Memorias del 2022 Congreso Nacional de Control Automático
12-14 de Octubre, 2022. Tuxtla Gutiérrez, México.

Model Calibration for a Hydraulic Network Using Genetic Algorithms

L. Gómez-Coronel * **I. Santos-Ruiz** * **L. Torres** ** **F.R. López-Estrada** * **J.A. Delgado-Aguiñaga** ***

* *Tecnológico Nacional de México, I.T. Tuxtla Gutiérrez, TURIX-Dynamics Diagnosis and Control Group, Carretera Panamericana km 1080 S/N, Tuxtla Gutierrez 29050, Mexico; M16270670@tuxtla.tecnm.mx, ildeberto.dr@tuxtla.tecnm.mx, frlopez@tuxtla.tecnm.mx*

** *Instituto de Ingeniería, Universidad Nacional Autónoma de México, 04510 Ciudad de México, México; ftorreso@iingen.unam.mx*

*** *Centro de Investigación, Innovación y Desarrollo Tecnológico CIIDETEC-UVM, Universidad del Valle de México, Periférico Sur Manuel Gómez Morín 8077, C.P. 45601, Tlaquepaque, Jalisco, Mexico; jorge.delgado@uvmnet.edu*

Abstract: This paper addresses the problem of parameter calibration in pipelines based on a Genetic Algorithm (GA). The parameters under consideration are the pipe roughness and the minor loss coefficient caused by fittings like valves, elbows, and couplings. These parameters cannot be directly measured, and their accuracy plays an essential role in successfully implementing leak diagnosis algorithms. The proposed GA generates calibrated values for both pipe roughness and minor loss coefficient by minimizing the root mean squared error (RMSE) in predicting pressures. The method was implemented in MATLAB and the calibration was validated in an experimental network by comparing the pressure heads measured at the nodes of the network and those from the calibrated model simulated with EPANET.

Keywords: Model calibration; Genetic algorithm; Optimization; Hydraulic network; Head loss.

1. INTRODUCTION

Water management companies often use hydraulic modeling and simulation to design efficient water distribution systems (WDS) and correctly manage them under operational conditions. Well-calibrated hydraulic modeling allows to implement efficient leak diagnosis strategies from the model-based approach. However, this task is not trivial because some pipeline parameters cannot be directly measured online or are time-varying, for example, the roughness coefficient and the minor loss coefficient. The reliability of such modeling depends on the accuracy of both physical and hydraulic parameters of the WDS. This work proposes a calibration strategy for a WDS employing genetic algorithms (GA).

Optimizing hydraulic models comprises computing the values of roughness and minor loss coefficients that provide an optimal fit between simulations and measurements. Previous investigations have addressed this task by using GA. For example, Do et al. (2016) showed how through multiple iterations of the GA it was possible to adjust the values of nodal demands and flow rates at unsensed nodes. It was highlighted that the GA should be implemented in combination with a decision support tool for the selection of optimal sensing nodes. Calibration of the hydraulic parameters of a WDS using heuristic methodological approaches, especially the GA, has also been addressed before. In (Drisya and Sathish Kumar, 2018) the Manning's roughness coefficient was estimated through a calibration process addressed as an optimization problem solved through the use of a genetic algorithm.

In hydraulic analysis, EPANET software (Rossman et al., 2020) has proven to be a powerful design tool for improving the performance of existing WDS as well as for new designs. By combinig the characteristics of EPANET together with the computing power of MATLAB, a powerful tool for hydraulic design and recalibration is available. Furthermore, in Heydari et al. (2020) EPANET software was used in combination with a GA in a MAT-

⋆ This research was supported by Tecnológico Nacional de México under grant 14515.22-P. The work of L. Gómez-Coronel is funded by CONACYT (Mexico) in the Scholarship Program for Postgraduate Studies.

Corresponding author: *ildeberto.dr@tuxtla.tecnm.mx*

LAB environment for the optimization of the full design cost by considering three different WDS. In Martínez-Bahena et al. (2018) an hydraulic model was designed in EPANET software considering some restrictions in terms of the mass and energy conservation laws. The model was optimized by using a GA aiming to find an optimal solution focused on the selection of new elements such as pressure reducing valves, tanks and pipes to solve a problem related to defficient water distribution while the redesign cost is minimized. The obtained results showed that the GA found a feasible solution for the problem and authors concluded that the GA approach could be considered an useful tool to redesign already existing WDS which do not operate in the desired way.

Nicolini and Falcomer (2020) developed a methodology to calibrate the hydraulic model for a WDS by the combination of both evolutionary algorithms and numerical modelling of the system. It was shown in this study that the GA allowed for tuning of the roughness coefficient of the pipe. Recently, in Santos-Ruiz et al. (2020) a nonlinear-optimization based method is proposed to estimate not only the roughness coefficient but also the minor loss coefficients by minimizing the fitting error in the well-known Colebrook-White equation. It is pointed out that the main limitation of the approach was that the loss model considered only elbow fittings without modeling losses in valves and other accessories. Following this direction, more recently, in Santos-Ruiz et al. (2021) a nonlinear optimization method was proposed to estimate the same hydraulic parameters, namely: the roughness and minor loss coefficients, in this case by using a Lambert W-function and by considering the roughness coefficient as aditional friction assumed as an equivalent pipe length. Since the calibration process for a WDS represents an important challenge nowadays for an efficient water managment, in this work a GA-driven optimization approach is presented for the sake of the improvement of those approaches recently proposed in Santos-Ruiz et al. (2020) and Santos-Ruiz et al. (2021).

This work is organized as follows: Section 2 presents the theoretical background. In Section 3, the proposed GA-based methodology is presented. Experimental results are provided in Section 4. Finally some conclusions and future perspectives are discussed in Section 5.

2. THEORETICAL BACKGROUND

2.1 Hydraulics fundamentals

The relationship between pressures and flows in hydraulic models is determined by pressure losses, which are classified into major losses (associated with flow turbulence and pipe roughness) and minor losses (associated with energy loss through pipe fittings). The relative roughness coefficient (ε_r) and the minor loss coefficient (K) are important parameters for an accurate hydraulic modeling in pipelines. In particular, the coefficient K is associated to fittings installed along the pipeline as valves, elbows, and tees, among others.

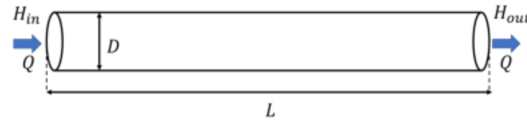

Fig. 1. Variables in the pipeline model

To analize the effect of those parameters, let us first consider a pipeline section of size L [m] and inner diameter D [m] transporting pressurized water, as shown in Fig. 1, where Q [m^3/s] is the flow rate, H_{in} [m] is the pressure head at upstream and H_{out} [m] at downstream, respectively. The major head-loss between upstream and downstream can be easily computed as: $h_f = H_{\text{in}} - H_{\text{out}}$, but also it can be modeled through the Darcy-Weisbach equation:

$$h_f = CQ^2, \tag{1}$$

where C [s^2/m^5] is a resistance coefficient computed as follows:

$$C = f(\varepsilon_r, \text{Re})\frac{8L}{g\pi^2 D^5}, \tag{2}$$

where g [m/s^2] is the acceleration due to gravity , $f(\varepsilon_r, \text{Re})$ [dimensionless] is the friction factor, and $\text{Re} = DQ/(A\nu)$ [dimensionless] is the Reynolds number, A [m^2] is the cross-section area of the pipe and ν [m^2/s] is the kinematic viscosity of the water at the operating temperature. The relation between f, Re, and ε_r in a turbulent regime (Reynolds number greater than 4000) is modeled by the well-known Colebrook-White equation:

$$\frac{1}{\sqrt{f}} = -2\,\log_{10}\left(\frac{\varepsilon_r}{3.7} + \frac{2.51}{\text{Re}\sqrt{f}}\right). \tag{3}$$

Since (3) can not be easily solved analytically, an iterative approach can be adopted instead. In addition, since the operational conditions in pipelines are continually changing, an update of the Reynolds number must be performed using measured data. The relative roughness coefficient ε_r remains approximately constant in the short term, but can vary over long periods, due to corrosion and the accumulation of solids on the pipe walls. This parameter is a normalized (dimensionless) version of the the absolute roughness coefficient ε [m], as defined by:

$$\varepsilon_r = \varepsilon/D. \tag{4}$$

The minor head losses due to fittings is modeled as:

$$h_L = K\frac{v^2}{2g}, \tag{5}$$

where K [dimensionless] is the minor loss coefficient computed as follows:

$$K = \sum_{i=1}^{n} k_i, \tag{6}$$

where k_i is the minor head-loss coefficient due to the i-th fitting, and n is the total number of fittings. It should be noted that the value of each k_i could be roughly approximated from datasheets but with the drawback that those values are not accurate because they are determined to satisfy only raw design requeriments. Nonetheless, from a practical point of view, an accurate hydraulic analysis and diagnosis (e.g., leak diagnosis) depends on the accuracy of the model parameters. For this reason, in this work a parameter estimation based on computational intelligence is proposed with the aim to provide values of the roughness coefficient and the minor loss coefficient as accurate as possible by using measurements of pressure head and flow rate. A genetic algorithm for estimating these parameters is presented below.

2.2 Genetic Algorithm

A genetic algorithm is a population-based stochastic optimization algorithm whose main operations are selection, crossover, and mutation (Mirjalili, 2019). In the initial step, a random population of feasible solutions is created, each comprises a set of parameter solution emulating the *genes* of a given individual. The whole set of genes is known as *chromosome*, similarly to the chromosomes of an alive organism. During the initialization stage of the GA-algorithm, the main objective is to spread the population across of a search space as uniformly as possible to increase the possibility of finding regions that provide good performance. Since natural selection is the main inspiration for GA, it uses a selection mechanism to assign selection probabilities to each individual proportionally to their fitness values. The creation of this new generation of individuals is simulated by combining two *parent* solutions to produce two new *children* solutions. The last evolutionary operator is known as a mutation, in which one or multiple genes are altered after creating children solutions. The operation of mutation preserves the diversity of the population by introducing another level of randomness. This operator prevents of similar solutions while reducing the probability of local solutions in the GA. The steps of the GA are summarized in Fig. 3.

3. GA-DRIVEN CALIBRATION METHODOLOGY

To perform the GA-driven calibration process, only measurable physical dimensions and topological parameters of the network are required. To do that, a .INP file containing the network layout is used in the EPANET-MATLAB Toolkit to recover these information, see (Eliades et al., 2016). Ideally, in addition to the pipe length, information about ε and K_i should also be available for each pipe segment; however, since these values are unknown, they are initialized to zero (by default, when there is no prior estimate). Then, the system under analysis (pipeline or network) is divided into several pipe segments, in which the value of ε and K_i are required to be estimated.

Subsequently, experiments are carried out to collect pressure measurements at the network nodes for different operating points (varying the working frequency of the pump). The flow measurement is also collected at each outlet point to be assumed as demand (consumption) in that node when solving the EPANET model in each iteration of the GA. Known physical parameters (e.g., lengths and diameters) and output flow rates are specified a priori in the .INP file, so that the GA only fits ε and K_i minimizing a cost function based on the calibration error.

The GA performs the following steps to calibrate the model (see Fig. 3):

(S1) *Generation of feasible solutions.* Each candidate solution is a vector containing estimated values of ε and K.

(S2) *Fitness test of each solution and elimination of the less fit.* This process is performed as follows: the values of ε and K are taken from each candidate solution and are updated in the .INP file and then a hydraulic simulation is performed. Values of the pressure head at predefined nodes coming from the simulation are then compared with the corresponding measurements allowing the calibration process to be validated. The cost function optimizes the root mean squared error (RMSE) between measured pressure head H_i and the estimated pressure head $\widehat{H}_i$ where suscript i stands for the i-th node as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(H_i - \widehat{H}_i\right)^2} \tag{7}$$

To ensure the reliability of the calibrated hydraulic model no matter the operational condition, this validation process is performed several times by updating the input value and by applying the RMSE criterion for all different data sets (with the pump at different working frequencies). The performance of every population member has been tested. However, since no cross and mutation has occurred, it is not expected that a next and single generation will provide a better solution to the problem, in this way, the fittest members of the population are selected to be the "parents" of the following generation.

(S3) *Generation of new solutions.* Cross and mutation processes implicitly take place during the creation of a new generation. The mutation process is necessary to ensure the diversity of the population. In this way, every possible combination in the search area is taken into account to locate the most feasible solutions and directly discard those that provide a poor performance.

This process is repeated until an established beforehand number of maximum generations is fulfilled, or until a satisfactory solution is found considering a predetermined error criteria. The content of this section has described how the calibration process was ejecuted for both pipeline configurations: single and branched configurations. This is interesting since the paramters to be calibrated are different in each case.

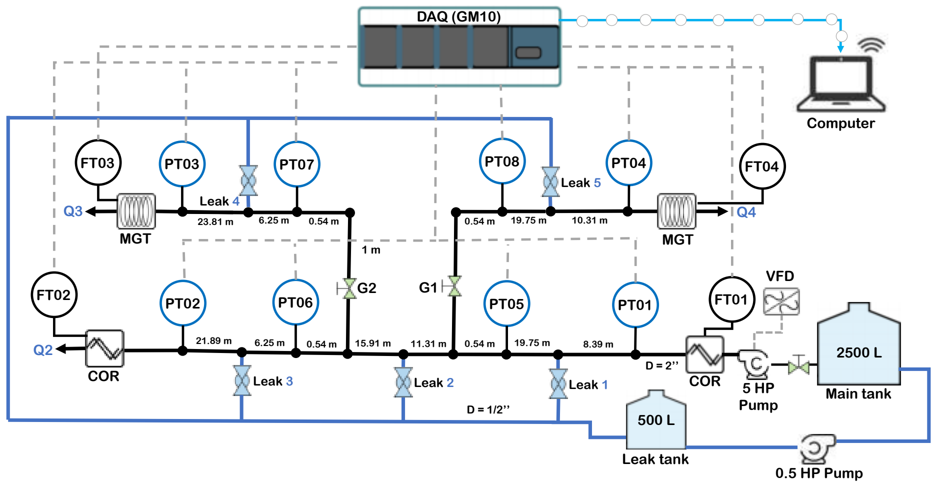

Fig. 2. P&ID of the pilot plant.

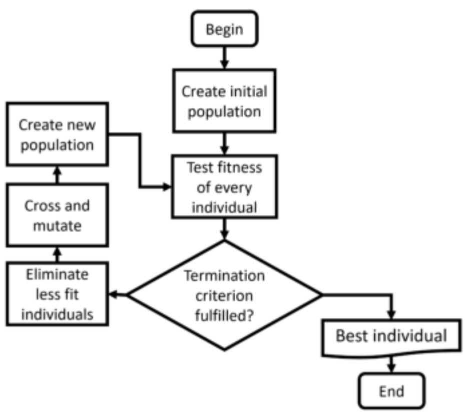

Fig. 3. Genetic algorithm

4. EXPERIMENTAL RESULTS

To test the proposed method, the experimental network (pilot plant) shown in Fig. 2 was used as a case study. This network is built with Schedule 80 PVC pipes with 48.6 mm inner diameter. The pilot plant is equipped with flow-rate sensors (FT01 at upstream and FT02, FT03 and FT04 at downstream) and pressure head sensors (from PT01 to PT07) whose measurements are collected by using a SCADA system connected via ethernet communication protocol. The collected measurements are stored as data sets in a personal computer and then are used as inputs for the GA to calibrate the model. This prototype is fed with drinking water from a 2500 L tank at upstream by using a 5 hp centrifugal pump. A variable frequency drive is used to vary the operating frequency of the pump in the range from 30 Hz to 60 Hz, this allows the input pressure head to be regulated during the experimentation. By manupulating the valves labeled with G1 and G2 the hydraulic system can operate as a branched network. Moreover, five valves labeled from "Leak 1" to "Leak 5", are installed at different positions to emulate leaks. The spreaded water caused by leaks is collected by using a second tank and the water is then pumped towards the main tank at upstream to be reused. On leak valves and T-type fittings there are no sensors, so the pressure head at those positions must be estimated somehow. In Fig. 4 a general framework for the calibration methodology is summarized.

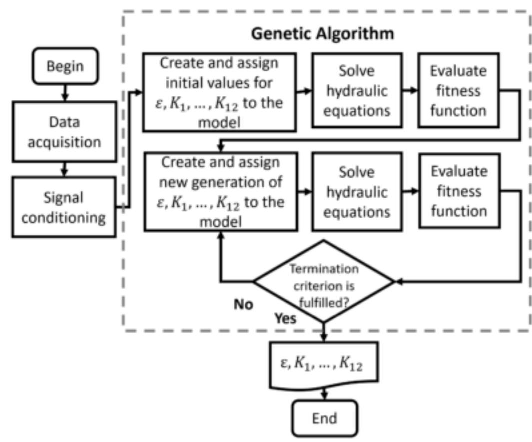

Fig. 4. Calibration workflow

By following the calibration process previously described, the pipeline system (see Fig. 2) is divided into several segments as shown in Fig. 5: where N_1 stands for the

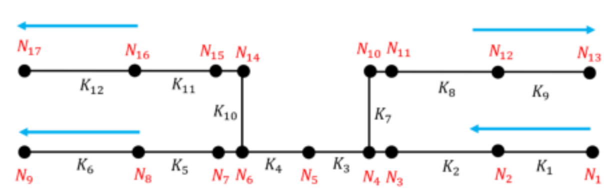

Fig. 5. Schematic diagram of the branched network

node at reservoir (tank), ε is considered as a constant value along the network whereas for each pipe segment a different value of K is considered while the calibration process takes place. Moreover, the pipes linking nodes

$N_3 - N_4$, $N_6 - N_7$, $N_{10} - N_{11}$ and $N_{14} - N_{15}$ are excluded from the analysis since K is assumed to be zero because they do not contain accessories. All nodes and their corresponding reference are summarized in Table 1:

Table 1. Physical reference of nodes

Node	Reference	Node	Reference
N_1	PT01, FT01	N_{10}	T-type fitting
N_2	Leak valve 1	N_{11}	PT08
N_3	PT05	N_{12}	Leak valve 5
N_4	T-type fitting	N_{13}	PT04, FT04
N_5	Leak valve 2	N_{14}	T-type fitting
N_6	T-type fitting	N_{15}	PT07
N_7	PT06	N_{16}	Leak valve 4
N_8	Leak valve 3	N_{17}	PT03, FT03
N_9	PT02, FT02		

After that, several experiments are performed for both, single and branched pipeline configurations, at different operating conditions. Every dataset is saved and processed in MATLAB environment using the .INP file and the GA-driven optimization algorithm. On the one hand, when the calibration process is run for the single pipeline configuration, the GA algorithm is set to optimize 7 variables: ε and $K_1, \ldots, K_6$. On the other hand, when the pipeline has the branched configuration the GA is set to optimize 13 variables: ε and $K_1, \ldots, K_{12}$.

For both configurations, single and branched, the GA performs the steps to calibrate the model presented in Section 3 (see Fig. 4): For the process described in step (S2), the values of ε and K are taken from each candidate solution and are updated in the .INP file and then a hydraulic simulation is performed. Values of the pressure head at nodes N_3, N_7, N_9, N_{11}, N_{13}, N_{15} and N_{17} coming from the simulation are then compared with the corresponding measurements allowing the calibration process to be validated. The cost function optimizes the root mean squared error (RMSE) as described by (7). To ensure the reliability of the calibrated hydraulic model no matter the operating condition, this validation process is performed several times by updating the input value and by applying the RMSE criterion for all data sets $(30\,\mathrm{Hz}, 35\,\mathrm{Hz}, \ldots, 60\,\mathrm{Hz})$ allowing the overall RMSE to be computed by adding $\mathrm{RMSE}_{30\,\mathrm{Hz}}, \mathrm{RMSE}_{35\,\mathrm{Hz}}, \ldots, \mathrm{RMSE}_{60\,\mathrm{Hz}}$, respectively.

4.1 Discussion of the results

For the single pipeline configuration the lowest RMSE is 9.3×10^{-3} m corresponding to a sampling rate of 50 Hz whereas for the two-branching configuration the corresponding lowest RMSE is 3.4×10^{-2} m at samplig rate of 60 Hz. Fig. 6 shows a comparison between measured and estimated values of pressure head for the single pipeline. Conversely Fig. 7 illustrates the comparison for the branched network. Table 2 shows the best-fitting RMSE for every tested operating point for both cases.

Table 2. Fitting RMSE.

(a) Single pipeline

AC Freq.	RMSE (m)
60 Hz	4.67×10^{-2}
55 Hz	2.17×10^{-2}
50 Hz	9.3×10^{-3}
45 Hz	3.15×10^{-2}
40 Hz	5.46×10^{-2}
35 Hz	12.95×10^{-2}
30 Hz	9.42×10^{-2}

(b) Branched network

AC Freq.	RMSE (m)
60 Hz	3.40×10^{-2}
55 Hz	5.04×10^{-2}
50 Hz	3.83×10^{-2}
45 Hz	3.87×10^{-2}
40 Hz	7.74×10^{-2}
35 Hz	9.11×10^{-2}
30 Hz	10.84×10^{-2}

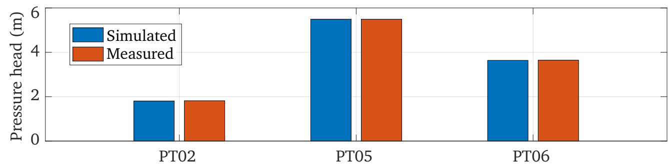

Fig. 6. Pressure heads in single pipeline configuration

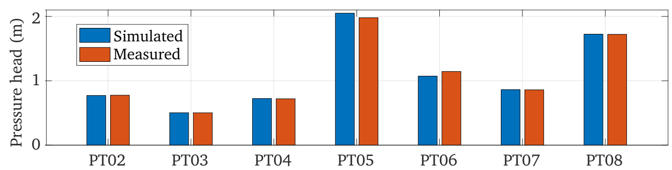

Fig. 7. Pressure heads in branched network configuration

Table 3. Calibrated model parameters

(a) Single pipeline

Parameter	Value
ε	2.99 µm
K_1	2.945
K_2	1.898
K_3	0.824
K_4	2.468
K_5	0.779
K_6	2.312

(b) Branched network

Parameter	Value
ε	2.93 µm
K_1	2.681
K_2	1.503
K_3	0.504
K_4	0.503
K_5	0.524
K_6	0.520
K_7	3.903
K_8	2.230
K_9	1.720
K_{10}	12.02
K_{11}	1.204
K_{12}	2.966

Calibrated values for each parameter and for both configurations are summarized in Table 3 where the reported ε value matches closely to empirical values for PVC pipe (Rossman et al. (2020)). The ε value was slightly different when changing the configuration from single to branched. However, these values should be the same, since the relative roughness is obtained from the absolute roughness and the internal diameter, which do not vary in the short term.

On the other hand, calibration processes for experiments without changes in the operating point were also executed, showing an almost perfect performance; however, since the interest from a practical point of view is focused on varying operating conditions, a trade-off between range of operation and accuracy was considered. It also was concluded that the best calibration process is achieved in the highest frequencies of the pump operation setup (from 45 Hz to 60 Hz). The parameters of the calibrated model shown in Table 3 correspond to the best-performing execution of the GA where even if the obtained ε and K provide a low-error match at different operation points, their physical feasibility, as well as parametrization of the K value of every accesory individually still needs to be addressed. After performing several calibration processes, a discrepancy was observed related to the fitting of the pressure head at positions of sensors: PT05, PT06, and PT02 (see Fig. 7). Such a discrepancy could be attributed to a poor calibration of the transducers in the form of a zero-point adjustment. Thus, it is proposed that the required zero-point recalibration value can be introduced into the optimization algorithm as a new variable that needs to be estimated for each poor-performance sensor. It should be noted that further research work must be performed following this direction. More specifically, the discrepancy showed in Table 3 regarding values of K among both configurations must be avoided. Such a discrepancy could be prevented by setting the GA with calibration values for the variable K for each accessory instead of a global K value for a pipeline segment. This strategy could be effective to avoid the estimation of non-realistic K values. Finally, a direct relation between the working frequency of the pump and the performance of the calibration process is observed. For high frequencies, the measurement noise increases possibly due to the high turbulence and makes it difficult to compare some measured pressures with the simulated values; however, it does not significantly affect the overall performance. Conversely, for a low frequency operation of the pump, the pipeline system does not operate to ensure the minimum-required pressure service.

It is highlighted that the noise level affecting the measurements clearly impacts the accuracy of the calibration. In addition it was also observed that the fitting error is attributed to the fluctuation of the flow regime in regions where the complety developed flow regime is not reached (unstable flow regime). In particular, it can be observed a less accurate estimation at the transducers PT05 and PT06 (see Fig. 7). This occurs because those sensors are installed near of a T-type fitting and a ball-type valve where a sudden change in the flow direction takes place. The high turbulence dynamics cause the pressure transducers to not perform reliably at these locations. Conversely, a better fit can be seen in the nodes corresponding to transducers where flow does not undergo fluctuant and sudden changes, in other words where the fully developed flow regime is reached.

5. CONCLUSION

This paper presented a GA-driven methodology to calibrate hydraulic parameters of a WDS. Experimental results evidenced a good performance by using a pipeline prototype which can adopt both: single and branched pipeline configurations and for which, the roughness and the minor loss coefficients (ε and K) have been estimated with accuracy no matter changes in the operation point, whereas previous investigations performed calibration for a single operating point. This work is relevant since future developments can extend this methodolgy to a large-scale problem of urban WDS always considering the computational effort and its inherent dependence on reliable measurements. The calibrated model will be a base tool for future investigations in leak diagnosis algorithms.

REFERENCES

Do, N.C., Simpson, A.R., Deuerlein, J.W., and Piller, O. (2016). Calibration of water demand multipliers in water distribution systems using genetic algorithms. *Journal of water resources planning and management*, 142(11), 04016044.

Drisya, J. and Sathish Kumar, D. (2018). Automated calibration of a two-dimensional overland flow model by estimating manning's roughness coefficient using genetic algorithm. *Journal of Hydroinformatics*, 20(2), 440–456.

Eliades, D.G., Kyriakou, M., Vrachimis, S., and Polycarpou, M.M. (2016). EPANET-MATLAB Toolkit: An Open-Source Software for Interfacing EPANET with MATLAB.

Heydari, S., Mamizadeh, J., Sarvarian, J., and Ahmadi, G. (2020). Optimization of water distribution networks using developed binary genetic algorithm and hydraulic model software. *Journal of Applied Research in Water and Wastewater*, 7(1), 30–35.

Martínez-Bahena, B., Cruz-Chávez, M.A., Ávila-Melgar, E.Y., Cruz-Rosales, M.H., and Rivera-Lopez, R. (2018). Using a genetic algorithm with a mathematical programming solver to optimize a real water distribution system. *Water*, 10(10), 1318.

Mirjalili, S. (2019). Genetic algorithm. In *Evolutionary algorithms and neural networks*, 43–55. Springer.

Nicolini, M. and Falcomer, L. (2020). Genetic algorithm for calibration and leakage identification in water distribution system. In *2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII)*, 273–276. IEEE.

Rossman, L.A., Woo, H., Tryby, M., Shang, F., Janke, R., and Haxton, T. (2020). EPANET 2.2 User Manual. Technical Report EPA/600/R-20/133, U.S. Environmental Protection Agency, Washington, DC.

Santos-Ruiz, I., López-Estrada, F.R., Puig, V., Torres, L., Valencia-Palomo, G., and Gómez-Peñate, S. (2021). Optimal estimation of the roughness coefficient and friction factor of a pipeline. *Journal of Fluids Engineering*, 143(5).

Santos-Ruiz, I., López-Estrada, F.R., Puig, V., and Valencia-Palomo, G. (2020). Simultaneous optimal estimation of roughness and minor loss coefficients in a pipeline. *Mathematical and Computational Applications*, 25(3), 56.