

Metamodel for nodal pressure estimation at near real-time in water distribution systems using artificial neural networks

Gustavo Meirelles Lima, Bruno Melo Brentan, Daniel Manzi and Edevar Luvizotto Jr

ABSTRACT

The development of computational models for analysis of the operation of water supply systems requires the calibration of pipes' roughness, among other parameters. Inadequate values of this parameter can result in inaccurate solutions, compromising the applicability of the model as a decision-making tool. This paper presents a metamodel to estimate the pressure at all nodes of a distribution network based on artificial neural networks (ANNs), using a set of field data obtained from strategically located pressure sensors. This approach aims to increase the available pressure data, reducing the degree of freedom of the calibration problem. The proposed model uses the inlet flow of the district metering area and pressure data monitored in some nodes, as input data to the ANN, obtaining as output, the pressure values for nodes that were not monitored. Two case studies of real networks are presented to validate the efficiency and accuracy of the method. The results ratify the efficiency of ANN as state forecaster, showing the high applicability of the metamodel tool to increase a database or to identify abnormal events during an operation.

Key words | artificial neural networks, monitoring, pressure estimation, water supply networks

Gustavo Meirelles Lima
Bruno Melo Brentan
Daniel Manzi
Edevar Luvizotto Jr (corresponding author)
School of Civil Engineering, Architecture and Urban
Design, Department of Water Resources,
University of Campinas-Unicamp-BR,
Campinas,
Brazil
E-mail: edevar@fec.unicamp.br

INTRODUCTION

The hydraulic simulations of water supply networks are widely applied for several uses, such as quality parameter determination (Sunela & Puust 2015), optimal network design (Mora-Melia *et al.* 2015), water leakage detection (Ishido & Takahashi 2014), and optimal operation (Brentan & Luvizotto Jr 2014).

A calibration process is required at previous use of these models to accurately reproduce the field conditions. The calibration process consists of adjusting some parameters which minimize the error between the observed and computed values. Typically, the pipe roughness has been explored as adjusting the parameter for network calibration, trying to minimize the difference between modeled and observed pressure (Alvisi & Franchini 2010; Giustolisi &

Berardi 2011; Roma *et al.* 2015). More complex calibration models involve the adjustment of hourly demand factors or even the emitters' coefficient, responsible to simulate nodal pressure guided by demand which can include at the objective function the minimization error between the observed and modeled flow (Cheng & He 2010; Khedr *et al.* 2015). However, the result of this process is a model with many uncertainties due to many degrees of freedom of the problem, since the relation between pressure monitoring flow points and the number of variables is very low (Alvisi & Franchini 2010). Thus, the modeled networks may not portray reality adequately, mainly as far as the different scenarios found in the observed data used for the calibration step.

The increase of data to minimize the uncertainties at calibration guarantees a higher reliability and applicability of the model, once the degrees of freedom for the system of equations involved decreases, determining thus the parameters with higher accuracy. In this sense, the calibration model at an extended period or with monitoring at several points corroborates this affirmation.

The use of hydraulic simulation models linked with optimization tools has emerged during the last decades as a fertile ground for new jobs, both for planning (Montalvo *et al.* 2014; Yoo *et al.* 2016) and for operation (Kougias & Theodossiou 2013; Price & Ostfeld 2014). Due to the topology of the search space, the use of bio-inspired algorithms is an alternative to solve hydraulic optimization problems, since these are not models that rely on derivative calculations. However, in some cases, a high number of simulations are necessary, bringing significant computational effort, which makes the use of hydraulic models impracticable (Rao & Alvarruiz 2007).

The metamodels are an alternative for the simulation models widely applied as optimization tools, once those models process input and output data generated by a base model. Initially, linear regression has been applied to estimate the correlation between input and output. However, actual approaches use artificial neural networks (ANNs) and machine learning theory, for more accurate results (Broad *et al.* 2010; Nazif *et al.* 2010; Razavi *et al.* 2012).

In addition, artificial intelligence, mainly the multi-layer perceptron (MLP) ANNs, have been widely applied to hydraulic engineering to estimate some parameters: Tiwari & Adamowski (2015) for short-term water demand forecasting; Cordoba *et al.* (2014) applied MLP to determine the chlorine concentration of water; and Rao & Alvarruiz (2007) presented an ANN structure to estimate the future state of a hydraulic network (pressure, flow, and tank levels at strategic locations) at near real-time to evaluate new operational conditions. Ping *et al.* (2014) presented a new development, where the nodal pressure data monitored during the last hours are used to determine the pressure of the next hour at the same nodes, using a support vector machine. This model presented good accuracy, but the authors highlighted that each node should use a specific forecaster, which increases the computational efforts mainly for the application at full nodes.

The usual application of metamodels surrogates the hydraulic simulator in order to save time at optimization problems. In these cases, the inputs are related to the operation of the systems, demands, or tanks level and the outputs are related to the hydraulic state of the networks, mainly in critical nodes or pipes of the network. Rao & Alvarruiz (2007) presented a metamodel based on ANNs, and the outputs are the pressure and flow in a set of nodes and pipes and tank level and power consumption by the pumps. The authors used as input control and hydraulic variables (pump and valve settings, tank level, and nodal demands). In the work presented by Broad *et al.* (2010), the development of a metamodel, also based on ANNs, used the 44 tank trigger levels as inputs. The outputs include minimum pressure over the control duration at the critical nodes.

Considering the efficiency of metamodel application for near real-time problems, this paper proposes the use of ANN for real time pressure estimation at all nodes of water networks at the current time using only the monitored pressure and flow. The present state forecaster is an innovative approach, because different to other applications, this purpose uses measured hydraulic parameters as input (only nodal pressure and inlet flow for the case studies). The objective of this metamodel is not to anticipate a future state of the network, but to obtain the knowledge of pressure, as an alternative to the full monitoring, which is uneconomical. The inlet flow at each sector and the monitored pressure of some nodes are used as input data to apply the method. Manifold scenarios are generated changing the pipes' roughness and nodal demands randomly to train the ANN. With this approach, the hydraulic model is not surrogated by the metamodel, but working together with the metamodel, creating an additional tool for water distribution system (WDS) management, since it can increase the knowledge of the hydraulic state of the system.

The method is evaluated in two real water distribution networks, which have different sizes and topologies: Campos do Conde II, located in Piracicaba, São Paulo, Brazil, and Cambuí system, a small city located in Minas Gerais, Brazil. For this second case, the data available from pressure monitoring done by Goulart (2015) are used to evaluate ANN performance with real data. The analysis is done at steady state and at extended period. The results point to it being a good tool to obtain the full monitoring

of nodal pressures, which can help the roughness calibration process, reducing uncertainty in the modeling process and the near real-time operation, and helping in the identification of anomalies and inadequate pressure zones.

ANNS AND MLP

Mathematical description

ANNs are structures which resemble a complex system of neurons, with each neuron receiving some input signals and generating an output signal computed using an activation function over the inputs. In other words, these are processing data units, which respond somehow to the input stimulus (Ding et al. 2011). Several architectures are proposed in the literature for ANN and each one can be more suitable to the problems featured, such as the size of data input, training data, or the available target data.

The MLP network has as its main feature the interconnection between the processing units, the perceptrons, which receive the data and after modulation by synaptic weights, pass along an activation function, responsible for generating the output of the unit. The advantage of increased interconnection is an increase in the adaptability of the ANN to the problem which leads to a more accurate mapping of inputs to outputs. Figure 1 shows a MLP, evidencing the input layer, two hidden layers, and an output layer.

An output s_k of MLP with two hidden layers can be written as:

$$s_k = \sum_{j=0}^n w_{rj} \cdot f \left(\sum_{i=1}^m v_{ji} \cdot x_i \right) = \sum_{j=0}^n w_{rj} \cdot f(\mathbf{v}^T \cdot \mathbf{x}) \quad (1)$$

where s_k is the r^{th} output of the neural network, w_{rj} is the synaptic weight of the output layer, referring to the output r for the input j of the previous layer, f is the activation function, v_{ji} is the synaptic weight for the first layer, referring to the output j for the input data x_i .

The training process of a MLP, also called learning stage, consists of determining the set of synaptic weights which minimize an error measurement between the output s and the observed data s^* . According to von Zuben (1996), the ANN training does not have the objective to determine rightly all weights, but to synthesize a surface

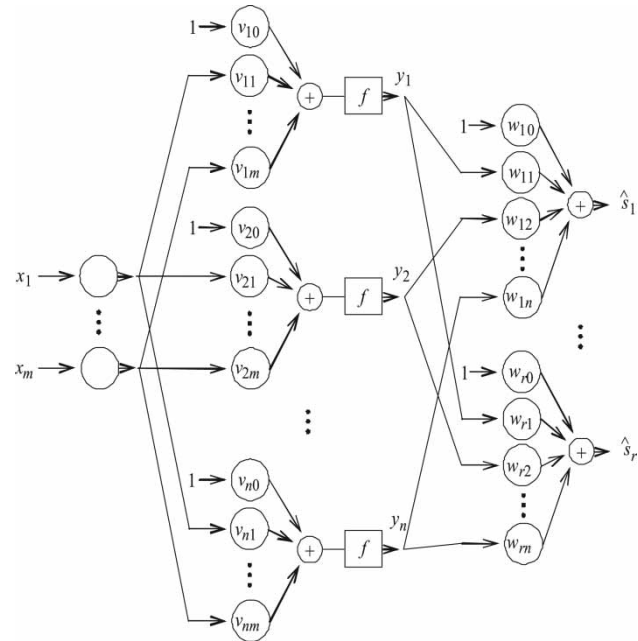


Figure 1 | Multilayer perceptron with two hidden layers (Source: von Zuben 1996).

that, when stimulated by similar data of the training, it minimizes the difference between the observed and the output data, with the ability to generalize; that means, to produce estimates with sufficient accuracy for nodes inside the training region which do not have data. Training the ANN involves minimizing the validation error on the training data via the backpropagation algorithm. After the linear combination of the weight vector and input data matrix, the result is passed along an activation function, usually a sigmoidal function that will process the output data.

ANN training

Hydraulic state of the network (pressure and flow) are affected by physical and operational changes, such as pipes' diameter, length, and roughness, tanks' water level, nodal demands, pumps and valves' status, and setting point. Since the topology of a network is well known and considering the variability of roughness and demand, and the uncertainty around them, the modification of these parameters can be an interesting way to generate new hydraulic states.

The database used to train the ANN should contain a wide range of pressure. Therefore, by defining a plausible

range for roughness and demand values, realistic random values can be generated for each pipe and node. The advantage of creating different scenarios of roughness and demand is the possibility to use a hydraulic model without previous calibration, which is a hard task on the computational analysis of water distribution systems.

The hydraulic parameters of interest, the pressure at nodes and the inlet flow, are obtained by hydraulic simulation using EPANET (Rossman 2000). In this way, the target database has as input vector the pressure at a set of monitored nodes for the steady state. In the extended period, in addition to the nodal pressure, the supply flow is used for input vector at ANN. The corresponding output vector is compounded by the pressure at all other nodes.

The process is repeated until the database has the previously established size. After training, a new database is generated to evaluate the ANN fitness. Figure 2 shows a flowchart of training data generation.

After the training, it is expected that ANN can reproduce the pressure response in all nodes using only the information collected from the monitoring pressure points, as shown in Figure 3.

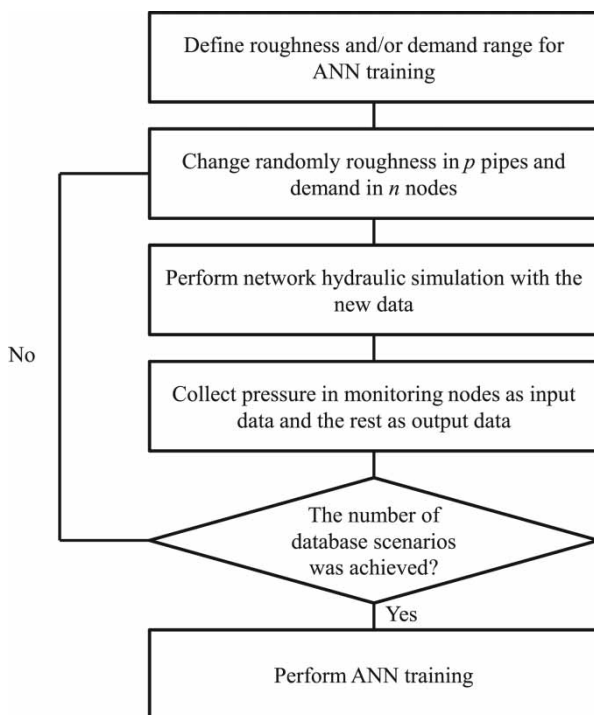


Figure 2 | ANN training flowchart.

ANN architecture definition

The success of ANN predictions requires finding the best architecture and model for the specific problem. Considering recent works, the ANN type used in this work is multilayer perceptron (MLP).

Once the main model was defined, preliminary tests were made to evaluate the use of hidden layers. In these tests, the number of neurons in the hidden layer was increased and the performance of the network was analyzed by the mean squared error. The tests generated several ANN architectures, changing not only the number of hidden layers but also the number of neurons for each hidden layer. For all cases, the activation function used was sigmoidal logarithmic. To obtain the synaptic weights at the training process, the conjugate gradient scaled optimization technique, developed by Moller (1993), was used. This algorithm combines the confidence region model used in the Levenberg–Marquardt method with the conjugate gradient technique, which reduces the processing time, avoiding the one-dimensional search process.

When the number of hidden layers increases, the computational time to process more synaptic weights also increases; however, the results do not show significant improvement. In this way, only one hidden layer with ten neurons was used. For each network studied, the number of neurons of the input and output layers was changed according to the number of monitored points and the total nodes.

In addition, for the extended period, one neuron was added in the input layer due to the flow monitoring. The use of these extra data is necessary to reduce the uncertainty of pressure estimation, since the knowledge of input flow restricts the pressure values on each node to a certain range related to the input flow. Table 1 shows these values for each case study.

For the ANN training in steady-state conditions, a database with 50,000 different scenarios was created and for the extended period 2,000 days, which represents 48,000 different scenarios. The roughness range selected to create these pressure and flow databases was 0.05 to 0.3 mm, which is the literature value for new and used cast iron pipes, respectively (Azevedo Netto 1998). The demand change was done using a multiplier factor in the range of 0.4 to 1.6. This

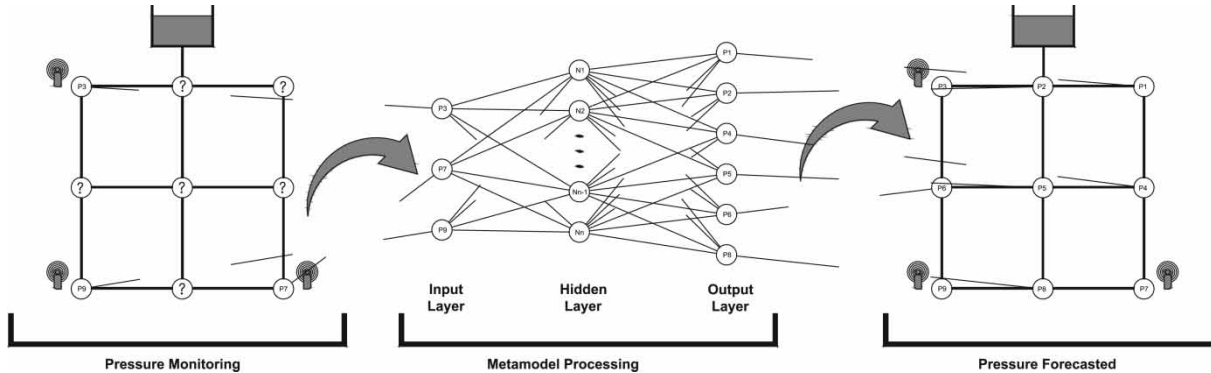


Figure 3 | Example of a trained ANN result.

Table 1 | Input and output layers for the ANN

Network	State	Monitoring nodes	Neurons on the input layer	Neurons on the output layer
Campos do Conde II	Steady	3	3	118
	Extended	3	96	2,832
CambuÍ	Steady – simulated	4	4	154
	Steady – monitored	4	4	4

factor is applied to reproduce a usual demand oscillation during a day (Arunkumar & Mariappan 2011).

To evaluate the ANN performance, a new database consisting of 1,000 scenarios for the steady-state case and 100 days for the extensive period was created. Therefore, the average and maximum errors (difference between estimated and real pressure) were obtained, and the standard deviation and the number of occasions that a node presented an error above 1 m (usual accuracy for pressure transducers used in water supply systems) for each situation.

CASE STUDY

The methodology presented in this work is applied in two real topologies. In order to verify the performance of ANN using different databases, the first network, Campos do Conde II, has the metamodel developed in three stages. The first stage fixed the nodal demand and varied pipes' roughness, generating the labeled database for pressures. At the second stage, the roughness is fixed and nodal demand varies, generating a new database. Finally,

roughness and demand vary randomly to generate a third database. The metamodel's tests at the second network are made in two stages. The first stage evaluates the randomly data generation, varying demand and roughness together, while the second stage uses real field measurement from Goulart (2015) to train an ANN.

Campos do Conde II

Campos do Conde II network is part of the water supply system of Piracicaba/SP, a medium-sized city in Brazil. It has 153 pipes, 121 nodes, and a reservoir. This is an extremely new network, with no consumers. Therefore, fictitious demands were generated to mobilize the system for the hydraulic analysis, creating different pressure zones. Considering a standard scenario with constant roughness and demands, the pressure on the system varies from 40.86 m to 87.38 m. Also, the average pressure amplitude in each node is 8.0 m, varying from the minimum of 0.16 m to the maximum of 18.4 m. Figure 4 shows a schematic of the network.

CambuÍ

This network represents a part of CambuÍ/MG system, a small Brazilian city. It has 167 pipes, 158 nodes, and a reservoir. This network was studied by Goulart (2015) for pipe roughness calibration. This system presents a pressure variation from 27.93 m to 86.26 m, and the average pressure amplitude in each node is 47.3 m, with the minimum of 27.4 m and the maximum of 77.3 m. Therefore, real data

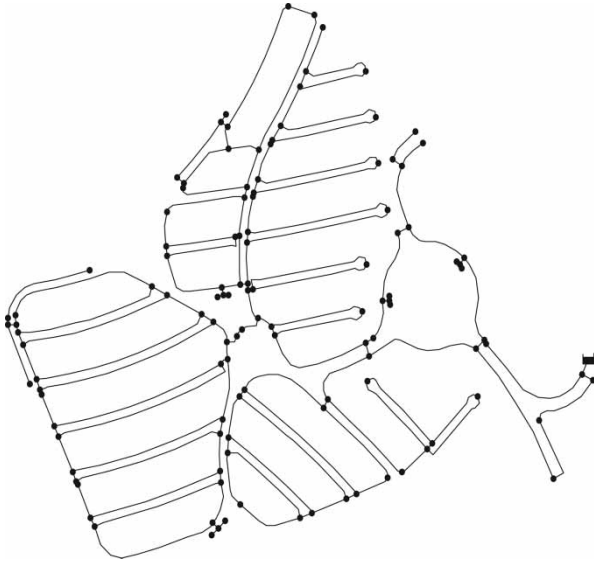


Figure 4 | Campos do Conde II network and pressure sensors' location.

from eight nodes are available to evaluate ANN performance. [Figure 5](#) shows the location of these points in the network.

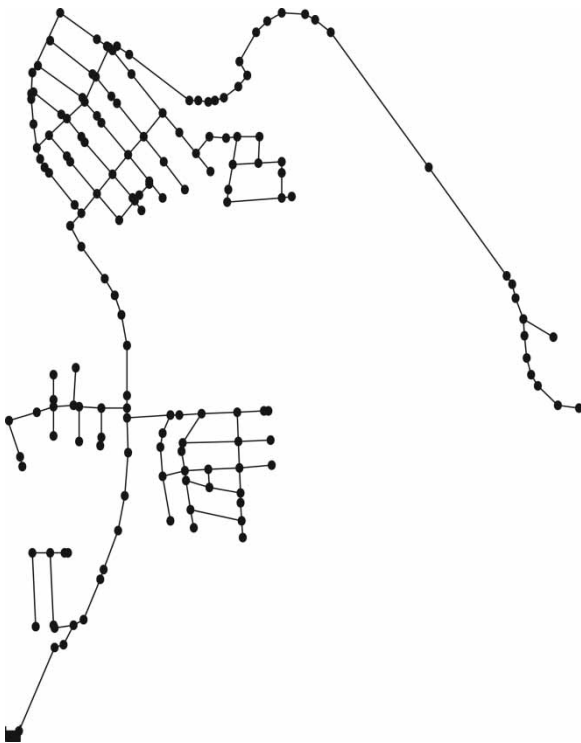


Figure 5 | Cambuí network and pressure sensors' location (Goulart 2015).

RESULTS

Campos do Conde II

First, an ANN was trained for steady-state conditions varying pipe roughness for database creation. Three nodes were used for monitoring (31, 72, and 97), which corresponds to 2.5% of the total, similar values usually observed in practice. [Figure 6](#) shows the average relative error in each node.

Despite the lower level of network monitoring, the results of estimated pressure show good agreement. Due to low velocities, the sensitivity of headloss with roughness variation is reduced, increasing the ANN accuracy. This point can be observed when nodes' demand are used for database creation. From [Figure 7](#) an increase in average error can be observed. This behavior is expected since a slight increment of velocity at some pipes increases headloss sensitivity. Furthermore, only 16 nodes have consumption (13.2% of total), creating scenarios where the total inflow is concentrated in some locations, making ANN training more difficult due to high pressure drops.

Finally, pipe roughness and nodes' demand are combined to create a new database. As expected, the error in estimated pressure increased ([Figure 8](#)), but the results are satisfactory, considering the uncertainty level existing in real pressure sensors.

[Table 2](#) and [Figure 9](#) summarize the results obtained for each database used for ANN training in the Campos do Conde II network.

For the extended period analysis, the same monitoring nodes were used, adding the inlet flow as input in ANN. [Figure 10](#) shows the results obtained, once again, with a low error. These results confirm the feasibility of using ANN for pressure estimation in near real-time operation, where the online information of pressure sensors would be used as input for ANN, which would respond to the pressure state of the network, allowing the identification of anomalies resulting from pipe bursts or pump problems, for example.

Cambuí

Considering the results obtained for Campos do Conde II, pipe roughness and node demand were used for the database creation, i.e., evaluating the worst possible scenario for ANN.

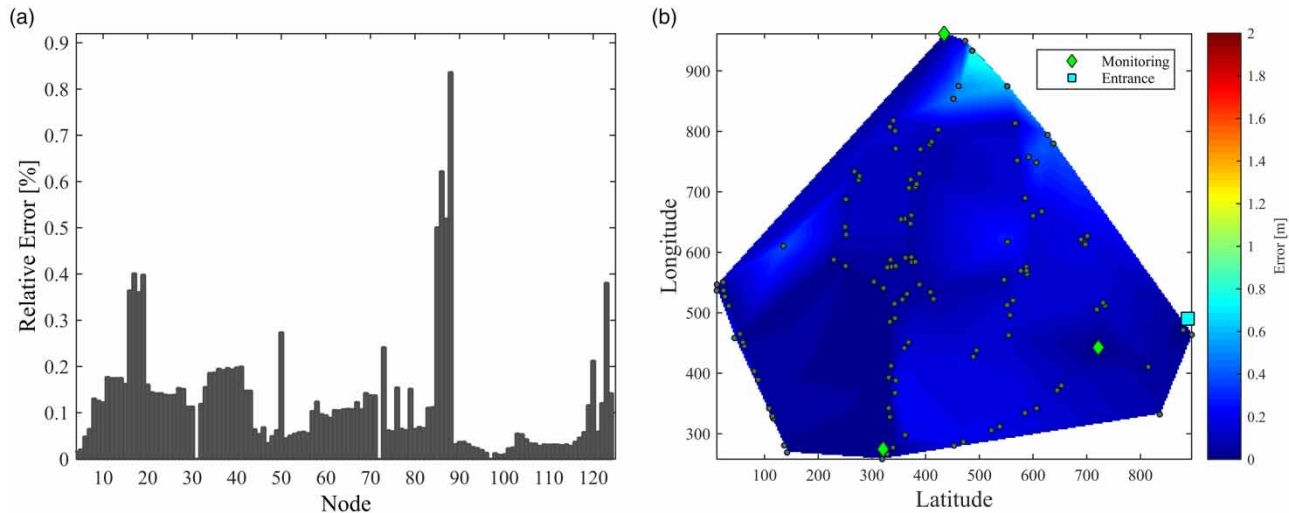


Figure 6 | Results for roughness database: (a) average error in each node; (b) error distribution.

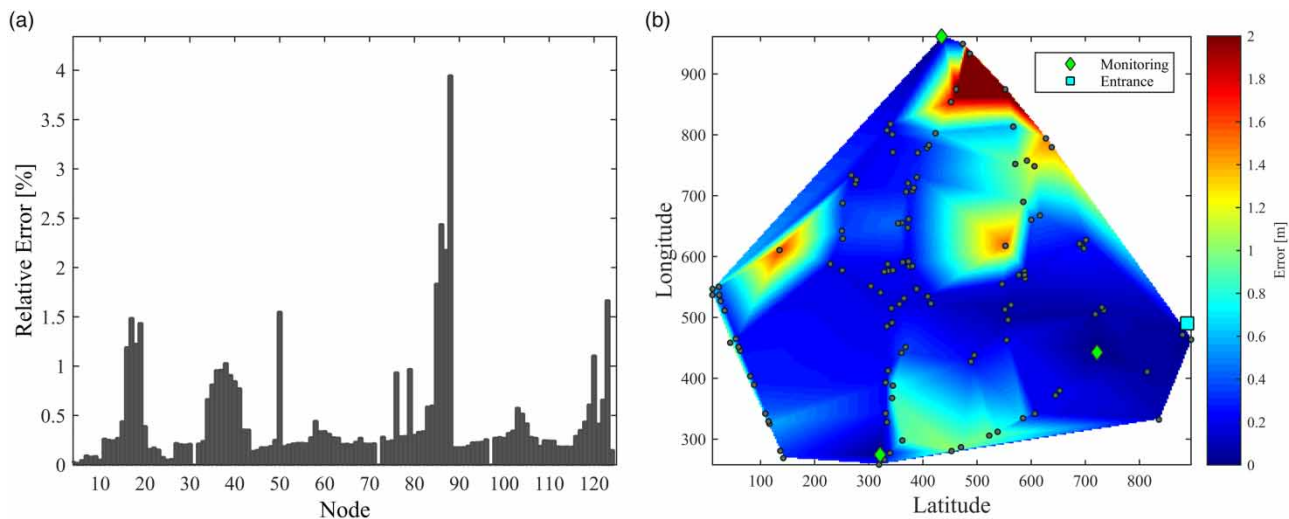


Figure 7 | Results for demand database: (a) average error in each node; (b) error distribution.

Four monitoring points were considered – nodes 7, 16, 46, and 96, one in each district of the sector. It can be observed from Figure 11 that ANN performance was better for this case, with an average error for estimated pressure of 0.105 m. This improvement can be explained due to a more uniform demand distribution, since all nodes have some consumption. Therefore, an increase in demand in one node has more chance to be balanced with a decrease in a neighbor node. Also, the absolute change in one node is very small, since the total inflow is uniformly distributed.

To validate the ANN, real data obtained from Goulart (2015) was used. The proposed methodology could not be

used, since the author obtained roughness values incompatible with the literature, reaching 10 mm, which indicates errors in network topology, as pipe diameter or valve closure were not known. Nevertheless, considering the eight monitoring points available, an ANN was trained to estimate pressure in four nodes – 11, 16, 51, and 110, using the data of the other four nodes. The database is composed of pressure measurements during more than 5 days, with a time interval of 1 minute, totaling 7,764 points. The first four days (6,300 points) were used to train the ANN, and the last day (1,464 points) was used to evaluate its performance. Figure 12 shows a good agreement between measured and estimated

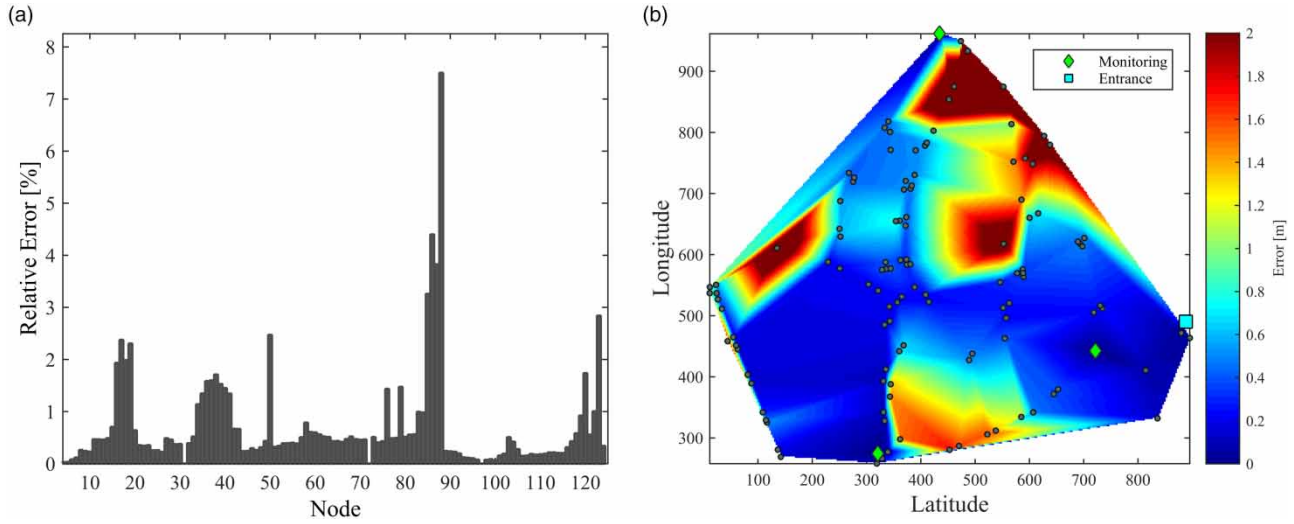


Figure 8 | Results for roughness and demand database: (a) average error in each node; (b) error distribution.

Table 2 | Results of ANN pressure forecast for the Campos do Conde II network

Data creation criteria	Steady-state		Extended period	
	Roughness	Demand	Roughness and demand	Roughness
Monitoring nodes	31, 72 and 97			
Average error [m]	0.066 (0.13%)	0.264 (0.45%)	0.392 (0.73%)	0.071 (0.05%)
Standard deviation [m]	0.054 (0.10%)	0.466 (0.80%)	0.719 (1.33%)	0.019 (0.04%)
Maximum error [m]	1.468 (0.84%)	1.752 (3.95%)	2.850 (7.50%)	0.523 (0.90%)
Network monitoring [%]	2.48			

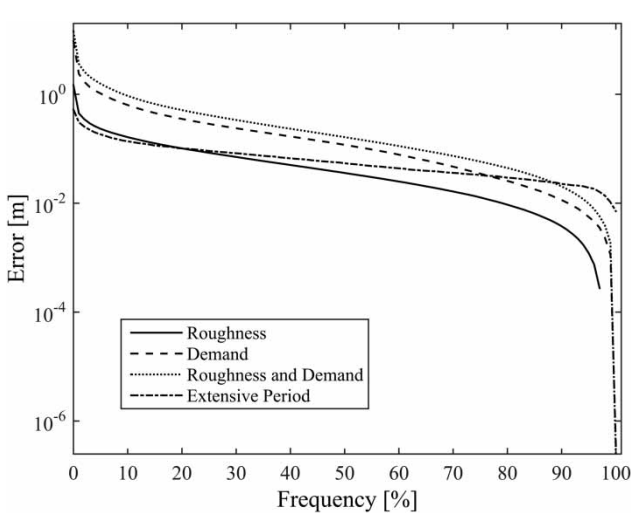


Figure 9 | Error frequency curve for the different scenarios of database creation.

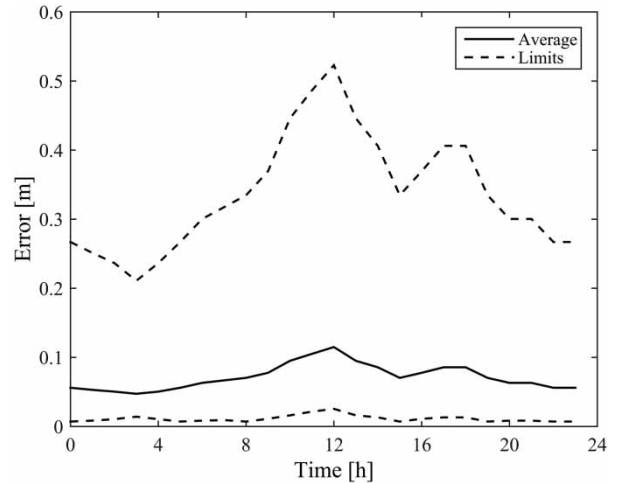


Figure 10 | Error of ANN pressure forecast in extended period for the Campos do Conde II network.

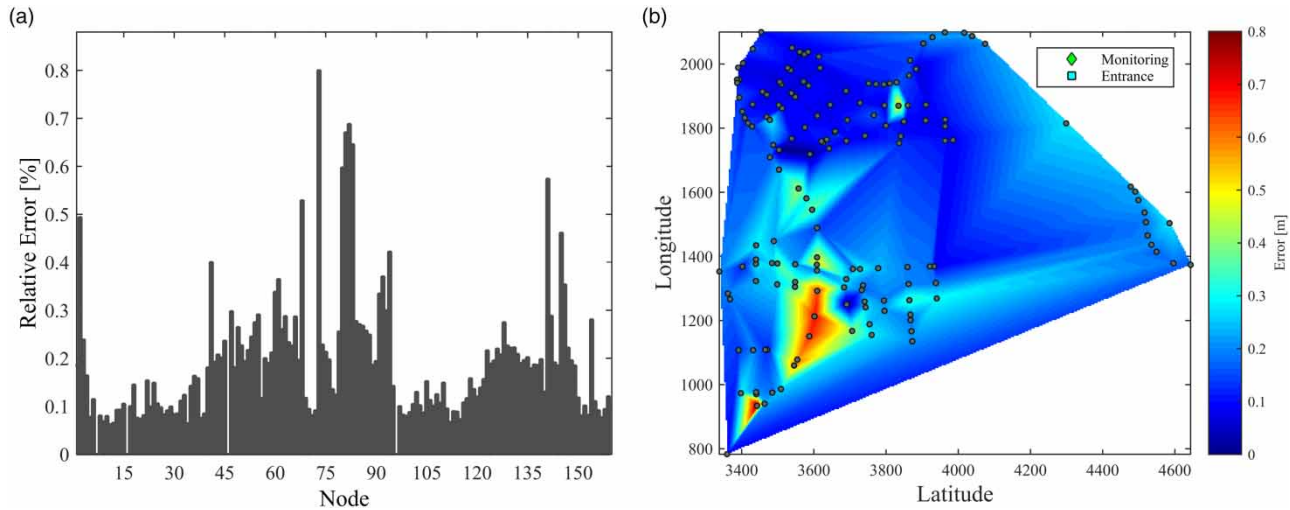


Figure 11 | Results for roughness and demand database: (a) average error in each node; (b) error distribution.

pressure, with only 5.4% of the database outside the trust region of ± 2 m defined by the [Water Research Centre \(1989\)](#).

However, if these points are closely observed, as shown in [Figure 13](#) for node 11 as an example, a sudden pressure change can be seen. This change is incompatible with the series behavior, indicating a measurement error. Despite

pressure sensors' uncertainty, the pressure signal behavior was maintained close to real curves, averaging an error of 0.72 m. This fact highlights the importance of a reliable database for ANN training. When accurately trained, if input data are not consistent with the real, ANN's response is very different from expected, indicating measurement

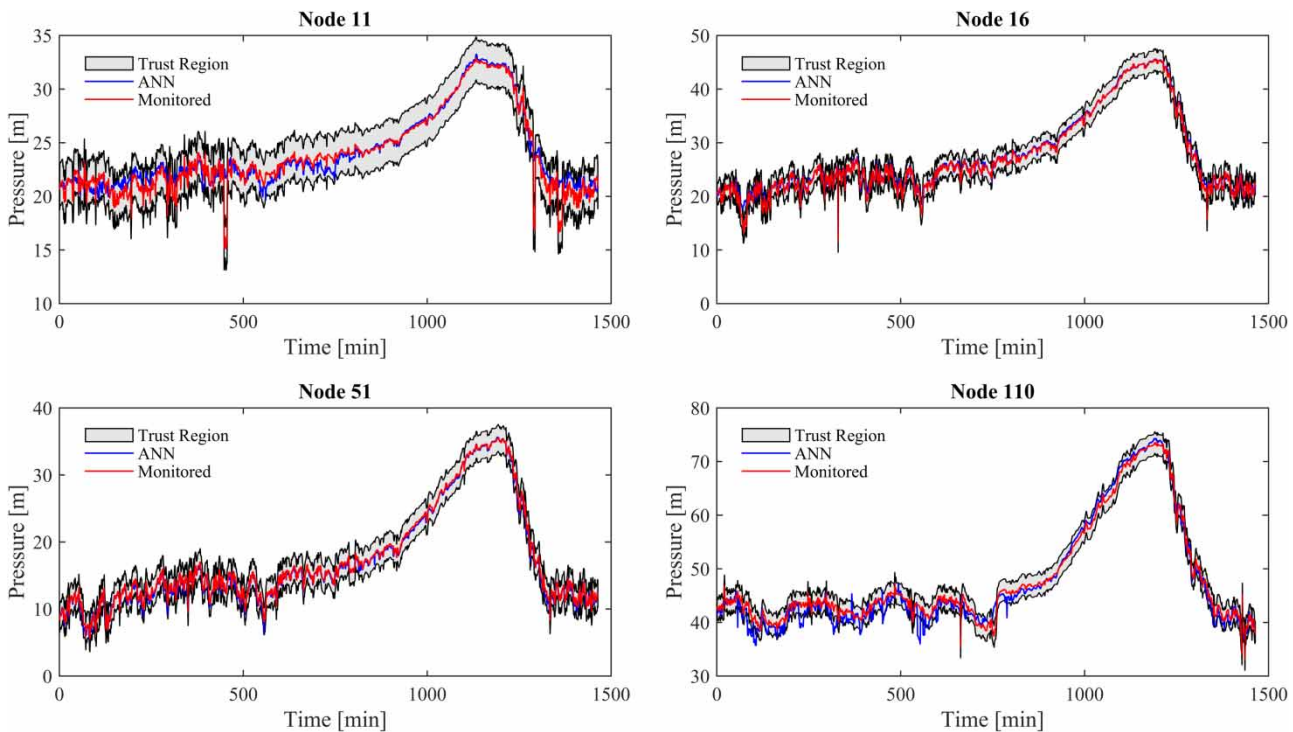


Figure 12 | Comparison between forecasted and measured pressure for the Cambuí network.

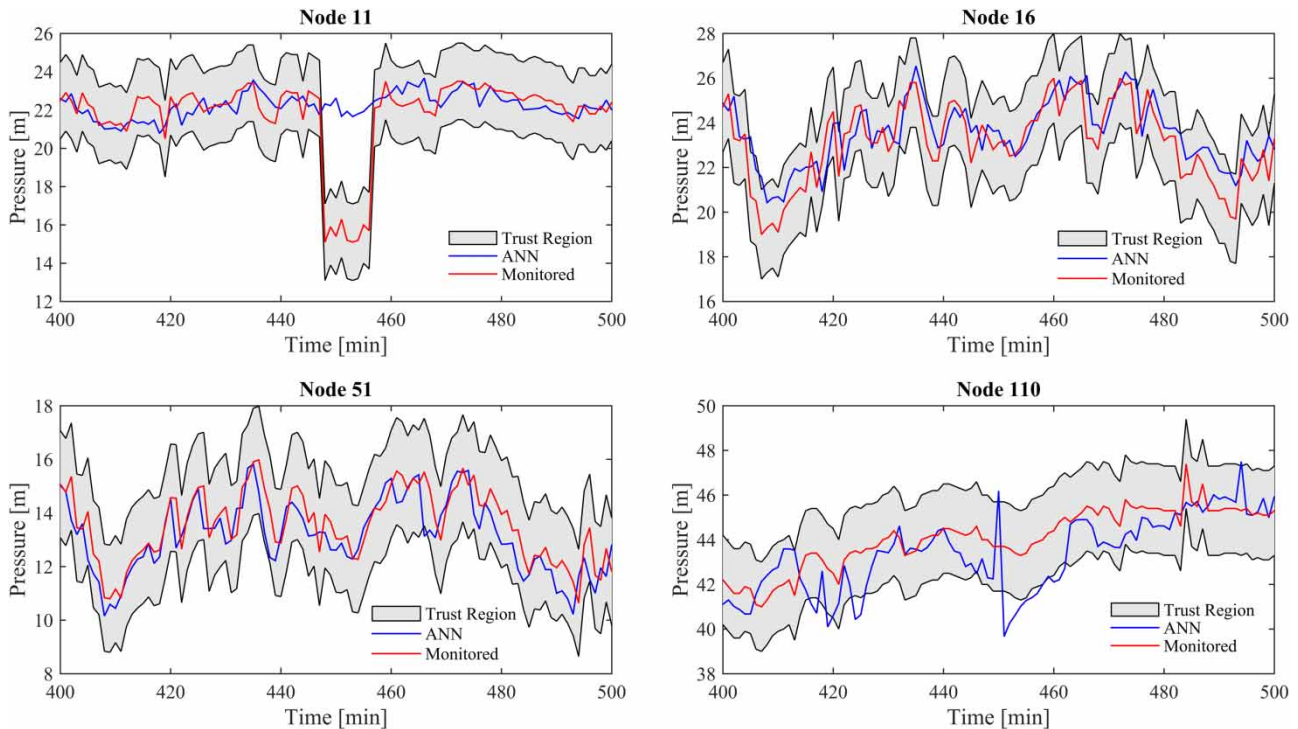


Figure 13 | Identification of pressure measurement failure.

errors. This prior identification prevents false alarms and unnecessary additional effort to identify the problem.

CONCLUSIONS

This paper evaluated the performance of ANN in pressure estimation for two different water supply networks, with different sizes and topologies. In general, the results were satisfactory both in the steady state and in the extended period, demonstrating the feasibility of the proposition. Therefore, if trained correctly, ANNs could be used to estimate pressure in near real-time operations, identifying zones with pressure problems, or in the calibration procedure, increasing the available samples, reducing the uncertainty in roughness definition. However, specific conditions were observed that must be evaluated carefully before using this tool, among them:

- There is no direct relationship between the number of monitoring nodes and the size of the network. This number depends largely on the topology and the system demand. It is expected that larger networks require a smaller percentage of monitored nodes.

- Using a wide range for roughness in ANN training causes more uncertainty in the process. Thus, the use of consistent lower and upper limits for the pipe material and lifetime is recommended.
- The selection of monitoring points should be done rigorously in order to obtain a more accurate estimation. It is recommended to use an optimization method to determine the optimal configuration of monitored nodes, minimizing the ANN training error.
- Nodes at dead-end pipes, such as 88, 120, and 123 of the Campos do Conde II network, produce larger errors, since they have only one flow path, increasing the sensitivity of roughness changes.
- Networks with low velocities have minor errors due to less susceptibility to headlosses. The velocity change between pipes can also hinder the ANN training.
- Monitored pressure values should be within the range used for ANN training. Otherwise, the roughness of the pipes or the demands of the nodes may result in unreliable values, indicating errors. Any change on the network, such as pipe substitution or installation of pumps and valves requires a new ANN training for the actual scenario.

- The use of a large database reduces the chance of pressure estimation errors, since the influence of one bad point is minimized by the other measurements.
- Although the use of hidden layers has not shown significant differences, previous tests are recommended with a reduced number of data, since the two networks studied present specific topologies.

REFERENCES

- Alvisi, S. & Franchini, M. 2010 Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics* **12** (4), 424–445.
- Arun Kumar, M. & Mariappan, V. E. N. 2011 Water demand analysis of municipal water supply using EPANET software. *International Journal on Applied Bioengineering* **5** (1), 9–19.
- Azevedo Netto, J. M. 1998 *Manual de Hidráulica*, 8th edn. Editora Edgard Blucher, São Paulo, Brazil, p. 670.
- Brentan, B. M. & Luvizotto Jr, E. 2014 Refining PSO applied to electric energy cost reduction in water pumping. *Water Research and Management* **4** (2), 19–30.
- Broad, D. R., Maier, H. R. & Dandy, G. C. 2010 Optimal operation of complex water distribution systems using metamodels. *Journal of Water Resources Planning and Management* **136** (4), 433–443.
- Cheng, W. & He, Z. 2010 Calibration of nodal demand in water distribution systems. *Journal of Water Resources Planning and Management* **137** (1), 31–40.
- Cordoba, G. A. C., Tuhovcáka, L. & Tausa, M. 2014 Using artificial neural network models to assess water quality in water distribution networks. In: *12th International Conference on Computing and Control for the Water Industry*, Perugia, Italy.
- Ding, S., Xu, X., Zhu, H., Wang, J. & Jin, F. 2011 Studies on optimization algorithms for some artificial neural networks based on genetic algorithm (GA). *Journal of Computers* **6** (5), 939–946.
- Giustolisi, O. & Berardi, L. 2011 Water distribution network calibration using enhanced GGA and topological analysis. *Journal of Hydroinformatics* **13** (4), 621–641.
- Goulart, T. D. C. 2015 *Studies for Calibration Algorithm Improvement and Application in Cambuí Water Distribution Network*. Master thesis, Universidade Federal de Itajubá – UNIFEI, Itajubá, MG, Brazil (in Portuguese).
- Ishido, Y. & Takahashi, S. 2014 Indicator for real-time leak detection in water distribution networks: design and simulation validation. In: *16th Conference on Water Distribution System Analysis – WDSA*, Bari, Italy.
- Khedr, A., Tolson, B. & Ziemann, S. 2015 Water distribution system calibration: manual versus optimization-based approach. *Procedia Engineering* **119**, 725–733.
- Kougias, I. P. & Theodossiou, N. P. 2013 Multiobjective pump scheduling optimization using harmony search algorithm (HSA) and polyphonic HSA. *Water Resources Management* **27** (5), 1249–1261.
- Moller, M. F. 1993 A scaled conjugate gradient algorithm for fast supervised learning. *Neural Network* **6**, 525–533.
- Montalvo, I., Izquierdo, J., Pérez-García, R. & Herrera, M. 2014 Water distribution system computer-aided design by agent swarm optimization. *Computer-Aided Civil and Infrastructure Engineering* **29** (6), 433–448.
- Mora-Melia, D., Iglesias-Rey, P. L., Martínez-Solano, F. J. & Ballesteros-Pérez, P. 2015 Efficiency of evolutionary algorithms in water network pipe sizing. *Water Resources Management* **26**, 4817–4831.
- Nazif, S., Karamouz, M., Tabesh, M. & Moridi, A. 2010 Pressure management model for urban water distribution networks. *Water Resources Management* **24**, 437–458.
- Ping, J., Wang, R., Sun, J. & Xiao, C. 2014 Pressure prediction of a water distribution network based on SVM. In: *ICPTT 2014 Creating Infrastructure for a Sustainable World*, Long Beach, CA, pp. 155–168.
- Price, E. & Ostfeld, A. 2014 Discrete pump scheduling and leakage control using linear programming for optimal operation of water distribution systems. *Journal of Hydraulic Engineering* **140** (6), 04014017.
- Rao, Z. & Alvarruiz, F. 2007 Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model. *Journal of Hydroinformatics* **9** (1), 15–24.
- Razavi, S., Tolson, B. A. & Burn, D. H. 2012 Review of surrogate modeling in water resources. *Water Resources Research* **48** (7), W07401.
- Roma, J., Pérez, R., Sanz, G. & Grau, S. 2015 Model calibration and leakage assessment applied to a real water distribution network. *Procedia Engineering* **119**, 603–612.
- Rossman, A. L. 2000 *EPANET 2.0 User's Manual*. Drinking Water Research Division, Risk Reduction Engineering Laboratory, US Environmental Protection Agency, Washington, DC.
- Sunela, M. I. & Puust, R. 2015 Real time water supply system hydraulic and quality modeling – a case study. In: *13th Computer Control for Water Industry Conference*, Leicester, UK.
- Tiwari, M. K. & Adamowski, J. F. 2015 An ensemble wavelet bootstrap machine learning approach to water demand forecasting: a case study in the city of Calgary, Canada. *Urban Water Journal* **14**, 185–201.
- Von Zuben, F. J. 1996 Parametric and non-parametric models of artificial neural networks and applications. Master Thesis, UNICAMP, Brazil (in Portuguese).
- WRC – Water Research Centre 1989 *Network Analysis – A Code for Practice*. WRC, Swindon.
- Yoo, D. G., Kang, D. & Kim, J. H. 2016 Optimal design of water supply networks for enhancing seismic reliability. *Reliability Engineering & System Safety* **146**, 79–88.

First received 2 March 2017; accepted in revised form 20 August 2017. Available online 8 December 2017