Leveraging Hydraulic Cyber-Monitoring Data to Support Primitive Condition Assessment of Water Mains

Ahmad Momeni, S.M.ASCE[1]; and Kalyan R. Piratla, A.M.ASCE[2]

Abstract: Buried water pipelines deteriorate in response to several variables such as pressure transients, corrosion, and pipeline material degradation, among others, that are dynamic processes, and it is therefore difficult to predict the pipeline condition without employing expensive sophisticated technologies. Such technologies are ad hoc in nature and may be worthwhile only for those pipelines that are known to be deteriorated and critical for the reliability of the water distribution network (WDN). Adopting cyber-monitoring methods for pipeline condition assessment, this paper presents and demonstrates a data-driven condition assessment platform that can serve as a primitive indicator of water pipeline conditions. Flow, pressure, and water consumption data collected in a synchronous manner are employed to predict pipeline roughness coefficients and effective internal diameters through a combination of hydraulic modeling, evolutionary algorithms, and neural networks utilizing two popular benchmark WDNs. The accuracy of pipeline condition prediction, measured using mean absolute percentage error (MAPE), ranged between 4.12% and 17.6% based on numerous scenarios in this study. Effective internal diameters were found to be more accurately predictable than pipeline roughness coefficients, and it was also found that pressure monitoring alone can suffice the requirements of the proposed framework in order to produce accurate pipeline condition prediction. It is recommended that future research be conducted over the robustness of this platform for other dynamic parameters such as leakages. **DOI: 10.1061/(ASCE)PS.1949-1204.0000596.**

Author keywords: Condition assessment; Hydraulic monitoring; Optimization.

Introduction and Background

Water distribution networks (WDNs) are crucial to the functionality of societies, yet are inevitably complex in behavior (Luciani et al. 2019; Momeni et al. 2018; Zhang et al. 2018). In order to ascertain a proper operational and financial maintenance of the WDNs, a paradigm shift is now required more than ever such that robust data-driven measures for condition assessment, rehabilitation, and replacement of deteriorating pipelines be taken into consideration in lieu of the conventional asset management. Conventional methods of water pipeline condition assessment primarily rely on onsite inspections that are not only ad hoc in nature but also prone to human errors (Moglia et al. 2006; Newton and Christian 2015; Wu et al. 2009). They cause geological and environmental disturbance to the water pipeline environment and are also expensive to carry out in large scale (Alegre et al. 2013; Bedjou et al. 2019; Grigg and Butler 2019). Condition assessment technologies including pit depth measurement, visual inspection, or ultrasonic testing are found to be either environmentally concerning, cost-ineffective, or inapplicable to all pipe types (Liu and Kleiner 2013). Although statistical inferences based on historical pipeline break records are often employed in preliminary condition assessment planning (Kamiński et al. 2017; Poulakis et al. 2003; Qi et al. 2018; Soldevila et al. 2019), their accuracy is not guaranteed, and availability of historical break data at sufficient granularity is also rare (Chen et al. 2017).

Also, recent sensor-based monitoring techniques include acoustic signal emissions as well as vibration analysis, electromagnetic inspection, transient signal analysis of pressure waves, infrared thermography, and radar frequency identification (Amoatey et al. 2018; Cody et al. 2020; El-Zahab and Zayed 2017; Ferrandez-Gamot et al. 2015; Gertler et al. 2010; Karney et al. 2009; Lin 2017; Perez et al. 2014; Soldevila et al. 2017b, 2018, 2019). Employing these approaches are also either confined to leakage detection in water pipelines (Fuentes and Pedrasa 2020; Kapelan et al. 2004; Poulakis et al. 2003; Shukla and Piratla 2020; Soldevila et al. 2017a; Yazdekhasti et al. 2020; Zhang et al. 2018) or laborious and often costly. Hybrid models where data-driven approaches that are dependent upon historical data are conjoined with expert recommendations are also recent, although somewhat unreliable (Kamiński et al. 2017; Park et al. 2016). These methods are often coupled with pattern recognizers or classification algorithms where machine learning methods are leveraged for the condition assessment of water pipelines (Abdulla et al. 2013; Ferrandez-Gamot et al. 2015; Kayaalp et al. 2017; Meseguer et al. 2015; Montáns et al. 2019). Other novel data-driven Internet-of-Things (IoT)–based methods entail merely monitoring purposes, albeit carried out remotely through water flow meters and wireless data collectors; nevertheless, some limitations include security issues, human interventions, or unaffordable cost and maintenance measures (Abdelhafidh et al. 2018).

With the increasing adoption of cyber monitoring tools for situational awareness of infrastructures (Giudicianni et al. 2020; Montáns et al. 2019; Soldevila et al. 2018; de Winter et al. 2019), it is possible to leverage hydraulic monitoring data from WDNs to infer the condition of the pipelines. For example, it is feasible to install pressure sensors at a few critical locations for near-real-time monitoring at a reasonable cost (Abdulshaheed et al. 2017; Raei et al. 2019; Soldevila et al. 2019). It may also be feasible to install

[1]Ph.D. Student and Graduate Research Assistant, Glenn Dept. of Civil Engineering, Clemson Univ., Clemson, SC 29634 (corresponding author). ORCID: https://orcid.org/0000-0002-4506-517X. Email: amomeni@clemson.edu

[2]Liles Associate Professor, Glenn Dept. of Civil Engineering, Clemson Univ., Clemson, SC 29634. Email: kpiratl@clemson.edu

Note. This manuscript was submitted on August 7, 2020; approved on June 10, 2021; published online on August 24, 2021. Discussion period open until January 24, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Pipeline Systems Engineering and Practice***, © ASCE, ISSN 1949-1190.

flowmeters, albeit at some inconvenience and greater cost (Pacheco et al. 2020). Furthermore, water consumption data can also be obtained using smart water meters on a near-real-time basis (Luciani et al. 2019; Zhang et al. 2021). It is hypothesized that water demand data in conjunction with the corresponding hydraulic data from the distribution pipeline network can be leveraged to predict the pipeline conditions (Shinozuka et al. 2005; Soldevila et al. 2016). In conventional computational models, the pipeline geometric and condition parameters are accurately defined and used to ensure appropriate pipe flows and network pressures for a given set of nodal demands because the pipelines are brand new. The research approach explored in this study is to reverse this conventional modeling approach such that the pipeline condition parameters can be predicted using measured pipe flow and network pressure data in conjunction with corresponding nodal demand data.

The underlying premise of this paper is to study and predict those dynamic parameters in WDNs whose nature is complicated due to their uncertain behavior and volatility over time including but not limited to pipeline roughness, effective hydraulic diameter, pipe wall thickness, and leakages (Braun et al. 2020; Jensen and Jerez 2018). Specific to this study, pipe roughness increases over time, causing hydraulic energy loss (Gao 2017) and internal pipe diameters are greatly reduced over time due to corrosion-related scaling (Mazumder et al. 2019). The prediction model in the presented study accounts particularly for these two dynamic parameters because the remaining useful service life of the pipeline is dependent on such dynamic parameters, and therefore it is imperative to be able to predict them with reasonable accuracy and least cost. Although availability of pipe flow, pressure, and nodal demand data in a synchronous manner through cyber monitoring is not currently ubiquitous, it is expected to be so in the coming years with increasing awareness of the benefits associated with cyber monitoring of critical infrastructures including time efficiency, considerable accuracy and cost effectiveness for the purpose of pipe rehabilitation, prioritization, or replacement, and it is in such context that this study is critical, novel, and highly relevant.

Objectives

Building upon the preliminary previous studies on this topic (Momeni et al. 2019; Momeni and Piratla 2020; Piratla and Momeni 2019), the objectives of this paper are to (1) present a computational model that takes water consumption data along with the corresponding pressure and pipe flow data from the distribution system as inputs to predict effective hydraulic diameters and pipe roughness coefficients as outputs; (2) incorporate genetic algorithms and cascade-forward neural networks in MATLAB 2019 into the model to both minimize the mean square error of the actual (harvested from smart meters) and predicted (generated from the prediction model) pressure and flow values and bypass hydraulic simulations respectively for higher temporal efficiency; and (3) conduct a comprehensive sensitivity analysis of placements and number of flow monitoring stations (FMS) and pressure monitoring stations (PMS) as well as the number of demand scenarios utilized in the optimization platform.

Methodology

Mathematical Formulation

The formulation of the computational model in this study consists of (1) equations for the functions of the optimization framework (Hamdia et al. 2021; Momeni et al. 2019; Momeni and Piratla 2020), and (2) the equations used to measure the accuracy of the results (Prayudani et al. 2019).

Optimization Framework

This section discusses the fundamental equations in the objective and constraint functions for the proposed genetic algorithm optimization platform where roughness coefficients and effective hydraulic diameters are predicted by minimizing the mean square error (Mirzal et al. 2012) of the predicted and actual pressure/flow values at the smart monitoring stations.

Decision Variables

Pipeline roughness (C) and effective internal diameters (D) are focused upon in the proposed prediction model. These sets C and D, which are characterized as follows, constitute the decision variables of the optimization framework:

$$C = \{c_1, c_2, \ldots, c_x\} \tag{1}$$

$$D = \{d_1, d_2, \ldots, d_x\} \tag{2}$$

where c = roughness coefficient for each pipe; d = effective internal diameter of each pipe; and x = number of pipelines in the WDN.

Objective Function

Pressure (P_k) and flow (F_k) measured at various monitoring stations in the WDN for a given set (k) of nodal demands (Q_k) are characterized as follows:

$$Q_k = \{q_{1k}, q_{2k}, \ldots, q_{yk}\} \tag{3}$$

$$P_k = \{p_{1k}, p_{2k}, \ldots, p_{mk}\} \tag{4}$$

$$F_k = \{f_{1k}, f_{2k}, \ldots, f_{nk}\} \tag{5}$$

where q = nodal demand; y = number of nodes in the WDN; p = pressure measured at monitoring stations located in the WDN; m = number of PMS placed in the WDN; f = pipeline flow measure at monitoring stations located in the WDN; and n = number of FMS placed in the WDN.

A genetic algorithm optimization framework is employed to determine the sets C and D using j sets of Q_k, P_k, and F_k. In other words, a number (j) of sets of demands (Q_k) at all the nodes and the corresponding pressure (P_k) and flow (F_k) values at the monitoring stations are required. For candidate (i) solution sets of C and D in the optimization process, pressures and pipe flows at the monitoring stations can be estimated as follows assuming all the dynamic condition parameters are known except for roughness and effective internal diameters:

$$g(Q_k, C_i, D_i) = P_{k,i} \tag{6}$$

$$h(Q_k, C_i, D_i) = F_{k,i} \tag{7}$$

where $g()$ and $h()$ are representative of the hydraulic simulations that could be carried out using software such as EPANET 2.0; C_i = candidate solution set of pipeline roughness values; D_i = candidate solution set of effective internal diameter values; i = candidate solution reference in the optimization algorithm; $P_{k,i}$ = estimated set of pressure values at all the monitoring stations for corresponding candidate solution sets C_i and D_i; and $F_{k,i}$ = estimated set of flow values at all the monitoring stations for corresponding candidate solution sets C_i and D_i.

The objective function in the optimization algorithm is to minimize Z in Eq. (8) where

$$Z = \text{Minimum}_{k=1\text{ to }j}\left\{\sum_{a=1}^{m}(p_{a,k,i} - p_{a,k})^2\right\} + \text{Minimum}_{k=1\text{ to }j}\left\{\sum_{b=1}^{n}(f_{b,k,i} - f_{b,})^2\right\} \quad (8)$$

where a = index for the pressure monitoring station; b = index for the flow monitoring station; $p_{a,k,i}$ = estimated pressure at PMS a for set of nodal demands k for candidate solution i; $p_{a,k}$ = actual measured pressure at PMS a for set of nodal demands k (obtained from set P_k); $f_{b,k,i}$ = estimated pipeline flow at FMS b for set of nodal demands k for candidate solution i; and $f_{b,k}$ = actual measured pipeline flow at FMS b for set of nodal demands k (obtained from set F_k).

Constraints

Three different constraints are considered in the optimization framework:

1. The minimum pressure head at all the nodes has to be greater than 10 m for any candidate solution to be considered viable.
2. A reasonable set of upper and lower boundaries are set for effective hydraulic diameter decision variables (D) as a percentage of the original values (i.e., when the pipelines were newly installed).
3. A reasonable set of upper and lower boundaries are set for pipeline roughness decision variables (C) as a percentage of the original values.

Specific details on the last two constraints are discussed in a subsequent section.

Accuracy Measurement

For the purpose of measuring the prediction accuracy of the optimization algorithm, a performance metric namely mean absolute percentage error (MAPE) is employed as shown in Eq. (9) to estimate the error between the predicted and actual values of the condition parameters (i.e., roughness coefficients or effective hydraulic diameters). MAPE is calculated as follows (de Myttenaere et al. 2016):

$$\text{MAPE} = \left(\frac{\sum_{i=1}^{x}\frac{\text{abs}(pr_i - \text{act}_i)}{\text{act}_i}}{x}\right) \times 100 \quad (9)$$

where pr_i = predicted value of either roughness coefficient or effective hydraulic diameter for pipeline i; act_i = actual value of either roughness coefficient or effective hydraulic diameter for pipeline i; and x = number of pipelines in the WDN.

Demonstration Scheme

The modified benchmark WDNs of Hanoi (Fujiwara and Khang 1990; Monsef et al. 2019) and GoYang (Kim et al. 1994; Poojitha et al. 2020) (Fig. 1) are used for demonstration purposes. The original hydraulic and geometric specifications of Hanoi and GoYang networks are presented in Tables 1 and 2, respectively. These networks were assumed to be made of metallic pipelines like many of our current WDNs. Expectedly, metallic pipelines would get rougher with age and could have corrosion-related scaling that would diminish their internal effective diameters. The original Hanoi and GoYang networks are modified in this study by randomly reducing pipe roughness (C) coefficients as well as their internal pipe diameters within certain ranges to make them representatives of real-world deteriorated WDNs.

The PMS and FMS are assumed to be placed at certain locations in the modified benchmark networks, and the corresponding demands are monitored at all nodes. The initial placement of PMS and FMS is carried out randomly on both the networks, and this placement scenario is referred to as the baseline scenario in this study. Eight PMS and nine FMS are placed at random locations in Hanoi WDN, whereas five PMS and seven FMS are randomly placed in GoYang WDN in the baseline scenario. Several other scenarios are subsequently explored to evaluate the sensitivity of the model results to the number and placement of PMS and FMS.

Model Calibration: Conventional versus Neural Network–Based Hydraulic Simulations

Initially, the popular hydraulic solver EPANET 2.0 is coupled with a genetic algorithm in the MATLAB programming interface for making the proposed prediction of the pipeline condition parameters. EPANET 2.0 is used in the place of $g()$ and $h()$ functions in the mathematical formulation presented in a previous section. Subsequently, artificial neural networks (ANNs) were trained and used to mimic EPANET 2.0 solver for rapid computation in the optimization algorithm given the high number of hydraulic computations required in the proposed prediction model. Nodal demands, pipeline roughness coefficients and pipeline diameters are inputs, and flows and pressures are outputs (targets) in the ANNs. In total, 250,000 sets of input and target data are used to train the ANN model in this study.

There are two essential hyperparameters that account for and tune the topological scheme of the neural network: (1) the number

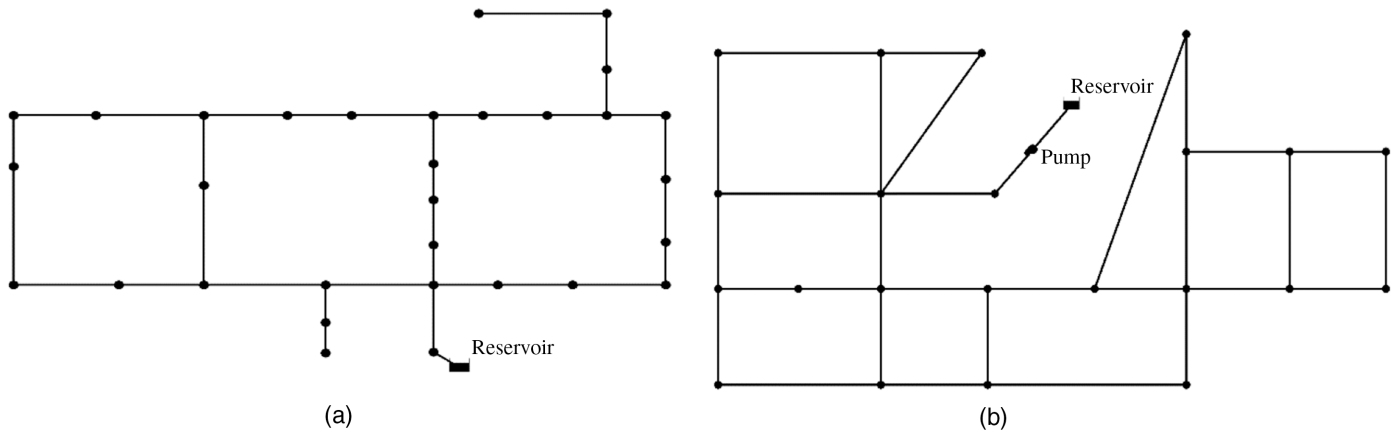

Fig. 1. Schematic layout of: (a) Hanoi WDN; and (b) GoYang WDN.

Table 1. Hanoi hydraulic and geometric specifications

Pipe/ node ID	Pipe diameter (mm)	Pipe roughness coefficient	Pipe length (m)	Nodal demands (cubic meters per hour)	Elevation (m)
1	1,371.6	130	100	890	0
2	1,524	130	1,350	850	0
3	1,219.2	130	900	130	0
4	1,371.6	130	1,150	725	0
5	1,219.2	130	1,450	1,005	0
6	914.4	130	450	1,350	0
7	762	130	850	550	0
8	914.4	130	850	525	0
9	762	130	800	525	0
10	1,371.6	130	950	500	0
11	914.4	130	1,200	560	0
12	762	130	3,500	940	0
13	609.6	130	800	615	0
14	1,066.8	130	500	280	0
15	914.4	130	550	310	0
16	1,066.8	130	2,730	865	0
17	1,066.8	130	1,750	1,345	0
18	1,371.6	130	800	60	0
19	1,066.8	130	400	1,275	0
20	1,371.6	130	2,200	930	0
21	914.4	130	1,500	485	0
22	762	130	500	1,045	0
23	1,219.2	130	2,650	820	0
24	914.4	130	1,230	170	0
25	762	130	1,300	900	0
26	609.6	130	850	370	0
27	1,066.8	130	300	290	0
28	762	130	750	360	0
29	762	130	1,500	360	0
30	609.6	130	2,000	105	0
31	914.4	130	1,600	805	0
32[a]	508	130	150	N/A	100
33	914.4	130	860	—	—
34	609.6	130	950	—	—

[a]Nodal characteristics for this ID are associated with the reservoir.

Table 2. GoYang hydraulic and geometric specifications

Pipe/node/ pump ID	Pipe diameter (mm)	Pipe roughness coefficient	Pipe length (m)	Nodal demands (liters per second)	Elevation (m)	Pump power (kW)
1	200	100	165	0	71.00	4.52
2	200	100	124	1.77	56.40	—
3	150	100	118	0.82	53.80	—
4	150	100	81	0.68	54.90	—
5	150	100	134	0.87	56.00	—
6	100	100	135	0.78	57.00	—
7	80	100	202	0.73	53.90	—
8	100	100	135	0.56	54.50	—
9	80	100	170	0.49	57.90	—
10	80	100	113	0.35	62.10	—
11	80	100	335	0.49	62.80	—
12	80	100	115	0.43	58.60	—
13	80	100	345	0.43	59.30	—
14	80	100	114	0.73	59.80	—
15	100	100	103	5.16	59.20	—
16	80	100	261	1.25	53.60	—
17	80	100	72	0.92	54.80	—
18	80	100	373	0.64	55.10	—
19	80	100	98	1.37	54.20	—
20	80	100	110	1.44	54.50	—
21	80	100	98	0.37	62.90	—
22	80	100	246	9.25	61.80	—
23[a]	80	100	174	N/A	71.00	—
24	80	100	102	—	—	—
25	80	100	92	—	—	—
26	80	100	100	—	—	—
27	80	100	130	—	—	—
28	80	100	90	—	—	—
29	80	100	185	—	—	—
30	80	100	90	—	—	—

[a]Nodal characteristics for this ID are associated with the reservoir.

of hidden layers, and (2) the number of nodes in each hidden layer (Wu et al. 2019). Specifying these two values along with functions including transfer and training functions plays a pivotal role in how the accuracy of training the network will turn out to be (Zhang and Duh 2020). Because the complexity of the model in this optimization procedure is substantially high, in order to capture the highest possible accuracy of the trained ANN model, a reliable method known as systematic experimentation has been used to adjust ANN parameters (Packianather et al. 2000). Specifically, parameters like transfer function, data division, and training function have been adjusted initially in a small-scale portion of the original 250,000 input data sets (up to only 200 data sets) based on the built-in outcome of the mean square error (MSE) in the MATLAB interface. Then, after reaching a minimum of MSE when fine-tuning the parameters, the number of data sets is increased up to an optimized point where there is no overfitting of data. A model that is not overfitted but is taking as much data as possible will yield an acceptable error range (Szegedy et al. 2014). Table 3 presents the detailed parametric values used for training the ANN model.

Similar to Eq. (9), to measure the prediction accuracy of the trained model, a prediction performance metric, i.e., MAPE, is utilized as shown in Eq. (10) to measure the accuracy of the trained ANN in percentage and is characterized as follows (de Myttenaere et al. 2016):

Table 3. Parametric identification for training ANN model

Parametric item	Parametric value
Training data sets	30,000 counts; one hidden layer
Validation data sets	10,000
Training function	"Trainrp," cascade forward net
Transfer function	"Purelin"
Performance	Mean squared error
Data division	Interleaved
Maximum fail parameter	5
Learning rate	0.01
Initial weight change	0.1
Increment to weight change	1.2
Decrement to weight change	0.5
Maximum weight change	50.0

$$\text{MAPE} = \left(\frac{\sum_{i=1}^{l} \left(\frac{\sum_{j=1}^{y} \frac{\text{abs}(pr_{ij} - \text{sim}_{ij})}{\text{sim}_{ij}}}{y} \right) \times 100}{l} \right) \tag{10}$$

where pr_{ij} = predicted value of either pressure (or flow rate) using the trained ANN model for node (or pipe) j in validation scenario i; sim_{ij} = simulated value of either pressure (or flow rate) calculated using EPANET 2.0 for node (or pipe) j in validation scenario i; y = number of nodes (or links) in the WDN; and l = number of validation scenarios.

Results and Discussion

This section discusses (1) the accuracy of the trained ANN model to be used in the prediction model is measured using MAPE metric for a set of validation scenarios; (2) sensitivity of the model as measured to the number of demand sets during optimization; (3) a baseline scenario established for each of the presented WDN benchmark networks to analyze the prediction model performance; and (4) a series of sensitivity analyses conducted to measure the model sensitivity to the number and locations of pressure and flow monitoring stations.

Neural Network Model Accuracy

According to Eq. (9), MAPE based on 10,000 validation scenarios for pressure at nodes and flow rates in links for Hanoi and GoYang networks are listed in Table 4.

Furthermore, conventional hydraulic simulations using EPANET 2.0 as well as ANN-based mimicking of hydraulic simulations are compared while separately predicting roughness and effective diameters for both Hanoi and GoYang networks. Tables 5 and 6 summarize this comparison in terms of the computation time as well as the prediction accuracy based on optimization runs completed on Clemson University's Palmetto Cluster (a supercomputing machine). It can be observed from Tables 5 and 6 that the ANN-based method either performed comparably to the conventional EPANET-based method or better in terms of MAPE while offering a significant advantage in terms of shorter computational time. Also, the training procedure and performance analysis of the cascade-forward neural networks for both Hanoi and GoYang WDNs are separately shown in Figs. 2 and 3. Consequently, the ANN-based method has been used for hydraulic simulations for rest of the analyses presented in this paper.

Choosing Optimal Number of Demand Sets

A certain number (j) of demand sets are used in the optimization algorithm for evaluating the prediction errors. A demand set is one set of nodal demands along with pipe flows and nodal pressures measured through smart monitoring devices placed throughout the distribution system. In order to choose the optimal number of demand sets to use, various sizes of demand sets are investigated for the resulting prediction accuracy as well as the computational time. The nodal demands are randomized within $\pm 20\%$ of the base nodal demands of the original benchmark networks to generate various demand sets of monitoring data from the deteriorated networks. As can be seen from the results presented in Table 7, the number of demand sets were varied from 5 through 5,000 in reasonably scaled increments. The monitoring locations associated with all the scenarios generated hereby are same as those in the baseline scenario of each of the networks.

The accuracy of prediction, measured using MAPE when roughness (R_{Comb}) and effective diameters (D_{Comb}) are predicted together, is used to compare different sizes of demand sets. ANN-based hydraulic simulations are used. Table 7 demonstrates that MAPE values for R_{Comb} or D_{Comb} do not necessarily increase with the size of the demand sets, whereas the computational time expectedly increases. Although it seems tempting to go with the smallest size of five, we have chosen to work with size of 200 (j) for rest of the analyses presented in this paper just to include more demand sets for better convergence at a reasonable computational time.

Table 4. Validation results of trained ANN model

Network	Node pressure MAPE (%)	Link flow rate MAPE (%)
Hanoi	3.36	4.70
GoYang	3.2	0.24

Prediction Model: Baseline Scenarios

Two benchmark WDNs are used to demonstrate and evaluate the primitive water pipeline condition prediction procedure proposed in this study.

Hanoi Network

According to the layout of the Hanoi network, there exist 34 links, in which case 34 decision variables are associated with roughness coefficient and 34 decision variables account for effective hydraulic diameters, totaling 68 decision variables in this section. In the baseline scenario, eight PMS and nine FMS are placed at random locations in Hanoi WDN. The Hanoi WDN is first modified by randomly reducing both pipeline roughness and effective internal diameter values to make it representative of a deteriorated WDN. Table 8 presents the range of reduction for both roughness and effective internal diameters for Hanoi WDN along with the search span across which the algorithm looks for the optimal solutions.

Fig. 4 represents the baseline scenario prediction results for Hanoi network. As can be seen in Fig. 4, condition parameters

Table 5. Computation time versus prediction MAPE for Hanoi using ANN and non-ANN methods

Parameter	Method	Computational time/run[a] (min)	MAPE (%)
Roughness	Conventional (EPANET 2.0)	~800	10.09 (Piratla and Momeni 2019)
	ANN-based	~20	10.84
Effective diameter	Conventional (EPANET 2.0)	~800	6.47 (Piratla and Momeni 2019)
	ANN-based	~20	4.56

[a]One run entails optimization for a population size of 500 and 1,000 generations.

Table 6. Computation time versus prediction MAPE for GoYang using ANN and non-ANN methods

Parameter	Method	Computational time/run[a] (min)	MAPE (%)
Roughness	Conventional (EPANET 2.0)	~700	11.01 (Momeni et al. 2019)
	ANN-based	~18	11.83
Effective diameter	Conventional (EPANET 2.0)	~700	9.95 (Momeni et al. 2019)
	ANN-based	~18	6.19

[a]One run entails optimization for a population size of 500 and 1,000 generations.

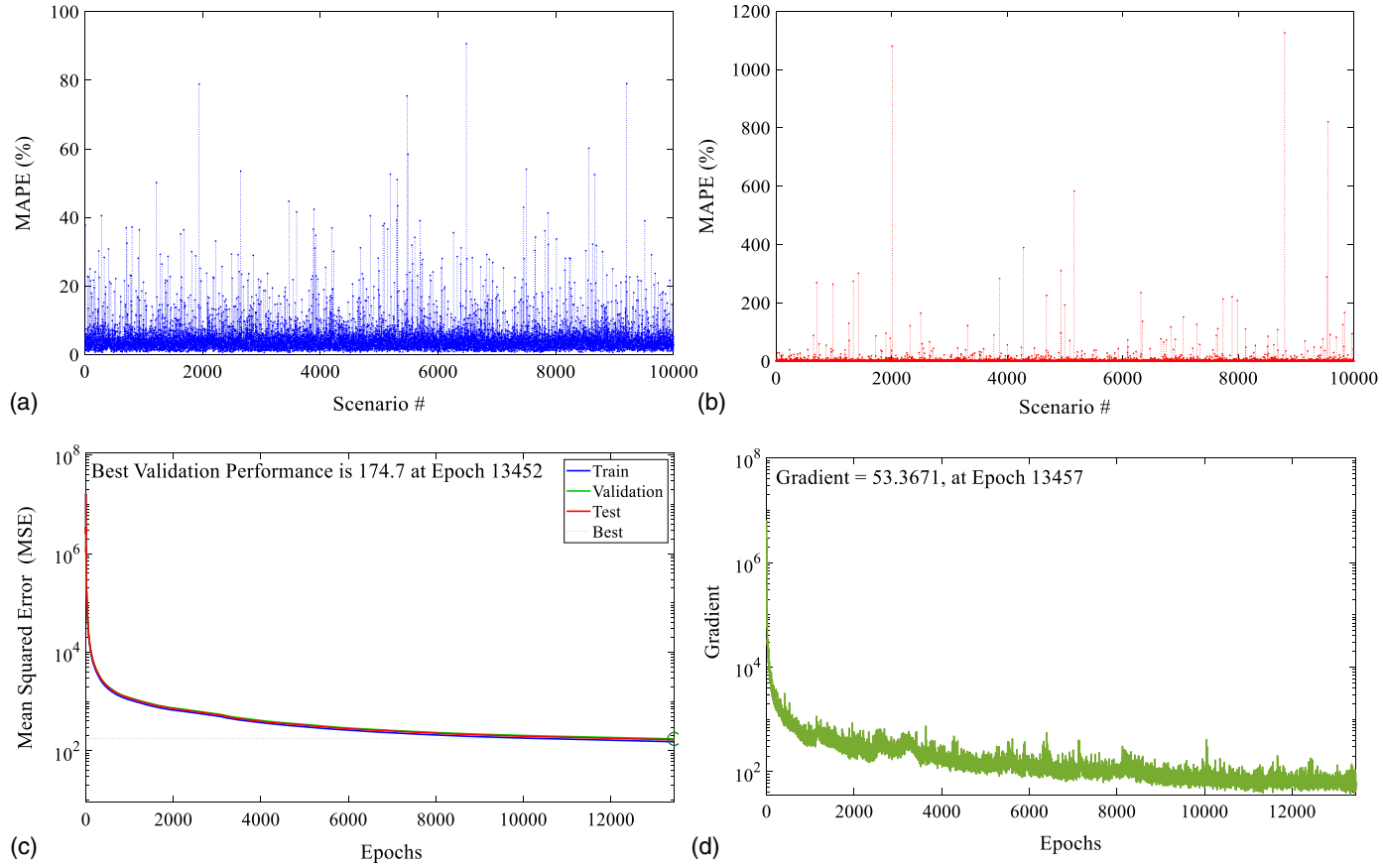

Fig. 2. Neural networks training procedure for Hanoi: (a) MAPE for 10,000 pressure validation scenarios; (b) MAPE for 10,000 flow validation scenarios; (c) mean squared error for ANN training performance; and (d) training gradient.

are predicted in the baseline scenario for the combination of both effective hydraulic diameters and roughness coefficients. Particularly, the MAPE values for diameter prediction equals 6.12%, and for roughness prediction, it equals 12.98% when both these parameters are predicted together using 68 decision variables in the optimization algorithm. It means that the mean error in prediction for effective diameter values is 6.12%, which is much lower than the 18.1% average reduction rate for effective diameters as per Table 8. Similarly, the mean error in prediction for roughness coefficients is 12.98%, which is much lower than the 41.7% average reduction rate of roughness values as per Table 8. Furthermore, it can be inferred from the MAPE values and Fig. 4 that the effective diameter values can be predicted more accurately than roughness values, which is likely due to the better influence of pipeline diameters than roughness values on the hydraulic features (i.e., nodal pressures and pipe flows) of WDNs.

Fig. 5 indicates the comparison between actual and predicted roughness coefficients when roughness alone is predicted assuming effective pipeline diameters are known. The corresponding MAPE value for roughness-only prediction of Hanoi network equals 10.84%. Similarly, Fig. 5 also indicates the comparison between actual and predicted effective diameter values when they are predicted individually, assuming roughness coefficients are known. The corresponding MAPE value for effective diameter-only prediction of Hanoi network equals 4.56%. The MAPE values for individual predictions are less than those in the combined prediction discussed previously, and this is likely due to the fewer unknowns in the optimization algorithm combined with more accurate characterization of the WDN condition.

GoYang Network

One of the main differences between GoYang and Hanoi networks is that GoYang has a pump linked up to the reservoir. GoYang consists of 30 pipes that make up the variables required for the prediction model in this scheme. Accordingly, this section embraces 60 decision variables, 30 of which establish the set of decision variables for roughness coefficients, and the other 30 are associated with effective hydraulic diameters. In the baseline scenario, five PMS and seven FMS are randomly placed in GoYang WDN. Like Hanoi, GoYang WDN was also first modified by randomly reducing roughness coefficients and internal diameters to make it representative of a deteriorated WDN. Table 9 presents the range of reduction for both roughness and effective internal diameters for GoYang WDN along with the search span across which the algorithm looks for the optimal solutions.

Fig. 6 illustrates the baseline scenario results through comparison of effective diameter and roughness predictions for GoYang in the combined prediction. Similar to the observation in Hanoi, diameter prediction is more accurate, with a MAPE value of 5.46%, than the roughness MAPE value of 14.66% when both these are predicted together in the optimization model using 60 decision variables. The diameter prediction MAPE of 5.46% is much lower than the 16.4% average reduction rate of effective diameters as per Table 9. Similarly, the roughness prediction MAPE of 14.66% is much lower than the 33.5% average reduction rate of roughness values as per Table 9.

Furthermore, Fig. 7 illustrates the comparison of actual and predicted values of roughness coefficients when they are separately predicted assuming the effective diameter values are known.

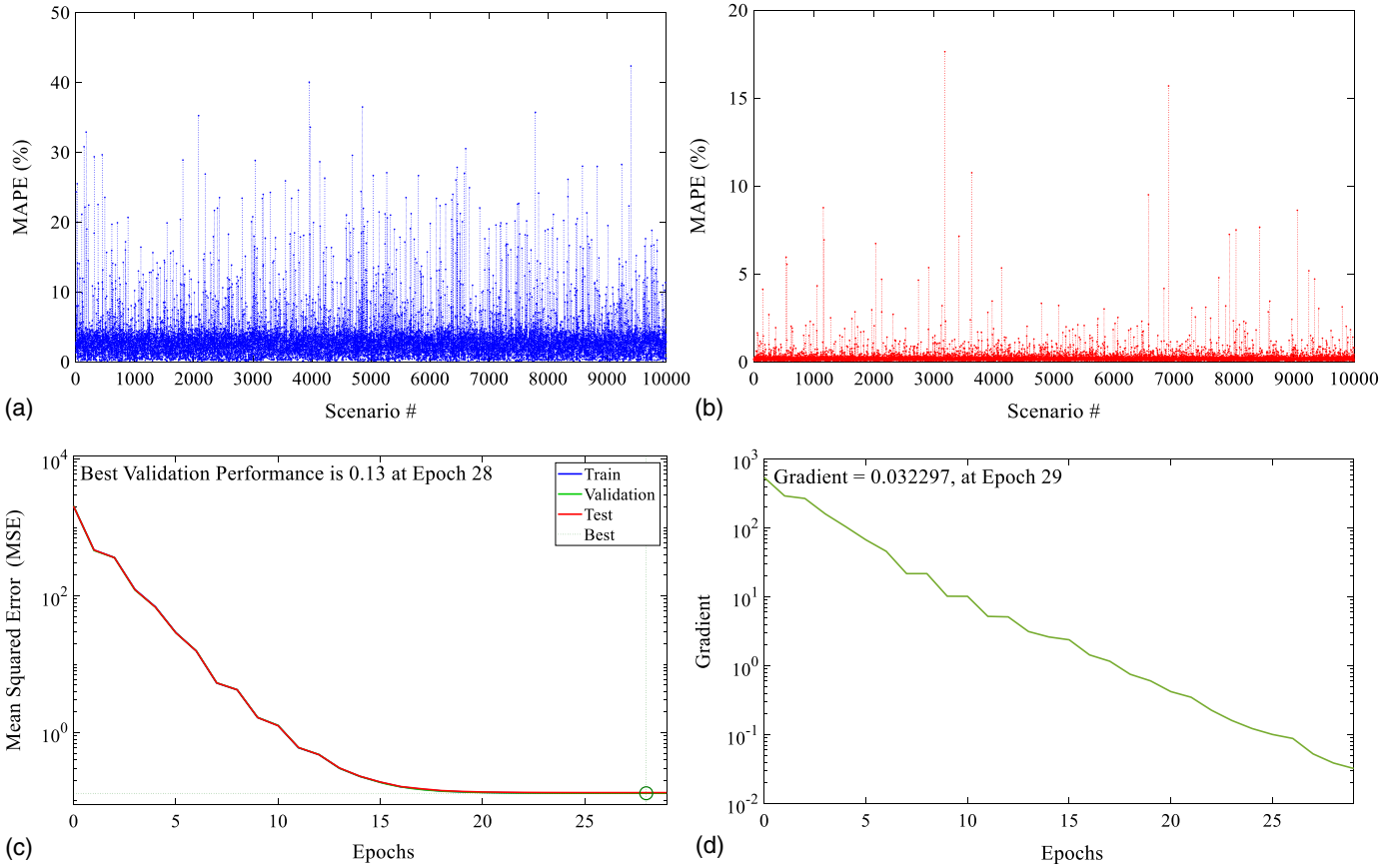

Fig. 3. Neural networks training procedure for GoYang: (a) MAPE for 10,000 pressure validation scenarios; (b) MAPE for 10,000 flow validation scenarios; (c) mean squared error for ANN training performance; and (d) training gradient.

Table 7. MAPE values for different demand scenarios

Network	Number of demand data sets	Computational time (min)	R_{Comb} (%)	D_{Comb} (%)
Hanoi WDN	5	23.33	9.34	6.78
	50	25.63	11.56	6.97
	100	28.72	10.21	6.22
	200	32.42	12.98	6.12
	400	39.60	12.09	6.75
	800	53.87	11.81	6.08
	1,000	60.93	10.23	7.53
	3,000	132.73	11.61	6.78
	5,000	206.6	11.08	6.82
GoYang WDN	5	18.16	17.74	4.95
	50	19.01	12.51	5.76
	100	20.16	14.76	5.25
	200	20.54	14.66	5.46
	400	24.04	14.76	5.68
	800	31.13	14.27	6.27
	1,000	33.34	12.69	4.96
	3,000	41.68	13.35	6.16
	5,000	150.1	16.27	5.77

Note: This applies to one run of optimization with a population size of 500 and 1,000 generations.

The corresponding MAPE value for this roughness-only prediction is 11.83%. Similarly, Fig. 7 also depicts the comparison of actual and predicted effective diameter values when they are separately predicted. The corresponding MAPE value for this effective diameter-only prediction is 5.19%. Similar to the case in Hanoi, the individual predictions resulted in better accuracies than in the combined predictions, which is likely due to lesser number of unknowns in the optimization model combined with more accurate characterization of the WDN condition.

Sensitivity Analyses

In this section, the sensitivity of the results to variations in the number and locations of pressure and flow monitoring stations across the two benchmark WDNs is presented. For this purpose, 26 scenarios have been created in addition to the baseline scenario that are categorized into three tiers (Tiers A through C). Table 10 presents the three tiers and the scenarios they comprise. Three specific research questions are answered through the sensitivity analyses:

- What is the sensitivity of the prediction accuracy to variation in the locations of FMS and PMS?
- What is the sensitivity of the prediction accuracy to the number of FMS?
- What is the sensitivity of the prediction accuracy to number of PMS?

These three questions are addressed in the following subsections.

Sensitivity to Variation in Locations of Flow and Pressure Monitoring Stations

Tier A, consisting of eight scenarios, is meant to analyze the sensitivity of MAPE values to variation in locations of PMS and FMS in both networks. As can be seen from Table 10, the locations of

Table 8. Hanoi original versus reduced values of condition parameters

Pipe ID	Diameter original values (mm)	Diameter reduced values (mm)	Diameter reduction rate (%)	Roughness original values	Roughness reduced values	Roughness reduction rate (%)
1	1,371.6	1,166.6	14.9	130	82	36.92
2	1,524	1,250.0	18.0	130	90	30.76
3	1,219.2	1,064.2	12.7	130	89	31.53
4	1,371.6	1,183.6	13.7	130	85	34.61
5	1,219.2	1,018.2	16.5	130	69	46.92
6	914.4	780.4	14.7	130	66	49.23
7	762	591.0	22.4	130	88	32.30
8	914.4	660.4	27.8	130	70	46.15
9	762	604.0	20.7	130	63	51.53
10	1,371.6	1,152.6	16.0	130	68	47.69
11	914.4	755.4	17.4	130	81	37.69
12	762	662.0	13.1	130	80	38.46
13	609.6	504.6	17.2	130	72	44.61
14	1,066.8	821.8	23.0	130	87	33.07
15	914.4	708.4	22.5	130	70	46.15
16	1,066.8	915.8	14.2	130	69	46.92
17	1,066.8	840.8	21.2	130	65	50
18	1,371.6	1,212.6	11.6	130	75	42.30
19	1,066.8	954.8	10.5	130	80	38.46
20	1,371.6	1,179.6	14.0	130	65	50
21	914.4	716.4	21.7	130	76	41.53
22	762	616.0	19.2	130	88	32.30
23	1,219.2	1,027.2	15.7	130	80	38.46
24	914.4	723.4	20.9	130	73	43.84
25	762	598.0	21.5	130	64	50.76
26	609.6	446.6	26.7	130	77	40.76
27	1,066.8	841.8	21.1	130	64	50.57
28	762	603.0	20.9	130	88	32.30
29	762	644.0	15.5	130	85	34.61
30	609.6	505.6	17.1	130	82	36.92
31	914.4	528.6	13.3	130	74	43.07
32	508	876.8	17.8	130	67	48.46
33	914.4	600.0	21.3	130	68	47.69
34	609.6	736.4	19.5	130	75	42.30

Note: Regarding search span, values range within 50% below original and 50% above original values.

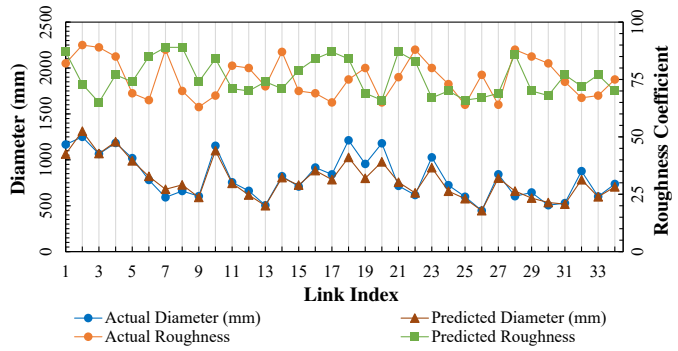

Fig. 4. Comparison of actual versus predicted pipeline conditions in the combined scenario for Hanoi WDN.

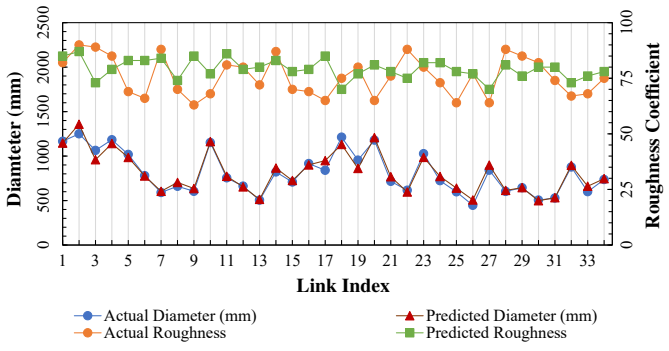

Fig. 5. Comparison of actual versus predicted pipeline conditions in the individual prediction scenario for Hanoi WDN.

PMS and FMS are randomly varied in each of the eight scenarios in Tier A keeping the numbers of PMS and FMS same as in the baseline scenario. The resulting MAPE values in each scenario for the combined prediction of effective diameter (D_{Comb}) and roughness (R_{Comb}) are also presented in Table 10.

As can be seen from Table 10, diameter prediction was found to be more accurate than roughness for all the Tier-A scenarios, which is consistent with the observations made for the baseline scenario. Furthermore, MAPE values did change when monitoring locations were varied in different Tier-A scenarios. It can be observed that the maximum variation in MAPE values among the nine scenarios (including baseline) for GoYang network are 2.61% and 4.83% for D_{Comb} and R_{Comb}, respectively. Similarly, the maximum variation in MAPE values among the nine scenarios for Hanoi network are 3.44% and 2.63% for D_{Comb} and R_{Comb}, respectively. From these results, it can be inferred that the variation in MAPE values in

Table 9. GoYang original versus reduced values of condition parameters

Pipe ID	Diameter original values (mm)	Diameter reduced values (mm)	Diameter reduction rate (%)	Roughness original values	Roughness reduced values	Roughness reduction rate (%)
1	200	182	9	100	80	20
2	200	187	6.5	100	56	44
3	150	133	11.34	100	62	38
4	150	136	9.34	100	56	44
5	150	130	13.34	100	73	27
6	100	89	11	100	79	21
7	80	68	15	100	71	29
8	100	90	10	100	46	54
9	80	65	18.75	100	59	41
10	80	61	23.75	100	67	33
11	80	65	18.75	100	43	57
12	80	64	20	100	73	27
13	80	65	18.75	100	79	21
14	80	61	23.75	100	74	26
15	100	84	16	100	51	49
16	80	66	17.5	100	75	25
17	80	64	20	100	63	37
18	80	70	12.5	100	78	22
19	80	70	12.5	100	66	34
20	80	61	23.75	100	60	40
21	80	68	15	100	54	46
22	80	64	20	100	76	24
23	80	69	13.75	100	80	20
24	80	68	15	100	46	54
25	80	60	25	100	55	45
26	80	68	15	100	75	25
27	80	62	22.5	100	69	31
28	80	61	23.75	100	74	26
29	80	67	16.25	100	76	24
30	80	68	15	100	78	22

Note: Regarding search span, values range within 50% below original and 50% above original values.

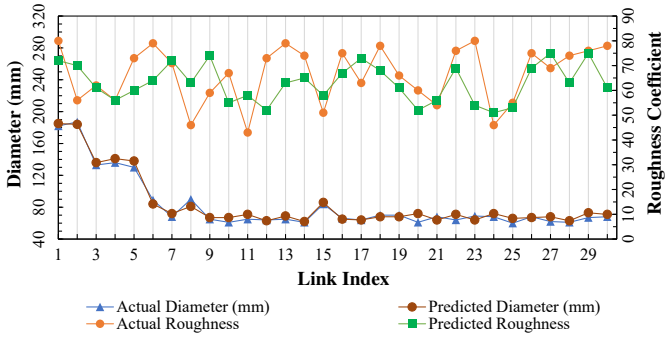

Fig. 6. Comparison of actual versus predicted pipeline conditions in the combined scenario for GoYang WDN.

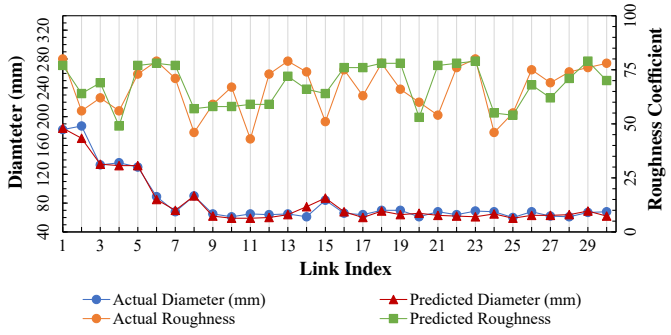

Fig. 7. Comparison of actual versus predicted pipeline conditions in the individual prediction scenario for GoYang WDN.

Tier-A scenarios is not highly significant, but those corresponding to diameter prediction were considerable compared with the least MAPE values for both networks. Clearly, optimizing the locations for placing FMS and PMS would yield best prediction accuracies for both the parameters.

Sensitivity to Variation in Number of Flow Monitoring Stations

Tier B, consisting of nine scenarios, is meant to analyze the sensitivity of MAPE values to variation in the number of FMS keeping the number and locations of PMS same as in the baseline scenario. Table 10 presents the number of FMS considered in different scenarios of Tier B for both the networks along with the resulting MAPE values. The number of FMS considered are 0, 4, 9, and 12 for Hanoi and 0, 4, 7, and 10 for GoYang. Fig. 8 presents the change in MAPE values for both condition parameters when the number of FMS are increased. The results from Tier-A scenarios are appropriately added to Fig. 8 to represent the MAPE variations for the baseline scenario. It can be observed from Fig. 8 that there is no consistent trend of MAPE values dropping when the number of FMS increased, which suggests that the accuracies of predicting roughness and diameters do not necessarily increase by adding more flow monitoring stations.

Furthermore, in order to evaluate the importance of flow monitoring on top of pressure monitoring, the scenario of zero FMS

Table 10. Additional scenarios considered for sensitivity analyses

Scenario	Tier	No. of FMS in Hanoi WDN	No. of PMS in Hanoi WDN	MAPE-R_{Comb} (%) in Hanoi WDN	MAPE-D_{Comb} (%) in Hanoi WDN	No. of FMS in GoYang WDN	No. of PMS in GoYang WDN	MAPE-R_{Comb} (%) in GoYang WDN	MAPE-D_{Comb} (%) in GoYang WDN
Baseline scenario	N/A	9	8	12.98	6.12	7	5	14.66	5.46
1	A	9	8	10.44	6.57	7	5	17.6	4.43
2		9	8	9.34	6.79	7	5	15.7	5.2
3		9	8	11.91	8.68	7	5	13.91	5.3
4		9	8	9.98	6.37	7	5	15.56	5.69
5		9	8	9.28	7.21	7	5	15.66	5.82
6		9	8	10.65	7.85	7	5	16.29	5.8
7		9	8	10.49	7.41	7	5	12.77	7.04
8		9	8	10.4	5.24	7	5	16.22	5.73
9	B	0	8	10.67	6.05	0	5	16.15	5.88
10		4	8	12.52	8.22	4	5	16.85	5.29
11		4	8	9.24	8.11	4	5	14.97	5.85
12		4	8	12.61	7.19	4	5	13.49	5.64
13		4	8	11.51	6.15	4	5	15.92	4.12
14		12	8	10.71	7.12	10	5	15.93	5.13
15		12	8	11.58	7.85	10	5	15.55	4.31
16		12	8	9.07	7.14	10	5	14.32	4.95
17		12	8	9.32	8.26	10	5	14.72	5.44
18	C	9	0	12.44	6.58	7	0	15.81	5.9
19		9	3	13.91	8.95	7	3	13.32	5.55
20		9	3	11.69	7.27	7	3	16.37	4.54
21		9	3	10.31	6.79	7	3	16.37	4.27
22		9	3	9.47	9.28	7	3	15.42	5.01
23		9	11	11.59	6.99	7	8	15.41	5.63
24		9	11	11.47	7.59	7	8	14.44	5.29
25		9	11	9.28	7.91	7	8	15.43	5.12
26		9	11	10.47	7.32	7	8	15.2	5.36

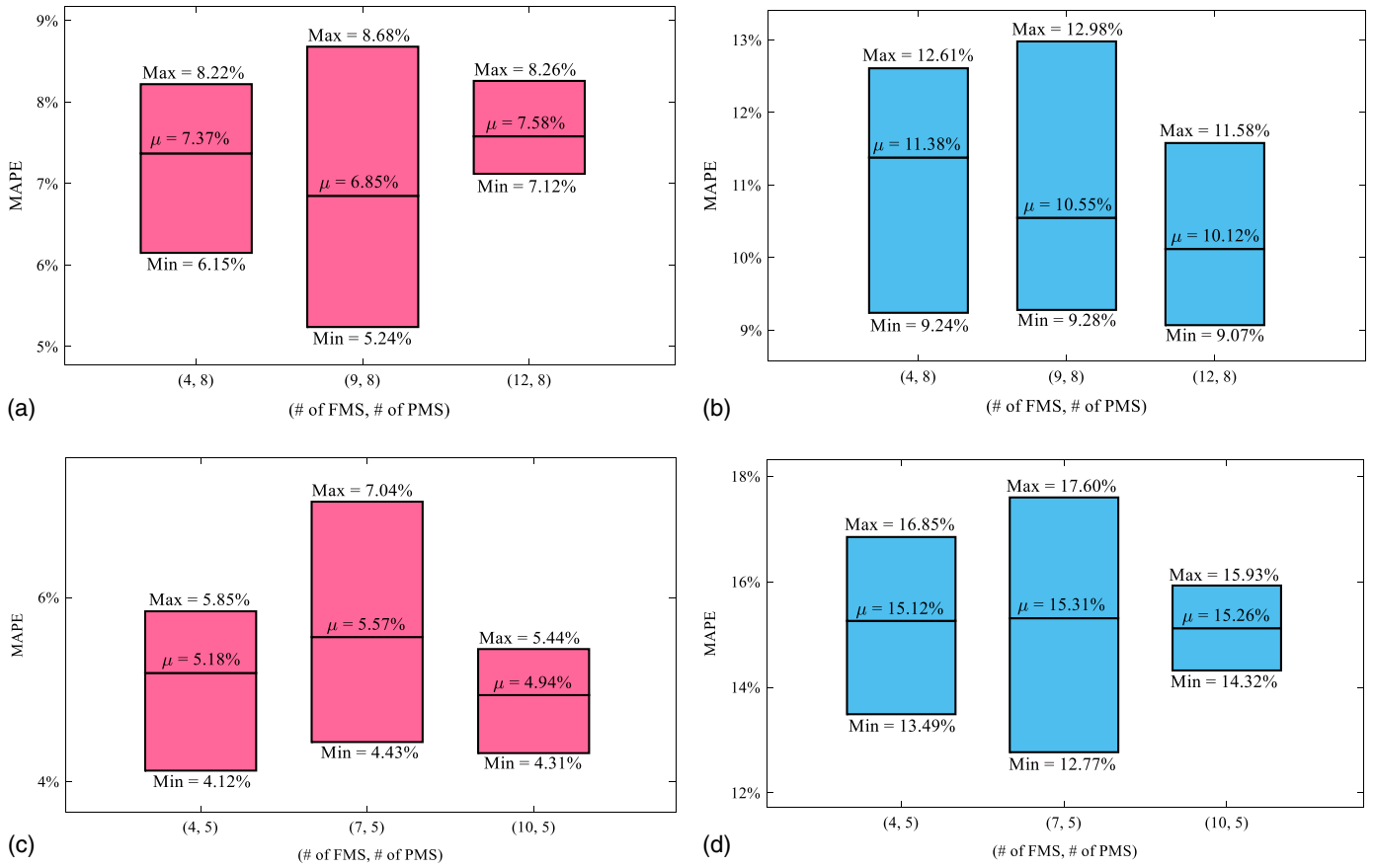

Fig. 8. MAPE variation with number of FMS: (a) Hanoi—diameter; (b) Hanoi—roughness; (c) GoYang—diameter; and (d) GoYang—roughness.

(Scenario 9 in Table 10) is separately evaluated. The ANN algorithm had to be trained again by eliminating the flow monitoring data as one of the sets of outputs in its prediction. In other words, pressure data are the only output in the ANN algorithm, with demands, roughness coefficients, and effective diameters as the inputs. The MAPE values for ANN validation resulting from the pressure-only scenarios where there exists only pressure as the output of the ANN model for Hanoi WDN is found to be 1% and for GoYang, it is 0.03%. The elimination of flow values as output variable in the ANNs improved their prediction accuracy.

Fig. 9 illustrates the actual and predicted values of both D_{Comb} and R_{Comb} for Hanoi. The MAPE value for D_{Comb} equals 5.70%, whereas that for R_{Comb} is equal to 7.73%. These values are slightly better than those obtained in the baseline scenario with nine FMS and eight PMS for Hanoi. These results suggest that flow monitoring may not add significant value on top of pressure monitoring for the kind of condition assessment prediction proposed in this study. It is likely that flow monitoring and pressure monitoring are redundant, and fewer computations resulting from the elimination of flow monitoring data may have resulted in slightly better accuracies, as observed in the results presented in Fig. 9.

Fig. 10 shows the actual and predicted values of both D_{Comb} and R_{Comb} for GoYang for the zero-FMS scenario. The MAPE value for D_{Comb} equals 6.48% whereas that for R_{Comb} is equal to 6.61%. These accuracies are comparable to those obtained in the baseline scenario with seven flow monitoring and five pressure monitoring stations. This reinforces the fact that pressure monitoring alone may suffice for predicting the condition assessment parameters investigated in this study.

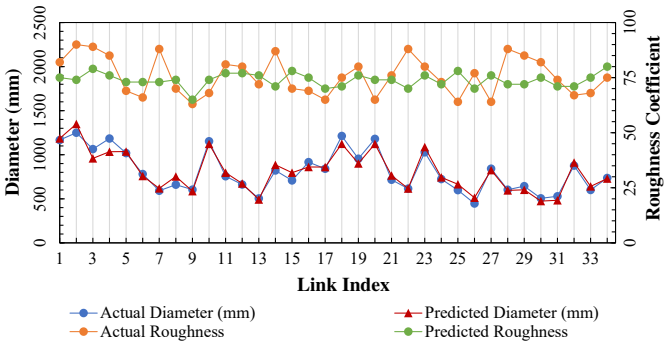

Fig. 9. Comparison of actual versus predicted pipeline conditions in the combined prediction scenario for Hanoi WDN using only pressure monitoring data.

Sensitivity to Variation in Number of Pressure Monitoring Stations

Contrary to Tier B, Tier C, consisting of nine scenarios, is meant to analyze the sensitivity of MAPE values to variation in the number of PMS while keeping the number and locations of FMS same as in the baseline scenario. Table 10 provides the number of PMS considered in different scenarios of Tier C for both the networks. The number of PMS considered are 0, 3, 8, and 11 for Hanoi and 0, 3, 5, and 8 for GoYang. Fig. 11 presents the change in MAPE values for both condition parameters when the numbers of PMS are increased for both the networks. The results from Tier-A scenarios are

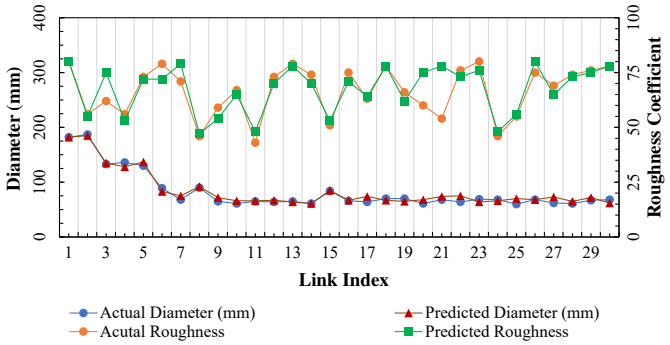

Fig. 10. Comparison of actual versus predicted pipeline conditions in the combined prediction scenario for GoYang WDN using only pressure monitoring data.

appropriately added to Fig. 11 to represent the MAPE variations for the baseline scenario.

Similar to the previous observations made on the sensitivity to the number of FMS, it can be observed from Fig. 11 that there is no clear trend of MAPE dropping when the numbers of PMS are increased for either of the networks. Furthermore, when the MAPE values corresponding to the scenarios with zero PMS are highest in Fig. 11, they are not significantly higher than the MAPE values in other scenarios. In other words, the prediction model did not suffer greatly by not having any pressure monitoring data on top of flow monitoring data.

To summarize the findings in this paper (1) a reliable ANN model was presented at a rough MAPE accuracy of 2.8% on average as an alternative for the conventional hydraulic simulator, (2) the performance of the reverse-engineering prediction model was found to be acceptable at an average MAPE accuracy of 5.79% and 13.79% for diameter and roughness predictions, respectively, using two benchmarks, and (3) the sensitivity analyses of the placements and numbers of pressure and flow monitoring stations indicated that not only do pressure stations play more essential a role than those of flow when it comes to the prediction of effective hydraulic diameters and roughness coefficients, but also a limited number of stations would suffice to acquire an acceptable prediction accuracy. In general, the study of the baseline scenarios along with sensitivity analyses offers novel insights into how cyber-monitoring data can be leveraged for primitive pipeline condition assessment.

Conclusions and Recommendations

This paper presented a water pipeline condition prediction framework that is driven by hydraulic monitoring of WDNs in conjunction with water consumption monitoring. Pressure and flow measured at multiple locations in the WDN for a given set of nodal demands are the inputs for the proposed framework. A MATLAB-based optimization framework is then used to demonstrate the proposed framework for the prediction of water pipeline roughness and effective internal diameters. Two popular benchmark WDNs are used for this demonstration purpose. For the first WDN, Hanoi, which is a gravity-driven system with 34 deteriorated pipelines, roughness and effective internal diameters are together predicted with a MAPE of 12.98% and 6.12%, respectively, using eight pressure monitoring and nine flow monitoring stations that are randomly located in the WDN. For the second WDN, GoYang, which is a pump driven system with 30 deteriorated pipelines, roughness coefficients and effective internal diameters are together predicted with MAPE of 14.66% and 5.46%, respectively, using five pressure monitoring and seven flow monitoring stations that are randomly located in the WDN. The MAPE values for predicting roughness and effective diameters individually assuming the other condition parameter is known are found to be slightly lower than when they are predicted together. These results suggest that the proposed framework could

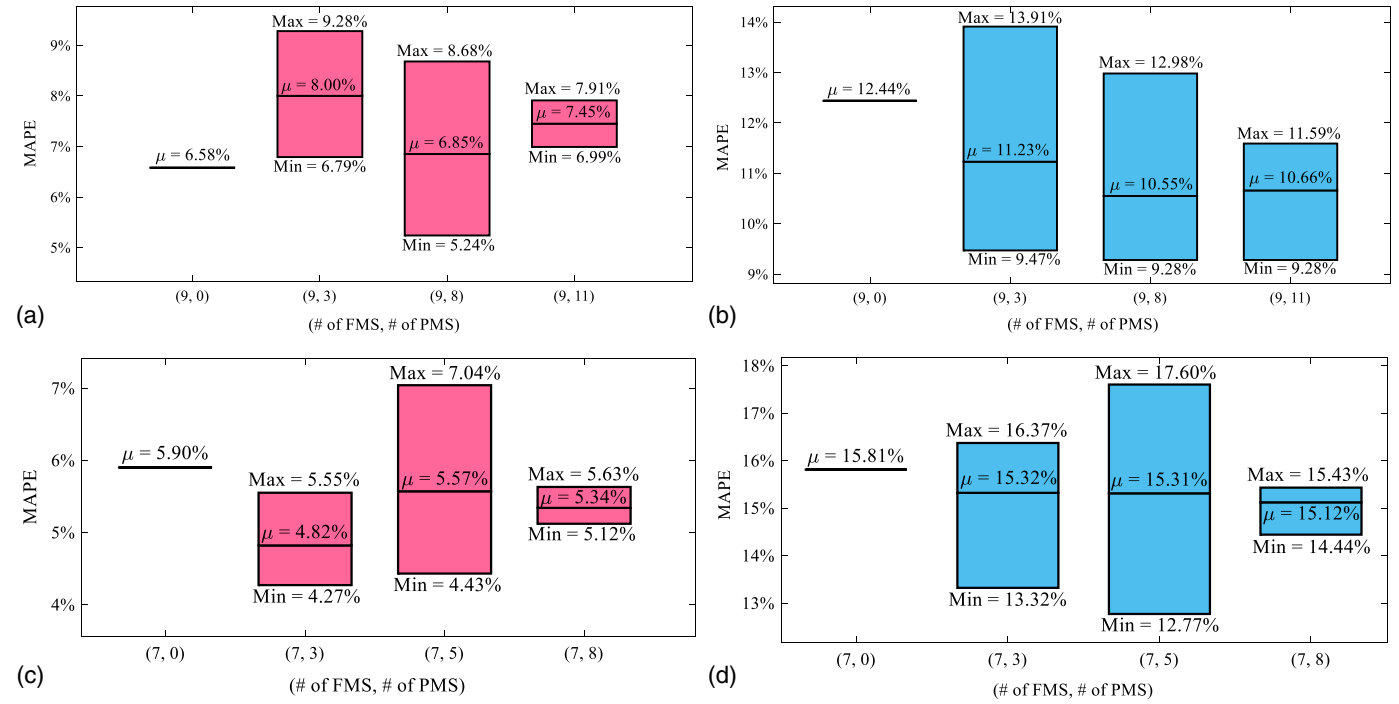

Fig. 11. MAPE variation with number of PMS for: (a) Hanoi—diameter; (b) Hanoi—roughness; (c) GoYang—diameter; and (d) GoYang—roughness.

work as a preliminary tool in inferring the conditions of deteriorated water pipelines.

Some key findings of this study include the following:

- The effective internal pipeline diameters were found to be more accurately predictable than the roughness based on hydraulic monitoring data, which is likely due to their greater influence on the flow and pressure variations in the WDN.
- Increasing the number of pressure or flow monitoring stations did not necessarily increase the prediction accuracies for either of the condition parameters.
- Placement of pressure or flow monitoring stations at optimal locations would lead to best prediction accuracies of the pipeline condition parameters.
- Prediction accuracies of both the condition parameters using only pressure monitoring data were comparable to those obtained using both pressure and flow monitoring data, which suggests that pressure monitoring alone may suffice the purpose of leveraging hydraulic fluctuations in WDNs for pipeline condition prediction.

Limitations of this study include (1) the assumption that all the pipeline features except roughness and effective internal diameter are known and can be accurately quantified, which may be optimistic given the potential for leakages, wall thinning, and many other issues; and (2) the assumption that water consumption data at all the WDN nodes can be captured synchronously with the corresponding pressure and flow monitoring data, which is less likely in the current scheme of water utility operations; however, this is definitely possible as more utilities adopt smart water meters and cyber monitoring of WDNs for better situational awareness.

The contribution of this study is the representation and calibration of a data-driven prediction model for uncertain hydraulic parameters such as effective internal diameters and roughness coefficients where manual in situ inspections for pipe rehabilitation or replacement accordingly can potentially be substituted for a cyber-monitoring framework at reasonable accuracy, cost, and time efficiency compared with previous studies in the literature discussed previously in the paper. This primitive pipeline condition assessment paradigm needs to be further investigated in the future for broadening the practical impact.

Future research in this field should focus on (1) expanding the variety of condition parameters that can be simultaneously predicted using hydraulic monitoring data; for example, presence of leakages may be added as an additional condition parameter; (2) although it may be possible to capture water consumption data using smart water meters, it may be less practical to be able to install smart water meters at all the nodes in the WDN and therefore future research should focus on using water consumption data from just a sample of nodes in the WDN; and (3) optimizing the locations for the placement of pressure monitoring stations and selecting nodes for installing smart water meters in order to achieve best pipeline condition prediction accuracies.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon a reasonable request.

Acknowledgments

This research was supported by the National Science Foundation (NSF) under Grant No. 1638321. The results and conclusion presented in this paper are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government. The authors are grateful to the NSF for this support.

Notation

The following symbols are used in this paper:

a = index for pressure monitoring stations;
act = actual values of condition parameters;
b = index for flow monitoring stations;
C, c = roughness coefficient;
D, d = pipe effective internal diameter;
D_{Comb} = combined predicted pipe diameter;
F, f = flow rate;
$g()$, $h()$ = representations of EPANET 2.0 hydraulic simulations;
i = candidate solution reference in the optimization procedure;
j = number of consumption datasets;
k = a given set of consumption data;
l = number of validation scenarios;
P, p = pressure;
pr = predicted values of condition parameters;
Q, q = nodal demands;
R_{Comb} = combined predicted roughness coefficient;
sim = simulated values of pressure or flow rate;
x = number of pipelines in the water distribution network; and
Z = objective function.

References

Abdelhafidh, M., M. Fourati, L. C. Fourati, and A. Abidi. 2018. "Remote water pipeline monitoring system IoT-based architecture for new industrial era 4.0." In *Proc., IEEE/ACS Int. Conf. on Computer Systems and Applications*, 1184–1191. New York: IEEE. https://doi.org/10.1109/AICCSA.2017.158.

Abdulla, M. B., R. O. Herzallah, and M. A. Hammad. 2013. "Pipeline leak detection using artificial neural network: Experimental study." In *Proc., Int. Conf. on Modelling, Identification and Control*, 328–332. New York: IEEE.

Abdulshaheed, A., F. Mustapha, and A. Ghavamian. 2017. "A pressure-based method for monitoring leaks in a pipe distribution system: A review." *Renewable Sustainable Energy Rev.* 69 (May): 902–911. https://doi.org/10.1016/j.rser.2016.08.024.

Alegre, H., S. T. Coelho, D. I. C. Covas, M. Do Céu Almeida, and A. Cardoso. 2013. "A utility-tailored methodology for integrated asset management of urban water infrastructure." *Water Sci. Technol. Water Supply* 13 (6): 1444–1451. https://doi.org/10.2166/ws.2013.108.

Amoatey, P. K., A. Bárdossy, and H. Steinmetz. 2018. "Inverse optimization based detection of leaks from simulated pressure in water networks. Part 2: Analysis for two leaks." *J. Water Manage. Model.* 26: C460. https://doi.org/10.14796/JWMM.C460.

Bedjou, A., A. Boudoukha, and B. Bosseler. 2019. "Assessment of wastewater asset management effectiveness in the case of rare data and low investments." *Int. J. Environ. Sci. Technol.* 16 (7): 3781–3792. https://doi.org/10.1007/s13762-018-2005-3.

Braun, M., O. Piller, A. Iollo, and I. Mortazavi. 2020. "A spectral approach to uncertainty quantification in water distribution networks." *J. Water Resour. Plann. Manage.* 146 (3): 04019080. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001138.

Chen, T. Y. J., J. A. Beekman, and S. D. Guikema. 2017. "Drinking water distribution systems asset management: Statistical modelling of pipe breaks." In *Proc., Pipelines 2017: Condition Assessment, Surveying,*

and Geomatics—Proc., Sessions of the Pipelines 2017 Conf., 173–186. Rseton, VA: ASCE. https://doi.org/10.1061/9780784480885.017.

Cody, R. A., B. A. Tolson, and J. Orchard. 2020. "Detecting leaks in water distribution pipes using a deep autoencoder and hydroacoustic spectrograms." *J. Comput. Civ. Eng.* 34 (2): 04020001. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000881.

de Myttenaere, A., B. Golden, B. Le Grand, and F. Rossi. 2016. "Mean absolute percentage error for regression models." *Neurocomputing* 192 (Jun): 38–48. https://doi.org/10.1016/j.neucom.2015.12.114.

de Winter, C., V. R. Palleti, D. Worm, and R. Kooij. 2019. "Optimal placement of imperfect water quality sensors in water distribution networks." *Comput. Chem. Eng.* 121 (Feb): 200–211. https://doi.org/10.1016/j.compchemeng.2018.10.021.

El-Zahab, S., and T. Zayed. 2017. "Leak detection model for pressurized pipelines using support vector machines." In Vol. 2 of *Proc., 6th CSCE-CRC Int. Construction Specialty Conf. 2017—Held as Part of the Canadian Society for Civil Engineering Annual Conf. and General Meeting*, 975–984. Hong Kong: Hong Kong Polytechnic Univ.

Ferrandez-Gamot, L., P. Busson, J. Blesa, S. Tornil-Sin, V. Puig, E. Duviella, and A. Soldevila. 2015. "Leak localization in water distribution networks using pressure residuals and classifiers." *IFAC-PapersOnLine* 48 (21): 220–225. https://doi.org/10.1016/j.ifacol.2015.09.531.

Fuentes, V. C., and J. R. I. Pedrasa. 2020. "Leak detection in water distribution networks via pressure analysis using a machine learning ensemble." In *Lecture notes of the institute for computer sciences, social-informatics and telecommunications engineering*, 31–44. New York: Springer. https://doi.org/10.1007/978-3-030-45293-3_3.

Fujiwara, O., and D. B. Khang. 1990. "A two-phase decomposition method for optimal design of looped water distribution networks." *Water Resour. Res.* 26 (4): 539–549. https://doi.org/10.1029/WR026i004p00539.

Gao, T. 2017. "Pipe roughness estimation in water distribution networks using head loss adjustment." *J. Water Resour. Plann. Manage.* 143 (5): 04017007. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000752.

Gertler, J., J. Romera, V. Puig, and J. Quevedo. 2010. "Leak detection and isolation in water distribution networks using principal component analysis and structured residuals." In *Proc., Conf. on Control and Fault-Tolerant Systems, SysTol'10—Final Program and Book of Abstracts*, 191–196. New York: IEEE. https://doi.org/10.1109/SYSTOL.2010.5676043.

Giudicianni, C., M. Herrera, A. Di Nardo, R. Greco, E. Creaco, and A. Scala. 2020. "Topological placement of quality sensors in water-distribution networks without the recourse to hydraulic modeling." *J. Water Resour. Plann. Manage.* 146 (6): 04020030. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001210.

Grigg, N. S., and J. Butler. 2019. "Distribution systems: Has asset management made a difference?" *J. Pipeline Syst. Eng. Pract.* 10 (2): 04019010. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000379.

Hamdia, K. M., X. Zhuang, and T. Rabczuk. 2021. "An efficient optimization approach for designing machine learning models based on genetic algorithm." *Neural Comput. Appl.* 33 (6): 1923–1933. https://doi.org/10.1007/s00521-020-05035-x.

Jensen, H. A., and D. J. Jerez. 2018. "A stochastic framework for reliability and sensitivity analysis of large scale water distribution networks." *Reliab. Eng. Syst. Saf.* 176 (Aug): 80–92. https://doi.org/10.1016/j.ress.2018.04.001.

Kamiński, K., W. Kamiński, and T. Mizerski. 2017. "Application of artificial neural networks to the technical condition assessment of water supply systems." *Ecol. Chem. Eng. S* 24 (1): 31–40. https://doi.org/10.1515/eces-2017-0003.

Kapelan, Z., D. Savic, and G. Walters. 2004. "Incorporation of prior information on parameters in inverse transient analysis for leak detection and roughness calibration." *Urban Water J.* 1 (2): 129–143. https://doi.org/10.1080/15730620412331290029.

Karney, B., D. Khani, M. Halfawy, and O. Hunaidi. 2009. "A simulation study on using inverse transient analysis for leak detection in water distribution networks." *J. Water Manage. Model.* https://doi.org/10.14796/jwmm.r235-23.

Kayaalp, F., A. Zengin, R. Kara, and S. Zavrak. 2017. "Leakage detection and localization on water transportation pipelines: A multi-label classification approach." *Neural Comput. Appl.* 28 (10): 2905–2914. https://doi.org/10.1007/s00521-017-2872-4.

Kim, J. H., T. G. Kim, J. H. Kim, and Y. N. Yoon. 1994. "A study on the pipe network system design using non-linear programming." *J. Korean Water Resour. Assoc.* 27 (4): 59–67.

Lin, C. C. 2017. "A hybrid heuristic optimization approach for leak detection in pipe networks using ordinal optimization approach and the symbiotic organism search." *Water* 9 (10): 812. https://doi.org/10.3390/w9100812.

Liu, Z., and Y. Kleiner. 2013. "State of the art review of inspection technologies for condition assessment of water pipes." *Measurement* 46 (1): 1–15. https://doi.org/10.1016/j.measurement.2012.05.032.

Luciani, C., F. Casellato, S. Alvisi, and M. Franchini. 2019. "Green Smart Technology for Water (GST4Water): Water loss identification at user level by using smart metering systems." *Water* 11 (3): 405. https://doi.org/10.3390/w11030405.

Mazumder, R. K., A. M. Salman, Y. Li, and X. Yu. 2019. "Reliability analysis of water distribution systems using physical probabilistic pipe failure method." *J. Water Resour. Plann. Manage.* 145 (2): 04018097. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001034.

Meseguer, J., J. M. Mirats-Tur, G. Cembrano, and V. Puig. 2015. "Model-based monitoring techniques for leakage localization in distribution water networks." *Procedia Eng.* 119 (1): 1399–1408. https://doi.org/10.1016/j.proeng.2015.08.1000.

Mirzal, A., S. Yoshii, and M. Furukawa. 2012. "PID parameters optimization by using genetic algorithm." Accessed April 4, 2012. http://arxiv.org/abs/1204.0885.

Moglia, M., S. Burn, and S. Meddings. 2006. "Decision support system for water pipeline renewal prioritization." *Electron. J. Inf. Technol. Constr.* 11 (18): 237–256.

Momeni, A., and K. R. Piratla. 2020. "A novel cyber-monitoring based asset management scheme for water distribution networks through fine-tuning genetic algorithm parameters." In *Proc., 37th Int. NO-DIG Conf. and Exhibition 2019*. London: International Society for Trenchless Technology.

Momeni, A., K. R. Piratla, and K. C. Madathil. 2019. "A novel computationally efficient asset management framework based on monitoring data from water distribution networks." In *Proc., ASCE Construction Research Congress*. Reston, VA: ASCE.

Momeni, A., V. Prasad, H. I. Dharmawardena, K. R. Piratla, and K. Venayagamoorthy. 2018. "Mapping and modeling interdependent power, water, and gas infrastructures." In *Proc., Clemson University Power Systems Conf.* New York: IEEE. https://doi.org/10.1109/PSC.2018.8664050.

Monsef, H., M. Naghashzadegan, A. Jamali, and R. Farmani. 2019. "Comparison of evolutionary multi objective optimization algorithms in optimum design of water distribution network." *Ain Shams Eng. J.* 10 (1): 103–111. https://doi.org/10.1016/j.asej.2018.04.003.

Montáns, F. J., F. Chinesta, R. Gómez-Bombarelli, and J. N. Kutz. 2019. "Data-driven modeling and learning in science and engineering." *C.R. Mec.* 347 (11): 845–855. https://doi.org/10.1016/j.crme.2019.11.009.

Newton, L. A., and J. Christian. 2015. "Challenges in asset management—A case study." Accessed February 13, 2015. https://www.researchgate.net/publication/44061438_Challenges_in_asset_management_-_a_case_study.

Pacheco, V. M. F., R. E. Valdés, E. B. Gil, A. N. Manso, and E. Á. Álvarez. 2020. "Techno-economic analysis of residential water meters: A practical example." *Water Resour. Manage.* 34 (8): 2471–2484. https://doi.org/10.1007/s11269-020-02564-x.

Packianather, M. S., P. R. Drake, and H. Rowlands. 2000. "Optimizing the parameters of multilayered feed forward neural networks through Taguchi design of experiments." *Qual. Reliab. Eng. Int.* 16 (6): 461–473. https://doi.org/10.1002/1099-1638(200011/12)16:6<461::AID-QRE341>3.0.CO;2-G.

Park, H., S. H. Ting, and H. D. Jeong. 2016. "Procedural framework for modeling the likelihood of failure of underground pipeline assets." *J. Pipeline Syst. Eng. Pract.* 7 (2): 04015023. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000222.

Perez, R., G. Sanz, V. Puig, J. Quevedo, M. A. Cuguero Escofet, F. Nejjari, J. Meseguer, G. Cembrano, J. M. Mirats Tur, and R. Sarrate. 2014.

"Leak localization in water networks: A model-based methodology using pressure sensors applied to a real network in Barcelona (applications of control)." *IEEE Control Syst. Mag.* 34 (4): 24–36. https://doi.org/10.1109/MCS.2014.2320336.

Piratla, K. R., and A. Momeni. 2019. "A novel water pipeline asset management scheme using hydraulic monitoring data." *Pipelines 2019: Multidisciplinary Topics, Utility Engineering, and Surveying—Proc., Sessions of the Pipelines 2019 Conf.*, 190–198. Reston, VA: ASCE. https://doi.org/10.1061/9780784482506.020.

Poojitha, S. N., G. Singh, and V. Jothiprakash. 2020. "Improving the optimal solution of GoYang network—Using genetic algorithm and differential evolution." *Water Sci. Technol. Water Supply* 20 (1): 95–102. https://doi.org/10.2166/ws.2019.139.

Poulakis, Z., D. Valougeorgis, and C. Papadimitriou. 2003. "Leakage detection in water pipe networks using a Bayesian probabilistic framework." *Probab. Eng. Mech.* 18 (4): 315–327. https://doi.org/10.1016/S0266-8920(03)00045-6.

Prayudani, S., A. Hizriadi, Y. Y. Lase, and Y. Fatmi. 2019. "Analysis accuracy of forecasting measurement technique on random K-nearest neighbor (RKNN) using MAPE and MSE." In Vol. 1361 of *Proc., Journal of Physics: Conf. Series*. Medan, Indonesia: IOP Publishing. https://doi.org/10.1088/1742-6596/1361/1/012089.

Qi, Z., F. Zheng, D. Guo, H. R. Maier, T. Zhang, T. Yu, and Y. Shao. 2018. "Better understanding of the capacity of pressure sensor systems to detect pipe burst within water distribution networks." *J. Water Resour. Plann. Manage.* 144 (7): 04018035. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000957.

Raei, E., M. E. Shafiee, M. R. Nikoo, and E. Berglund. 2019. "Placing an ensemble of pressure sensors for leak detection in water distribution networks under measurement uncertainty." *J. Hydroinf.* 21 (2): 223–239. https://doi.org/10.2166/hydro.2018.032.

Shinozuka, M., J. Liang, and M. Q. Feng. 2005. "Use of supervisory control and data acquisition for damage location of water delivery systems." *J. Eng. Mech.* 131 (3): 225–230. https://doi.org/10.1061/(ASCE)0733-9399(2005)131:3(225.

Shukla, H., and K. Piratla. 2020. "Leakage detection in water pipelines using supervised classification of acceleration signals." *Autom. Constr.* 117 (Sep): 103256. https://doi.org/10.1016/j.autcon.2020.103256.

Soldevila, A., J. Blesa, R. M. Fernandez-Canti, S. Tornil-Sin, and V. Puig. 2019. "Data-driven approach for leak localization in water distribution networks using pressure sensors and spatial interpolation." *Water* 11 (7): 1500. https://doi.org/10.3390/w11071500.

Soldevila, A., J. Blesa, S. Tornil-Sin, E. Duviella, R. M. Fernandez-Canti, and V. Puig. 2016. "Leak localization in water distribution networks using a mixed model-based/data-driven approach." *Control Eng. Pract.* 55 (Oct): 162–173. https://doi.org/10.1016/j.conengprac.2016.07.006.

Soldevila, A., J. Blesa, S. Tornil-Sin, R. M. Fernandez-Canti, and V. Puig. 2018. "Sensor placement for classifier-based leak localization in water distribution networks using hybrid feature selection." *Comput. Chem. Eng.* 108 (Jan): 152–162. https://doi.org/10.1016/j.compchemeng.2017.09.002.

Soldevila, A., R. M. Fernandez-Canti, J. Blesa, S. Tornil-Sin, and V. Puig. 2017a. "Leak localization in water distribution networks using Bayesian classifiers." *J. Process Control* 55 (Jul): 1–9. https://doi.org/10.1016/j.jprocont.2017.03.015.

Soldevila, A., R. M. Fernandez-Canti, J. Blesa, S. Tornil-Sin, and V. Puig. 2017b. "Leak localization in water distribution networks using model-based Bayesian reasoning." In *Proc., 2016 European Control Conf.*, 1758–1763. New York: IEEE. https://doi.org/10.1109/ECC.2016.7810545.

Szegedy, C., W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus. 2014. *Intriguing properties of neural networks*. New York: Cornell Univ.

Wu, J., X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng. 2019. "Hyperparameter optimization for machine learning models based on Bayesian optimization." *J. Electron. Sci. Technol.* 17 (1): 26–40. https://doi.org/10.11989/JEST.1674-862X.80904120.

Wu, S., S. Hrudey, S. French, T. Bedford, E. Soane, and S. Pollard. 2009. "A role for human reliability analysis (HRA) in preventing drinking water incidents and securing safe drinking water." *Water Res.* 43 (13): 3227–3238. https://doi.org/10.1016/j.watres.2009.04.040.

Yazdekhasti, S., K. R. Piratla, J. Sorber, S. Atamturktur, A. Khan, and H. Shukla. 2020. "Sustainability analysis of a leakage-monitoring technique for water pipeline networks." *J. Pipeline Syst. Eng. Pract.* 11 (1): 04019052. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000425.

Zhang, C., A. C. Zecchin, M. F. Lambert, J. Gong, and A. R. Simpson. 2018. "Multi-stage parameter-constraining inverse transient analysis for pipeline condition assessment." *J. Hydroinf.* 20 (2): 281–300. https://doi.org/10.2166/hydro.2018.154.

Zhang, Q., F. Zheng, Y. Jia, D. Savic, and Z. Kapelan. 2021. "Real-time foul sewer hydraulic modelling driven by water consumption data from water distribution systems." *Water Res.* 188 (Jan): 116544. https://doi.org/10.1016/j.watres.2020.116544.

Zhang, X., and K. Duh. 2020. "Reproducible and efficient benchmarks for hyperparameter optimization of neural machine translation systems." *Trans. Assoc. Comput. Ling.* 8 (Jul): 393–408. https://doi.org/10.1162/tacl_a_00322.