Urban Water Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nurw20

Unified parameter optimisation approach for leakage detection and extended-period simulation model calibration

Zheng Yi Wu [a]

[a] Applied Research Group, Bentley Systems, Incorporated , 27 Siemon Company Drive, Suite200W, Watertown, CT, 06795, USA

Published online: 07 Apr 2009.

To cite this article: Zheng Yi Wu (2009) Unified parameter optimisation approach for leakage detection and extended-period simulation model calibration, Urban Water Journal, 6:1, 53-67, DOI: 10.1080/15730620802541631

To link to this article: http://dx.doi.org/10.1080/15730620802541631

Urban Water Journal
Vol. 6, No. 1, March 2009, 53–67

Taylor & Francis
Taylor & Francis Group

Unified parameter optimisation approach for leakage detection and extended-period simulation model calibration

Zheng Yi Wu*

Applied Research Group, Bentley Systems, Incorporated, 27 Siemon Company Drive, Suite200W, Watertown, CT 06795, USA

(*Received 4 April 2008; final version received 30 September 2008*)

Leakage detection is of great interest and significance for water utilities. There are different field testing apparatuses developed for leakage detection, but no systematic modelling approach was available for predicting the most likely leakage hotspots that might guide leakage detection crew to quickly locate exact leakage locations. In this paper, leakage is represented as pressure-dependent emitter flow at a node in a water distribution model. Leakage detection is formulated as a nonlinear parameter identification problem to search for the possible emitter node locations and the emitter coefficients while minimising the same objective function as model calibration, namely the distance between the field-observed and the model-simulated flows and hydraulic grades. Thus the leakage detection optimisation is developed as an integral component of the unified model parameter identification framework. The comprehensive methodology is applied to the water system of a district meter area (DMA) in the UK. It illustrates the effectiveness of the unified approach for both leakage hotspot detection and extended-period simulation model calibration using the real field data.

Keywords: water distribution system; hydraulic model; leakage detection; water loss; model calibration; genetic algorithm and optimisation

Introduction

Leakage detection in a water distribution system is a vital but difficult task for water utilities around the world. There are leakage detection apparatuses such listening sticks, flow testing and acoustic loggers. Applying these detection instruments is time-consuming to find leaks without informed guidance about where the leakages are most likely to occur in a large water system. Leakage is more difficult to find than ever by using acoustic loggers due to (1) implementing pressure management in order to reduce leakages and (2) replacing ferrous mains with plastic pipes, both actions resulting in less acoustical response. Background and burst estimate (BABE) analysis (Lambert 1994) indicates that three types of leakages exist in a water distribution system: (1) reported leakage, (2) unreported leakage and (3) background leakage. The reported leakage usually results in the water loss of the largest flow rate in a short time. This type of leakage is caused by a pipe burst or break and the flow emerges from the surface so that it is easily perceived and thus usually reported by customers. Such an incident is treated as an emergency handled in 24 hours as required by governmental regulations. The second type of leakage is hidden underground. Although such a leak may be at a lower flow rate than the reported ones, it can continuously run for a long time until the leaking pipe is detected and mended. Background leakages happen at the joints and fittings of pipelines, and continuously run at a low flow rate, and are found to be the dominant component of unavoidable annual real losses (UARL) and are not economically recoverable (Lambert *et al*. 1999). It is the type of unreported leakage that must be efficiently detected and consequently fixed to improve system integrity so that water loss can be cost-effectively managed for sustainable development.

Using a well-developed and widely-adopted hydraulic network model to assist in leakage detection is an attractive approach. There has been a large amount of investment not only in network information management but also in constructing computer models that have been used for analysing and simulating different scenarios of system conditions. To construct a practically useful EPS model, leakage, as a significant component of water consumption in many systems, must be correctly quantified as a nodal demand, so-called leakage demand. Modelling leakage requires the determination of which nodes should be assigned how much leakage demand to in addition to the actual

*Email: zheng.wu@bentley.com

ISSN 1573-062X print/ISSN 1744-9006 online

DOI: 10.1080/15730620802541631
http://www.informaworld.com

consumption demand, so that the field-observed flows and pressures match well with the model-simulated values. This is essentially a parameter optimisation problem similar to model calibration. Many different optimisation methods have been applied to solve the model calibration problem. Comprehensive reviews have been given by Omsebee (2006) and Lansey (2006) on the applications of optimisation techniques in water distribution system analysis. Over last decade, inverse transient analysis has been one of the popular research topics for leakage detection. Productive research outcomes have been reported in numerous publications (Vítkovský *et al.* 2000, Wang *et al.* 2002, Kapelan *et al.* 2004, Beck *et al.* 2005, Covas and Ramos 2005, Lee *et al.* 2005, Taghvaei *et al.* 2006). The range of validity of inverse transient analysis method has been carefully studied by Nixon *et al.* (2006). It is found that the applicability of the inverse transient method is limited to the instantaneous small-amplitude disturbance and the simple reservoir–pipe–valve type system or reservoir–pipe–reservoir type system. The challenges remain for applying the inverse transient modelling method to a real water system. Wu and Sage (2006, 2007) developed a leakage detection model that uses emitters to emulate leakages at nodes and applied a competent genetic algorithm (GA) to simultaneously optimise the emitter locations and the corresponding emitter coefficients. The approach has been successfully applied to identifying the leakage hotspots of a small case problem and a real district meter area. In this paper, the leakage detection model is unified as an integral part of the solution method for both leakage detection and model calibration. The generalised problem is formulated as parameter identification optimisation and solved by using a competent genetic algorithm (Wu and Simpson 2001). A hydraulic model of a district meter area in the UK has been used as a case study to demonstrate the typical steps in undertaking leakage detection and EPS model calibration using field-observed data.

Unified formulation

Both leakage detection and model calibration within the EPS context are formulated as an implicit nonlinear optimisation problem of parameter identification, given as:

$$\text{Search for:} \quad \vec{X} = (x_i); \quad i = 1, \ldots, N \tag{1}$$

$$\text{Minimise:} \quad F(\vec{X}) \tag{2}$$

$$\text{Subject to:} \quad \underline{x_i} \le x_i \le \overline{x_i} \tag{3}$$

where $\vec{X}$ is the vector of the parameters to be identified in a hydraulic model; x_i is the i-th model parameter; $\overline{x_i}$ is the upper limit of the i-th model parameter; $\underline{x_i}$ is the lower limit of the i-th model parameter; N is the number of model parameters and $F(\vec{X})$ is the objective function that is defined as a distance function measuring the goodness-of-fit between the field observed values and the model-simulated values. The objective function definitions are given as below.

Objectives

A model calibration solution is evaluated by the discrepancy between the model-simulated and the field-measured junction HGLs and pipe flows. In order to equivalently consider both HGL and flow contribution to the dimensionless fitness, the goodness-of-fit score is calculated by using two user-specified conversion factors, namely the hydraulic head per fitness point and the pipe flow per fitness point, which convert the head difference and flow difference into a dimensionless fitness value. Three fitness functions are then defined as:

(1) Objective type I: minimise the sum of difference squares

$$F(\vec{X}) = \sum_{t=1}^{T} \frac{\sum_{nh=1}^{NH} w_{nh} \left(\frac{Hs_{nh}(t) - Ho_{nh}(t)}{Hpnt} \right)^2 + \sum_{nf=1}^{NF} w_{nf} \left(\frac{Qs_{nf}(t) - Qo_{nf}(t)}{Qpnt} \right)^2}{NH + NQ} \tag{4}$$

(2) Objective type II: minimise the sum of absolute differences

$$F(\vec{X}) = \sum_{t=1}^{T} \frac{\sum_{nh=1}^{NH} w_{nh} \left| \frac{Hs_{nh}(t) - Ho_{nh}(t)}{Hpnt} \right| + \sum_{nf=1}^{NF} w_{nf} \left| \frac{Qs_{nf}(t) - Qo_{nf}(t)}{Qpnt} \right|}{NH + NQ} \tag{5}$$

(3) Objective type III: minimise the maximum absolute difference

$$F(\vec{X}) = \arg\max_{t,nh,nf} \left\{ \left| \frac{Hs_{nh}(t) - Ho_{nh}(t)}{Hpnt} \right|, \left| \frac{Qs_{nf}(t) - Qo_{nf}(t)}{Qpnt} \right| \right\} \tag{6}$$

where $Ho_{nh}(t)$ designates the observed hydraulic grade of the nh-th junction at time step t, $Hs_{nh}(t)$ is the model-simulated hydraulic grade of the nh-th junction at time step t, $Qo_{nf}(t)$ is the observed flow of the nf-th link at time step t, $Qs_{nf}(t)$ is the simulated flow of the nf-th link at time step t, $Hpnt$ notes the hydraulic head per fitness point while $Qpnt$ is the flow per fitness point (both are user-specified coefficients to convert pressure and flow differences into dimensionless values, and also can be applied as weighting factors for pressure and flow calibration), NH is the number of observed hydraulic grades and NQ is the number of observed pipe discharges, and W_{nh} and W_{nf} represent the normalised weighting factors for observed hydraulic grades and flows respectively. They are given as:

$$W_{nh} = w\left(Hloss_{nh} \Big/ \sum Hloss_{nh}\right) \tag{7}$$

$$W_{nf} = w\left(Qo_{nf} \Big/ \sum Qo_{nf}\right) \tag{8}$$

where $w()$ is a function which can be linear, square, square root, log or constant and $Hloss_{nh}$ is the head loss from water source to observation data point nh. For multiple water resources, an average head loss can be calculated for an observed HGL point. An optimised calibration can be conducted by selecting one of the three objectives above and the weighting factors between junction hydraulic heads and pipe flows. The model parameters are calculated by using a genetic algorithm while minimising the selected objective function.

It is the objective function that quantifies the optimality of a possible solution and that drives the GA to search for good model parameters. There is no real right or wrong answer to the question as to which objective/fitness type is better than another. It is a matter of which objective function fits the requirement of a particular application. Three objective formulations given by Equations (4)–(6) are valid for model parameter identification. The objective function of Equation (5) calculates the sum of absolute differences of all the data points between the simulated and the observed values as the fitness for any given calibration solution. It evenly spreads the influence over all observed data sets. The objective function given by Equation (4) calculates the sum of difference squares of all the data points between the simulated and the observed values as fitness for a given calibration solution. Similar to Equation (5), which considers all observed data sets, but the objective function of Equation (4) emphasises the bad data points because of squaring the difference values. The objective function given by Equation (6) calculates the maximum difference of all the data points between the simulated and the observed values as fitness for each solution. The maximum difference values for different calibration solutions may occur at different data points, but the GA will always retain the maximum difference of all the data points as fitness for each corresponding solution. From a GA search perspective, the different methods for evaluating fitness may result in different solutions, but a well calibrated model should be obtained by applying one of three fitness functions provided that sufficient field data are of good quality and calibration is appropriately set up.

Leakage detection model

Leakage is pressure dependent, the greater the pressure, the greater the leakage. It can be modelled as emitter flow given as:

$$Q_i = K_i P_i^{\alpha} \tag{9}$$

where Q_i is the leakage aggregated at node i, P_i is the nodal pressure at node i, α is the exponent which can be flexibly specified in a hydraulic model although 0.5 is used for leaks as default, and K_i is the emitter coefficient. It is the emitter coefficient K_i that serves as the leakage indicator. If the emitter coefficient is greater than zero, it indicates that there is a leak at a node. There is no leak at a node if the emitter coefficient is zero. It is well known that the leakage flow exponent can be in a wide range due to the physical conditions of leakage holes, but it is the emitter coefficient that primarily determines whether or not a leakage flow exists. There are often hundreds and thousands of nodes in a real water distribution model. It is not possible to efficiently solve for an optimisation problem with thousands of decision variables. However, by the nature of leakage hotspots, there are a limited number of spots that experience leakages in a real system. Thus a new leakage detection optimisation model is formulated to identify a given number of leakage nodes and corresponding emitter coefficients. In practical application, all the nodes in a large water distribution model can be aggregated into a number of node or demand groups; for instance, all the nodes in one DMA can be aggregated into one demand group. Each demand group is specified to identify a given maximum number of leakage nodes within the prescribed maximum emitter coefficient. The actual numbers of leakage nodes and locations are determined by the optimised node indices with positive emitter coefficients. Therefore, the leakage detection model is given as:

$$\begin{aligned} &\text{Search for:} \quad \vec{X} = (LN_i^n, K_i^n); \\ &LN_i^n \in \mathbf{J}^n; n = 1, \ldots, NGroup; i = 1, \ldots, NLeak^n \end{aligned} \tag{10}$$

$$\text{Minimise: } F(\vec{X}) \tag{11}$$

$$\text{Subject to: } 0 \le K_i^n \le \overline{K^n} \tag{12}$$

$$P_i^n > 0 \tag{13}$$

$$\sum_{n=1}^{NGroup} NL_{dup}^n = 0 \tag{14}$$

where LN_i^n is the leakage node index for leakage node i within demand group n, K_i^n is the emitter coefficient for leakage node i in group n, $\mathbf{J}^n$ is the set of nodes within node group n, $NGroup$ is the number of node groups, $NLeak^n$ is the number of the specified leakage nodes to be identified for node group n, $\overline{K^n}$ is the maximum emitter coefficient for node group n, P_i^n is the pressure head at the detected leakage node i within group n and NL_{dup}^n is the number of the duplicated nodes that are identified as leakage emitters in one solution for group n.

The constraints given by Equations (13) and (14) are required for the GA to search for good solutions of the leakage detection problem. During the optimisation process, nodal pressure may become negative when too much demand is assigned to a node. It is necessary to ensure that a leakage detection solution with the nodes of negative pressure heads is not selected for producing the next generation of solutions during GA optimisation because a negative nodal pressure will result in a undesired artificial flow into a system. On the other hand, it is highly desirable not to have duplicated nodes identified as leakage nodes in one solution. This is guaranteed by satisfying the constraint given by Equation (14), which states that the identified leakage nodes are unique for one solution. Both constraints are handled by using a penalty function given as:

$$F_{leak}(\vec{X}) = F(\vec{X}) + f_{penalty}\left[\left|\sum_{n=1}^{NGroup}\sum_{i=1}^{NLeak^n} \min(0, P_i^n)\right| + \sum_{n=1}^{NGroup} NL_{dup}^n\right] \tag{15}$$

where $f_{penalty}$ is the penalty factor prescribed as one of the optimisation parameters. For leakage detection solutions that meet both constraints, the fitness values are the same as the model calibration objectives, otherwise an extra penalty term is calculated and added to the solution fitness by Equation (15).

Hydraulic model calibration

Hydraulic model calibration within the EPS context requires identifying not only the static model parameters but also the dynamic model parameters. A previously developed GA-based framework (Wu *et al*. 2002) has been extended by taking into account the adjusting for pressure-dependent demand and the pattern coefficients of the volume-based demands. Therefore, four types of model parameters are considered for hydraulic model calibration including:

- Pipe roughness factor f_i for pipe group i: either all the pipes within one group are set to the same roughness coefficient or the original roughness coefficients are multiplied with the same multiplier;
- Pressure-dependent demand adjustment k_j for demand group j;
- Demand pattern adjustment multiplier $m_{j,t}$ for junction group j at time step t: junction demands within the same demand group are multiplied with the same demand adjustment multiplier;
- Link operation and control status $s_{k,t}$ for link k (pipes, valves and pumps) at time step t.

The extended model calibration formulation is formulated as:

$$\text{Search for: } \vec{X} = (f_i, k_j, m_{j,t}, s_{k,t}) i = 1, \ldots, PGroup; \; j = 1, \ldots, NGroup; k = 1, \ldots, NStatus. \tag{16}$$

$$\text{Minimise: } F(\vec{X}) \tag{17}$$

$$\text{Subject to: } \underline{f_i} \le f_i \le \overline{f_i} \tag{18}$$

$$\underline{m_{j,t}} \le m_{j,t} \le \overline{m_{j,t}} \tag{19}$$

$$s_{k,t} \in \{0, 1\} \tag{20}$$

$$\underline{k_j} \le k_j \le \overline{k_j} \tag{21}$$

where $\vec{X}$ represents a set of model parameters, $\overline{f_i}$ and $\underline{f_i}$ are the upper and lower limits of roughness factor for pipe group i, $\overline{m_{j,t}}$ and $\underline{m_{j,t}}$ are the upper and lower limits for demand adjustment multiplier for junction group j at time step t; $\overline{k_j}$ and $\underline{k_j}$ are the upper and lower limits for emitter coefficient for junction group j.

The formulation given by Equations (1) to (21) is a unified nonlinear optimisation problem for both leakage detection and model calibration, which is solved within the common solution framework using a competent genetic algorithm optimisation technique. A modeller can undertake a model calibration task for any combination of four types of model parameters. The effective model parameters are obtained by minimising one of the objective functions given as Equations (4) to (6). An optimisation run is conducted with user-selected snapshots of field data. Each

snapshot of data was collected at the same time and consisted of (1) calibration target data of node hydraulic grades and pipe flows, and (2) calibration boundary data of tank levels, control element status and pump speeds. Multiple snapshots of data can be selected for one optimisation run. However, one snapshot of data requires one steady-state simulation to produce modelling results for fitness evaluation. Hence, the more snapshots are selected for one run, the longer it takes to complete.

Integrated framework implementation

The implementation of the leakage detection method is the extension of the previously developed model calibration framework (Wu *et al*. 2002) as shown in Figure 1. The information flows in both directions between the end-user and the data storage, thus enabling engineers to seamlessly manage the data and optimisation runs by exploiting the powerful combination of a competent GA optimiser (using fast messy GA) and a hydraulic network simulator, both being embodied in one modelling system. It consists of a user interface, an evaluation module, a competent GA optimiser, a hydraulic simulation model and a database system.

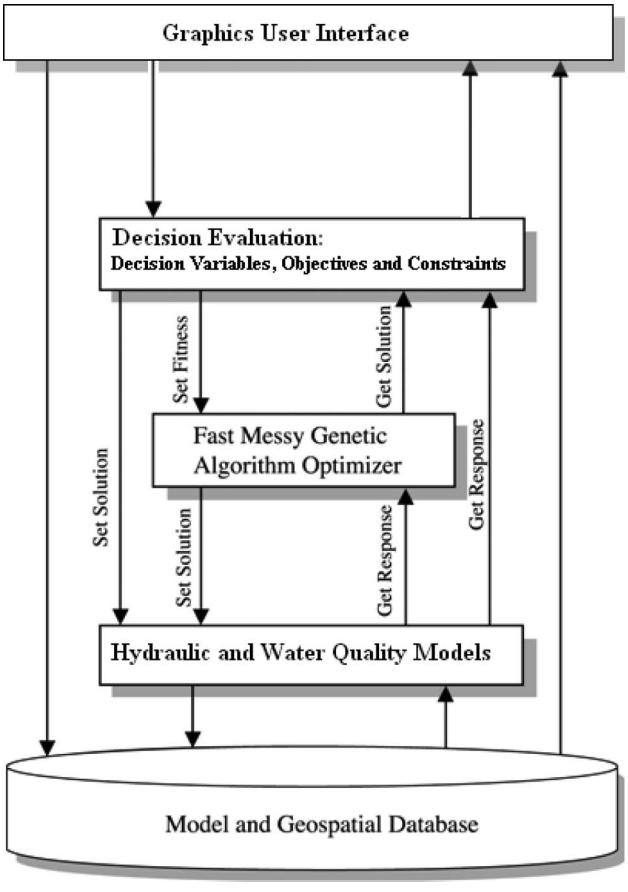

Figure 1. Integrated solution framework for leakage detection and model calibration.

A user interface permits users to enter the field-observed data, load the hydraulic model, and select corresponding boundary conditions and optimisation criteria. All the input data and the results are consistently stored in a modelling database along with the simulation model. This allows engineers to revisit the optimisation runs at any time in the future, helps modellers to better manage a modelling project and facilitates the maintenance of model accuracy in the long term.

A parameter optimisation run of either leakage detection or model calibration is initiated by passing a set of optimisation run data to the evaluation module. The dataset is processed and prepared for the user-desired criteria. For a leakage detection optimisation run, both leakage nodes and corresponding emitter coefficients are optimised by the GA optimiser. For model calibration, modellers may proceed with a parameter adjustment run by either interactively adjusting model parameters (engineers manually set a value for each parameter) or automatically optimising calibration solutions. Without activating the GA optimiser, the user-estimated model parameters are submitted to the hydraulic simulator. It predicts the hydraulic responses of HGL and pipe flows that are passed back to the evaluation module. The goodness-of-fit is calculated and reported to a user. Modellers can estimate the parameters and iterate the process to enhance model calibration. In contrast, calibration can proceed with the GA optimiser, which automatically generates and progressively improves the calibration solutions. Each trial solution of either leakage detection or optimisation-based model calibration, along with the selected data sets, corresponding loading and boundary conditions, is submitted to the hydraulic network analyser for predicting the hydraulic responses. The model-simulated results are passed back to the evaluation module and used to calculate the fitness values of the solutions that are evolved by the GA optimiser over generations. The integrated framework provides a versatile optimisation modelling tool for leakage detection and EPS model calibration.

Case study

The case study is to illustrate the application of the unified optimisation approach for leakage detection and model calibration. With the unified methodology and integrated tool, it is possible to set up one optimisation run to optimise all the model parameters, including leakage node locations, leakage emitter coefficients, pipe roughness coefficients and demand

adjustments for the system, but the results can make it difficult for engineers to understand the system characteristics with such ad hoc run. Instead, to illustrate how leakage detection and model calibration can be achieved for large systems, this study is intended to decompose a complex parameter identification or calibration problem into a number of small tasks. Thus the leakage detection and model calibration can be completed by undertaking a number of steps as follows.

- System evaluation for leakage detection;
- Leakage detection optimisation;
- Roughness and peak flow pattern calibration;
- Extended-period flow pattern calibration.

System evaluation

The system, as shown in Figure 2, is a district meter area (DMA) in northern England and owned by United Utilities Water (UUW). The model is built to contain all the pipes with diameters of 50 mm and above. A total of 924 pipes, 701 junctions, 178 control valves and one reservoir are included in the DMA model. Field-observed pressures were collected by installing six pressure loggers throughout the system and one flow logger at the inlet of the DMA. Pressure and flow data were collected over a long period. The data were extracted and classified into different datasets. Each dataset is a collection of the field data that represent one snapshot of the measured junction pressures, inflow and DMA boundary conditions at the same time. A total of 96 snapshots of data are prepared in time steps of 15 minutes over 24 hours from midnight to midnight. The extracted data were stored in an MS Excel file and automatically imported into the model database by using a model-building utility tool.

Figure 2. Case study system with six pressure logger locations at junctions J-192, J-267, J-656, J-322, J-431 and J-132.

Before undertaking any optimisation run for identifying possible leakage hotspots, the system evaluation is conducted by comparing the field-observed inflows with the model-simulated inflow of actual consumptions over 24 hours. Figure 3 illustrates the flow differences between the observed and the simulated. The differences represent the sum of water losses, including leakages and other likely NRW consumptions, which are to be identified by the leakage detection optimisation modelling.

Leakage detection

Using the integrated optimisation modelling tool, the leakage detection study proceeded with defining the detection scope by aggregating the nodes into different demand groups. The aggregation permits engineers to make use of the system-specific information and the available knowledge. For each demand group, a maximum number of leakage nodes, the maximum emitter coefficient and the increment are specified for a leakage detection run.

Since no detailed knowledge of the leakage condition is acquired for the case study system, one demand group is made to contain all the nodes. Leakage detection optimisation runs are undertaken for the low-demand hours. This is because flow velocity is very small during the low-demand hours so that the roughness value is the least sensitive to the flow change. As indicated in Figure 3, the inflow into the system reaches the minimum flow from 2:30 a.m. to 3:30 a.m., thus a total number of four snapshots at 2:30 a.m., 2:45 a.m., 3:00 a.m. and 3:15 a.m. are used for leakage detection. The Darcy-Weisbach head loss equation is used with a default pipe roughness of 0.0015 mm for all the pipes.

Four optimisation runs are set up for the demand group with the same optimisation options including: (1) select Detect Leakage Node as demand adjustment operation; (2) specify a maximum number of 10 leakage nodes; (3) set the maximum emitter coefficient of 0.5 L/s/(m H_2O)$^{0.5}$; (4) choose an emitter coefficient increment of 0.01 and (5) use an emitter exponent of 0.5 for the leakage flow emitters for this system. Although the exponent may vary from 0.5 to 2.5 due to different pipeline materials, the exponent is found to be close to 0.5 for detectable leaks and bursts on metal pipes (Lambert 2002). Both the maximum number of

leakage nodes and the maximum emitter coefficient are specified to allow the GA optimiser to have adequate freedom to search for the most likely leakage hotspots. If the number of optimised leakage nodes and/or the emitter coefficient is the same as the upper limit, the optimisation model needs to be rerun with a relaxed upper bound, which is an increased number of leakage nodes and/or the maximum emitter coefficient.

Each optimisation run is constructed to use one field data snapshot at the low-demand hours. All the optimisation runs are executed in batch mode, in which the optimisation runs are completed one after another sequentially without human interaction. This is a beneficial function for performing numerous optimisation runs by using the computing resources available after work or overnight. All the optimisation runs have been executed by using the objective function of Minimise Difference Squares given as Equation (4), head loss per fitness point of 0.1 meter and flow per fitness point of 0.1 L/s. The other optimisation parameters include random seed of 0.5, population size of 100, maximum era of 6, era generation number of 150, mutation probability of 0.01, cut probability of 0.017 and splice probability of 0.8.

After completing all the optimisation runs, the best solution attained by each of four leakage detection runs is exported from the GA-based optimisation tool to the hydraulic model database. Leakage detection solutions, represented by a set of the identified leakage nodes with positive emitter coefficients, are shown in Table 1. Solutions 2 and 3, obtained for leakage detection runs at 2:45 a.m. and 3:00 a.m., are exactly

Figure 3. Inflow comparison between the field-observed and model-simulated before leakage detection and model calibration.

Table 1. Optimised leakage detection solutions.

Solution 1		Solution 2		Solution 3		Solution 4	
Leakage nodes and areas	Emitter coefficients $(L/s/(m\ H_2O)^{0.5})$	Leakage nodes and areas	Emitter coefficients $(L/s/(m\ H_2O)^{0.5})$	Leakage nodes and areas	Emitter coefficients $(L/s/(m\ H_2O)^{0.5})$	Leakage nodes and areas	Emitter coefficients $(L/s/(m\ H_2O)^{0.5})$
J-700	0.32	J-626	0.35	J-626	0.35	J-98	0.36
J-598	0.24	J-598	0.29	J-598	0.29	J-104	0.20
J-97	0.20	J-353	0.04	J-353	0.04	J-273	0.19
J-322	0.08	J-700	0.02	J-700	0.02	J-322	0.16
N/A	N/A	J-321	0.02	J-321	0.02	J-700	0.11
Area A	0.44	Area A	0.29	Area A	0.29	Area A	0.56
Area B	0.32	Area B	0.08	Area B	0.08	Area B	0.11
Area C	0.08	Area C	0.35	Area C	0.35	Area C	0.16

the same but different from the other solutions. To illustrate the locations of the detected leakage nodes, the leakage detection solutions are colour-coded in the leakage map. As shown in Figure 4, a good similarity of leakage node areas emerged from all the different runs. Four solutions consistently identify the same areas A, B and C as annotated in Figure 4 as likely leakage hotspots for the system. The total optimised emitter coefficients for areas A, B and C are also summed up and presented in Table 1 for each solution. Area A is consistently predicted as the leakage hotspot with similar emitter coefficient. The optimised emitter coefficients for areas B and C vary from 0.08 to 0.35 L/s/(m H_2O)$^{0.5}$, which indicates some uncertainty that may exist in the model and field data for these areas. However, the aggregated emitter coefficients by possible leakage area show better consistency than the emitter coefficient at each individual node. The flow comparisons at leakage detection hours are given in Table 2. The good flow balance achieved by the leakage detection serves as the base for further model calibrations of steady-state and extended-period simulation. The detected leakage emitter locations and coefficients are used for the hydraulic model calibration as follows.

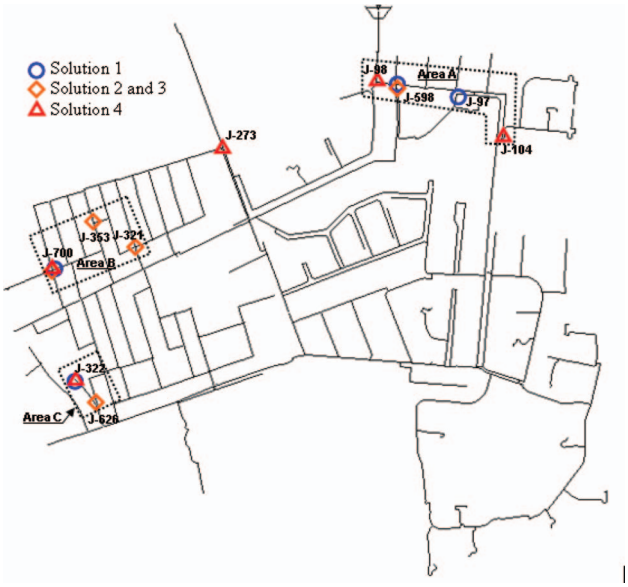

Figure 4. Leakage detection results for the case study system.

Table 2. Comparisons of flows for leakage detection solutions.

Solutions	Simulated inflow (L/s)	Observed inflow (L/s)	Difference (L/s)
Solution 1	8.59	8.93	0.34
Solution 2	8.48	8.89	0.41
Solution 3	8.95	8.80	0.15
Solution 4	8.50	8.96	0.46

Roughness and peak flow pattern calibration

Pipeline roughness is one of the primary parameters to be calibrated for the hydraulic model. It is one of the pipeline physical attributes and is unlikely to change over the extended-period simulation. Pipe roughness is often calibrated by using the snapshot of data at peak-demand hours when the flow velocity is high and consequently results in sensible pipeline head losses with roughness adjustment during the calibration process. In spite of the detected leakage emitters, there still exists the uncertainty of demand at peak hours, so both demand and pipe roughness must be calibrated at the time steps of peak demand.

There are more than 900 pipes in the model. Not every single pipe can be assigned an independent variable for roughness calibration. To reduce the dimension of the roughness optimisation, the pipes with similar characteristics and on the same flow path are aggregated into one roughness group. Aggregating the pipes on the same flow path will ensure the roughness adjustment is sensitive to the hydraulic head loss to the observed pressure point where the simulated and observed pressure values are used to calculate the fitness of a solution. For the case study system, nine roughness groups are made as shown in Figure 5. For each roughness group, pipe roughness is adjusted by multiplying the original roughness value of 0.0015 mm by the multiplier that is optimised within the prescribed range from 0.5 to 10.0 and the increment of 0.1.

Figure 5. Roughness grouping for steady-state calibration.

Roughness calibration runs were created and undertaken using the field data collected at high demand hours at 7:30 a.m., 7:45 a.m. and 8:00 a.m.

The adjusted roughness values are given in Table 3 while the flow and pressure comparisons of the optimisation run are given in Table 4, which exemplifies the good agreement between the field-observed and model-simulated values. The Darcy-Weisbach roughness height e values change from 0.0015 mm to 0.011 mm for roughness group 2, 3, 7 and 9, which have demonstrated a greater roughness increase than the other groups of pipes. The increased roughness values indicate that the pipes in these areas are probably deteriorated more than the others. The optimised demand adjustment factors are 1.2, 1.3 and 1.2 for the time steps of 7:30 a.m., 7:45 a.m. and 8:00 a.m. respectively. The demand pattern coefficients at the corresponding time steps are adjusted by multiplying the original pattern coefficient by the optimised demand adjustment factor. The independent validation of the calibrated model parameters over 24 hours shows good agreement. Pressure head comparisons are illustrated for six nodes as shown in Figures 6–11 for the pressure logger locations. This demonstrates that high efficacy has been achieved for the roughness and peak-demand calibration of the case system. The remaining task is to calibrate the extended flow patterns for the rest of the demand hours.

Table 3. Optimised roughness factors for each group.

Roughness groups	Optimised roughness multipliers	Default D-W e (mm)	Optimised D-W e (mm)
Group 1	4.30	0.0015	0.00645
Group 2	7.30	0.0015	0.01095
Group 3	8.10	0.0015	0.01215
Group 4	1.50	0.0015	0.00225
Group 5	3.70	0.0015	0.00555
Group 6	2.40	0.0015	0.00360
Group 7	7.80	0.0015	0.01170
Group 8	0.80	0.0015	0.00120
Group 9	8.80	0.0015	0.01320

Table 4. Comparison of the field-observed and the model-simulated attributes for time 7:30 a.m.

Attributes	Elements	Observed attributes	Simulated attributes	Differences
HGL (m)	J-192	120.48	120.35	0.14
	J-656	118.37	119.05	0.68
	J-322	119.59	118.44	1.15
	J-132	118.23	118.98	0.75
	J-431	118.69	118.96	0.27
	J-267	119.63	119.45	0.18
Flows (L/s)	P-872	17.29	17.52	0.23

Extended-period flow pattern calibration

The primary goal for good ESP modelling is to ensure that the model-simulated flows match well with the field-observed flows over an extended period, typically 24 hours. For this model, the available flow data are the DMA inflows extracted every 15 minutes. To calibrate the inflow balance for the extended-period time steps, the model is updated with the previously calibrated parameters, including the optimised emitter coefficients, pipeline roughness and the demand pattern coefficients at peak hours. The updated model serves as a new baseline for EPS flow balance calibration.

To proceed with EPS flow balance calibration, the demand calibration is undertaken to optimise the demand pattern adjustment factor or multiplier with the specified range from 0.5 to 10.0 and an increment of 0.1. Each run is created to optimise the demand multiplier for one time step by using the field dataset collected at the corresponding time. The resulting multipliers are used to adjust the demand pattern coefficients. For instance, an optimisation run for optimising the demand multiplier at 10:00 a.m. is created by just using the field data snapshot at 10:00 a.m. The optimised demand factor is used to multiply the coefficients of all demand patterns at 10:00 a.m.

Multiple optimisation runs are created for optimising the demand of all the time steps except the time steps that were previously applied for leakage detection, roughness and peak demand calibration. A total of 89 demand optimisation runs are created and completed in batch mode. This completes the extended flow pattern calibration and results in the new demand patterns given in Table 5. Using the optimised demand patterns, an EPS run is conducted and the results of the simulated inflows are compared with the field observed inflows as shown in Figure 12. Before any leakage detection run is undertaken, there are big flow differences between the observed flows and the model-simulated flows of water consumption over 24 hours. The simulated flow matches better with the observed flows after leakage detection than that before leakage detection at the low-demand hours, but overshoots the observed flows for the other hours. It clearly indicates that a good agreement has been achieved with the extended-period flow balance calibration. The EPS flow pattern calibration, together with the model parameters identified by leakage detection optimisation, peak demand and roughness calibration, represents a complete procedure for parameter identification for this case system. This study demonstrates the essential steps for practical engineers to apply the unified optimisation approach for

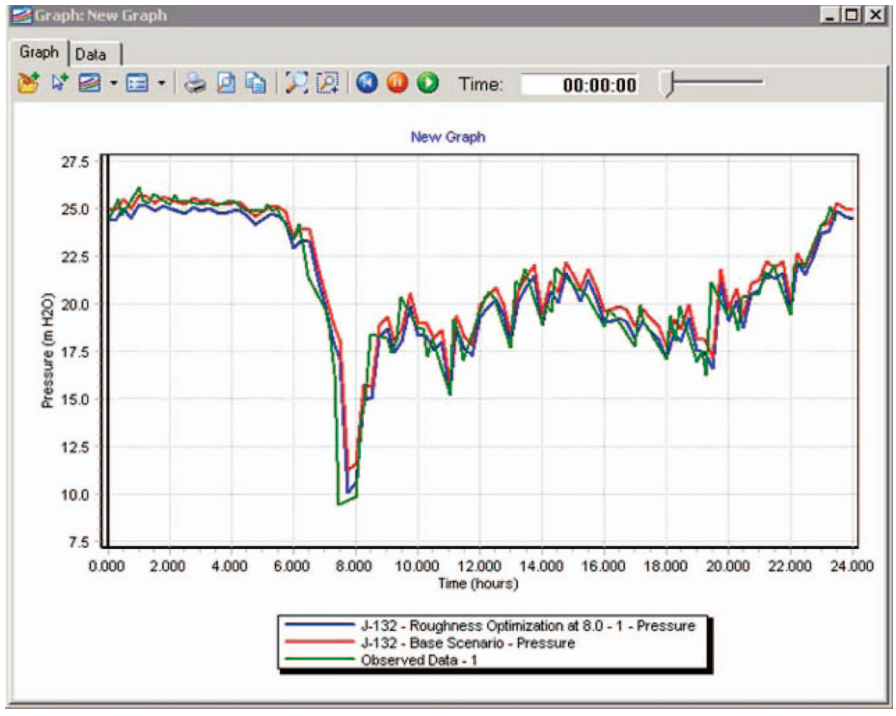

Figure 6. Pressure head comparisons between the observed and the simulated at node J-132.

Figure 7. Pressure head comparisons between the observed and the simulated at node J-322.

accomplishing the leakage detection and model calibration of a real water system. It offers a viable tool to assist modellers to predict leakage hotspots and construct effective hydraulic models for steady-state and extended-period simulation of a water distribution system.

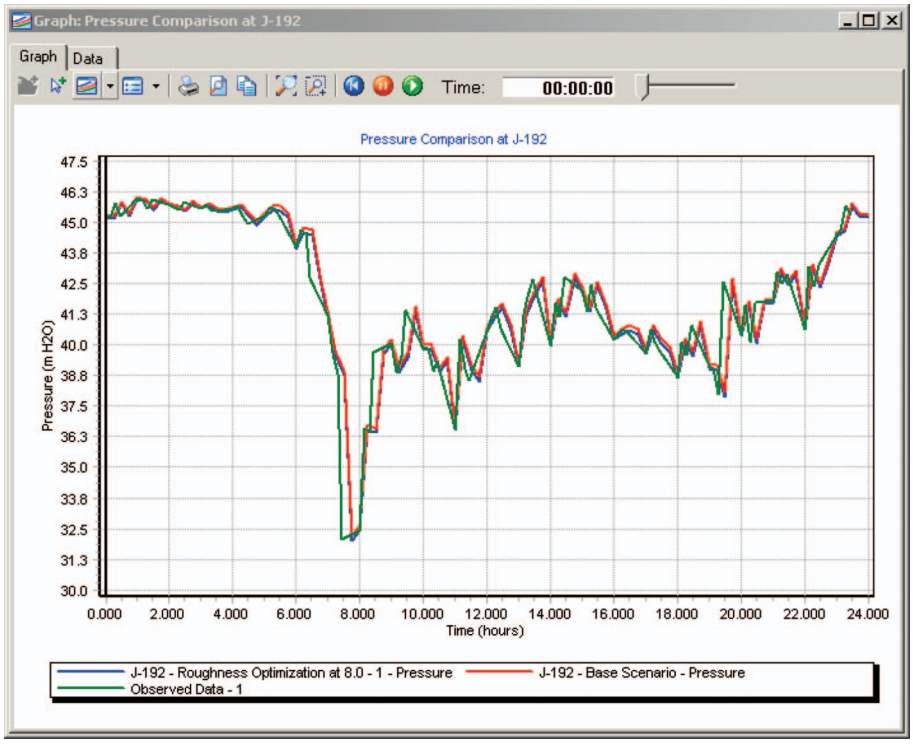

Figure 8. Pressure head comparisons between the observed and the simulated at node J-192.

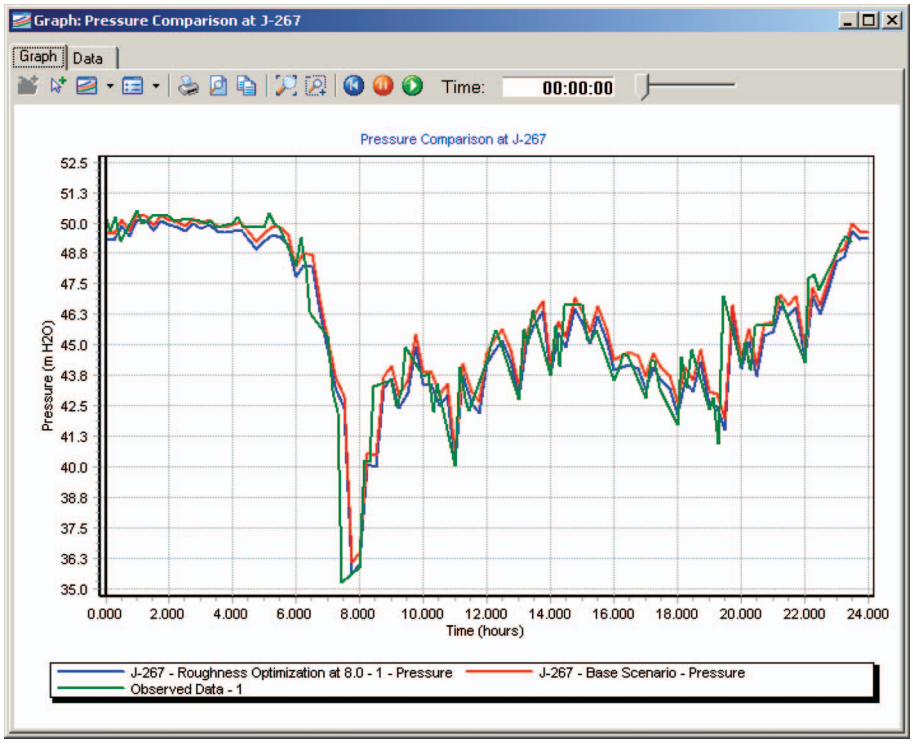

Figure 9. Pressure head comparisons between the observed and the simulated at node J-267.

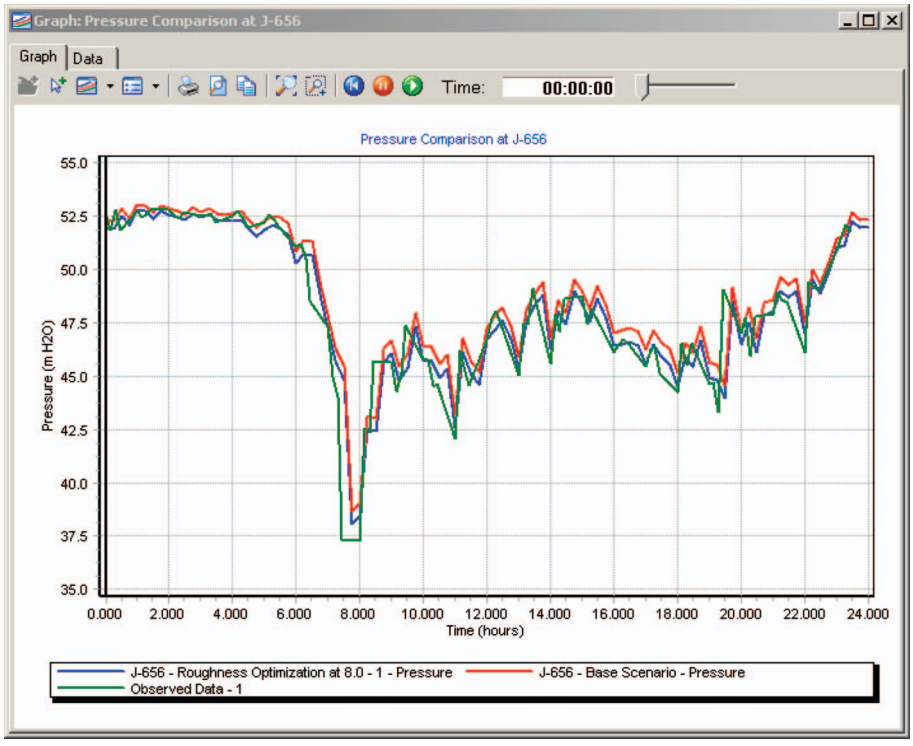

Figure 10. Pressure head comparisons between the observed and the simulated at node J-656.

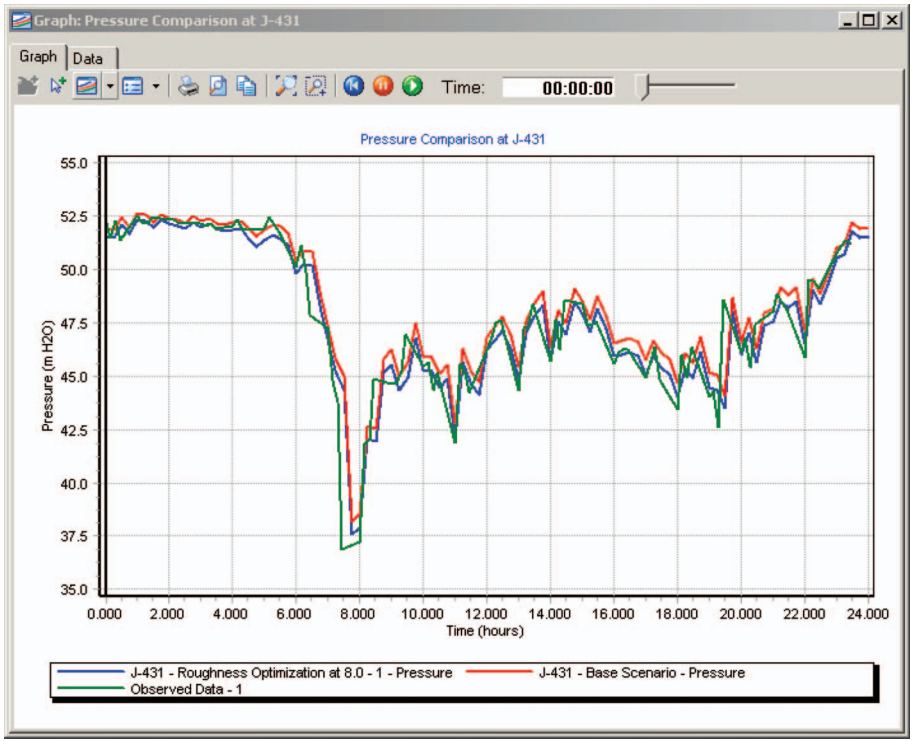

Figure 11. Pressure head comparisons between the observed and the simulated at node J-431.

Table 5. Optimised demand pattern coefficients over 24 hours.

Time	Pattern I	Pattern II	Time	Pattern I	Pattern II	Time	Pattern I	Pattern II
0:00	0.00204	0.00504	8:00	0.00601	0.03152	16:00	0.00408	0.01238
0:15	0.00213	0.00496	8:15	0.00540	0.02763	16:15	0.00467	0.01294
0:30	0.00192	0.00426	8:30	0.00489	0.02396	16:30	0.00486	0.01348
0:45	0.00191	0.00542	8:45	0.00490	0.02280	16:45	0.00501	0.01404
1:00	0.00183	0.00578	9:00	0.00485	0.02164	17:00	0.00511	0.01458
1:15	0.00183	0.00558	9:15	0.00504	0.02025	17:15	0.00531	0.01485
1:30	0.00171	0.00621	9:30	0.00526	0.01887	17:30	0.00530	0.01510
1:45	0.00174	0.00597	9:45	0.00517	0.01748	17:45	0.00544	0.01537
2:00	0.00177	0.00571	10:00	0.00524	0.01609	18:00	0.00540	0.01562
2:15	0.00191	0.00554	10:15	0.00521	0.01586	18:15	0.00567	0.01554
2:30	0.00191	0.00753	10:30	0.00480	0.01562	18:30	0.00519	0.01545
2:45	0.00180	0.00519	10:45	0.00478	0.01539	18:45	0.00523	0.01537
3:00	0.00168	0.00505	11:00	0.00485	0.01516	19:00	0.00522	0.01528
3:15	0.00191	0.00433	11:15	0.00498	0.01557	19:15	0.00542	0.01491
3:30	0.00171	0.00537	11:30	0.00464	0.01597	19:30	0.00466	0.01453
3:45	0.00172	0.00357	11:45	0.00451	0.01638	19:45	0.00485	0.01416
4:00	0.00175	0.00276	12:00	0.00415	0.01678	20:00	0.00441	0.01377
4:15	0.00180	0.00273	12:15	0.00480	0.01612	20:15	0.00509	0.01308
4:30	0.00196	0.00295	12:30	0.00449	0.01545	20:30	0.00462	0.01238
4:45	0.00212	0.00345	12:45	0.00435	0.01479	20:45	0.00425	0.01169
5:00	0.00232	0.00394	13:00	0.00407	0.01412	21:00	0.00367	0.01100
5:15	0.00261	0.00611	13:15	0.00372	0.01314	21:15	0.00455	0.01024
5:30	0.00312	0.00828	13:30	0.00357	0.01215	21:30	0.00391	0.00949
5:45	0.00313	0.01045	13:45	0.00400	0.01117	21:45	0.00417	0.00874
6:00	0.00341	0.01262	14:00	0.00384	0.01019	22:00	0.00349	0.00799
6:15	0.00444	0.01534	14:15	0.00370	0.01019	22:15	0.00339	0.00733
6:30	0.00412	0.01806	14:30	0.00378	0.01019	22:30	0.00280	0.00666
6:45	0.00436	0.02078	14:45	0.00395	0.01019	22:45	0.00267	0.00600
7:00	0.00461	0.0235	15:00	0.00388	0.01019	23:00	0.00241	0.00532
7:15	0.00532	0.02903	15:15	0.00413	0.01074	23:15	0.00262	0.00484
7:30	0.00554	0.02986	15:30	0.00397	0.01128	23:30	0.00225	0.00434
7:45	0.00624	0.03325	15:45	0.00407	0.01184	23:45	0.00189	0.00424

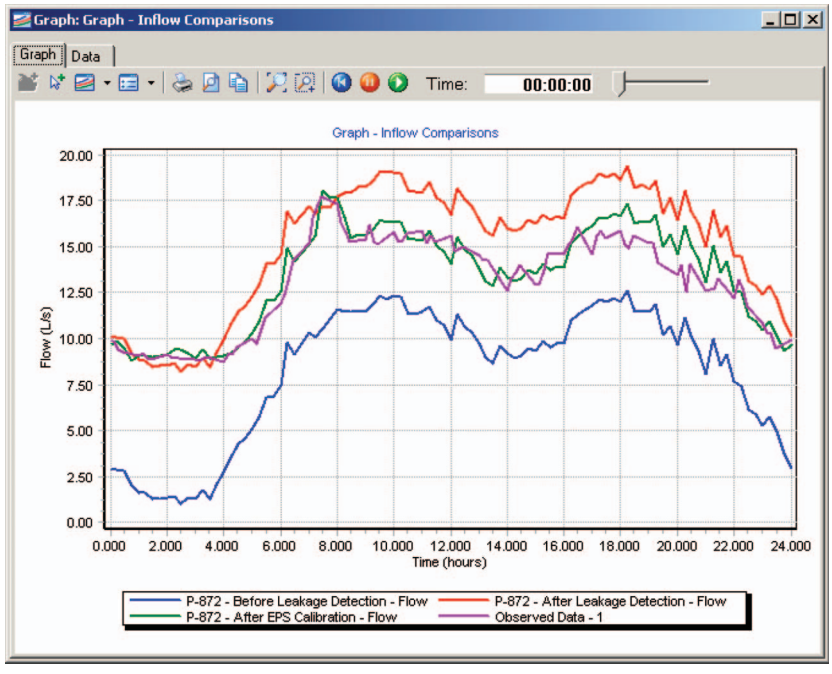

Figure 12. Flow comparisons after extended-period flow calibration.

Discussion

Although many methods have been developed for parameter identification for a water distribution model, great challenges still remain for most methods to be efficiently applied to the real systems. One of the most difficult issues is the uncertainty caused by the limited field data in comparison with the large dimension of overall model parameters to be identified. There is no ultimate method enabling the engineers to address the uncertainties that may exist in both field data and model parameters. It is the author's perspective that the pragmatic approach to minimising the uncertainty is to decompose the problem into a number of sub-tasks. Each task is solved by using relevant field datasets. This has been demonstrated by applying the unified methodology to the case study.

The study based on the real system has demonstrated that the leakage detection method has consistently predicted the leakage hotspots for different optimisation runs using the field data at minimum night flow hours. Those predicted locations represent the most likely leakage hotspot areas for the field crew to investigate. It should be noted that, when undertaking leakage hotspot predictions, there are likely to be instances where hotspots are predicted sporadically after running a number of simulations, rather than always at the same location. The model user will need to decide whether site investigation of the sporadic hotspots should be carried out. Engineering judgement for this may develop with experience. It must also be highlighted that, to achieve reliable predictions, the leakage hotspot prediction tools rely on reliable domestic and commercial demands being set up in the model to an appropriate level of detail. If the number of demand patterns or profiles used is insufficient, or demand profiles assigned are not representative of users, the tools may just highlight these deficiencies in demand allocation rather than identify true hotspots. A further consideration required by the model user relates to the accuracy of pressure loggers used during field testing. Providing equipment calibration is maintained, standard loggers used for modelling purposes are normally accurate to 0.1% of their full range (i.e. typically accurate to 0.1 m). As the leakage prediction tools predict hotspots based on differences in total heads between different monitoring locations in a DMA, logger accuracy is an important consideration, particularly when running simulations during the night when head losses across DMAs are often small.

Demand and roughness are two types of primary model parameters to be calibrated for real water systems. But both types of parameters may compensate each other, known as compensating error (Walski *et al*. 2001), which is another type of uncertainty that occurs in practical model calibration. Although the detected leakage emitters reduce the demand uncertainty to a large extent especially at low-demand hours, calibrating the demand and roughness at peak hours is an important step to further reduce the possible compensating error. In general, the model calibration based on the detected leakage emitters is useful for simulating the system hydraulics of the current conditions prior to any leakage hotspot being repaired. The model needs to be recalibrated after actual leakage hotspots are located and fixed. This indicates that the model calibration is not a one-off task, but a repetitive task for adjusting the model to emulate the latest system conditions.

In addition to the data limitation as discussed above, a hydraulic model is just an approximate mathematic representation of a real water system. One cannot expect to achieve absolutely correct or accurate modelling results, but just useful results for guiding an informed engineering decision-making process. It is the perception that no correct model exists, but that only a useful model can be constructed for engineering systems, and that motivates the development of the unified formulation for water distribution leakage detection and extended-period model calibration. It is not a panacea method for calibrating a hydraulic model, but strives to provide a pragmatic approach for engineers to undertake the practical challenges in solving leakage detection and building effective models of water distribution systems.

Conclusions

A unified approach has been developed for leakage detection and model calibration. Both types of tasks are generalised as a parameter identification problem with the same objective functions. The generalised formulation is solved by using fast messy GA and implemented in an integrated solution method. A real water system is demonstrated as an example of applying the method.

The results of the case study illustrate that integrated optimisation modelling is a powerful tool for identifying the most likely leakage areas to support water companies in forming cost-effective leakage reduction strategies. Although it is impossible to exactly locate the water losses or leakages in a distribution system by just using the integrated optimisation–simulation approach, the model-predicted leakage hotspots help engineers to narrow down the possible areas of water loss (including leakages, unmetered and illegal consumption, etc.) and thus enable more efficient leakage reduction programmes. The case study results illustrate that the

identified leakages are the important uncertain demands to improve the model accuracy. The unified optimisation approach is also flexible at conducting steady-state and extended-period model calibration. Based on the results of leakage detection optimisation at minimum night flow hours, roughness and demand are calibrated for the peak demand hours. The calibration results further streamline the flow pattern optimisation at any desired time step for EPS model simulation. Therefore, the unified method provides a systematic modelling tool for water engineers to efficiently construct an effective hydraulic model and reduce water losses in a cost-effective manner. It consequently facilitates the sustainable development of urban water systems.

Acknowledgements

United Utilities Water and Atkins Water and Environment Group in the UK are gratefully acknowledged for providing the hydraulic network model and field data used in this paper. The results presented, however, do not represent or imply the opinions of the organisations.

References

Covas, D., Ramos, H., and Almeida, A.B. de, 2005. Standing wave difference method for leak detection in pipeline systems. *ASCE Journal of Hydraulic Engineering*, 131 (12), 1106–1116.

Kapelan, Z., Savic, D., 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.

Lambert, A.O., 1994. Accounting for losses–background and burst estimates concepts. *Journal of the Institution of Water and Environmental Management*, 8 (2), 205–214.

Lambert, A.O., 2002. International report on water losses management techniques. *Water Science and Technology: Water Supply*, 2 (4).

Lambert, A.O., Brown, T.G., Takizawa, M., and Weimer, D., 1999. Review of performance indicators for real losses from water supply systems. *Journal of Water Supply: Research and Technology-AQUA*, 48, 227–2237.

Lansey, K., 2006. The evolution of optimizing water distribution systems applications. *In*: *Proceedings CD of the 8th Annual International Water Distribution Systems Analysis Symposium*, 27–30 August, Cincinnati, OH, USA.

Lee, P.J., Vítkovský, J.P., Lambert, M.F., Simpson, A.R., and Liggett, J.A., 2005. Frequency domain analysis for detecting pipeline leaks. *ASCE Journal of Hydraulic Engineering*, 131 (7), 596–604.

Nixon, W., Ghidaoui, M.S., and Kolyshkin, A.A., 2006. Range of validity of the transient damping leakage detection method. *ASCE Journal of Hydraulic Engineering*, 132 (9), 944–957.

Ormsbee, L., 2006. The history of water distribution network analysis: the computer age. *In*: *Proceedings CD of the 8th Annual International Water Distribution Systems Analysis Symposium*, 27–30 August, Cincinnati, OH, USA.

Taghvaei, M., Beck, S.B.M., and Staszewski, W.J., 2006. Leak detection in pipelines using cepstrum analysis. *Measurement Science and Technology*, 17, 367–372.

Vítkovský, J.P., Simpson, A.R., and Lambert, M.F., 2000. Leak detection and calibration using transients and genetic algorithms. *ASCE Journal of Water Resources Planning and Management*, 126 (4), 262–265.

Wang, X.-J., Lambert, M.F., Simpson, A.R., Liggett, J.A., and Vítkovský, J.P., 2002. Leak detection in pipelines using the damping of fluid transients. *ASCE Journal of Hydraulic Engineering*, 128 (7), 697–711.

Walski, T., Chase, D., and Savic, D., 2001. *Water distribution modelling*. Waterbury, CT: Haestad Press.

Wu, Z.Y. and Sage, P., 2006. Water loss detection via genetic algorithm optimization-based model calibration. *In*: *Proceedings CD of the 8th Annual International Water Distribution Systems Analysis Symposium*, 27–30 August, Cincinnati, OH, USA.

Wu, Z.Y. and Sage, P., 2007. Pressure dependent demand optimization for leakage detection in water distribution systems. *In*: *Proceedings of CCWI2007*, 3–5 Sept., Leicester, UK.

Wu, Z.Y. and Simpson, A.R., 2001. Competent genetic algorithm optimization of water distribution systems. *Journal of Computing in Civil Engineering, ASCE*, 15 (2), 89–101.

Wu, Z.Y., Walski, T., Mankowski, R., Cook, J., Tryby, M., and Herrin, G., 2002. Calibrating water distribution model via genetic algorithms. *In*: *Proceedings of the AWWA IMTech Conference*, 16–19 April, Kansas City, MI.