991). "Parameter estimation for water distribution networks." *J. Water Resour. Plann. Manage.*, 117(1), 126–144.

Lansey, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J., and Haan, C. T. (2001). "Calibration assessment and data collection for water distribution networks." *J. Hydraul. Eng.*, 127(4), 270–279.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., and Teller, A. H. (1953). "Equations of state calculations by fast computing machines." *J. Chem. Phys.*, 21, 1087–1091.

Ormsbee, L. E. (1989). "Implicit network calibration." *J. Water Resour. Plann. Manage.*, 115(1), 243–257.

Ormsbee, L. E., and Wood, D. J. (1986). "Explicit pipe network calibration." *J. Water Resour. Plann. Manage.*, 112(1), 166–182.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1990). *Numerical recipes: The art of scientific computing*, Cambridge University Press, Cambridge, U.K.

Rahal, C. M., Sterling, M. J. H., and Coulbeck, B. (1980). "Parameter tuning for simulation models of water distribution networks." *Proc. Institution of Civil Engineers*, Part 2, Institute of Civil Engineers, London, 751–762.

Rossman, L. A. (2000). *Epanet2 users manual*, U.S. Environmental Protection Agency, Washington, D.C.

Shamir, U., and Howard, C. D. D. (1968). "Water distribution systems analysis." *J. Hydr. Div.*, 94(1), 219–234.

Vitkovsky, J. P., Liggett, J. A., Simpson, A. R., and Lambert, M. F. (2003). "Optimal measurement site locations for inverse transient analysis in pipe networks." *J. Water Resour. Plann. Manage.*, 129(6), 480–492.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S. (2003). "A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters." *Water Resour. Res.*, 39(8), 1–16.

Walski, T. M. (1983). "Technique for calibrating network models." *J. Water Resour. Plann. Manage.*, 109(4), 360–372.

Walski, T. M., et al. (1987). "Battle of the Network Models: Epilogue." *J. Water Resour. Plann. Manage.*, 113(2), 191–203.

J. Hydraul. Eng. 2007.133:927-936.