



HAL
open science

A Hybrid Method for Calibration of Unknown Partially/Fully Closed Valves in Water Distribution Systems

Nhu Cuong Do, Olivier Piller, Angus Simpson, Jochen Deuerlein

► **To cite this version:**

Nhu Cuong Do, Olivier Piller, Angus Simpson, Jochen Deuerlein. A Hybrid Method for Calibration of Unknown Partially/Fully Closed Valves in Water Distribution Systems. 1st International WDSA / CCWI 2018 Joint Conference, Jul 2018, Kingston, Canada. 8 p. hal-01890935

HAL Id: hal-01890935

<https://hal.science/hal-01890935>

Submitted on 9 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A HYBRID METHOD FOR CALIBRATION OF UNKNOWN PARTIALLY/FULLY CLOSED VALVES IN WATER DISTRIBUTION SYSTEMS

Nhu Cuong Do^{1,2}, Angus Simpson³, Jochen Deuerlein⁴, Olivier Piller⁵

¹University of Saskatchewan, Saskatoon, Saskatchewan, Canada

²Thuyloi University, Hanoi, Vietnam

³University of Adelaide, Adelaide, South Australia

⁴3S Consult GmbH, Karlsruhe, Germany

⁵Irstea UR ETBX, Dept. of Water, Cestas, France

¹*nhu.do@usask.ca*

ABSTRACT

Water distribution systems (WDSs) are an essential part of the urban infrastructure, which consist of a labyrinth of networks of pipes, tanks, pumps and monitoring systems. In order to provide water of appropriate quantity, quality and pressure to customers, a WDS needs to be stable and reliable. Isolation valves, which are installed along the pipelines, are designed to provide system reliability. By closing related isolation valves, a pipe or a portion of the network can be isolated for inspection, maintenance and replacement without disrupting other parts of the system. However, these isolation valves may cause adverse operating conditions in the network if some of the valves in the system are active (partially or fully closed) while their statuses are unknown (due to poor or non-existent documentation, valves left closed inadvertently, errors in data transfer or valve mechanical failure, etc.). This can be observed by a disagreement between the measured pressure and flow values from the real system and results from its corresponding hydraulic simulation model. Calibration of these unknown valve statuses is, therefore, a necessary step that has to be implemented to ensure the reliability of the system.

This paper introduces a hybrid method for the identification of unknown partially/fully closed valves in a water distribution network. An optimization problem is formulated for unknown valve statuses and solved by application of three sequentially applied methods, which include: a local sensitivity analysis, the application of genetic algorithms and the application of the Levenberg-Marquardt algorithm. In order to eliminate the valves that are unable to be localized by the measurement data, first, the sensitivity of the flow rates and nodal heads at measurement locations with respect to the change in the minor losses of the valves is computed. Following, a genetic algorithm combined with an extended period simulation is applied to preliminarily identify the locations of the partially/fully closed valves and their degree of opening of the valves. Finally, the application of the Levenberg-Marquardt algorithm has been implemented to correct the results from the GA model.

A water distribution network is used to demonstrate the applicability of the proposed methodology. Results show that the model is able to provide relatively good approximation of the locations of unknown partially/fully closed valves as well as their corresponding settings.

Keywords: Partially closed valves, Genetic Algorithms, Levenberg - Marquardt, Sensitivity analysis, Water distribution systems, Calibration

1 Introduction

Water distribution systems (WDS) are responsible for supplying water of appropriate quantity, quality and pressure to customers. This continuous delivery of water is highly dependent on the stability and reliability of the pipeline network, in which isolation valves play an important task in the maintenance process of the system, both preventive and corrective.

Isolation valves serve to separate some portions of the network where the maintenance work can be performed without disrupting service of the entire network. These valves are usually placed at the ends of a pipe, around junctions or at critical locations of the WDS. According to [1] and [2], at the least $(N_p - 1)$ isolation valves are required to be installed around a junction to which N_p links are connected while an ideal (or fully valved) system requires two valves at the opposite ends of each pipe (i.e. $2 * N_p$ valve rule). Given that rule of thumb, there are often a large number of isolation valves in every distribution network.

Although being installed to ensure the reliability of the WDS, isolation valves, by themselves, may cause adverse operating conditions in the network if one or some of the valve statuses are unaccounted for or unknown. Problems related to these valves possibly come from a number of reasons such as: missing valves in a hydraulic model due to poor or non-existent documentation, valves left closed inadvertently, errors in data transfer, valve mechanical failure or temporarily closed valves during inspection/rehabilitation times without adequate reporting or have been missed for reopening. This can be observed by a disagreement between the measured pressure and flow values from the real system and results from its corresponding hydraulic simulation model. Calibration of these unknown valve statuses is, therefore, a necessary step that has to be implemented.

The most challenging aspect for the calibration problem of unknown valve statuses in WDSs is that it is highly underdetermined. As the number of uncertain parameters (i.e the status of each isolation valve in the network) is larger than the total number of independent observations (i.e. pressure and flow measurements), solving the underdetermined problem usually leads to non-unique solutions [3]. In such a case, different combinations of the valve statuses can result in a good agreement between observed and modeled behavior of the system, which also means that the correct locations of unknown valve statuses could not be identified.

In order to overcome this difficulty, [4] has proposed a hybrid model to calibrate and, more importantly, identify the locations of unknown partially/fully closed valves in a water distribution network. By integrating measurement data into the demand driven model EPANET [5] through three sequential steps: a sensitivity analysis, the application of a Genetic Algorithm (GA) model and the application of the Levenberg-Marquardt (LM) algorithm, the regions of partially closed valves and their settings can be approximated. The research work presented in this paper evaluates the outputs from the hybrid model. New developments focus on the calibration and identification of unknown partially closed valves given the uncertainty from observation data.

2 Methodology

This section provides an overview of the methodology proposed by [4] (see [4] for the detailed methodology). In their paper, the problem of finding partially closed valves is considered as the problem of finding accumulated minor losses in the pipes where these valves are located. The proposed solution approach is mainly based on the minimization of the difference between observed and simulation model values for flows and pressures and consists of three sequential stages (Figure 1): (1) a local sensitivity analysis, which constructs the sensitivity matrix of flow rates and pressure heads at measurement locations with respect to (wrt) the variations of the minor loss in the valves. This matrix is used to remove valves that are unobservable; (2) the application of a Genetic Algorithm

(GA) model during an extended period simulation (e.g. 24, 48 hours) to reduce the size of the search region as well as to provide a preliminary estimate of the settings of the valves; and (3) the application of the LM algorithm to localize the regions of partially closed valves and correct the statuses of the valves.

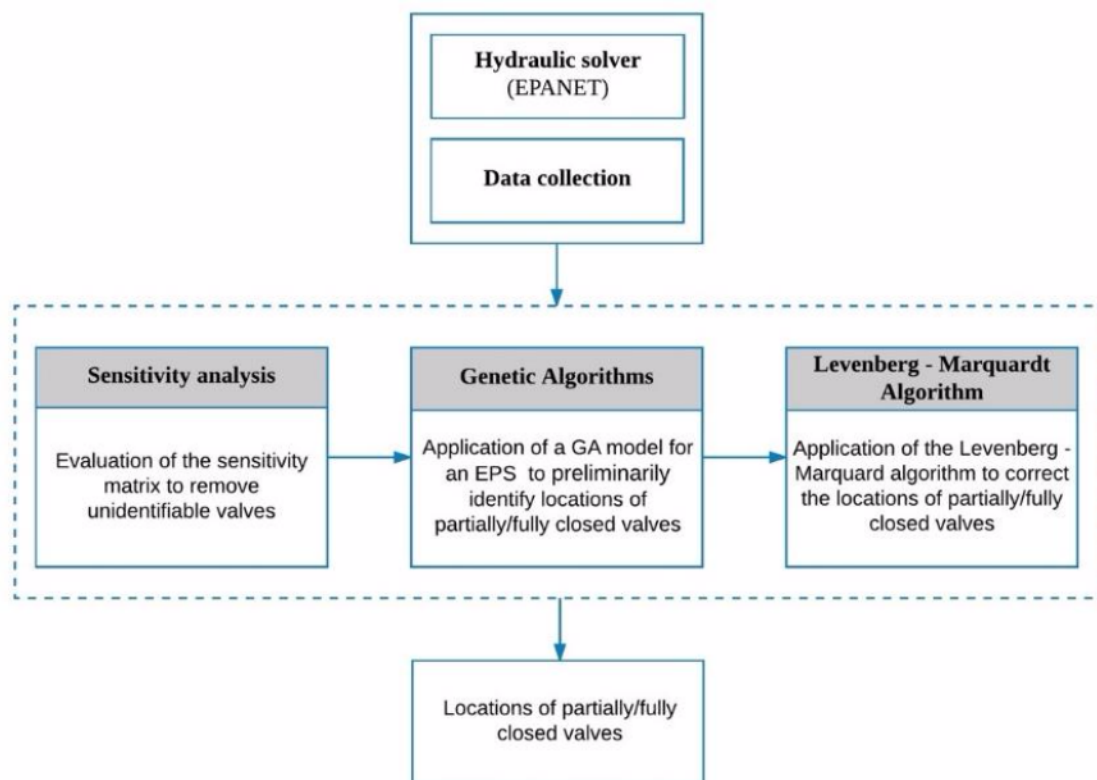


Figure 1. Process for identifying partially closed valves in WDS [4]

2.1 Application of local sensitivity analysis

The local sensitivity analysis is firstly implemented in order to identify “unobservable valves”. An unobservable valve is defined as a valve in a pipe for which different degrees of opening of the valve have no impact on the observations (i.e. the values at measured locations). This can be identified by computing the sensitivity of the nodal heads ($\nabla_m h$) and pipe flow rates ($\nabla_m q$) in the network wrt variation of minor loss coefficient (m) in pipes, expressed as:

$$\nabla_m h = [A^T(F + M)^{-1}A]^{-1}A^T(F + M)^{-1}B \quad (1)$$

$$\nabla_m q = (F + M)^{-1}A[A^T(F + M)^{-1}A]^{-1}A^T(F + M)^{-1}B - (F + M)^{-1}B$$

where: - A and A^T are the unknown head node incidence matrix and its transpose matrix.

- B is the diagonal matrix with diagonal elements $q_i|q_i|$ (q_i is the flow rate in the i^{th} pipe).

- $F = \nabla_q g(r(q), q)$ is the partial derivative of the friction head loss function $g(r(q), q)$ wrt the flow rates q (where $r(q)$ is the pipe resistance factors).

- M is the diagonal matrix whose diagonal elements are $2m_i|q_i|$.

Based on the local sensitivity analysis, i.e. Eq. (1), the sensitivity matrix of size $(N_H + N_Q, n_p)$ is obtained where the number of rows is the total number of measured locations $[N_H + N_Q]$ (N_H is the number of pressure measurements and N_Q is the number of flow measurements) and the number of columns is equal to the number of pipes n_p in the network.

A pipe flow rate or a nodal head that is not sensitive to the variation of the minor loss of a valve in a pipe in the network results in a zero value at its corresponding element in the sensitivity matrix. Therefore, the pipes that are insensitive to all of the measured locations of the network will have all zero values in their column of the sensitivity matrix. These corresponding pipes are unobservable and need to be removed from the set of unknowns in order to reduce the search space size of the problem.

2.2 Application of Genetic Algorithms

The objective function of the GA model is formulated by the sum of squared residuals between the measured and simulated values of pipe flow rates and nodal heads at measurement locations during a selected period of time, expressed as:

$$\text{Min } G(K) = \frac{1}{2} \sum_{t=1}^{N_T} \left[\sum_{i=1}^{N_Q} w_i (Q_i^{\text{Sim}} - Q_i^{\text{Meas}})^2 + \sum_{j=1}^{N_H} w_j (H_j^{\text{Sim}} - H_j^{\text{Meas}})^2 \right] + \text{pen} \quad (2)$$

where: - $Q_i^{\text{Sim}}, H_j^{\text{Sim}}$ are the simulated flow rate and nodal head for the i^{th} pipe and j^{th} node.

- $Q_i^{\text{Meas}}, H_j^{\text{Meas}}$ are the measured flow rate and head for the i^{th} pipe and j^{th} node, respectively.

- N_Q, N_H are number of head and flow measurement sites in the network.

- w_i, w_j are weighting factors applied to different terms to ensure they have similar magnitude.

- N_T is the number of time steps during the simulation period.

- pen is a penalty term (pen), which is added into the objective function to constrain cases in which the EPANET solver generates negative pressures at some nodes of the network.

For the selection of the number of decision variables, a preprocessing step is implemented in order to reduce the search space and regularize the problem. According to [4], the GA model should consider only a single pipe for each series of pipes in the network. If the method detects a minor loss in that pipe, it will be necessary to physically check all the valves along the series of the pipes in this area. The decision variables for the GA model, therefore, are the minor loss setting values K (the loss coefficients of pipes in EPANET) for the remaining pipes as selected from the sensitivity analysis and the preprocessing steps. The range of the decision variables is $0 \leq K \leq K_{max}^{GA}$, where the value of $K = 0$ represents the situation of no minor loss in the pipe, i.e. all the valves in that pipe are fully opened, and the value of $K = K_{max}^{GA}$ represents the situation for a closed valve in the pipe. In this case, the status of the pipe in the EPANET solver is switched to “Closed”, which does not allow any flow through the pipe.

Tournament selection, two-point crossover and bitwise mutation were chosen for the GA model. In addition, an integer coding scheme has been applied. The values of the loss setting K , therefore, are taken as discrete values, with the increment of ΔK being selected based on the size of specific network. The pipes with positive K values from the GA model are used as initial values for and will be refined by the LM algorithm. Pipes with $K = 0$ found from the GA model are assumed to be fully opened and not considered in the LM model.

2.3 Application of Levenberg-Marquardt Algorithms

The application of the LM method is to fine-tune the output from the GA model. The objective of the LM model is similar to the objective function of the GA model, i.e. Eq. (2), without the penalty term. The minor losses in the pipes are estimated with the following expression:

$$\left[\begin{array}{l} m^{i+1} = m^i - \left(\sum_{t=1}^{N_T} [J_{i,t}^T J_{i,t} + \lambda_i S_{i,t}] \right)^{-1} \sum_{t=1}^{N_T} J_{i,t}^T r_{i,t}; K^{i+1} = 2gA^2 m^{i+1} \\ \text{if } K^{i+1} \leq 0, \quad \text{set } K^{i+1} = 0 \rightarrow m^{i+1} = 0 \\ \text{if } K^{i+1} \geq K_{max}^{LM}, \quad \text{set } K^{i+1} = K_{max}^{LM} \rightarrow m^{i+1} = m_{max}^{LM} \end{array} \right. \quad (3)$$

where: - m^i, m^{i+1} is the minor loss coefficient associated with the minor loss setting K^i, K^{i+1} at iteration i and $(i + 1)$, respectively.

- $J_{i,t}$ and $J_{i,t}^T$ are the Jacobian matrix and its transpose at time $T = t$ and at iteration i . The Jacobian matrix is a selection of the rows of $\nabla_m h$ and $\nabla_m q$ (in Eq. 1) corresponding to the measurement locations.

- λ_i is the damping factor for the LM algorithm. This value was selected based on [6], in which λ_i starts at 0.0001 and is changed throughout the iterations based on the least squares criterion (i.e. (Eq. 2)): if the residual sum of squares decreases sufficiently and the iteration matrix is invertible, λ_i is set to be decreased by 60% (i.e. multiplied by a factor of 0.4) for the next iteration. Otherwise, λ_i is forced to be increased by a factor of 10 and a new least squares value is recalculated based on the new value of λ_i . Within the iteration, this process is repeated until the least squares value decreases.

$$- S_{i,t} = \text{diag}(J_{i,t}^T J_{i,t}).$$

- $r_{i,t}$ is the weighted residual at time $T = t$ and at iteration i .

- K_{max}^{LM} is the maximum value of the minor loss setting. For the case study in this paper, $K_{max}^{LM} = 500,000$ was selected, which has been shown to be sufficient to represent an almost closed pipe in the network.

Two stopping criteria were applied for the LM model: either the change of the K values after every iteration is less than $\|\Delta K_i\| = 0.01$ or the maximum number of iterations is reached $N_{iter} = 200$. N_{iter} is used to terminate the method if the function in (Eq. 2) cannot converge to a solution.

3 Case study

The case study used to evaluate the model is shown in Figure 2.a, which is modified from the Net3 network example of the EPANET software. The network has 2 reservoirs, 3 tanks, 92 nodes, 115 pipes and 2 pumps. It is assumed that the valve system of the network follows the $2 * N_p$ valve rule [2], corresponding to 230 valves across the network. Among them, two valves are partially closed, in which a valve at pipe 179 has the setting values of KP179=6500 and the other at pipe 231 with the setting of KP231=2890. These valves are required to be localized by the proposed methodology.

Figure 2.a also plots the locations of 18 measurement devices (10 flow measurements and 8 pressure measurements), which are assumed to be placed within the network in order to record data at every hour. Data for a period of 48 hours are assumed to be available to be used for the model. The input dataset is synthetically generated based on the determined network, where all the network parameters are fully known, as follows: (1) a demand pattern is assigned to nodal demands; (2) EPANET is run to record nodal heads and flow rates at selected measurement locations; (3) to introduce the measurement errors, each nodal head or flow rate has a random error added to it in an allowable range of 1% of its value.

The total inflow ($\sum Q_t^{in}$) and outflow ($\sum Q_t^{out}$) at a time step t are measured by the flow measurements at all sources, including two reservoirs (the river and the lake) and three tanks. Hence, the demand multiplier of the demand pattern at each time step, which is required to be known in EPANET before executing the proposed model, can be computed as:

$$DMF_t = \frac{\sum Q_t^{In} - \sum Q_t^{Out}}{\sum_{i=1}^{nj} D_i^0} \quad (4)$$

where: - D_i^0 is the base demand of i^{th} node in the network.

Figure 2.b shows the results from the local sensitivity analysis, in which 15 out of 115 pipes are unobservable (represented as the green lines), and the pipes selected for the GA model after preprocessing. By considering only a single pipe in each series of pipes, the application of the GA model is limited to 70 decision variables. For this case study, the following parameters were set for the GA model: the probability of crossover $P_c = 0.75$, the probability of mutation $P_m \approx 1/L = 0.015$ ($L = 70$ is the length of a chromosome), population size $N = 1000$, the number of generations was selected as $N_{stop} = 1500$ and the range of the decision variables was selected from 0 to 10,000 with an increment of 1000.

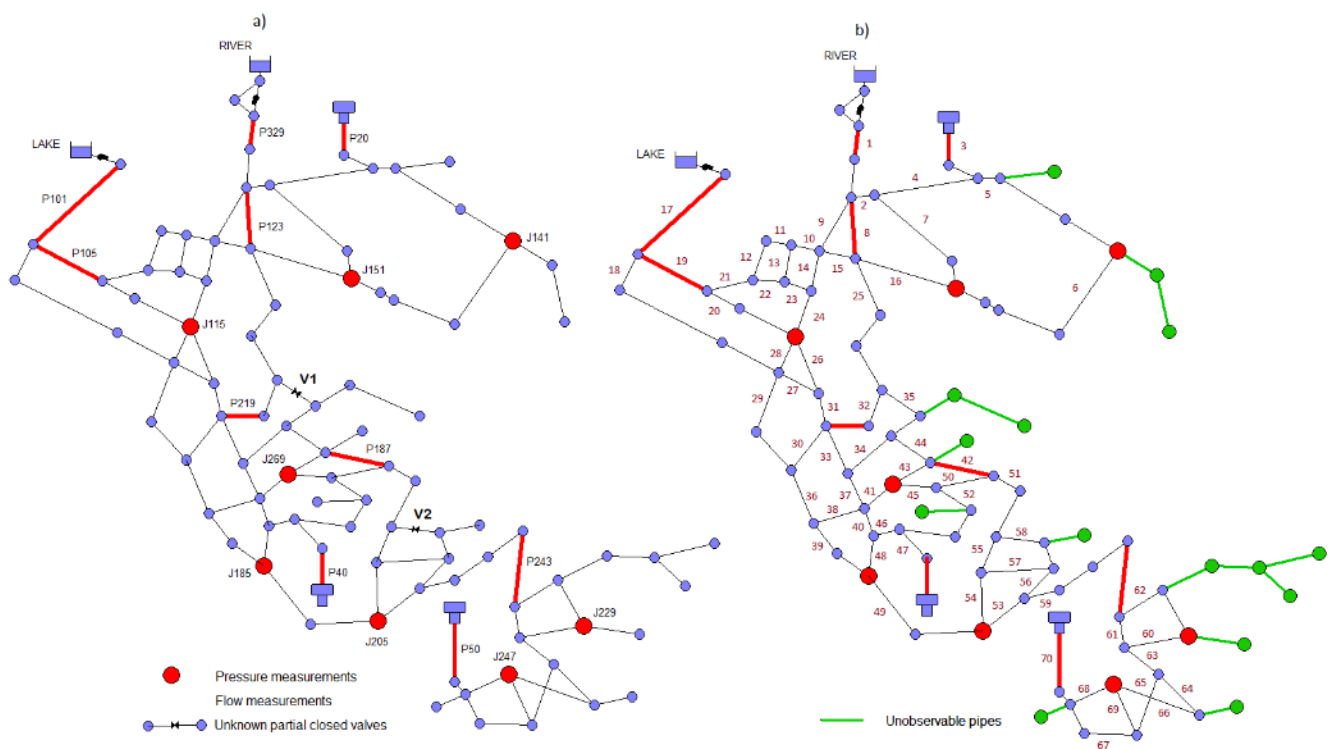


Figure 2. (a) Network characteristics and locations of two partially closed valves, (b) Results from local sensitivity analysis (unobservable pipes are represented by green color) and preprocessing step.

Figure 3.a shows the output from the GA model. It is seen that the model has significantly reduced the size of the search space, from 70 pipes to 13 pipes that possibly contain minor losses. The minor loss values of 13 pipes are shown in Table 1. The minor losses of the other 57 pipes are all equal to zero, i.e. all the valves in these pipes are fully opened.

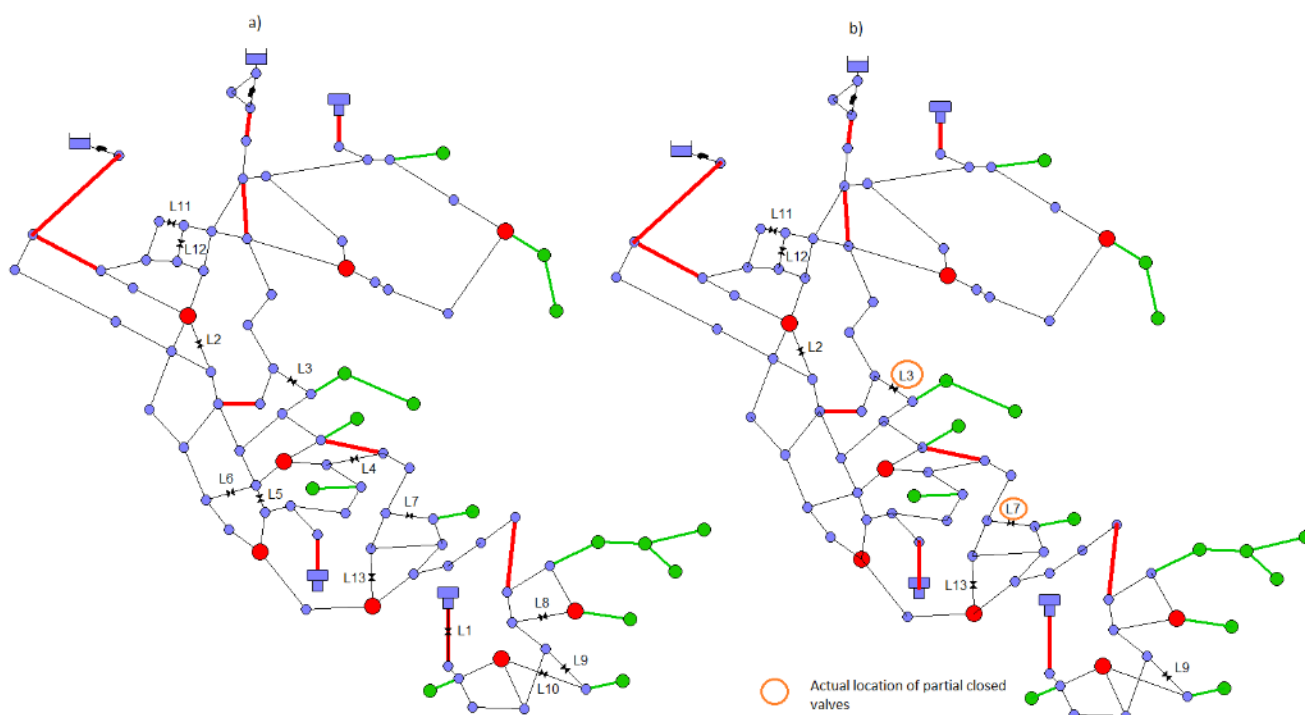


Figure 3. (a) Results from GA model, (b) Results from LM model.

The results from the LM model is presented in Figure 3.b and Table 1. This fine-tuning step has additionally eliminated 6 out of 13 pipes from the search space, which means only 7 pipes in the system need to be examined. Due to measurement errors and possibly the non-uniqueness of solutions as explained in [4], it is unlikely that the model converges to the actual solution. However, if the main aim is to identify the locations of partially/fully closed valves, the results from the LM model are acceptable in this case. With the application of the proposed methodology, the partially/fully closed valve problem has been simplified from the search of 115 pipes (i.e. 230 valves) to 7 pipes, which is equivalent to the search of only 14 valves in the network.

Table 1: Minor loss setting values obtained from GA and LM models

Locations	Pipe ID	Valve setting (K) from GA model	Valve setting from LM model	Actual valve setting values
L1	50	9,000	0	0
L2	114	9,000	Closed	0
L3	179	4,000	3,951.41	6,500
L4	191	1,000	0	0
L5	207	9,000	0	0
L6	209	7,000	0	0
L7	231	1,000	1,652.56	2,890
L8	271	9,000	0	0
L9	275	8,000	32.17	0
L10	281	8,000	0	0
L11	299	1,000	3,319.54	0
L12	303	1,000	3,549.89	0
L13	319	6,000	66.89	0

4 Conclusions

The work in this paper has evaluated the performance of the model proposed by [4] for the calibration of unknown partially/fully closed valves in water distribution networks under uncertainty. In particular, the model has been applied to a new network case study, in which inaccuracies from observations have been incorporated. Results show that although the actual solution is unlikely to be found due to the highly underdetermined feature of the problem as well as the errors from measurement devices, the methodology still can identify very well the locations of partially closed valves in the system. The proposed methodology, therefore, could be used to solve the problem of finding unknown partially/ fully closed valves in practice.

5 References

- [1] TM. Walski, JS. Weiler and T. Culver 2006, 'Using criticality analysis to identify impact of valve location', paper presented at Water Distribution Systems Analysis Symposium 2006.
- [2] H. Jun and GV. Loganathan 2007, 'Valve-controlled segments in water distribution systems', *Journal of Water Resources Planning and Management*, vol. 133, no. 2, pp. 145-155.
- [3] TM. Walski, DV. Chase, DA. Savic, W. Grayman, S. Beckwith & E. Koelle 2003, 'Advanced water distribution modeling and management', Haestad Method Inc., USA
- [4] NC. Do, AR. Simpson, JW. Deuerlein and O. Piller 2018, 'Locating inadvertently partially closed valves in water distribution systems', *Journal of Water Resources Planning and Management*, DOI: 10.1061/(ASCE)WR.1943-5452.0000958.
- [5] LA. Rossman, 2000, 'EPANET 2: Users manual', US Environmental Protection Agency (EPA), USA.
- [6] O. Piller and JE. van Zyl 2014, 'Modeling Control Valves in Water Distribution Systems Using a Continuous State Formulation', *Journal of Hydraulic Engineering*, vol. 140, no. 11, p. 04014052.